

The Impact of Culture Time on Genistin and Genistein Accumulation in *Flemingia macrophylla* (Willd.) Prain Seedlings and the Differentially Expressed Genes

Qiao Ling Yu^{1,2,3} , Guang Xiu Wu^{3,4}, Kai Wen Tan^{3,4}, Ying Liang^{2,3}, Shuang Shuang Qin^{2,3}, Fan Wei³, Jin Yuan Cai⁵, Ling Yun Chen⁴, Guo Yong Xie⁴, Ming Zhai⁶ and Kun Hua Wei^{1,2,3,7}

Abstract

Background: Genistein and genistin are the main active constituents of *Flemingia macrophylla* (Willd.) Prain and the amount of active constituents in the roots varies depending on the cultivation time. The harvesting period has a great influence on its quality, and there are fewer studies on the optimal harvesting time of *F. macrophylla*, which is urgently needed to explore the optimal harvesting period to improve its use value.

Objectives: *F. macrophylla* (Willd.) Prain is an important medicinal plant distributed in China. This study aimed to enhance the accumulation of active ingredients genistin and genistein in *F. macrophylla*, and explore the mechanism of biological synthesis of these active compounds at molecular levels.

Materials and Methods: Transcriptome sequencing and high-performance liquid chromatography (HPLC) technology were applied to *F. macrophylla* tissue culture seedlings for 30 days after culturing (DAC), 45 DAC, 60 DAC, and 75 DAC. Then, gene expression, representative active substances, and important phenotypic indexes were analyzed.

Results: The highest fresh weight and root thickness were attained at 45 DAC, the highest plant height at 60 DAC, and the maximum rooting rate at 75 DAC. The genistin content reached the highest level at 45 DAC, and the content of genistein reached the highest level at 30 DAC. The total content of genistin and genistein achieved the maximum levels at 45 DAC. A total of 7,735 differently expressed genes (DEGs) in the three comparable groups, including T45d vs T30d, T60d vs T30d, and T60d vs T45d, were identified. Through the Kyoto Encyclopedia Genes and Genomes (KEGG) enrichment analysis of DEGs, isoflavonoid biosynthesis and phenylpropanoid biosynthesis pathways were significantly enriched in two groups, including T45d vs T30d and T60d vs T30d. A total of 41 DEGs involved in the production of isoflavones were discovered in the three groups.

Conclusion: The optimum culture time for the accumulation of genistein and genistin in *F. macrophylla* seedlings ranged from 30 to 45 days. Forty-one DEGs in the genistein and genistin synthesis pathways were found through transcriptomics analysis. This study sheds light on the biosynthesis of genistin and genistein in *F. macrophylla*.

Keywords

Flemingia macrophylla, cultural time, tissue culture seedling, phenotypic index, RNA-sequencing

Received 26 September 2024; accepted 06 November 2024

¹Key Laboratory of State Administration of Traditional Chinese Medicine for Production and Development of Cantonese Medicinal Materials/ Guangdong Engineering Research Center of Good Agricultural Practice and Comprehensive Development for Cantonese Medicinal Materials, School of Chinese Materia Medica, Guangdong Pharmaceutical University, Guangzhou, Guangdong, China

²Department of Pharmacology, School of Pharmacy, Guangxi Medical University, Nanning, Guangxi, China

³Guangxi Key Laboratory of Medicinal Resources Protection and Genetic Improvement/Guangxi Engineering Research Center of TCM Resource Intelligent Creation, Guangxi Botanical Garden of Medicinal Plants, Nanning, Guangxi, China

⁴Department of Resources Science of Traditional Chinese Medicine, School of Traditional Chinese Pharmacy, China Pharmaceutical University, Nanjing, Jiangsu, China

⁵School of Food and Chemical Engineering, Liuzhou Institute of Technology, Liuzhou, Guangxi, China

⁶Department of Pharmacy, Medical College, Jiaying University, Meizhou, Guangdong, China

⁷Yunfu Traditional Chinese Medicine Resource Germplasm Bank Management Center, Yunfu, Guangdong, China

Corresponding author:

Kun Hua Wei, School of Traditional Chinese Medicine Resources, Guangdong Pharmaceutical University, Guangzhou, Guangdong 510006, China.

E-mail: divinekh@163.com



Introduction

Flemingia macrophylla (Willd.) Prain is a subshrub or shrub widely distributed in Yunnan, Guizhou, Guangdong, Hainan, Guangxi (China), and other South Asian regions, such as India, Vietnam, and Malaysia. The species belongs to Papilionoideae, a subfamily of the Fabaceae. As a Chinese herbal medicine, the dried roots and stems of *F. macrophylla* contain flavonoids, saponins, anthraquinones, triterpenes, and volatile and are used for the treatment of rheumatic bone pain, lumbar muscle strain, menstrual irregularities, dysmenorrhea, and leucorrhea, and other gynecological diseases (EBCMM, 1999). Flavonoids are the primary individual compounds in the *F. macrophylla*. For example, genistin and genistein possess potent biological activities, including antioxidant, hormone-like, antibacterial, anticancer, and cardiovascular protection effects (Kato et al., 2000; Lee et al., 2005; Li et al., 2014; Ruiz-Larrea et al., 1997; Weng et al., 2016; Yanagihara et al., 1993; Zhang et al., 2017; Zhu et al., 2018). Therefore, *F. macrophylla* is the main raw material of Chinese patent medicines such as “Fuke Qianjin tablet” and “Fuke Qianjin capsule,” and so on.

Genistein and genistin are the active components of soybean isoflavones; both belong to phytoestrogen in structure, and both have a variety of pharmacological effects (Glazier & Bowman, 2001; Zhang et al., 2008). At the same time, both have advantages, including abundant sources, stable nature, low side effects, and high medicinal value. The best way to obtain genistein is to extract it from plants, such as soybean, clover, pueraria, and other plants (Zhang et al., 2008). The isoflavones found in soybeans with the highest quantity and highest action are genistein and its glycoside, genistin (Zhang et al., 2008). The 4',5,7-trioxisoflavone-7-glycoside, or the 7-position glucoside of genistein, is the chemical name for genistin (Zhang et al., 2017). The chemical structure of genistein is similar to the mammalian estrogen 17 β -estradiol in spatial arrangement, and this structure makes genistein have estrogen-like effects, which can be combined with the estrogen receptor *in vivo* (Glazier & Bowman, 2001). This also determines its pharmacological activity, such as improvement of menopausal syndrome, antitumor, and improvement of metabolic syndrome (Yu et al., 2010).

The great therapeutic value of *F. macrophylla* and increasing market demands lead to excessive mining and a decline in natural resources. Although cultivation systems have been established in Guangxi, Yunnan, and other regions, the production of *F. macrophylla* has not yet met the market demand due to the long growth cycle, low yield, difficult mining, and the low market price of the medicinal materials, as well as the low enthusiasm of pharmaceutical farmers and the significant difficulties in promotion (Zhao, 2002). Plant tissue culture technology has the advantages of controllable culture conditions, high reproduction efficiency, and fast growth rate, supplying a measure to solve the problem of shortage of medicinal

plant resources efficiently and quickly. At the same time, it is crucial for us to increase the yield and quality of *F. macrophylla*. The concentration of flavonoids in *F. macrophylla* is somewhat influenced by the time of harvest. The content of these three components in the roots of the *F. macrophylla* is highest in the 6.5–7.5 years old *F. macrophylla*, but the 3.5-year-old *F. macrophylla* is more suitable for harvesting when taking planting costs into account, according to the results of Niu Yingfeng's study. Genistein, genistin, and soybean glycosides are primarily distributed in the roots of *F. macrophylla*, with the highest accumulation occurring in October and the lowest accumulation occurring in February (Wu et al., 2023). However, 3.5 years is a better harvesting window when planting costs are taken into account.

