horticulturae

MDPI

Article

Transcriptomics Analysis Reveals Differences in Purine and Phenylpropanoid Biosynthesis Pathways Between *Camellia sinensis* var. Shuchazao and *Camellia ptilophylla*

Waqar Khan, Peng Zheng, Binmei Sun * and Shaoqun Liu *

College of Horticulture, South China Agricultural University, Guangzhou 510642, China; waqar.khan399@scau.edu.cn (W.K.); zhengp@scau.edu.cn (P.Z.)
* Correspondence: binmei@scau.edu.cn (B.S.); scauok@scau.edu.cn (S.L.)

Abstract: Tea production and quality are largely determined by the many genetic and biochemical characteristics that occur in tea plant cultivars. Worldwide, tea is consumed for its pleasing and refreshing effects due to its caffeine content. The present study performed transcriptomics analyses of two tea species (*Camellia sinensis* var. Shuchazao (SCZ) and *Camellia ptilophylla* (CAF)) and identified diversity in the gene expression levels and major regulatory transcription factors (TFs) for the characterization of purine alkaloids and phenylpropanoid biosynthesis pathways. The RNA-seq analysis of two species (SCZ and CAF) revealed the differences in caffeine and catechins synthesis. In the purine alkaloid biosynthesis pathway, the S-adenosyl methionine (SAM) and adenosine monophosphate (AMP) pathway genes were significantly related to xanthosine synthesis in contrasting purine alkaloids among (*Camellia sinensis* var. Shuchazao (SCZ) and *Camellia ptilophylla* (CAF)). The significant expression of *SAMS-5, PPAT-2, IMPDH-2, TCS-2, TCS-3, XMT-1, XMT-13,* and *XDH-4* in the xanthosine degradation pathway in CAF is attributed to higher theobromine content as compared to SCZ. Moreover, the transcription factors (TFs) AP2/ERF (20%), WRKY (12%), NAC (11%), and MYB (8%) were significantly correlated. The upregulated expression of caffeine synthesis genes in SCZ was correlated with MYB and AP2/ERF transcription factors. This study provides the basis for differences in the genetic mechanism in purine alkaloids, phenylpropanoid, and flavonoid biosynthesis pathways, which would be helpful in the development and selection of tea plant species with high or low caffeine concentrations. This study also provides a road map for future genetic improvement in tea species and cultivars.

Keywords: *Camellia sinensis*; *Camellia ptilophylla*; transcriptomics; purine; phenylpropanoid; pathways

check for updates

Academic Editor: Xinchao Wang

Received: 3 December 2024
Revised: 22 December 2024
Accepted: 24 December 2024
Published: 26 December 2024

Citation: Khan, W.; Zheng, P.; Sun, B.; Liu, S. Transcriptomics Analysis Reveals Differences in Purine and Phenylpropanoid Biosynthesis Pathways Between *Camellia sinensis* var. Shuchazao and *Camellia ptilophylla*. *Horticulturae* **2025**, *11*, 8. https://doi.org/10.3390/horticulturae11010008

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Camellia sinensis, or the tea plant, is a vital leaf crop for economic purposes [1], and the chemical constituents of the tender shoots can have considerable effects on the tea's flavor [2]. It was earlier believed that *Camellia ptilophylla* solely contained theobromine rather than caffeine, although recent research refutes this theory [3]. *Camellia ptilophylla* containing various purine base compositions was analyzed using the random amplified polymorphic DNA (RAPD) technique, and it was reported that only theobromine-containing plants were related, while the plants containing caffeine and theobromine showed less relation to *Camellia ptilophylla* plants [4]. Tea is the most favored non-alcoholic beverage after water due to its rich flavor and many health advantages [5]. Tea, coffee, and cocoa

are all non-alcoholic beverages containing purine alkaloids with different concentrations, such as 3,7-dimethylxanthine (theobromine), 1,3,7-trimethylxanthine (caffeine), and 1,3,7,9-tetramethyluric acid (theacrine). Theacrine is found in some tea plants and has different roles. Not only are the tea quality and taste influenced by purine alkaloids, but these alkaloids have additional health advantages [6]. Various secondary metabolites like caffeine, catechins, terpenoids, and theanine are the primary sources of tea's taste, flavor, and health benefits [7–9].

