Deciphering Alkaloid and Flavonoid Biosynthesis in Different Tissues of Sophora tonkinensis via Transcriptomics and Metabolomics

Ying Liang
National Center for Traditional Chinese Medicine (TCM) Inheritance and Innovation, Guangxi Botanical Garden of Medicinal Plants

Guili Wei
National Center for Traditional Chinese Medicine (TCM) Inheritance and Innovation, Guangxi Botanical Garden of Medicinal Plants

Ximei Liang
National Center for Traditional Chinese Medicine (TCM) Inheritance and Innovation, Guangxi Botanical Garden of Medicinal Plants

Meiqiong Tang
National Center for Traditional Chinese Medicine (TCM) Inheritance and Innovation, Guangxi Botanical Garden of Medicinal Plants

Hong He
National Center for Traditional Chinese Medicine (TCM) Inheritance and Innovation, Guangxi Botanical Garden of Medicinal Plants

Danfeng Tang
National Center for Traditional Chinese Medicine (TCM) Inheritance and Innovation, Guangxi Botanical Garden of Medicinal Plants

Yang Lin
National Center for Traditional Chinese Medicine (TCM) Inheritance and Innovation, Guangxi Botanical Garden of Medicinal Plants

Linxuan Li
National Center for Traditional Chinese Medicine (TCM) Inheritance and Innovation, Guangxi Botanical Garden of Medicinal Plants

Shuangshuang Qin
qin_double@126.com

National Center for Traditional Chinese Medicine (TCM) Inheritance and Innovation, Guangxi Botanical Garden of Medicinal Plants

Fan Wei

National Center for Traditional Chinese Medicine (TCM) Inheritance and Innovation, Guangxi Botanical Garden of Medicinal Plants

Research Article

Keywords: Sophora tonkinensis, transcriptomics, metabolomics, alkaloids, flavonoids

Posted Date: May 2nd, 2024

DOI: https://doi.org/10.21203/rs.3.rs-4311231/v1

Additional Declarations: No competing interests reported.

Abstract

Background *Sophora tonkinensis* Gagnep is an important Chinese herbal medicine, with alkaloids and flavonoids being the primary pharmacological substances. However, theinformation on the metabolic pathways of alkaloids and flavonoids in *S. tonkinensis* are not well understood.

Results: This research delves into the molecular regulation of the accumulation of alkaloids and flavonoids in *S. tonkinensis*. By examining four tissues (seeds, leaves, stems, and roots), through high-throughput transcriptome sequencing (RNA-seq) and liquid chromatography-mass spectrometry (LC-MS), it identified differential metabolites and genes related to alkaloid and flavonoid biosynthesis. Transcriptome analysis identified 2,727 differentially expressed genes (DEGs). Among these DEGs, 35 were related to alkaloids and 48 to flavonoids. Furthermore, metabolome analysis revealed 296 differentially expressed metabolites (DEMs), comprising 23 alkaloid-related DEMs and 23 flavonoid-related DEMs. Additionally, the weighted gene co-expression network analysis highlighted the potential of *StCAO* (evm.model.3.924) in regulating alkaloid biosynthesis and *StCHIs* (evm.model.3.2047, evm.model.1.2104, and evm.model.1.2101) in flavonoid biosynthesis. To reinforce these findings, qPCR validation confirmed the consistency of expression trends for 12 selected DEGs across the roots, stems, leaves, and seeds of *S. tonkinensis*.

Conclusion: This study provides a detailed exploration of the regulatory mechanisms of alkaloid and flavonoid accumulation and associated key genes in different *S. tonkinensis* tissues. These findings provide insights for future investigations into the regulatory mechanisms of alkaloids and flavonoids in *S. tonkinensis*.

Background

Sophora tonkinensis Gagnep. (family Leguminosae), a medicinal plant native to China found in karst areas, is listed in the national protection Category II[1]. In Chinese, the dried rhizomes and roots are referred to as “Shan-Dou-Gen,” and have been traditionally used to treat conditions like heat toxicity accumulation, sore throat, swollen gums, and mouth and tongue soreness[2]. Over 150 compounds, including over 80 flavonoids, 40 volatile oils, and 20 alkaloids, have been identified in *S. tonkinensis*[3], with alkaloids and flavonoids being the primary pharmacological substances[4]. Notable alkaloids include large amounts of matrine and oxymatrine[4] and small amounts of sophocarpine, harmine, sophoranol, and sophoramol. Matrine and oxymatrine are recognized for their antibacterial, anti-inflammatory, and anti-cancer properties [5-7]. *S. tonkinensis*, aside from its roots and rhizomes, also yields matrine and oxymatrine from its aerial stems and leaves[1]. *S. tonkinensis* also contains over 100 flavonoids, including dihydroflavonoids, isoflavonoids, flavonols, and chalcones, with diverse pharmacological activities like anti-tumor, anti-inflammatory, antioxidant, antibacterial, antiviral, and cardiovascular and cerebrovascular protection[8]. While research on *S. tonkinensis* has primarily focused on germplasm breeding[9], cultivation optimization [10], and pharmacological effects [11, 12],

investigations into the metabolic pathways of alkaloids and flavonoids remain limited, with existing studies mainly concentrating on extraction, detection, and activity analysis [13, 14].

The fast advancements in systems biology and high-throughput sequencing technologies in recent years have made multi-omics an indispensable research method. This approach provides insights into the dynamic changes in plant growth and development at both the system and cellular levels. The combination of metabolomics and transcriptomics (RNA-seq) has become an effective tool for studying and identifying post-genomic processes and the molecular basis of various plants [15]. The combined analysis of metabolomics and RNA-seq reveals the mechanisms and dynamic changes in the synthesis of diverse plant metabolic compounds, including *Mesona chinensis*[16], *Plumbago zeylanica* [17], *Bupleurum chinense*[18], and *Gynostemma pentaphyllu*m[19].

This study employed a combination of transcriptomics and metabolomics to systematically investigate the metabolic variations of alkaloids and flavonoids in the roots, stems, leaves, and seeds of *S. tonkinensis*. Furthermore, the study identified key enzyme genes closely associated with the biosynthesis of alkaloids and flavonoids, laying a foundation for exploring the molecular regulatory mechanisms behind the differential accumulation of alkaloids and flavonoids in the various plant tissues of *S. tonkinensis.* Additionally, it contributes to understanding the formation mechanisms of pharmacological effects and quality.

Results

Transcriptome alterations overview

Transcriptome sequencing was conducted on *S. tonkinensis* roots, stems, leaves, and seeds (Fig. 1) to determine the molecular regulation in different tissues. The acquired raw reads ranged from 45.71 to 51.09 million, with respective Q20 and Q30 values exceeding 97% and 93% (Table S1), suggesting the acquisition of high-throughput and high-quality RNA-Seq data. Following the removal of low-quality reads, a further analysis was performed on 44.36 to 49.33 million clean reads. Out of all the clean reads, between 69.19% and 92.05% were found to map to the genome of *S. tonkinensis*. The RNA-Seq dataset with the accession number PRJNA1052132 is archived in the NCBI SRA database. Additionally, the number of DEGs across various *S. tonkinensis* tissues was observed (Fig. 2A). In all comparison groups, a comprehensive set of 2,727 DEGs were screened (Fig. 2B and Table S2). Of these, 543 DEGs were commonly expressed in all groups, accounting for 3.55%. The remaining six groups showed specific genes for each comparison group (Fig. 2B), indicating that the expression of these genes was activated by *S. tonkinensis* in response to various tissue functional structures. To ascertain if the RNA-Seq data were reliable, we selected 12 DEGs, including two related to alkaloid biosynthesis, two related to flavonoid biosynthesis, two related to transcription factors, two related to photosynthesis, and four randomly selected DEGs. Their expression levels were assessed using qPCR, revealing a positive correlation between the expression profiles detected by qPCR and the RNA-Seq results (Fig. 3).

DEGs involved in the alkaloid biosynthesis and flavonoid biosynthesis pathways

Thirty-five DEGs involved in alkaloid biosynthesis were identified to further understand the differential expression genes associated with the biosynthetic pathways of flavonoids and alkaloids in distinct *S. tonkinensis* tissues (Fig. 4A). These 35 DEGs encompass ten tropinone reductases (*TRs*), eight copper amine oxidases (*CAOs*), seven polyphenol oxidases (*PPOs*), three tyrosine aminotransferases (*TATs*), two dependent decarboxylase conserves (*DDCs*), two histidinol-phosphate aminotransferases (*HisCs*), one aspartate aminotransferase (*ASP5*), one aspartate aminotransferase (*GOT2*), and one lysine decarboxylase (*LDC*).