Transcriptome sequencing is currently an emerging technology for molecular-level research; the results inferred from transcriptome can accurately reflect the gene expression pattern, facilitate the study of active substance biosynthesis and regulation mechanisms such as plant development, and support molecular breeding and planting of the plant, among other uses (Niu et al., 2016). In order to create a theoretical framework for fruit development, Sun et al. (2022) undertook transcriptome sequencing analysis of green ripening fruit and ripening fruit, created the blue fruit transcriptome database, and investigated the functional genes and regulatory networks of the fruit in various developmental stages. Yan et al. (2023) studied the transcriptome data of fruits (seed and pericarp) of 80, 110, and 140 days betel nut (*Areca catechu* L.), and screened the differently expressed genes (DEGs) closely related to the biosynthesis of the main active substances of betel nut. A large amount of gene information related to the synthesis of each pathway was preliminarily obtained, which provides a theoretical basis for the functional identification of the betel nut gene, the analysis of the secondary metabolic pathway, and the study of the regulation mechanism. However, transcriptome sequencing for *F. macrophylla* has not been reported.

In this study, we aimed to enhance the content of the active components in *F. macrophylla* through the evaluation of different culture times. Additionally, we conducted transcriptome sequencing and differential gene expression analyses to explore the molecular mechanisms underlying the changes in genistein and genistin content.

Materials and Methods

Plant Materials and Culture Condition

Healthy shoots were transferred to MS medium supplemented with 1.5 mg/L 6-benzyl adenine (6-BA), 0.5 mg/L indoleacetic acid (IAA), and 0.2 mg/L Kinetin (KT). The culture conditions were 25°C, 50% humidity, 1,000 lx of light, and 12 h/d of light. All indexes were measured after 30, 45, 60, and 75 days after culturing (DAC) separately.

Determination of Plant Growth and Biomass

The plant height, rooting rate, root length, fresh weight, and leaf number were measured at four harvest stages, and the fresh and dry weight of the root was weighed. The rest part was further divided into two groups, one was frozen in liquid nitrogen and then saved in an ultra-low temperature refrigerator at -80°C that was used for RNA-sequencing (RNA-Seq) analysis, and the other part was dried in an oven at 40°C and crushed with mortars, and the powder was sieved through a 60-mesh screen and be used to the determination of genistein and genistin contents.

Determination of the Content of Genistein and Genistin

Sample Pretreatment

The powder of the seedling roots was accurately weighed into a 50 mL stoppered flask, into which then 10 mL of methanol was added and soaked at room temperature for 30 min; then, the samples were extracted for 30 min by using an ultrasonic extraction. After the mixture was further filtered and evaporated, it was redissolved with 1 mL of chromatographic methanol, filtered through a $0.22\ \mu\text{m}$ microporous membrane, and injected into a 1.5 mL sample vial.

Chromatographic Condition

The contents of genistin and genistein in the roots of *F. macrophylla* seedlings at different culture times were detected by using high-performance liquid chromatography (HPLC). A Stable Bond Analytical C_{18} ($4.6\ \text{mm} \times 250\ \text{mm}$, $5\ \mu\text{m}$) column was used with methanol as mobile phase A and 0.1% phosphoric acid water as mobile phase B and each sample was repeated three times for gradient elution (0–35 min A: 25–65%, B: 35–75%, 35–37 min A: 65–35%, B: 35–75%, 37–40 min A: 25%, B: 75%). The contents of genistein and genistin were calculated by using the regression equation of the relevant standards.

Preparation of the Standard Solutions

A total of 2.78 mg and 2.34 mg of genistin and genistein control were accurately weighed, respectively, and then dissolved

in chromatographic methanol and diluted with chromatographic methanol to 10 mL.

Building of Standard Curves for Genistein and Genistin

The solutions of genistein and genistin standard were taken, and the samples were injected according to Table 1. For linear regression with the standard solution concentration as the abscissa and the corresponding peak area as the ordinate, the regression equation was shown in Table 2, and each component had a good linear relationship.

Determining Genistein and Genistin Contents

The peak area values of genistein and genistin were determined in each treatment group, and their contents were calculated using the respective regression equations.

Transcriptome Sequencing

Extraction and Sequencing of RNA

Total RNA was extracted from different samples using a total RNA extraction kit, and the samples including the roots of *F. macrophylla* were cultured for 30, 45, and 60 days. Nanodrop 2000 and gel electrophoresis were employed to quantify and test the total RNA, and the RIN value was detected using Agilent 2100. When the samples passed the necessary steps, the libraries were built and sequenced as follows: (a) mRNA was isolated from the total RNA by using magnetic beads with Oligo (dT) and paired with A-T bases in ploy A. (b) mRNA was randomly broken by adding the fragmentation buffer and the small fragments of about 300 bp were isolated by screening with magnetic beads. (c) The small fragments of mRNA were used as a template and the single-stranded cDNA was synthesized by reverse transcriptase using six-base random primers (random hexamers). (d) The end-repair mix was added to flatten the sticky ends of the duplex's cDNA structure, and then an "A" base was added to the 3' end to join the Y-shaped joint. (e) Purpose bands were recovered with 2% agarose gel and 15 cycles of PCR amplification were performed. TBS380 (Picogreen) was quantified according to the

Table 1. Genistein and Genistin Label Production Table.

Sample Volume (μL)	V_1	V_2	V_3	V_4	V_5	V_6
Genistin	2	5	8	10	15	20
Genistein	1	2	4	6	15	20

Table 2. Table of Regression Equations for Genistein and Genistin.

Component	Regression Equation	Correlation Coefficient	Linearity Range (μg)
Genistin	$S_1 = 5E + 06X_1 - 35,006$	$R^2 = 0.9999$	0.556–4.170
Genistein	$S_2 = 9E + 06X_2 - 309,270$	$R^2 = 0.9999$	0.117–3.450

data proportion mixed and placed on the device for testing. Bridge-type PCR amplification was performed on the cBot to generate the clusters and the PE libraries were sequenced by the Illumina platform to get the length of 150 bp.

Bioinformatics and DEGs Analysis

To obtain the final transcript sequences, the clean reads were obtained by filtering the raw data and subjecting it to *de novo* assembled by using the reference technique (Grabherr M G) and Trinity software (Version v2.8.5) (<https://github.com/trinityrnaseq/trinityrnaseq/wiki>). After removing redundant transcripts, functional annotation of transcripts was performed by aligning to NR, Swiss-Prot, Pfam, Clusters of Orthologous Groups (COG), Gene Ontology (GO), and Kyoto Encyclopedia Genes and Genomes (KEGG) databases. The gene expression level was quantified by the FPKM (fragments per kilobase per million) using RSEM (Version 1.3.1) (<http://deweylab.biostat.wisc.edu/rsem/>) software. DEGs between two samples were identified by using edgeR analytic software (Version 3.24.3) (<http://bioconductor.org/packages/stats/bioc/edgeR/>) with the parameter $|\log_2FC| > 1$ and $p < 0.05$. The function of DEGs was annotated by aligning them to KEGG and GO databases.

Validation by qRT-PCR

The expression levels of nine DEGs were validated by using qRT-PCR. Primer Premier 5 (Zhai et al., 2008) were used to

design primers. The actin gene (gene_2356) was used as the endogenous control. Three biological replicates for each treatment were used. The qRT-PCR was performed in a 21.0 μ L reaction system (10 μ L ChamQ Universal SYBR qPCR Master Mix, 0.8 μ L Primer A and Primer B, 1 μ L Template cDNA, and 9.2 μ L ddH₂O). The reaction conditions for qRT-PCR included three phases: 95°C, 30 s, 1 cycle; 95°C, 10 s $\rightarrow$ 60°C, 30 s, 40 cycles; 95°C, 15 s $\rightarrow$ 60°C, 60 s $\rightarrow$ 95°C, 0.01 s, 1 cycle. Quant Studio 3 Real-Time PCR instrument (Thermo Fisher Scientific) was employed to monitor qRT-PCR and the $2^{-\Delta\Delta Ct}$ method was used to calculate cycle threshold values and obtain curves. The primers are listed in Table 3.

Data Analyses

The experimental data were analyzed using Excel 2019 and SPSS 25 statistical software.

Results

Effects of Culture Time on Growth Indexes and Phenotypic Indexes of Tissue Culture Seedlings

F. macrophylla seedlings were collected at four culture times, T30d, T45d, T60d, and T75d (Figure 1). Different culture times were not significant for differences in phenotypic indexes, but the seedlings grew higher with more fibrous

Table 3. Primer Information Table for Quantitative Reverse Transcription Polymerase Chain Reaction (qRT-PCR).

Gene	Primer Base Sequences (5'–3')
gene_2356S	TAGCCCGCATTCGCCCTAAGAT
gene_2356A	CTGTTGGGTCTGTTGCCTGAAT
gene_2754S	TTGGTCACCCTCTTGAGCCTTAC
gene_2754A	GCTTCCTGTGATTGCCTTTCTC
gene_7205S	TGGAAGGCGGTGACTGAGGAAT
gene_7205A	CTTGACAGGCTGAACGGTGAT
gene_18009S	GCGTTGTTGACGAGGATGTTTC
gene_18009A	AGAGTAGTGCCTTGATGTTGGTGAG
gene_15924S	TTCCTCCTCGGGATAAGGTAGT
gene_15924A	CTCTTGGGTCATTGCTGCTTC
gene_20861S	TAGATGTGGTGGATGAGGATGAT
gene_20861A	AACAACAAAGGGTGGGACAAGG
gene_10598S	TGAAGTTCGTATTCCCAAGCAG
gene_10598A	AATAACCACCGCCAGACGAGAC
gene_4275S	GGACTCGACAATGGCTTTACC
gene_4275A	CCTCCCAACGACACCAACTCTT
gene_5858S	TCCGAACGCTATGGTCCTTTGT
gene_5858A	TCGTAAGTGAGACGCCTAATGG
gene_5859S	TCCGAACGCTATGGTCCTTTGT
gene_5859A	TCGTAAGTGAGACGCCTAATGG

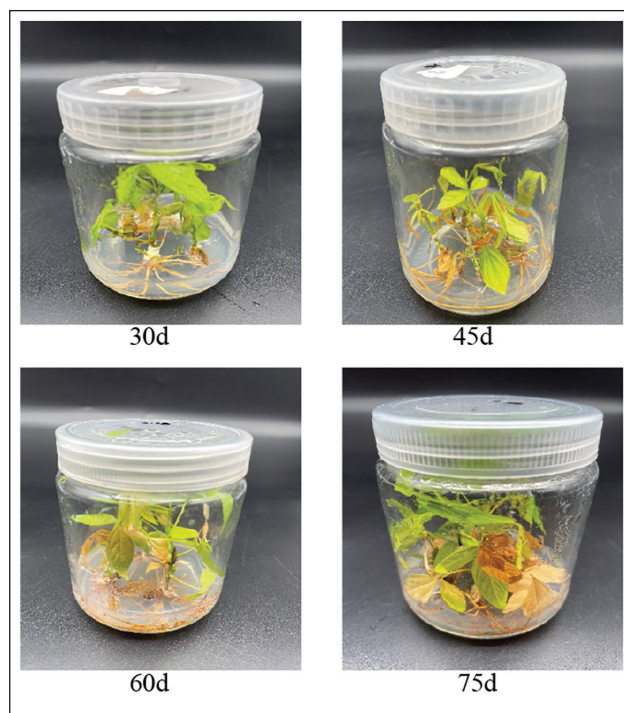


Figure 1. The Effects of Different Culture Times on the Growth Status of *Flemingia macrophylla*.

Table 4. Effect of Different Growth Times on Phenotypic Indexes of *Flemingia macrophylla*.

Growth Time (d)	Take Root Rate (%)	Root Fresh Heavy (g)	Root Dry Heavy (g)	Thick Root (mm)	The Average Root Length (cm)	Number of Leaves	Plant Height (cm)
30	76.35 ± 0.10a	1.34 ± 0.11a	0.24 ± 0.02a	1.00 ± 0.21a	11.24 ± 2.21a	134.67 ± 30.92a	2.57 ± 0.56a
45	76.29 ± 0.14a	1.76 ± 0.45a	0.33 ± 0.08a	1.28 ± 0.06a	10.06 ± 2.97a	103.67 ± 37.82a	2.62 ± 0.51a
60	86.49 ± 0.02a	1.69 ± 0.45a	0.28 ± 0.08a	0.98 ± 0.17a	14.33 ± 2.19a	142.67 ± 39.50a	4.04 ± 0.20a
75	89.02 ± 0.06a	1.74 ± 0.04a	0.34 ± 0.02a	0.92 ± 0.10a	15.32 ± 1.82a	128.33 ± 27.61a	3.95 ± 1.21a

Note: Each index is represented as mean ± standard deviation, $n = 3$. Different case letters in each group indicate significant differences in content between different treatments ($p < 0.05$).

roots and withered yellow leaves as the cultivation time increased (Figure 1). The highest rooting rate (89.02%) and the longest root length (15.32 cm) were observed in the treatment of 75 DAC. Roots' fresh weight and dry weight reached a maximum of 1.76 g and 0.34 g at 45 DAC and 75 DAC, respectively. The maximum root thickness and plant height were 1.28 mm and 4.04 cm after 45-day and 60-day cultures. At 60 DAC, the highest number of leaves was 142.67. Considering various phenotypic indexes comprehensively, it is appropriate to cultivate sterile seedlings in a culture flask for 30–45 days.

Changes of Genistein and Genistin Contents During Culturing of *F. macrophylla*

An HPLC was used to determine genistein and genistin contents in four samples. The highest genistin concentration of seedlings

was found at 45 DAC (8.65 mg/g), and the lowest content was at 30 DAC (5.72 mg/g) (Table 4). In contrast, the genistein content increased from 1.82 mg/g (75 DAC) to 3.11 mg/g (30 DAC), reaching its peak at 30 DAC (Table 5). Therefore, cultivation times significantly affected the accumulation of genistein and genistin in tissue culture seedlings. The total content of genistein and genistin reached the highest content of 10.93 mg/g at 45 DAC. The results also indicated that the *F. macrophylla* had little free isoflavones (genistein), and the isoflavones were mainly found in the form of glycoside (genistin).

Analysis of the Transcriptome Sequencing

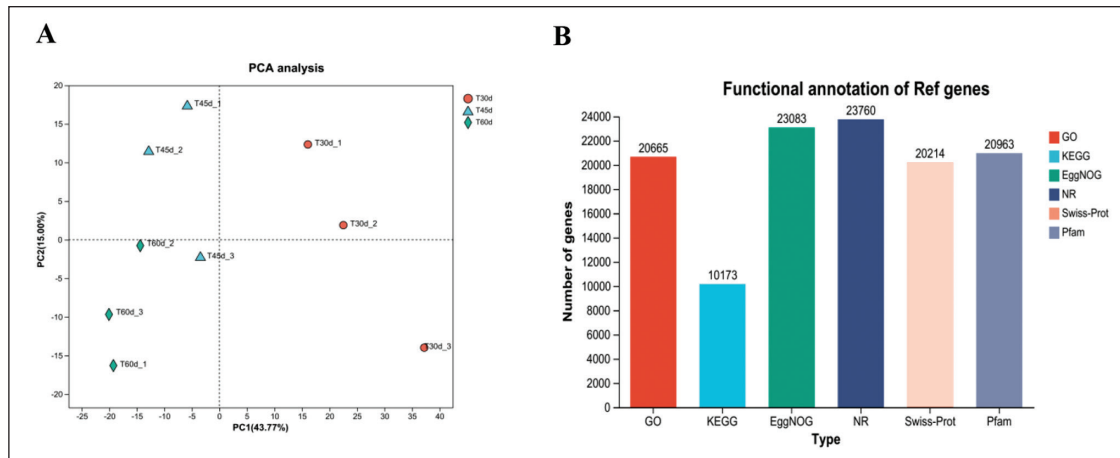
Transcriptome Sequencing Quality Analysis and Functional Annotation

Samples from 30 DAC, 45 DAC, and 60 DAC were selected for RNA-Seq. After filtering low-quality reads, a total of

Table 5. Effect of Different Growth Times on Isoflavone Content of *Flemingia macrophylla*.

Growth Time (d)	Genistin (mg/g)	Genistein (mg/g)	Total Content (mg/g)
30	5.72 ± 0.10d	3.11 ± 0.066a	8.83 ± 0.17b
45	8.65 ± 0.26a	2.28 ± 0.046b	10.93 ± 0.30a
60	6.12 ± 0.14c	2.03 ± 0.044c	8.14 ± 0.19d
75	7.14 ± 0.17b	1.82 ± 0.028d	8.96 ± 0.20b

Note: Each index is represented as mean ± standard deviation, $n = 3$. Different case letters in each group indicate significant differences in content between different treatments ($p < 0.05$).