Furthermore, 48 DEGs involved in flavonoid biosynthesis were identified (Fig. 4B), including 11 chalcone synthases (*CHSs*), seven hydroxycinnamoyl transferases (*HCTs*), six chalcone isomerases (*CHIs*), three flavonol synthases (*FLSs*), three spermidine hydroxycinnamoyl transferases (*SHTs*), two caffeoyl-CoA O-methyltransferases (*CCoAOMTs*), two cinnamate 4-hydroxylases (*CYP73As*), two flavonoid 3',5'-hydroxylases (*CYP75As*), two dihydroflavonol reductases (*DFRs*), two flavonoid 3',5'-methyltransferases (*FAOMTs*), one anthocyanidin reductase (*ANR*), one cinnamoyl-CoA O-methyltransferase (*CCOMT*), one flavonoid 3'-monooxygenase (*CYP75B1*), one coumaroylquinate (coumaroylshikimate) 3'-monooxygenase (*CYP98A*), one omega-hydroxypalmitate O-feruloyl transferase (*HHT1*), one leucoanthocyanidin dioxygenase (*LDOX*), one NAD(P)H-dependent 6'-deoxychalcone synthase (*NADH*), and one phloretin 2'-O-glucosyltransferas (*PGT1*).

By examining the expression levels, we found that *StCAO* (evm.model.3.924) in the alkaloid biosynthesis pathway exhibited significantly enhanced expression in the roots, stems, leaves, and seeds, particularly in the seeds where its expression surpasses that of other genes significantly. Additionally, the expression level of *StCHI* (evm.model.3.2047) was significantly higher in the root, stem, leaf, and seed in the flavonoid biosynthetic pathway. Among them, the expression level in the seed was 1.76, 1.98, and 3.07 times greater than that in the root, stem, and leaf, respectively. In the seed, *StCHI* (evm.model.3.2047) expression was significantly upregulated relative to that of other genes.

Alkaloid and flavonoid content in different *S. tonkinensis* tissues

As shown in Fig. 5, the concentrations of total alkaloids, matrine, and oxymatrine in seeds were significantly elevated relative to those in other tissues ($P<0.05$), registering values of 19.88 mg/g, 169.48 μg/g, and 9.43 mg/g, respectively. Similarly, the concentrations of total flavonoids and genistin in seeds were significantly elevated in comparison to other tissues ($P<0.05$), measuring 14.02 mg/g and 27.13 μg/g, respectively. The concentration of genistein in leaves reached 15.05 μg/g, which was significantly elevated relative to those in other tissues ($P<0.05$). These results highlight the variations in alkaloid and flavonoid contents across different *S. tonkinensis* tissues.

Overview of the metabolome changes

The results of metabolic compound detection in these samples have undergone quality control (QC) analysis. The QC sample relative standard deviation (RSD) evaluation plot is shown in Fig. S1. The overall data is considered qualified and reliable for subsequent analysis when the RSD is 0.3, with a cumulative proportion of ion peaks reaching 82.01%. PCA revealed that the proportion of total variance explained by the first principal component (PC1) was 43.90 %, distinguishing samples based on different tissues of *S. tonkinensis*. Conversely, 31.50 % of the total variance was attributed to the second principle component (PC2) (Fig. 6A). Samples from the same tissue cluster closely together, indicating good data condition. The roots and stems exhibit closer proximity, indicating higher similarity in chemical composition. In contrast, leaves and seeds are more distant from each other and the roots and stems, indicating lower chemical composition similarity and unique metabolic characteristics. The PCA results underscore differences in the metabolic products of different *S. tonkinensis* tissues.

The Pearson correlation coefficient analysis revealed consistent cumulative metabolite values among the six biological replicates in 24 *S. tonkinensis* samples (Fig. 6B). Additionally, the correlation coefficient within group samples surpasses that between inter-group samples, affirming the reliability of the obtained differential metabolites. These findings underscore the robust correlation and reliability of experimental results across different tissues of *S. tonkinensis*.

In total, 4,138 metabolites were identified from four tissues, with 2,282 in positive and 1,856 in negative ion modes (Table S3). Employing PLS-DA model analysis, differentially expressed metabolites (DEMs) were screened using significance criteria of $P<0.05$ and VIP value >1. The results revealed 1,553 DEMs between roots and stems (432 upregulated and 1,121 downregulated), 1,658 DEMs between roots and leaves (584 upregulated and 1,074 downregulated), 1,761 DEMs between roots and seeds (4 upregulated and 867 downregulated), 1,634 DEMs between stems and leaves (861 upregulated and 773 downregulated), 1,831 DEMs between stems and seeds (1,173 upregulated and 658 downregulated), and 1,846 DEMs between leaves and seeds (1,149 upregulated and 697 downregulated (Fig. 6C). These results highlight the diverse metabolic composition across different tissues of *S. tonkinensis*. Further analysis of DEMs among roots, stems, leaves, and seeds identified 18 major DEMs (Fig. 6D), including sparteine, homoferreirin, O-desmethyltramadol, dihydrofolic acid, deoxypyridinoline, artocarpesin, ID14326, 12,20-dioxo-leukotriene B4, CDP-ethanolamine, lucuminoside, 4-hydroxyandrostenedione glucuronide, 6-Epi-7-isocucurbic acid glucoside, isopropyl beta-D-glucoside, ribostamycin, (3S,7E,9R)-4,7-megastigmadiene-3,9-diol 9-[apiosyl-(1->6)-glucoside], kuwanon B, wyerone, isovitexin 2"-(6'''-feruloylglucoside) 4'-glucoside. Among them, sparteine is associated with alkaloid biosynthesis.

Integrated transcriptome and metabolome analysis

The histogram illustrates KEGG pathway enrichments of DEGs and DEMs. In root_vs_stem involving 126 DEGs and 79 DEMs, 20 metabolic pathways were enriched, with significant enrichment ($P<0.01$) observed in phenylpropanoid biosynthesis and flavonoid biosynthesis pathways (Fig. 7A). For leaf_vs_root involving 130 DEGs and 84 DEMs, 21 metabolic pathways were enriched, with significant enrichment ($P<0.01$) in the pathways related to the biosynthesis of isoflavonoids, phenylpropanoids,

flavonoids, and betalains (Fig. 7B). Root_vs_seed, with 129 DEGs and 84 DEMs, exhibited enrichment in 20 metabolic pathways, showing significant enrichment (P<0.01) in pathways related to the biosynthesis of phenylpropanoids, flavonoids, and isoflavonoids (Fig. 7C). Leaf_vs_stem, featuring 129 DEGs and 82 DEMs, showed enrichment in 21 metabolic pathways, with significant enrichment (P<0.01) in flavonoid biosynthesis, phenylpropanoid biosynthesis, tryptophan metabolism, isoflavonoid biosynthesis, and alanine, aspartate and glutamate metabolism pathways (Fig. 7D). In leaf_vs_seed, encompassing 132 DEGs and 69 DEMs, 21 metabolic pathways were enriched, with significant enrichment (P<0.01) in isoflavonoid biosynthesis, phenylpropanoid biosynthesis, and flavonoid biosynthesis pathways (Fig. 7E). Seed_vs_stem, with 125 DEGs and 79 DEMs, showed enrichment in 20 metabolic pathways, exhibiting significant enrichment (P<0.01) in flavonoid biosynthesis, phenylpropanoid biosynthesis, tryptophan metabolism, isoflavonoid biosynthesis, and alanine, aspartate and glutamate metabolism pathways (Fig. 7F). The findings of this study emphasize the critical metabolic pathways that are implicated in the biosynthesis of phenylpropanoid, flavonoid, and isoflavonoid in *S. tonkinensis*.

Co-expression network analysis

Using RNA-seq and content data, we employed weighted gene co-expression network analysis (WGCNA) to examine genes related to the biosynthesis and metabolism of various flavonoids and alkaloids. The dendrogram illustrates the fifteen identified distinct modules (Fig. 8A). These modules, represented by distinct colors, revealed significant associations with the content of alkaloids, flavonoids, matrine, oxymatrine, genistein, and genistin, notably the “magenta,” “tan,” “salmon,” “magenta,” “tan,” and “green-yellow” modules (r>0.7, P<0.05) (Fig. 8B).