**Figure 2.** Polymerase Chain Reaction (PCA) Analysis Chart (A), Distribution Map of the Functional Annotation (B).**Table 6.** Quality Statistics of the Transcriptome Sequencing Data.

Sample	Raw Reads	Clean Reads	Error Rate (%)	Q30 (%)	GC Content (%)
T30d-1	48,495,972	45,816,876	0.025	94.29	44.59
T30d-2	46,431,128	43,970,738	0.0251	94.17	45.16
T30d-3	49,187,150	46,687,174	0.0249	94.34	44.75
T45d-1	49,476,984	46,914,694	0.0251	94.19	44.82
T45d-2	43,285,128	41,342,906	0.0245	94.76	44.79
T45d-3	45,681,526	43,676,038	0.0251	94.16	44.69
T60d-1	44,677,434	42,842,656	0.0248	94.44	45.59
T60d-2	43,124,720	41,235,806	0.0249	94.33	44.45
T60d-3	43,819,664	42,067,836	0.0246	94.71	44.91

3,946,547,424 clean reads were obtained for the nine samples. The average error rate in nine samples was 0.0251%, the proportion of bases with sequencing quality above 99.9% (Q30) $\geq 94.16\%$ and GC was 44.45–45.59% (Table 6). Principal component analysis (PCA) revealed clear grouping for samples during various cultivation times (Figure 2A). A total of 26,058 genes were annotated using GO, KEGG, COG, NR, Swiss-Prot, and Pfam databases. Among them, the smallest number of genes could be annotated to the KEGG database, and the largest number could be annotated to the NR database (Figure 2B).

Identification and Verification of DEGs

In differential gene expression analyses, genes with $|\text{fold change}| > 2$ and $p\text{-adjust} < 0.05$ were treated as DEGs. A total of 4,318 DEGs were identified in T60d vs T30d, followed by T45d vs T30d (2,954) and T60d vs T45d (463) (Figure 3A). There were more downregulated DEGs than upregulated DEGs in three cultivation times. For example, 1,095 DEGs were upregulated and 3,223 downregulated for the 4,318 DEG (Figure 3A). Results also indicated that the number of genes involved in the growth regulation and active

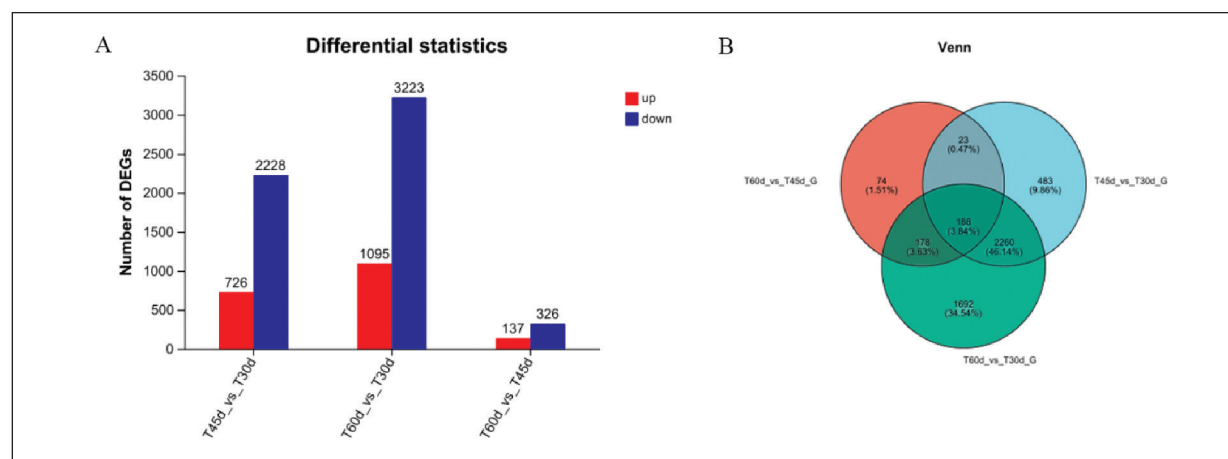


Figure 3. Statistics of the Number of Differential Genes (A), Venn Plot of Differently Expressed Genes (B).

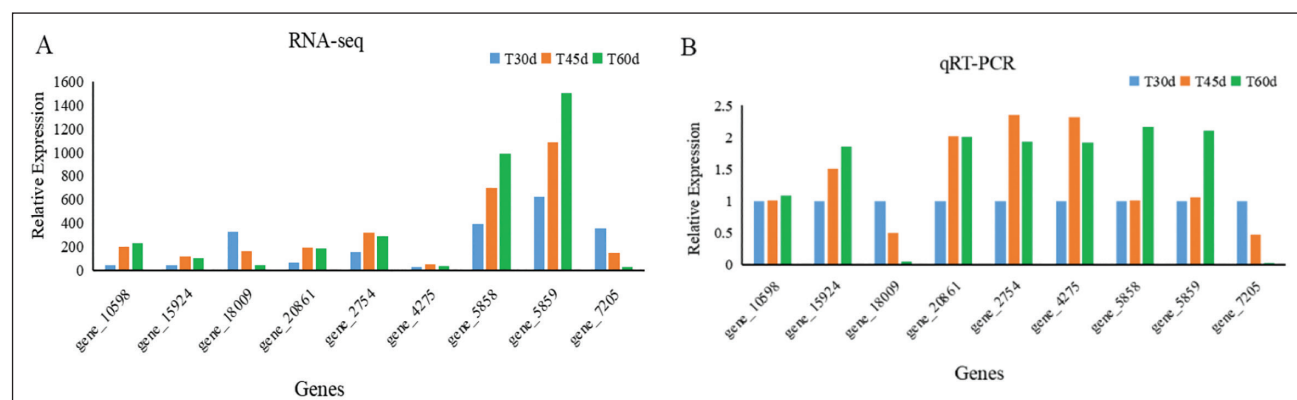


Figure 4. Bar Graph of the Relative Expression. (A) Relative Expression of RNA-Sequencing (RNA-seq), (B) Relative Expression of Quantitative Reverse Transcription Polymerase Chain Reaction (qRT-PCR).

component accumulation increased as the culture time increased. This implied that the accumulation of active ingredients in *F. macrophylla* might be through the negative regulation of genes to process of growth and development. The Venn diagram results revealed 188 shared DEGs among T45d vs T30d, T60d vs T30d, and T60d vs T45d, 2,260 shared DEGs between T45d vs T30d and T60d vs T30d, 178 shared DEGs between T60d vs T30d and T60d vs T45d, 23 shared DEGs between T60d vs T45d and T45d vs T30d (Figure 3B). These shared DEGs may contain key genes for the genistein and genistin synthesis pathways.

To verify expression level inferred from transcriptome, we performed qRT-PCR for nine candidate DEGs. Results of qRT-PCR indicated that all of these nine candidates displayed an expression pattern roughly in line with the transcriptome (Figure 4A), proving the validity of the transcriptome sequencing results (Figure 4B).

GO Function Annotation and Enrichment Analysis of the Differently Expressed Genes

GO functional annotation and enrichment analysis of DEGs could effectively reflect the biological functions of DEGs

during cultivation time. A total of 22,574 DEGs were annotated in three categories of the GO database (Figure 5A): biological process, cellular component, and molecular function. As a result, the number of DEGs annotated in the GO database of T60d vs T30d was at the top of the list, followed by T45d vs T30d and T45d vs T30d. Functional enrichment of DEGs was helpful for further mining and identification of functional genes. The GO functional enrichment analysis could reflect the primary biological functions of DEGs at various culture times. In our study, comparing the transcriptome data of tissue culture seedlings with different culture times, p -adjust < 0.05 was used as a condition and found 2,522 DEGs between T45d and T30d, enriched on 158 GO pathways, 3,686 DEGs between T60d and T30d, 179 GO pathways, 400 DEGs between T60d and T45d, and enriched on 31 GO pathways. GO enrichment analysis revealed that the top 20 terms were DNA replication initiation, regulation of cyclin-dependent protein serine/threonine kinase activity, regulation of protein kinase activity, phloem or xylem histogenesis, regulation of protein serine/threonine kinase activity, and other functions in T45d vs T30d group. The most enriched term was found in the hydrogen peroxide metabolic process and the hydrogen peroxide catabolic process. The majority of

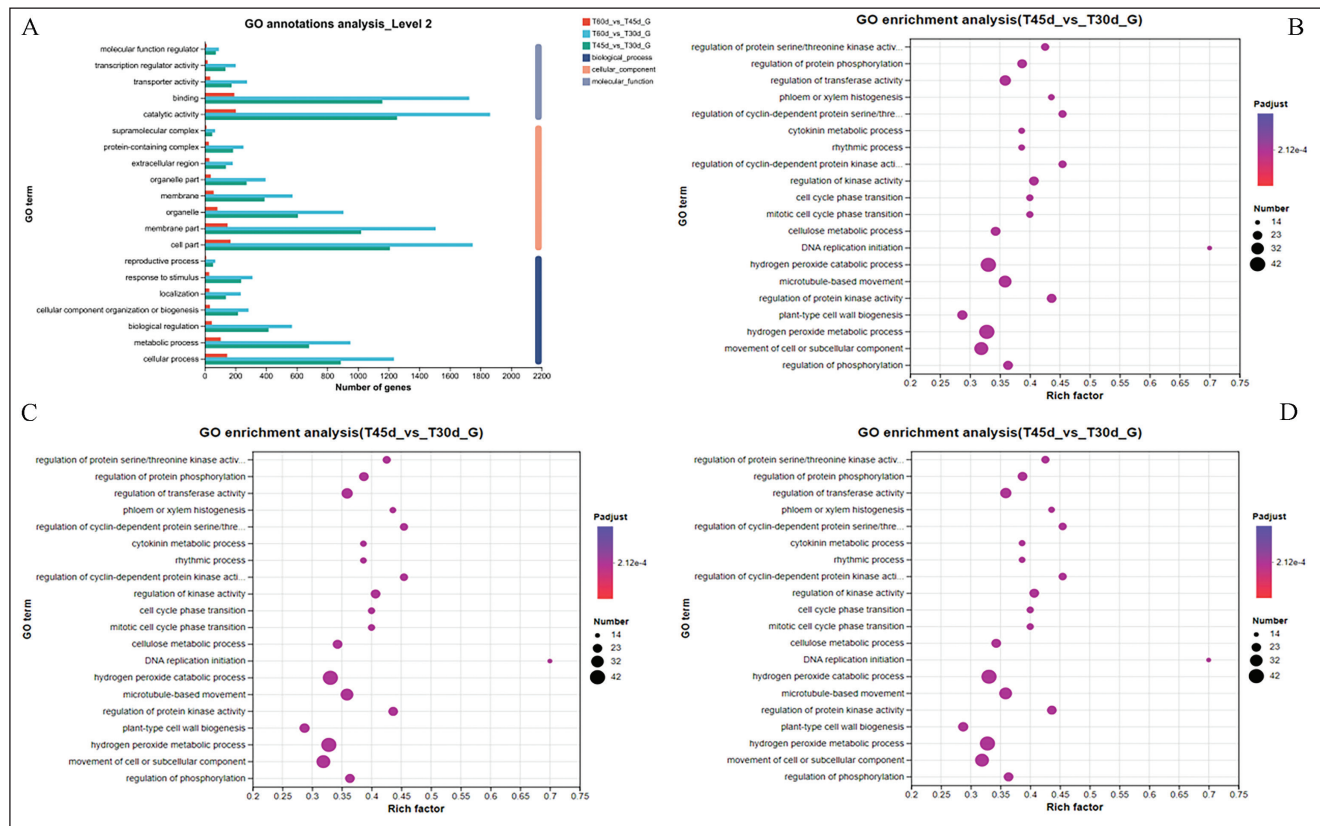


Figure 5. GO Functional Enrichment Analysis of Different Genes in Different Culture Times. (A) Gene Ontology (GO) Annotation Map, (B) GO Enrichment Analysis Bubble Map (T45d vs T30d), (C) GO Enrichment Analysis Bubble Map (T60d vs T30d), (D) GO Enrichment Analysis Bubble Map (T60d vs T45d).

DEGs in the T60d vs T30d group were mainly enriched in cell wall biogenesis, DNA replication initiation, regulation of cyclin-dependent protein kinase activity, regulation of cyclin-dependent protein serine/threonine kinase activity, regulation of protein kinase activity, and other functions. DEGs in T60d vs T45d group were primarily enriched in leucoanthocyanidin reductase activity, oxidoreductase activity, acting on CH or CH₂ groups, NAD or NADP as acceptor, momilactone-A synthase activity, phenol-containing compound metabolic process, and xyloglucan: xyloglucosyl transferase activity, among other functions (Figures 5B–5D).

KEGG Enrichment Analysis of the Differently Expressed Genes

KEGG enrichment analysis identified the top 20 enriched pathways in each group. Among these, 18 enriched KEGG pathways, including isoflavonoid biosynthesis, phenylpropanoid biosynthesis, ascorbate and alternate metabolism, starch and sucrose metabolism, flavonoid biosynthesis, photosynthesis-antenna proteins, tryptophan metabolism, and DNA replication were detected in T45d vs T30d group.

In T60d vs T30d group, phenylpropanoid biosynthesis, isoflavonoid biosynthesis, starch and sucrose metabolism,

flavonoid biosynthesis, sesquiterpenoid and triterpenoid biosynthesis, tryptophan metabolism, pyruvate metabolism, DNA replication, ABC transporters, cutin, suberine, and wax biosynthesis, cysteine and methionine metabolism, glycolysis/gluconeogenesis and monoterpene biosynthesis were detected. However, only three significantly enhanced pathways were identified in the T60d and T45d groups, including fatty acid elongation, phenylpropanoid biosynthesis, and flavonoid biosynthesis. KEGG enrichment analysis revealed that isoflavonoid biosynthesis was identified in T45d vs T30d and T60d vs T30d, whereas phenylpropanoid biosynthesis was not significantly enriched in T60d vs T45d. These results indicated that the genes regulated isoflavone synthesis and accumulated active components were activated within 45 DAC, and reminded that the best period for tissue culture seedling collection was at 45 DAC (Figure 6).

Key DEGs Related Closely to Biosynthesis Pathways of the Genistein and Genistin

Genistein and genistin are crucial active components in *F. macrophylla*. The DEGs aligned into KEGG pathways related to biosynthetic pathways of genistein and genistin were significantly enumerated, but the accumulation of their active