Based on eigengene connectivity (kME) values in the co-expression network, co-expression subnetworks were generated using the top 30 node genes from the "magenta," "tan," and "green-yellow" modules. In the “magenta” module, evm.model.8.1855 (no functional annotation information) exhibited the highest kME value, followed by copper amine oxidase (*StCAO*, evm.model.3.924). Additionally, tyrosine aminotransferase (*StTAT*, evm.model.2.3509), involved in alkaloid biosynthesis, was located in this network (Fig. 8C). In the “tan” module, squalene monooxygenase (*StSQLE*, evm.model.5.860) had the highest kME value. Additionally, chalcone isomerase (*StCHI*, evm.model.1.2104 and evm.model.1.2101), contributing to flavonoid synthesis, occupied a central role in this network (Fig. 8D). In the “green-yellow” module, chlorophyll a-b binding protein (*StCAP10A*evm.model.1.2462) exhibited the highest kME value and displayed robust correlations with other node genes. Every node in this network was associated with photosynthesis, suggesting its crucial role in the network (Fig. 8E). These findings indicate that *StCAO*, *StTAT*, *StCHI*, and *StCAP10A* are involved in regulating alkaloid and flavonoid compound metabolism in different tissues of *S. tonkinensis*.

Discussion

DEGs and DEMs were identified in different *S. tonkinensis* tissues

According to the transcriptome data, 2727 DEGs were identified across the roots, stems, leaves, and seeds of *S. tonkinensis* (Fig. 2). The DEG analysis revealed associations with specific processes, including photosynthesis, plastid nucleoid, phenylpropanoid metabolic process, oxidoreductase activity, secondary metabolic process, nucleoside transmembrane transporter activity, auxin-activated signaling pathway, response to reactive oxygen species, starch metabolic process, and thylakoids (Fig. S2). These findings suggest that DEGs in different tissues may be influenced by both internal and external environments[20]. The DEGs exhibited significant induction effects in processes such as isoflavonoid biosynthesis, phenylpropanoid biosynthesis, plant hormone signal transduction, nitrogen metabolism, flavonoid biosynthesis, isoquinoline alkaloid biosynthesis, amino sugar and nucleotide sugar metabolism, tropane, piperidine, and pyridine alkaloid biosynthesis, and biosynthesis of various secondary metabolites (Fig. S3). Secondary metabolites undergo dynamic changes throughout the plant's growth process, responding to diverse environmental conditions[21]. The synthesis of secondary metabolites in plants is intricately regulated by intracellular genes. The accumulation of chemical components affects the activities of related regulatory enzymes and the expression of associated genes[22, 23].

Based on the metabolome data, we identified 296 DEMs in the experimental group (Fig. 6). Regarding secondary metabolite accumulation, oxysolavetivone and 1-ethyl-4-methyl-2-[2-(3-methylbutylamino)-2-oxoethyl]pyrrole-3-carboxylic acid significantly accumulated in the roots of *S. tonkinensis*. Indomethacrylic acid, N-[2-(cyclohexen-1-yl)ethyl]-N'-[[1-(hydroxymethyl)cyclopropyl]methyl]oxamide and N-diethyl-m-toluamide showed significant accumulation in the stem, whereas 3-genistein-8-c-glucoside, vitexin-2'-o-rhamnoside, vitexin-2''-o-rhamnoside, and leucylphenylalanine were significantly accumulate in the leaves. Matrine significantly accumulated in the seeds, and oxymatrine significantly accumulated in both roots and seeds (Fig. S4). Notably, genistein, matrine, and oxymatrine displayed a consistent change trend in different tissues of *S. tonkinensis*, aligning their relative abundance values and content changes in various tissues (Fig. 5). These findings indicate a crucial role played by the abundance values of different genes in *S. tonkinensis*, influencing the content of different components.

Procrustes analysis helps visualize the overall correlation between the transcriptomes and metabolomes of diverse tissues. Based on the correlation between their loads, it facilitates the integration of multiple data types, including microbiome, metabolome, and transcriptome, by assessing the distribution of two or more datasets within the same system[24]. In this study, we used Procrustes and histogram analyses to assess the overall correlation between the transcriptomes and metabolomes of different tissues in *S. tonkinensis*. Additionally, we conducted KEGG pathway enrichment analyses for both DEGs and DEMs. The results highlighted the significance of flavonoid, phenylpropanoid, and isoflavonoid biosynthesis as crucial metabolic pathways in *S. tonkinensis*.

Through comprehensive analysis of both transcriptomics and metabolomics, we elucidated the metabolic pathways within the roots, stems, leaves, and seeds of *S. tonkinensis*. This detailed exploration offers profound insights into the biosynthetic mechanisms responsible for the active ingredients in *S. tonkinensis*. The molecular-level clarification of *S. tonkinensis*'sbiosynthetic

mechanisms offers a theoretical foundation and reference for further research on the metabolic engineering of active ingredients in *S. tonkinensis*.

Metabolites and genes involved in alkaloid biosynthesis in *S. tonkinensis*

Alkaloid metabolism in *S. tonkinensis* has been associated with important therapeutic effects, including anti-inflammatory[25], antiviral[26], and anti-cancer properties[27]. Matrine, a key alkaloid in *S. tonkinensis*, primarily originates from the biosynthesis pathway of quinolizidine alkaloids (QAs)[27]. However, the detailed QA biosynthesis mechanism remains largely unknown[28], hampering our understanding of alkaloid metabolism in *S. tonkinensis*. This study identified 35 DEGs involved in alkaloid biosynthesis, including eight unigenes encoding CAO (Fig. 4A), aligning with previous research on CAO in *S. tonkinensis* under drought stress[29]. Among these, the eight CAO-encoding unigenes were expressed in all tissues, with four highly expressed in leaves, two in seeds, and two in stems. Notably, *StCAO* (evm.model.3.924) demonstrated significantly upregulated expression in distinct tissues of *S. tonkinensis*, particularly in seeds where its expression was significantly upregulated relative to that of other genes. Consistent with this expression pattern, the alkaloid content in seeds was significantly higher than in roots, stems, and leaves of *S. tonkinensis* (Fig. 5).

Additionally, this study preliminarily identified 23 alkaloid synthesis-related DEMs in different tissues of *S. tonkinensis*, including morphine, alkaloid RC, aquifoliunine EI, tetrahydroharmol, cinchonidine, 5-methoxycanthin-6-one, 3-hydroxyquinidine, 3-hydroxyquinine, and matrine (Fig. S5). Cinchona alkaloids and morphinans exhibited the highest content among the four tissues, followed by harmala alkaloids, lupin alkaloids, rhoeadine alkaloids, strychnos alkaloids, betalains, indolonaphthyridine alkaloids, erythrina alkaloids, benzophenanthridine alkaloids and camptothecins. Notably, lupin alkaloids belong to QAs, an area of biosynthesis research less developed than in some economically important plants[30].

Utilizing transcriptome analysis and gene co-expression analysis, we identified specific metabolic pathways associated with new genes[31]. By using this approach, we were able to select genes with expression patterns similar to alkaloid content, revealing potential alkaloid biosynthesis genes such as CAO (*StCAO*, evm.model.3.924) and tyrosine aminotransferase (*StTAT*, evm.model.2.3509). Historically, CAO has been implicated in the second step in QA biosynthesis, converting cadaverine to 5-aminopentanal and subsequently cyclizing it to 1-piperideine[30, 32]. This indicates the role of the *StCAO* gene as a potential indicator of alkaloid accumulation in *S. tonkinensis*, providing theoretical and genetic resources for developing new varieties with high alkaloid content.

Metabolites and genes involved in flavonoid biosynthesis in *S. tonkinensis*

S. tonkinensis is a valuable medicinal plant, particularly due to its richness in flavonoids, a class of plant secondary metabolites encompassing monomeric flavonols, flavanones, flavones, and flavonols playing crucial roles in many biological processes[33]. However, the molecular-level exploration of flavonoids in *S. tonkinensis* has been hindered by limited information. In our study, we present a transcriptome assembly for *S. tonkinensis*, unveiling 48 DEGs associated with the biosynthesis of flavonoid

compounds (Fig. 4B). Eleven unigenes encoding CHS, the initial enzyme in the plant flavonoid synthesis pathway and a key enzyme in secondary metabolic pathways, were expressed across all tissues. Among them, six unigenes exhibited high expression in roots, two in seeds, two in leaves, and one in stems. CHS holds significant physiological importance for plants[34]. Additionally, our research identified six CHI-encoding unigenes with generally higher expression levels in root and seed tissues. Notably, *StCHI* (evm.model.3.2047) displayed significantly higher expression levels in various *S. tonkinensis* tissues, particularly in seeds where its expression surpassed that of other genes. This aligns with the observed higher levels of total flavonoids and genistin content in *S. tonkinensis* seeds compared to roots, stems, and leaves (Fig. 5). Previous research has highlighted the pivotal regulatory and catalytic roles of CHS and CHI in flavonoid biosynthesis, emphasizing their impact on compound accumulation and diversity[35]. Therefore, the transcriptional changes in genes encoding these enzymes in *S. tonkinensis* tissues may suggest their important roles in flavonoid metabolism.

Levels of flavonoid accumulation are correlated with the expression levels of other genes involved in flavonoid synthesis[36]. FLS is a key enzyme that converts dihydroflavonol to flavonol, promoting flavonol accumulation[37]. In our study, FLS expression levels displayed variability, with two unigenes exhibiting higher expression in stems and leaves, while one unigene in roots. Our research highlights distinctive expression patterns of genes involved in flavonoid biosynthesis across the roots, stems, leaves, and seeds of *S. tonkinensis*. The combined expression of these enzyme-encoding genes is likely associated with changes in flavonoid compound concentrations and related metabolites in different tissues of *S. tonkinensis*.

Conversely, this study initially identified 23 flavonoid-related DEMs in different tissues of *S. tonkinensis*, including butein, galangin, L-epicatechol, tricetin, and naringenin chalcone (Fig. S6). Furthermore, we consider phenylpropanoid biosynthesis and flavonoid biosynthesis as crucial metabolic pathways for the differential metabolites in different *S. tonkinensis* tissues. The phenylpropanoid pathway constitutes the initial three steps in the flavonoid biosynthetic pathway, converting phenylalanine to 4-coumaroyl-CoA through phenylalanine ammonia-lyase (*PAL*) and 4-coumaroyl-CoA ligase (*4CL*). Subsequently, it sequentially converts 4-coumaroyl-CoA into flavonoids via *FLS*, *DFR*, *CHS*, *CHI*, and *ANR*[38].

In co-expression network analysis, we identified genes exhibiting similar expression patterns to flavonoid content, including potential flavonoid biosynthesis genes such as *CHI* (*StCHIs*, evm.model.1.2104 and evm.model.1.2101). *CHI* is a crucial rate-limiting enzyme in flavonoid biosynthesis in plants, catalyzing the intramolecular cyclization of chalcones into specific 2S-flavanones[39]. Additionally, we observed a close association between squalene monooxygenase (*StSQLE*, evm.model.5.860) and genes involved in flavonoid biosynthesis in *S. tonkinensis*. Squalene monooxygenase, a rate-limiting enzyme in the cholesterol biosynthetic pathway, catalyzes squalene cyclization[40, 41]. It is speculated that the epoxidation of squalene by *StSQLE* may indirectly impact the activity or expression levels of enzymes involved in the flavonoid biosynthesis of *S. tonkinensis*, thus influencing flavonoid synthesis.

Conclusions

This study compares DEGs and DEMs associated with alkaloid and flavonoid biosynthesis in different tissues of *S. tonkinensis* using transcriptome and metabolome datasets. We identified 2,727 DEGs, including 35 alkaloid-related DEGs and 48 flavonoid-related DEGs. Additionally, 296 DEMs were identified, including 23 alkaloid-related DEMs and 23 flavonoid-related DEMs. The analysis of metabolome data revealed the significance of phenylpropanoid biosynthesis and flavonoid biosynthesis as pivotal metabolic pathways in *S. tonkinensis*, aligning with the results of the integrated analysis. Furthermore, through WGCNA, we highlighted *StCAO* (evm.model.3.924) as a key gene regulating alkaloid metabolism and *StCHI* (evm.model.3.2047, evm.model.1.2104, and evm.model.1.2101) as crucial genes regulating flavonoid metabolism in *S. tonkinensis*. These findings provide insights for future investigations into the regulatory mechanisms of alkaloids and flavonoids in *S. tonkinensis*.

Material And Methods

Plant materials

S. tonkinensis seeds, leaves, stems, and roots were used in this study. All plant parts were sourced from the *S. tonkinensis* planting base in Huanjiang County, Hechi City, Guangxi Zhuang Autonomous Region, China (24°53'53" N, 107°53'29" E). The plant materials were formally identified by Professor Kunhua Wei and have been deposited in the Germplasm Repository of the Guangxi Botanical Garden of Medicinal Plants, Nanning, China, with voucher number YYZW20220070. Planting commenced at the end of March 2019, and sampling at the end of September 2022. The region experiences an annual average temperature of approximately 20–22 ℃, with January being the coldest at around 10 ℃ and July being the hottest at approximately 28 ℃. The roots, stems, leaves, and seeds of *S. tonkinensis* were labeled Root, Stem, Leaf, and Seed, respectively. Replicates comprised five healthy plants of similar size and growth, with three and six biological replicates per transcriptome sample and metabolome sample, respectively. All samples were rinsed with distilled water and rapidly placed in liquid nitrogen, followed by storage in a freezer at –80 ℃ for subsequent experimental analysis.

RNA isolation, library construction, and Illumina sequencing

Extraction of total RNA was performed on the tissue sample and the NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) was employed to assess RNA purity and concentration. Agilent 2100/Lab Chip GX was utilized to assess the RNA integrity. After testing the samples, a library was constructed by enriching eukaryotic mRNA with magnetic beads with oligo(dT). Thereafter, a fragmentation buffer treatment was then applied to fragment the mRNA. The first and the second cDNA strands were synthesized using mRNA as a template, after which the cDNA was purified. For sequencing, end repair, A-tail addition, and adapter ligation were performed on the purified double-stranded cDNA. AMPure XP beads were utilized to select fragment size, and the cDNA library was finally enriched through PCR.

Following library quality control, the Illumina Nova Seq6000 sequencing platform was used to sequence all mRNA transcripts, generating raw reads. Low-quality reads containing adapters, an N ratio over 10%, and a Q10 base quality value surpassing 50% of the entire read were removed, yielding high-quality clean reads.

Read mapping and annotation

Following the quality control process, mapped reads were obtained by aligning the clean reads with the *S. tonkinensis* genome in preparation for further analysis. TopHat2 software (http://tophat.cbcb.umd.edu/) was utilized to perform sequence alignment analysis[42]. The StringTie software (http://ccb.jhu.edu/software/stringtie/) [43]was utilized to assemble and concatenate the mapped reads.

Annotations for gene functions were based on the following databases: euKaryotic Orthologous Groups (KOG), Swiss-Prot (a manually annotated and reviewed protein sequence database), Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), evolutionary genealogy of genes: Non-supervised Orthologous Groups (eggNOG), Clusters of Orthologous Groups (COG), Protein family (Pfam), and National Center for Biotechnology Information (NCBI) non-redundant protein sequences (NR).

Differential gene expression analysis

The fragments per kilobase of transcript per million fragments (FPKM) value was used to quantify the levels of gene expression. Subsequently, differentially expressed genes (DEGs) between the two groups were identified by differential gene expression analysis. DESeq2 software[44] was employed for differential analysis, with the screening threshold set at an up- and down-regulation fold change > 2 and a significance level of $P < 0.05$.

Quantitative real-time PCR analysis

Quantitative real-time PCR (qPCR) was employed to determine the expression of selected genes. Twelve genes were selected, and specific primers were generated for them using the retrieved sequences. The reference gene used was evm.model.5.2407. The qPCR reaction was conducted on a qTOWER2.2 Real-Time PCR system (AnalytikJena, Jena, Germany) using 2×SYBR Green Supermix (Bio-Rad, Hercules, CA, USA). Table S4 contains a list of the qPCR primers. Three biological replicates per sample were used for analysis and the $2^{-\Delta\Delta Ct}$ method was employed to determine the relative level of expression[45].

Metabolite extraction

A 6 mm diameter grinding bead and 50 mg of the sample were placed in a 2 mL centrifuge tube. Then, 400 μL of extraction solution (methanol: acetonitrile=1:1(v:v)) was added in a low-temperature environment. The tissue was ground with a freezing tissue grinder for 6 min at −10 ℃ and 50 Hz. After 30 min of low-temperature ultrasound extraction (5 ℃, 40 KHz), the sample was left at −20 ℃ for another half an hour. The supernatant was transferred to an insert-coupled sample vial for analysis after

15 minutes of centrifugation at 4 ℃ and 13,000 ×*g*. Additionally, a quality control sample was prepared by transferring and mixing 20 μL of supernatant from each sample.

Mass spectrometry parameter settings and data collection

Chromatography conditions involved using an ACQUITY UPLC HSS T3 column (100 mm × 2.1 mm, 1.8 μm; Waters, Milford, MA, USA) with a column temperature of 40 ℃ and a flow rate of 0.4 mL/min. The injection volume was 3 μL. Mobile phase A comprised 95% water + 5% acetonitrile (with 0.1% formic acid) while mobile phase B included 47.5% acetonitrile + 47.5% isopropanol + 5% water (with 0.1% formic acid).