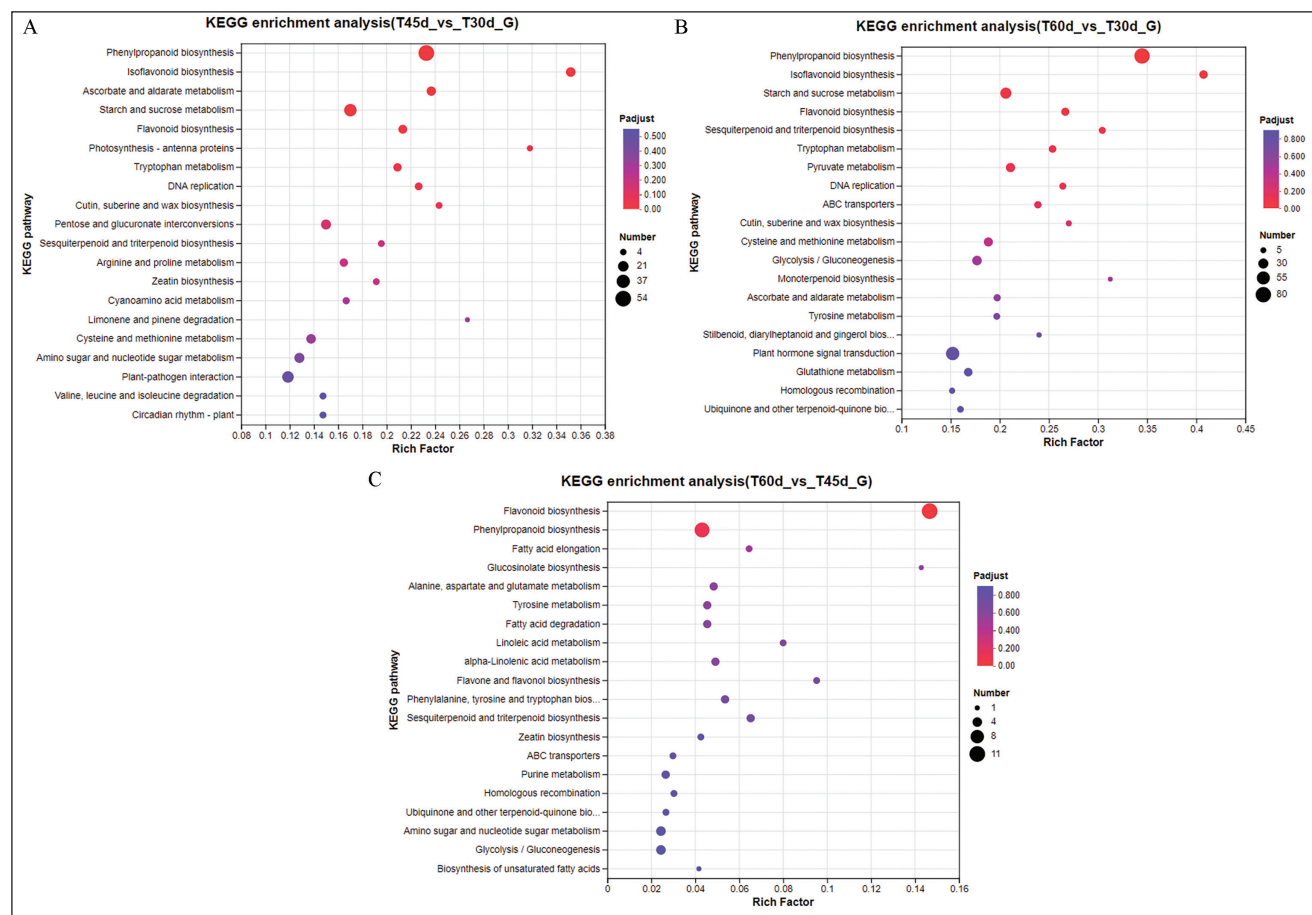


Figure 6. Bubble Plot of Kyoto Encyclopedia Genes And Genomes (KEGG) Enrichment Analysis. (A) T45d vs T30d, (B) T60d vs T30d, (C) T60d vs T45d.

components varied between different culture times. The accumulation of genistin content showed a trend of first increasing and then decreasing and then rising with the increase of culture time, and the accumulation of genistein was the highest (8.65 mg/g) at 45 DAC. However, the accumulation of genistein showed a decreasing trend with the increase of culture time, and the total accumulation of genistein and genistin was the highest at 45 DAC. A total of 41 DEGs were found to participate in 2-hydroxyisoflavone dehydratase, isoflavone-glucosyltransferase, chalcone flavanone isomerase, phenylalanine ammonia-lyase, and cinnamic acid 4-hydroxylase (C4H) pathway. There were 31 upregulated DEGs and 10 downregulated DEGs in T45d vs T60d, 23 upregulated DEGs and 18 downregulated DEGs in T30d vs T60d, and 19 DEGs upregulated and 22 downregulated in T30d vs T45d. These results indicated that genes regulating genistein and genistin biosynthesis mainly upregulated their expression with increasing culture time. Their expression levels were subsequently examined (Figure 7) to visualize the expression patterns of these DEGs. The majority of DEGs on the genistein and genistin synthesis pathway are shown in Table S1, and the order of their upregulated was T30d > T45d > T60d.

Discussion

Little research has been done yet on the best period to cultivate tissue culture seedlings of *F. macrophylla*. *F. macrophylla* tissue culture seedlings were cultivated in this study for 30, 45, 60, and 75 days. After comparing the phenotypic indices and active ingredient contents of the tissue culture seedlings after each of the four cultivation times, it was discovered that the best cultivation period was between 30 and 45 days. Using transcriptomics technology, the synthesis of genistein and genistin in tissue culture seedlings at 30 DAC, 45 DAC, and 60 DAC was found to be primarily regulated by the phenylpropanoic acid biosynthesis, flavonoid biosynthesis, and isoflavonoid biosynthesis pathways of Phenylalanine lyase (PAL), C4H, and 4-coumaroyl CoA ligase (4CL), chalcone synthase (CHS), chalcone isomerase (CHI), CYP93C, HIDH, IF7GT, and IF7MAT and the regulation of related DEGs.

F. macrophylla is known for its high medicinal and nutritional value, and it is a national folk herb that is frequently utilized in southwest ethnic communities of China (Chen et al., 2016). This is used as food and medicine. It is one of the popular medicinal materials on the market for

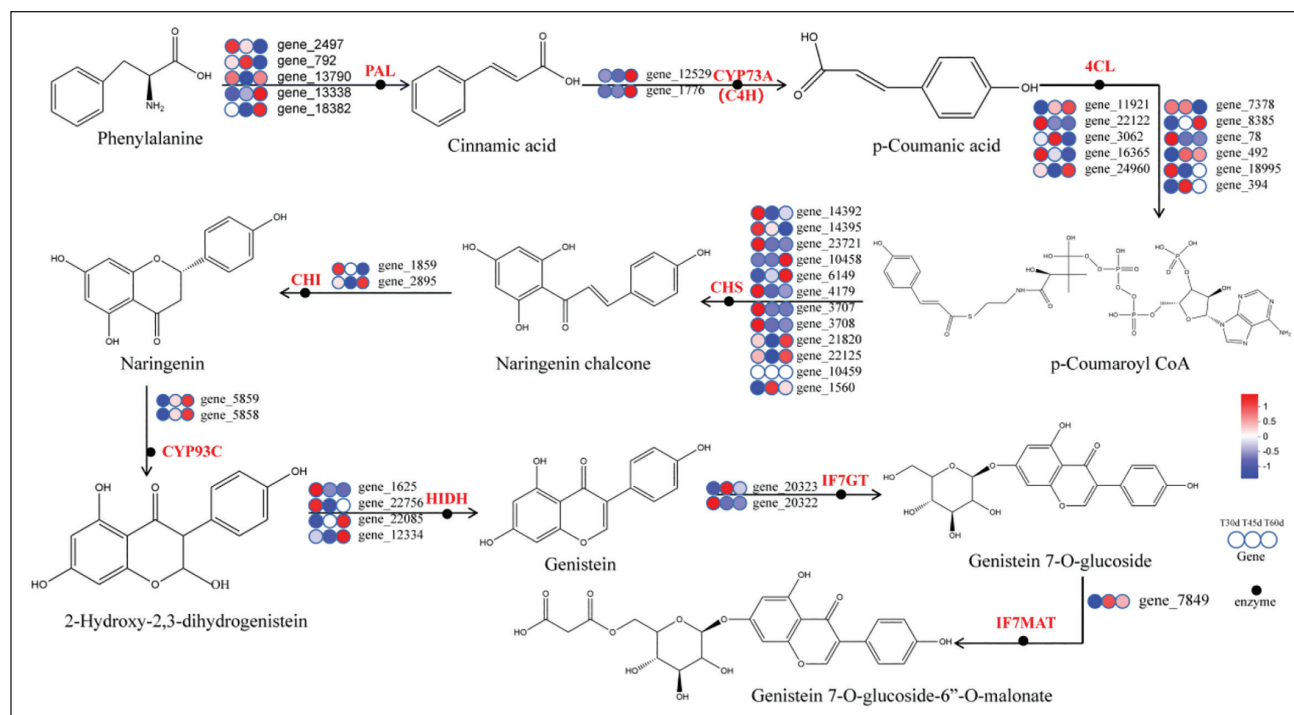


Figure 7. Plot of Significantly Altered Differently Expressed Genes (DEGs) in Genistein and Genistin Synthesis Pathways.

anti-inflammatory analgesia, and it plays an important role in people's treatment of diseases. However, the shortage of *F. macrophylla* resources has restricted the healthy development of its industry (Liang et al., 2017). Therefore, plant tissue culture can provide a great number of *F. macrophylla* seedlings in a short time, and adventitious roots used to extract compounds can be obtained, which alleviates the resource shortage.

Cultivation time may affect the synthesis and accumulation of active ingredients, as well as plant growth status. Plants' phenotypic indexes directly indicate the plants' growth status and biomass. A limited number of studies have investigated the phenotypic indexes and the change of active component content of different culture times. In this study, *F. macrophylla* seedling's culture at various times reveals that as culture time increased, the leaf area, the number of yellow leaves, and the development of the root system increased with culture time. However, there was no discernible difference in the phenotypic indexes, suggesting that the best culture time should be between 30 days and 45 days.

Genistein and genistin are the most abundant secondary metabolites in *F. macrophylla*, and their accumulation is generally influenced by harvest time. Additionally, previous studies have proved that genistein is the most active component in *F. macrophylla*, which belongs to the isoflavone subfamily of legumes (Niu et al., 2016). Secondary metabolites are typically the active ingredients in medicinal plants, and their biosynthesis and accumulation are influenced by various conditions, including environmental stress (Su et al., 2022),

light quality (Guo, 2022), altitude (Wu et al., 2021), moisture (Zhang, 2021), and others. In this study, the lowest genistein content (3.11 mg/g) in *F. macrophylla* was found at 30 days, while the maximum genistein content (8.65 mg/g) was at 45 days. The highest content between genistein and genistin was found at 45 DAC, which was 10.93 mg/g. As a result, the accumulation of active components in medicinal plants is not always inversely related to their growth rate. The main medicinal part of *F. macrophylla* is its root. Based on the phenotype of tissue culture seedlings at different cultivation times, even though the 45 DAC seedlings had a low rooting rate and a relatively short root length, their root fresh weight, root dry weight, and root thickness all reached the maximum value, indicating that the root growth of 45 DAC was better, and the active ingredient content was also higher. This suggested that the growth condition of the 45 DAC roots was better, leading to the highest accumulation of active ingredients at this time.

Currently, more than 10,000 medicinal plants are known to exist in China. However, the paucity of genomic data has hampered future research into many non-model plants' genetic pathways and functional genes (Liu, 2019). Transcriptomics is widely applied to prove insights into the genes' function and allow for the detection of genes with variable expression (Jia et al., 2015). The transcriptome data analysis revealed (Figure 8) that the isoflavonoid, flavonoid, and phenylpropanoic acid biosynthesis pathways were primarily responsible for the synthesis of genistein and genistin active components in *F. macrophylla* tissue culture seedlings.

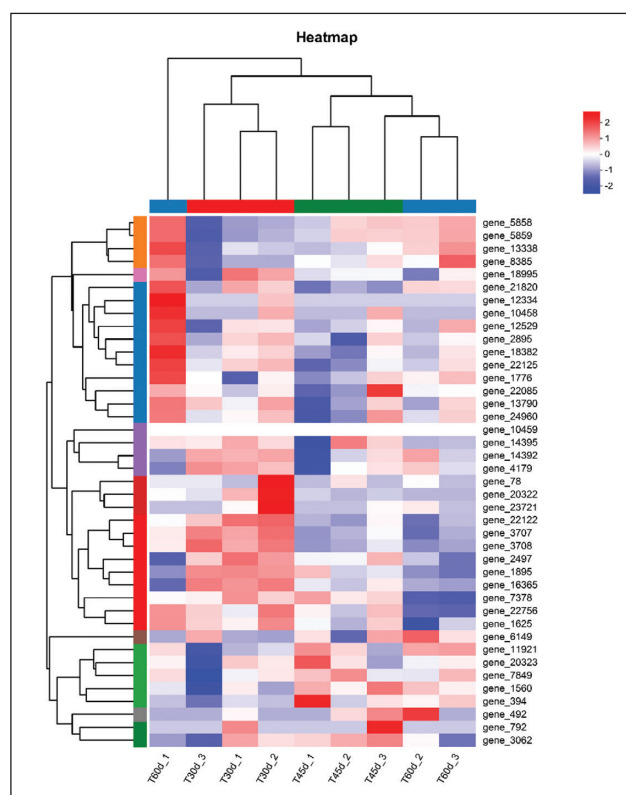


Figure 8. Heat Map of Differently Expressed Genes (DEGs) in the Synthesis Pathway of Genistein and Genistin at Different Incubation Times.

PAL, C4H, and 4CL catalyze the generation of coumaroyl coenzyme A from phenylalanine in the phenylpropionic acid biosynthesis pathway. This pathway is primarily influenced by PAL, C4H (CYP73A), and 4CL, which are regulated by 18 DEGs, such as gene_2497, gene_792, and gene_13790. Subsequently, in the pathway of flavonoid biosynthesis, coumaroyl coenzyme A produces chalcone through the action of CHS, and naringenin through the action of CHI. This process is primarily influenced by CHS and CHI, which are primarily controlled by gene_14392, gene_14395, and gene_1859, among other 14 DEGs. Finally, the isoflavone biosynthesis pathway is where naringenin enters the branch of isoflavone biosynthesis under the action of CYP93C to produce 2-Hydroxy-2,3-dihydrogenistein. The action of HIDH produces genistein, then Genistein-7-O-glucoside is formed by the action of IF7GT, and Genistein-7-O-glucoside-6"-malonate is formed by the action of IF7MAT. The primary regulators of these enzymes are 10 DEGs, including gene_5759, gene_1625, and gene_7849, with an additional 10 DEGs being regulated. Numerous enzymes, including PAL, C4H, 4CL, CHS, CHI, and others, were involved in each step of the isoflavone biosynthesis pathway in earlier research (Zhao et al., 2023). In the phenylpropanoid biosynthesis pathways, the PAL is the essential enzyme and rate-limiting enzyme that influences the expression of genes downstream in the isoflavone synthesis pathway (Wang et al.,

2019). Li et al. (2007) reported that C4H exhibited significant action across a range of plant tissues and that its expression may facilitate the synthesis of isoflavones. The final important enzyme in the phenylalanine pathway, 4CL, catalyzes the production of different hydroxyl cassia bark acids and is involved in the synthesis of phenylalanine and other end products (Zhao et al., 2023). One of the important enzymes that isoflavones express is CHS (Wang et al., 2019). Most plants employ CHS to catalyze the synthesis of isoflavones by 4-coumaroyl-CoA, which forms chalcones, other plants can use caffeoyl-CoA or acyl-CoA from cassia bark as a substrate to generate flavonoid flavanes (Martin, 1993). In order to create the matching tricyclic product, CHI catalyzes the cyclization reaction in molecules, specifically the chalcone of the second ring (Tian et al., 2014). Some genes also regulated these enzymes, and most were involved in the modulation of genistein and genistin biosynthesis through genetic negative regulation of related enzymes. Therefore, it is speculated that the three pathways and the related enzymes and related genes in the three pathways regulated their growth and development, resulting in the optimal growth status of 45 DAC and the highest accumulation of genistein and genistin. Analysis of the three comparison groups found that both isoflavonoid biosynthesis and phenylpropanoid biosynthesis pathways were significantly enriched in both T45d vs T30d and T60d vs T30d, whereas isoflavonoid biosynthesis was not

significantly enriched in the T60d vs T45d group. It indicates that the genes-regulated isoflavone synthesis increased significantly with culture time within 45d of culture, while there was a considerable reduction in the genes that control isoflavone synthesis at 45 DAC. Additionally, this is generally in line with the findings of the determination of genistin and genistein.

Conclusion

The optimum culture time for the accumulation of genistein and genistin in *F. macrophylla* seedlings ranged from 30 to 45 days. Forty-one DEGs in the genistein and genistin synthesis pathways were found through transcriptomics analysis. This study sheds light on the biosynthesis of genistin and genistein in *F. macrophylla*.

Abbreviations

6-BA: 6-benzyl adenine; 4CL: 4-coumaroyl CoA ligase; C4H: Cinnamic acid 4-hydroxylase; CHI: Chalcone isomerase; CHS: Chalcone synthase; COG: Clusters of orthologous groups; DAC: Days after culturing; DEGs: Differently expressed genes; FC: Fold change; *F. macrophylla*: *Flemingia macrophylla* (Willd.) Prain; GO: Gene ontology; HPLC: High-performance liquid chromatography; IAA: Indoleacetic acid; KEGG: Kyoto Encyclopedia Genes and Genomes; KT: Kinetin; NR: Nonredundant protein sequence database; PAL: Phenylalanine lyase; PCR: Polymerase chain reaction; PFKM: Fragments per kilobase per million; qRT-PCR: Quantitative reverse transcription polymerase chain reaction; RNA-seq: RNA-sequencing.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethical Approval and Informed Consent

Not applicable.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was funded by the Construction of Southern Medicine Germplasm Resource Base for Guangdong Northern (20231206), the Guangxi Science and Technology Project (GuiKe ZY20198018) and China Agriculture Research System (CARS-21), Young and Middle Teachers' Basic Ability Improvement Project of Guangxi (2022KY1700).