Mass spectrometry data were obtained utilizing an electrospray ionization source (ESI) in both negative and positive ion modes on a Q-Exactive HF-X mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA). The positive ion spray voltage was adjusted to 3500 V, whereas the negative ion spray voltage was −3500 V. The gas flow rates for the sheath and auxiliary were 50 and 13 arb, respectively. The capillary temperature was 325 ℃ The scanning range used for a full scan was 70-1050 m/z, with a resolution of 60,000, and the collision energy was set at 40 eV.

Conjoint analysis

KEGG pathway enrichment analysis involves evaluating selected differential gene sets and differential metabolic sets for enrichment. *P*-values were corrected via the Benjamini–Hochberg (BH) method, considering pathway enrichment significant when the corrected *P*-value was < 0.05.

Phytochemical determination

Alkaloid content determination: The sample was dried, crushed, and sieved through an 80-mesh screen. Approximately 0.1 g of the sample was weighed, and 1 mL of 80% ethanol was added. After mixing thoroughly, the mixture was transferred to an EP tube, subjected to ultrasound extracted for 60 min, and centrifugated for 10 minutes at 8,000 ×*g* and 25 ℃. We retrieved the supernatant for testing. Following the alkaloid content kit (Suzhou Michy Biomedical Technology Co., Ltd. Item No. M0122A) steps, the solution was vigorously shaken and allowed to settle for 40 minutes at 25 ℃, and the lower chloroform layer was absorbed with 200 μL, and the absorbance at 416 nm was measured.

Flavonoid content determination: The sample was dried to a constant weight, crushed, passed through a 40-mesh screen, and approximately 0.1 g was weighed. Subsequently, 1 mL of 60% ethanol was added, and shaken for 2 h at 60 ℃, followed by 10 min of centrifugation at 10,000 ×*g* and 25 ℃. We retrieved the supernatant for testing. Following the flavonoid content kit (Suzhou Michy Biomedical Technology Co., Ltd. Item No. M0118A) steps, the solution was mixed, and allowed to settle for 15 min at 25 ℃, and the absorbance at 510 nm was measured.

Matrine and oxymatrine content determination: We referred to the Chinese Pharmacopoeia (Commission, 2020) for the high-performance liquid chromatography (HPLC) method

under the item [Content determination] of *S. tonkinensis*. Following the grinding and weighing of 0.1 g of the sample, 1.0 ml of methanol was applied. After grinding into a slurry using a grinding instrument, ultrasonic extraction was conducted for 1 h. Centrifugation yielded the supernatant, which underwent HPLC analysis after filtration via a 0.45 mm filter. An Agilent 1100 high-performance liquid chromatography instrument (wavelength 220 nm) was utilized for the HPLC analysis. The following conditions were applied: chromatographic column, Compass C18 (2) reversed-phase chromatographic column (250 mm × 4.6 mm, 5 μm); injection volume, 10 μL; flow rate, 0.8 mL/min; column temperature, 30 ℃. The mobile phase comprised acetonitrile: mixed solution (3.4 g of potassium dihydrogen phosphate mixed with 500 mL of water, with 900 μL of triethylamine) = 1:9. To generate calibration curves, reference standards of matrine and oxymatrine were prepared by diluting with methanol to appropriate concentrations. Three biological and three technical replicates were utilized for each sample analysis. The reference standard chromatogram is shown in Fig. S7.

Genistin and genistein content determination: The sample was ground into powder. Approximately 0.1 g of the sample was weighed, and 1.0 mL of 70% methanol was added. The sample was ground into a slurry using a grinding instrument followed by a two-hour ultrasonic extraction. The needle filter was employed to filter the supernatant after centrifugation. An Agilent 1100 high-performance liquid chromatography (wavelength 260 nm) was utilized for HPLC analysis. The following conditions were applied: chromatographic column, Compass C18 (2) reversed-phase chromatographic column (250 mm × 4.6 mm, 5 μm); injection volume, 10 μL; flow rate, 1.0 mL/min; column temperature, 30 ℃. The mobile phase comprised methanol and 1%acetic acid water. Calibration curves were generated by preparing reference standards of genistin and genistein through dilution with methanol to the appropriate concentrations. Three technical and three biological replicates were used to examine each sample. The reference standard chromatogram is shown in Fig. S8.

Statistical analysis

One-way ANOVA with least significant difference (LSD) multiple comparisons was used for testing, and SPSS software version 22.0 was utilized to perform the statistical analyses. The format of means ± standard errors of the means was used to express the results. $P < 0.05$ was established as the significance criterion. Student's *t*-test was used to calculate the *P*-value for the significance of samples and qPCR assay results.

Abbreviations

S. tonkinensis: *Sophora tonkinensis*; DEGs: Differentially expressed genes; DEMs: Differentially expressed metabolites; KOG: euKaryotic Orthologous Groups; GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; eggNOG: Non-supervised Orthologous Groups; COG: Clusters of Orthologous Groups; NR: non-redundant protein sequences; NCBI: National Center for Biotechnology Information; Pfam: Protein family; LC-MS: Liquid chromatography-mass spectrometry; FPKM: The

fragments per kilobase of transcript per million fragments; PCA: principal component analysis; qRT-PCR: Quantitative reverse-transcription PCR; HPLC: High-performance liquid chromatography.

Declarations

Ethics approval and consent to participate

The experiments did not involve endangered or protected species. The data collection of plants was carried out with permission of related institution, and complied with national or international guidelines and legislation.

Consent for publication

Not applicable.

Data availability

The original contributions presented in the study are publicly available. The data have been deposited in the NCBI Sequence Read Archive database under accession number: PRJNA1052132.

Competing interests

The authors declare no competing interests.

Funding

This research was supported by the National Natural Science Foundation of China (82060689, 82260747, 82260749), Guangxi Natural Science Foundation (2020GXNSFBA159018), Guangxi Science and Technology Project (GuiKe AD22080012), Guangxi elite team of medicinal plant conservation.

Author contributions

SQ and FW conceived and designed the research. YL conducted the experiments, analyzed data and wrote the manuscript. GW and XL revised the manuscript. MT, HH, DT, YL and LL assist during the experiments. All authors read and approved the manuscript.

Acknowledgments

We appreciate the cooperation of all the people that help in anything related with this research.

Approval, accordance and (for human subjects) informed consent statements for reporting experiments on live vertebrates and/or higher invertebrates

Not applicable

Appropriate permission and/or credit for reproduced images

Not applicable

References

1. Liang Y, Wei F, Qin S, Li M, Hu Y, Lin Y, et al. Sophora tonkinensis: response and adaptation of physiological characteristics, functional traits, and secondary metabolites to drought stress. Plant Biol (Stuttg). 2023; 25(7):1109-1120.
2. Chinese Pharmacopoeia Commission. Pharmacopoeia of the People's Republic of China (part I). Beijing: China Medical Science Press; 2020.
3. Chen Z, Mo X, Shen C, Jiang L, Cai J, Wei K. Research progress on extraction and biological activity of effective components of *Sophora Tonkinensis* Gapnep. Applied Chemical Industry. 2018; 47(6):1237-1240.
4. Nie A, Zhao X, Gao M, Chao Y, Li X, Gui X, et al. Discussion and consideration on safety of *Sophorae Tonkinensis* Radix et Rhizoma and its rational use. Chinese Traditional and Herbal Drugs. 2018; 49(17):4152-4161.
5. Zhang S, Guo S, Gao XB, Liu A, Jiang W, Chen X, et al. Matrine attenuates high-fat diet-induced in vivo and ox-LDL-induced in vitro vascular injury by regulating the PKCα/eNOS and PI3K/Akt/eNOS pathways. J Cell Mol Med. 2019; 23(4):2731-2743.
6. Gao Y, Wu C, Huang J, Huang Z, Jin Z, Guo S, et al. A new strategy to identify ADAM12 and PDGFRB as a novel prognostic biomarker for matrine regulates gastric cancer via high throughput chip mining and computational verification. Comput Biol Med. 2023; 166:107562.
7. Hao X, Yan W, Yang J, Bai Y, Qian H, Lou Y, et al. Matrine@chitosan-D-proline nanocapsules as antifouling agents with antibacterial properties and biofilm dispersibility in the marine environment. Front Microbiol. 2022; 13:950039.
8. Fu Y, Yu D, Wang S, Yang L, Deng Z. Research progress on pharmacological effects and mechanisms of flavonoids from Sophorae Tonkinensis Radix et Rhizoma. Chinese Traditional and Herbal Drugs. 2022; 53(19):6234-6244.
9. Wei KH, Xu JP, Li LX, Cai JY, Miao JH, Li MH. In vitro Induction and Generation of Tetraploid Plants of Sophora tonkinensis Gapnep. Pharmacogn Mag. 2018; 14(54):149-154.
10. Liang Y, Lin Y, Qin S, Wei F, Wei G, Miao J, et al. Influence of soil nutrient on the yield and quality of *Sophora tonkinensis* in producing area. Biotic Resources. 2023; 45(3):259-266.
11. Jin X, Lu Y, Chen S, Chen D. UPLC-MS identification and anticomplement activity of the metabolites of *Sophora tonkinensis* flavonoids treated with human intestinal bacteria. J Pharm Biomed Anal. 2020; 184:113176.
12. He L, Liu J, Luo D, Zheng Y, Zhang Y, Wang G, et al. Quinolizidine alkaloids from *Sophora tonkinensis* and their anti-inflammatory activities. Fitoterapia. 2019; 139:104391.

13. Wu C, He L, Yi X, Qin J, Li Y, Zhang Y, et al. Three new alkaloids from the roots of *Sophora tonkinensis*. J Nat Med. 2019; 73(3):667-671.
14. Zhang J, Liu Y, Fang J. The biological activities of quinolizidine alkaloids. The Alkaloids. Chemistry and biology. 2023; 89:1-37.
15. Morgenthal K, Weckwerth W, Steuer R. Metabolomic networks in plants: Transitions from pattern recognition to biological interpretation. Biosystems. 2006; 83(2-3):108-117.
16. Tang D, Huang Q, Wei K, Yang X, Wei F, Miao J. Identification of Differentially Expressed Genes and Pathways Involved in Growth and Development o*f Mesona chinens*is Benth Under Red- and Blue-Light Conditions. Front Plant Sci. 2021; 12:761068.
17. Vasav AP, Pable AA, Barvkar VT. Differential transcriptome and metabolome analysis of *Plumbago zeylanica* L. reveal putative genes involved in plumbagin biosynthesis. Fitoterapia. 2020; 147:104761.
18. He Y, Chen H, Zhao J, Yang Y, Yang B, Feng L, et al. Transcriptome and metabolome analysis to reveal major genes of saikosaponin biosynthesis in *Bupleurum chinens*e. Bmc Genomics. 2021; 22(1):839.
19. Zhou Y, Yao L, Huang X, Li Y, Wang C, Huang Q, et al. Transcriptomics and metabolomics association analysis revealed the responses of *Gynostemma pentaphyllu*m to cadmium. Front Plant Sci. 2023; 14:1265971.
20. Thonglim A, Bortolami G, Delzon S, Larter M, Offringa R, Keurentjes J, et al. Drought response in Arabidopsis displays synergistic coordination between stems and leaves. J Exp Bot. 2023; 74(3):1004-1021.
21. Verma N, Shukla S. Impact of various factors responsible for fluctuation in plant secondary metabolites. J Appl Res Med Aromat Plants. 2015; 2(4):105-113.
22. Zifkin M, Jin A, Ozga JA, Zaharia LI, Schernthaner JP, Gesell A, et al. Gene expression and metabolite profiling of developing highbush blueberry fruit indicates transcriptional regulation of flavonoid metabolism and activation of abscisic acid metabolism. Plant Physiol. 2012; 158(1):200-224.
23. Li Y, Wu H. The research progress of the correlation between growth development and dynamic accumulation of the effective components in medicinal plants. Chinese Bulletin of Botany. 2018; 53(3):293-304.
24. Bauermeister A, Mannochio-Russo H, Costa-Lotufo LV, Jarmusch AK, Dorrestein PC. Mass spectrometry-based metabolomics in microbiome investigations. Nat Rev Microbiol. 2022; 20(3):143-160.
25. You L, Yang C, Du Y, Wang W, Sun M, Liu J, et al. A systematic review of the pharmacology, toxicology and pharmacokinetics of matrine. Front Pharmacol. 2020; 11:1067.
26. Zou J, Zhao L, Yi P, An Q, He L, Li Y, et al. Quinolizidine alkaloids with antiviral and insecticidal activities from the seeds of *Sophora tonkinensis* Gagnep. J Agric Food Chem. 2020; 68(50):15015-15026.

27. Huang J, Xu H. Matrine: bioactivities and structural modifications. Curr Top Med Chem. 2016; 16(28):3365.
28. Schultz CJ, Goonetilleke SN, Liang J, Lahnstein J, Levin KA, Bianco-Miotto T, et al. Analysis of genetic diversity in the traditional chinese medicine plant 'Kushen' (*Sophora flavescens* Ait.). Front Plant Sci. 2021; 12:704201.
29. Liang Y, Wei K, Wei F, Qin S, Deng C, Lin Y, et al. Integrated transcriptome and small RNA sequencing analyses reveal a drought stress response network in *Sophora tonkinensis*. Bmc Plant Biol. 2021; 21(1):566.
30. Frick KM, Kamphuis LG, Siddique KHM, Singh KB, Foley RC. Quinolizidine alkaloid biosynthesis in lupins and prospects for grain quality improvement. Front Plant Sci. 2017; 8:87.
31. Yonekura-Sakakibara K, Fukushima A, Saito K. Transcriptome data modeling for targeted plant metabolic engineering. Curr Opin Biotechnol. 2013; 24(2):285-290.
32. Golebiewski WM, Spenser ID. Biosynthesis of the lupine alkaloids. II. Sparteine and lupanine. Can J Chem. 1988; 66(7):1734-1748.
33. Tang M, Li Z, Luo D, Wei F, Kashif MH, Lu H, et al. A comprehensive integrated transcriptome and metabolome analyses to reveal key genes and essential metabolic pathways involved in CMS in kenaf. Plant Cell Rep. 2021; 40(1):223-236.
34. Yuan Y, Zhang J, Liu X, Meng M, Wang J, Lin J. Tissue-specific transcriptome for *Dendrobium officinale* reveals genes involved in flavonoid biosynthesis. Genomics. 2020; 112(2):1781-1794.
35. Furumura S, Ozaki T, Sugawara A, Morishita Y, Tsukada K, Ikuta T, et al. Identification and functional characterization of fungal chalcone synthase and chalcone isomerase. J Nat Prod. 2023; 86(2):398-405.
36. Gichuki DK, Li Q, Hou Y, Liu Y, Ma M, Zhou H, et al. Characterization of flavonoids and transcripts involved in their biosynthesis in different organs of *Cissus rotundifolia* Lam. Metabolites. 2021; 11(11):741.
37. Kumari G, Nigam VK, Pandey DM. The molecular docking and molecular dynamics study of flavonol synthase and flavonoid 3'-monooxygenase enzymes involved for the enrichment of kaempferol. J Biomol Struct Dyn. 2023; 41(6):2478-2491.
38. Chen J, Tang W, Li C, Kuang D, Xu X, Gong Y, et al. Multi-omics analysis reveals the molecular basis of flavonoid accumulation in fructus of Gardenia (*Gardenia jasminoides* Ellis). Bmc Genomics. 2023; 24(1):1-588.
39. Shah F, Baharum SN, Goh HH, Leow TC, Ramzi AB, Oslan SN, et al. Molecular cloning and in silico analysis of chalcone isomerase from *Polygonum minus*. Mol Biol Rep. 2023; 50(6):5283-5294.
40. Yoshioka H, Coates HW, Chua NK, Hashimoto Y, Brown AJ, Ohgane K. A key mammalian cholesterol synthesis enzyme, squalene monooxygenase, is allosterically stabilized by its substrate. Proc Natl Acad Sci U S a. 2020; 117(13):7150-7158.
41. Chua NK, Coates HW, Brown AJ. Squalene monooxygenase: a journey to the heart of cholesterol synthesis. Prog Lipid Res. 2020; 79:101033.

42. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 2013; 14(4):R36.
43. Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol. 2015; 33(3):290-295.
44. 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.
45. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 2001; 25(4):402-408.

Figures

Figure 1

Phenotypic characteristics of the entire plant (A), roots (B), stems (C), leaves (D), and seeds (E) of *S. tonkinensis*.