Supplementary Material

Supplementary material for this article is available online.

ORCID iD

Qiao Ling Yu  <https://orcid.org/0009-0003-0032-4195>

References

- Chen, P., Li, M., & Qin, H. G. (2016). Progress in quality control of Moghania. *Chinese Pharmaceutical Journal*, 51(08), 601–606.
- Editorial Board of Chinese Materia Medica [EBCMM]. (1999). State administration of traditional Chinese medicine. In *Chinese materia medica* (vol. 41, pp. 474–478). Shanghai Science and Technology Press.
- Glazier, M. G., & Bowman, M. A. (2001). A review of the use of phytoestrogens as a replacement for traditional estrogen replacement therapy. *Archives of Internal Medicine*, 161(9), 1161–1172. <https://doi.org/10.1001/archinte.161.9.1161>
- Guo, X. Z. (2022). *Study on accumulation of pharmacodynamic components and function of AIHY5 gene in Atractylodes lancea (Thunb.) DC. by light quality*. China Academy of Chinese Medical Sciences.
- Jia, C. L., Zhang, Y., Zhu, L., & Zhang, R. (2015). Application progress of transcriptome sequencing technology in biological sequencing. *Molecular Plant Breeding*, 10, 2388–2394.
- Kato, K., Takahashi, S., Cui, L., Toda, T., Suzuki, S., Futakuchi, M., Sugiura, S., & Shirai, T. (2000). Suppressive effects of dietary genistin and daidzin on rat prostate carcinogenesis. *Japanese Journal of Cancer Research*, 91(8), 786–791. <https://doi.org/10.1111/j.1349-7006.2000.tb01014.x>
- Lee, C. H., Yang, L., Xu, J. Z., Venus Yeuny, S. Y., Huang, Y., & Chen, Z. Y. (2005). Relative antioxidant activity of soybean isoflavones and their glycosides. *Food Chemistry*, 90(4), 735–741. <https://doi.org/10.1016/j.foodchem.2004.04.034>
- Li, L., Zhao, Y., & Ma, J. L. (2007). Recent progress on key enzymes: PAL, C4h, 4CL of phenylalanine metabolism pathway. *Chinese Journal of Bioinformatics*, 4, 187–189.
- Li, P. S., Liu, Y. Y., Fan, B. D., Wei, Q., Li, C. N., Xu, T. H., & Liu, T. H. (2014). Research progress of chemical composition and pharmacological effectiveness of chickpeas. *Chinese Journal of Experimental Traditional Medical Formulae*, 20(11), 235–238.
- Liang, Y., Wei, K. H., Qiao, Z., Li, L. X., Li, M. J., Zhang, Z. Y., & Miao, J. H. (2017). Overview of the study of *Flemingia macrophylla*. *Guangxi Journal of Traditional Chinese Medicine*, 40(02), 5–8.
- Liu, H. B. (2019). Applications of RNA-seq technology on medicinal plants. *Chinese Traditional and Herbal Drugs*, 24, 5346–5354.
- Martin, C. R. (1993). Structure, function, and regulation of the chalcone synthase. *International Review of Cytology*, 147, 233–284. [https://doi.org/10.1016/S0074-7696\(08\)60770-6](https://doi.org/10.1016/S0074-7696(08)60770-6)
- Niu, Y. F., Li, X. H., Li, H. T., Zhang, Z. L., Guan, Y. H., & Zhang, L. X. (2016). Distribution and accumulative dynamic variation of bioactive components in *Moghania macrophylla*. *Chinese Medical Herald*, 13(28), 17–20.
- Ruiz-Larrea, M. B., Mohan, A. R., Paganga, G., Bolwell, G. P., & Rice Evans, C. A. (1997). Antioxidant activity of phytoestrogen isoflavones. *Free Radical Research*, 26(1), 63–70.
- Su, Q., Zhang, K., Huang, S. Q., Zhu, X. Q., Liu, X. P., & Jiang, N. (2022). Effects of adversity stress on accumulation of medical components in plants. *Journal of Hubei Minzu University (Natural Science Edition)*, 40(02), 129–134+180.
- Sun, Y., Wang, J. G., Zang, D. D., Zhao, H. T., & Liu, S. H. (2022). Transcriptome analysis of *Lonicera caerulea* fruit at different developmental stages. *Biotechnology Bulletin*, 38(12), 204–213.

- Tian, L. (2014). Cloning and expression profile analysis of transcription factors genes regulating isoflavone synthesis in soybean. In *Chinese academic of agriculture* [Unpublished sciences dissertation].
- Wang, C. L., Wang, Z. L., Yang, J. L., Gou, S. X., & Chen, W. J. (2019). Biosynthesis, regulation, physiological and ecological functions of application of soybean isoflavones: A review. *Chinese Agricultural Science Bulletin*, 35(19), 34–40.
- Weng, C. Y., Yang, Y., & Li, X. B. (2016). The advance of studies on the anti-atherosclerosis effect of genistein. *Guide of Chinese Medicine*, 14(07), 33–35.
- Wu, G. Y., Zhang, M., Yao, S. C., Huang, D., Qin, H. Q., & Wang, X. X. (2023). Transcriptomics analysis reveals the differential genes in roots of *Stemona tuberosa* Lour. with different growth years. *Molecular Plant Breeding*, 1–31.
- Wu, P., Sun, Z., Yang, L. M., & Han, M. (2021). Study on the effect of altitude on the accumulation of main active components and expression of key enzyme genes of astragalus Mongolians. *Chinese Traditional and Herbal Drugs*, 52(13), 4031–4038.
- Yan, H. Q., Yu, J., Zheng, W., Li, Y., Liu, L. Y., & Li, Z. G. (2023). Analysis of transcriptome and screening differential genes of *Areca catechu* L. at different developmental stages. *Molecular Plant Breeding*, 21(14), 4595–4605.
- Yanagihara, K., Ito, A., Toge, T., & Numoto, M. (1993). Antiproliferative effects of isoflavones on human cancer cell lines established from the gastrointestinal tract. *Cancer Research*, 53(23), 5815–5821.
- Yu, L. H., Liu, K., & Liu, B. L. (2010). Studies on pharmacological effect of genistein. *Pharmaceutical and Clinical Research*, 18(03), 299–302.
- Zhai, Z. H., Chen, X. N., & Wang, J. (2008). Primer design with primer Premier 5.0. *Medical Education Research and Practice*, 04, 695–698.
- Zhang, P., Zheng, W. J., & Zhong, Y. (2008). Research progress of genistein. *Journal of Pharmaceutical Research*, 02, 103–106.
- Zhang, Q. (2021). *Effects of water regulation on pharmacodynamic components and biosynthetic gene expression of Astragalus membranaceus (Fish) Bge. Var. mongholicus (Bge.) Hsiao*. Jilin Agricultural University.
- Zhang, Q. Y., Ma, F. C., Wang, S., Wu, C. M., & Guo, P. (2017). Pharmacological activities of genistin: Research advances. *Journal of International Pharmaceutical Research*, 44(04), 315–318+325.
- Zhao, X. (2002). *Flemingia macrophylla*—Excellent forage garage germplasm resource in the tropical area of Yunnan. *Journal of Grassland and Forage Science*, 04, 31–32+42.
- Zhao, Y. M., Liu, L. Y., Lu, M. Q., He, B., Long, C. F., Gong, X. J., & Huang, Y. C. (2023). Research advance on pathway of isoflavone biosynthesis and relevant key enzymes. *Food and Fermentation Industries*, 2, 1–9.
- Zhu, Y. Y., Yao, Y., Shi, Z. X., Everaert, N., & Ren, G. X. (2018). Synergistic effects of bioactive anticarcinogens from soybean on anti-proliferative activity in MDA-MA-231 and MCF-7 human breast cancer cells in vitro. *Molecules*, 23(7), 1557. <https://doi.org/10.3390/molecules23071557>