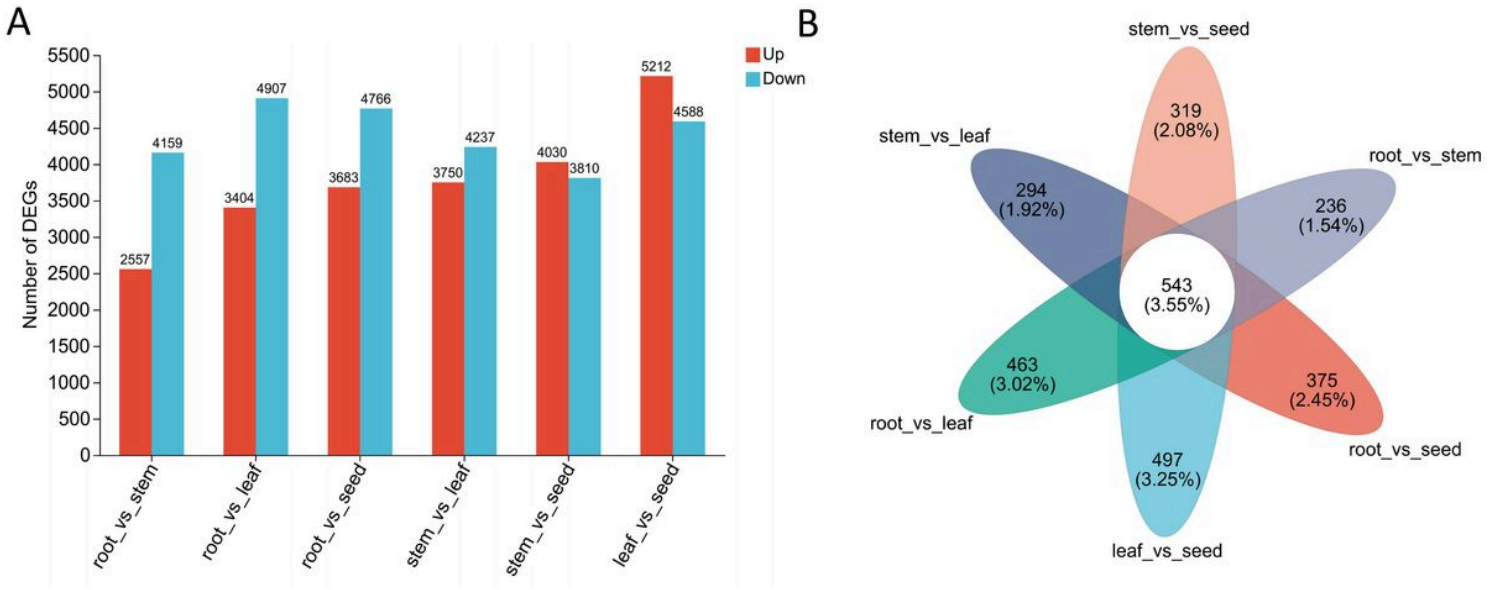

Figure 2

Analysis of differentially expressed genes (DEGs) in various *S. tonkinensis* tissues (A) Comparison of up- and down-regulated DEGs in various plant parts. (B) DEGs in various plant parts are depicted in a Venn diagram.

Figure 3

Verification of selected 12 DEGs. (A) RNA-seq heatmap analysis. (B) qPCR value difference analysis. Different letters on the bars indicate the significant difference at the 0.05 level.

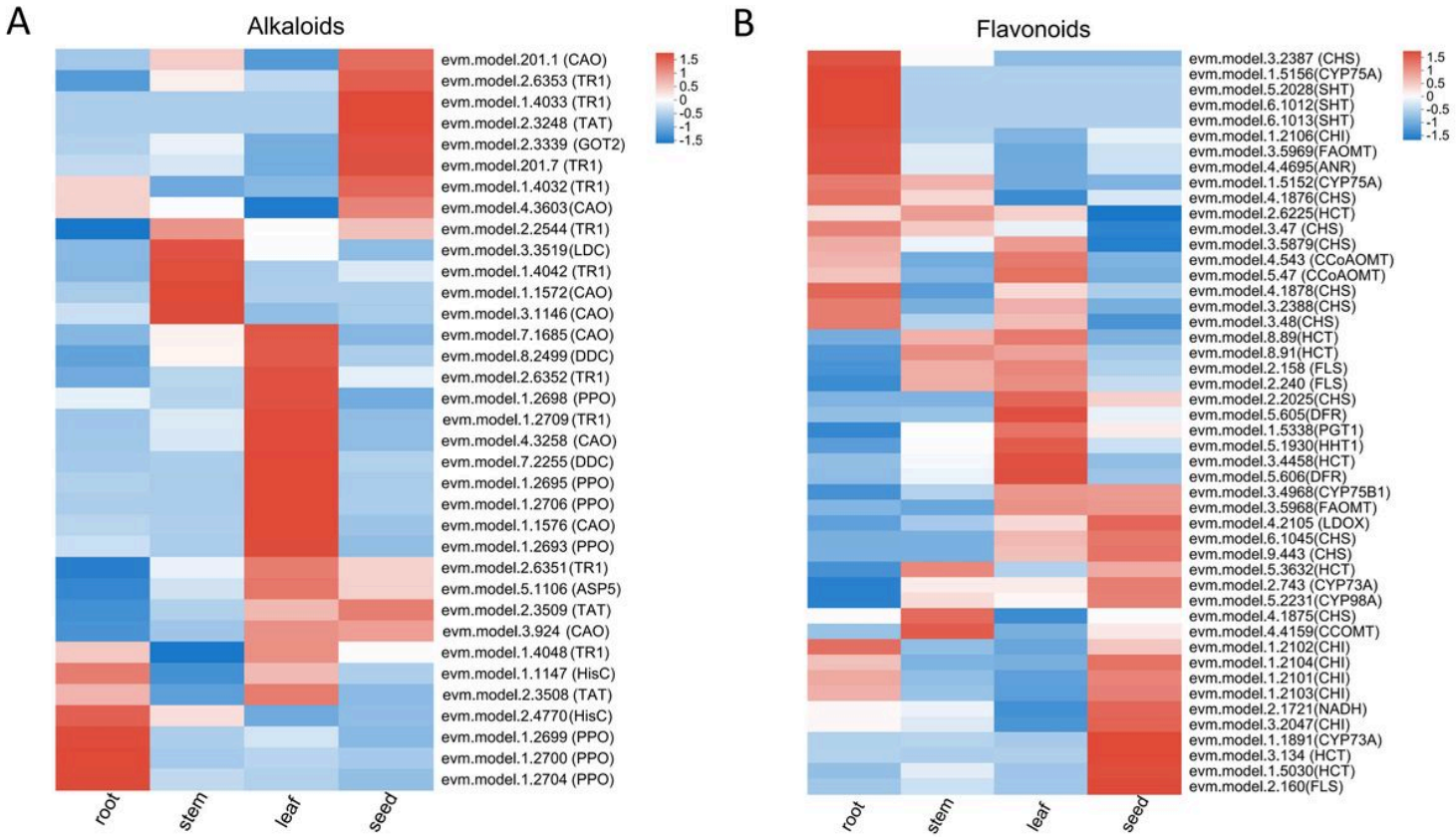

Figure 4

DEGs associated with alkaloid and flavonoid biosynthesis in different tissues of *S. tonkinensis.* (A) Genes related to alkaloid biosynthesis. (B) Genes related to flavonoid biosynthesis.

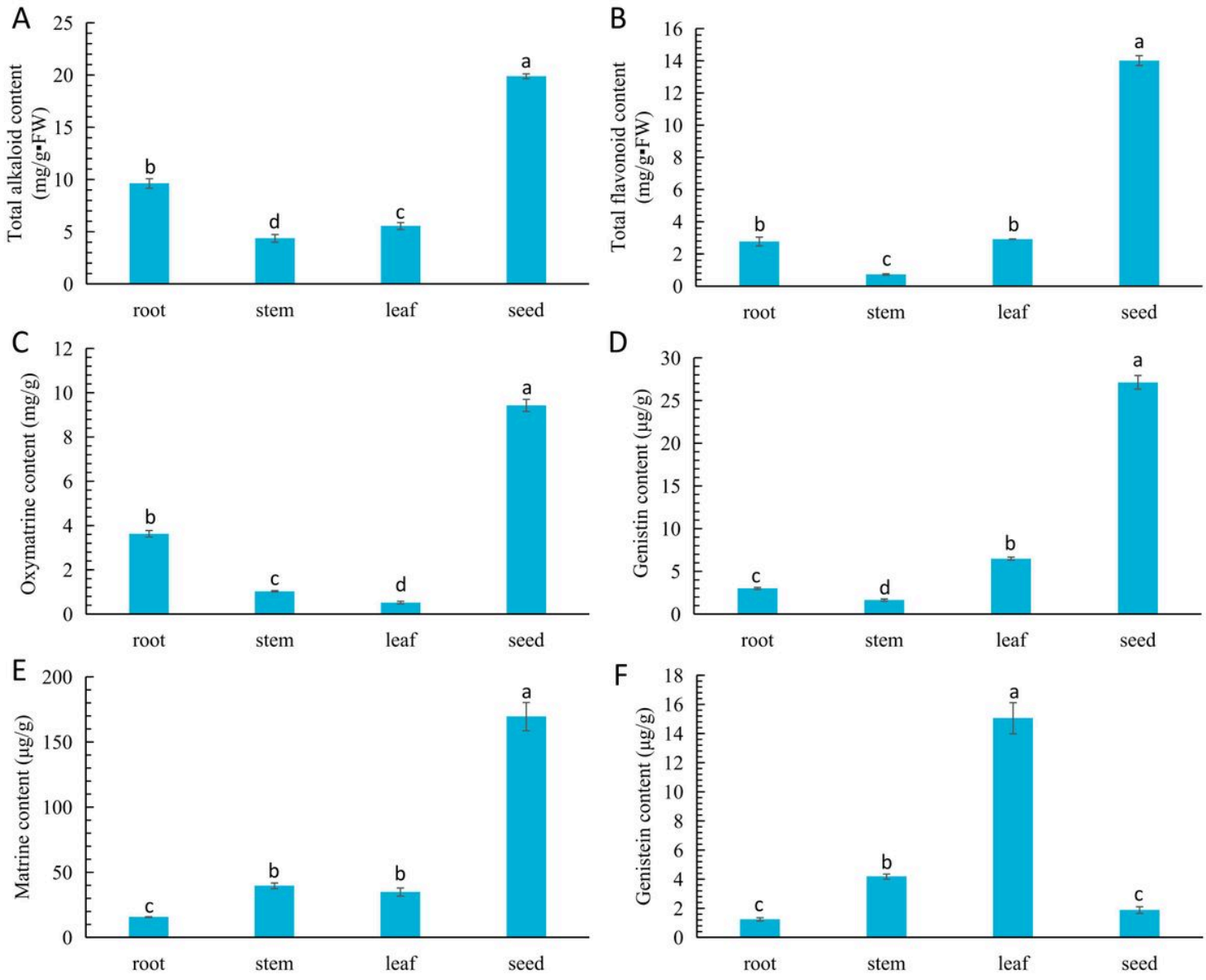

Figure 5

Analysis of alkaloid and flavonoid compounds in distinct *S. tonkinensis* tissues. (A) Total content of alkaloid, (B) flavonoid, (C) Oxymatrine, (D) Genistin, (E) Matrine, and (F) Genistein content. Different letters on the bars indicate the significant difference at the 0.05 level.

Figure 6

Principal component analysis (PCA) and differentially expressed metabolite (DEM) analysis of different tissues of *S. tonkinensis*. (A) Metabolite PCA of LC-MS in each tissue. (B) Cluster heatmap between samples of different tissues. (C) DEMs that are up-and down-regulated in certain tissues. (D) DEMs in various tissues are depicted as a Venn diagram.

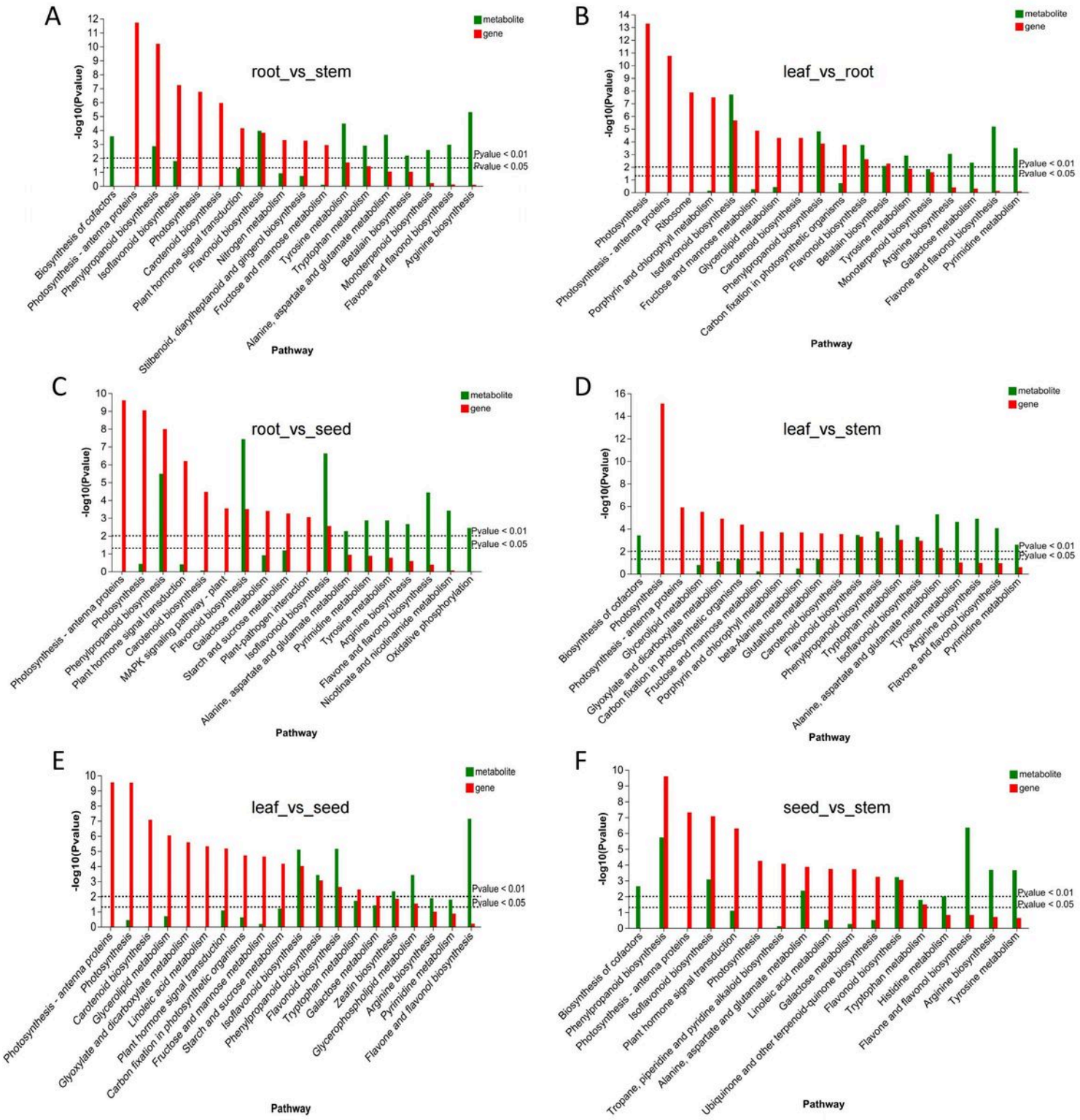

Figure 7

KEGG pathway enrichment analysis of DEGs and DEMs in the transcriptome and metabolome. (A) root_vs_stem KEGG pathway enrichment analysis. (B) leaf_vs_root KEGG pathway enrichment analysis. (C) root_vs_seed KEGG pathway enrichment analysis. (D) leaf_vs_stem KEGG pathway enrichment analysis. (E) leaf_vs_seed KEGG pathway enrichment analysis. (F) seed_vs_stem KEGG pathway enrichment analysis. The metabolic pathways are shown along the horizontal axis, whereas the vertical

axis denotes the enriched *P*-values of DEGs (red) and DEMs (green), expressed using $-\log^{P}$. The genes and metabolites enriched in the pathways with *P*-values≤0.05 are marked, and the top 10 pathways are selected for each comparison.

Figure 8

Networks of co-expression of transcripts that are involved in the biosynthesis and metabolism of alkaloids, flavonoids, oxymatrine, matrine, genistin, and genistein. (A) Fifteen modules discovered by weighted gene co-expression network analysis (WGCNA) are shown by a hierarchical cluster tree and color bands. The genes that are not classified into specific modules are shown by gray modules. Every branch in the tree corresponds to a certain gene. (B) Module–trait correlation analysis. Columns correspond to distinct chemical compounds, whereas rows correspond to modules. The left box indicates the number of genes present in each module. Intersections of rows and columns display correlation coefficients and P-values between chemical compounds and modules. (C) Co-expression subnetwork analysis of magenta modules related to alkaloid and oxymatrine accumulation. (D) Co-expression subnetwork analysis of tan modules related to flavonoid, genistin and matrine accumulation. (E) Co-expression subnetwork analysis of greenyellow modules related to genistein accumulation.

Supplementary Files

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

- SupplementaryFiguresS1S8.docx
- SupplementaryTablesS1S4.xlsx