In most tea plants, the primary purine alkaloid among the reported metabolites is caffeine. Caffeine is known worldwide for its capacity to activate the central nervous system [10]. Theacrine and theobromine are typical purine alkaloids and are also found in various tea plants varieties and species. Purine alkaloids also have very insignificant quantities of pyrimidine bases [11–13]. The position and number of methyl group on the purine ring differentiate the chemical structures based on the skeletons of xanthine and uric acid [14]. In China's National Germplasms Hangzhou Tea Repository (CNGTR), more than 580 tea germplasms were preserved and analyzed for their caffeine content, which varies from 1.2–5.9%, with an average value of 4.2% [15]. In different tea varieties and species, the purine alkaloid content varies substantially; the range of caffeine content in tea plants is 25–45 milligrams per gram, while the range of theobromine is 2.4–4.3 milligrams per gram [12,16]. However, consuming too much caffeine may have negative consequences on health, including headaches, sleeplessness, and changes in intraocular pressure [6,17].

Previous studies reported that in caffeine biosynthesis, S-adenosyl methionine (SAM) is an active methyl donor. The organic compound xanthine is thought to be a very significant substance in the synthesis of caffeine [18]. The SAM → SAH → adenosine → adenine → AMP → IMP → XMP → xanthosine pathway was proposed as the primary route from S-adenosylmethionine (SAM) to xanthosine; it was additionally suggested that SAM was the source for the purine ring in caffeine [19]. However, other pathways, like AMP, GMP, and de novo, can also synthesize xanthine. The primary route for the synthesis of caffeine from xanthosine and the four steps reactions are xanthosine → 7-methylxanthosine → 7-methylxanthine → theobromine → caffeine. One nucleosidase step and three methylation steps are involved [11]. In the caffeine biosynthesis pathway in tea plants, the important enzyme is tea caffeine synthase (TCS) which promotes the methylation of N-3 and N-1.

The phenylpropanoid pathway is one of the specialized metabolic pathways that has been studied the most. The plant shikimate pathway is the earliest step in phenylpropanoid biosynthesis. *Cinnamate 4-hydroxylase*, *4-coumaroyl CoA-Ligase*, and *PAL* catalyze the first three steps of the route, which are required to form the basis for all other branches and the metabolites that follow. The small gene families are responsible for the subsequent steps downstream of *PAL* in all species that have been studied, resulting in phenylpropanoid monomers (except those from *cinnamate 4-hydroxylase*) [20]. The next stage directs the resultant fragrant CoA esters through various biosynthetic pathways through the *4-CL* ligases gene family. In phenylpropanoid biosynthesis, the most significant branch point is most likely *4-coumaroyl CoA*. Approximately 6000 compounds make up the most significant class of polyphenolic metabolites, which is produced by the substantial branch in phenylpropanoid metabolism known as flavonoid metabolism [21,22]. The synthesis of a particular flavonoid, which starts with the formation of chalcones, is signified by the inclusion of p-coumaroyl-CoA into the flavonoid biosynthesis pathway [23]. *Chalcone synthase* (*CHS*), a polyketide synthase, is the key enzyme in the flavonoid biosynthetic pathway, performing the conversion of one molecule of p-coumaroyl-CoA and three molecules of malonyl-COA to naringenin chalcone by the activity of acetyl-CoA carbooxylase [24–26].

The present study compared the *Camellia sinensis* var. Shuchazao (SCZ) (moderate to high caffeine content) and *Camellia ptilophylla* (CAF) (very low caffeine content) species to

investigate the mechanisms of alkaloid, phenylpropanoid, and flavonoid biosynthesis in two tea species through RNA-seq study.

2. Materials and Methods

2.1. Plant Materials

The leaf samples were collected from *Camellia sinensis* var. Shuchazao (SCZ) and *Camellia ptilophylla* (CAF) plants grown in Guangzhou, Guangdong Province, China, in the same environmental condition for both the species. The leaf sample were stored at −80 °C for RNA extraction and transcriptomics analysis.

2.2. RNA Extraction and Sequencing

The collected samples were divided in two groups: *Camellia sinensis* var. Shuchazao, SCZ (SCZ-1, SCZ-2, and SCZ-3) and *Camellia ptilophylla*, CAF (CAF-1, CAF-2, CAF-3). Fresh tea leaves were used for transcriptome sequencing, and three biological replicates were taken for each sample. Liquid nitrogen was used to grind the samples into a fine powder. RNA was extracted according to the instructions of the kit (Hipure plant RNA micro kit B, R4151-02B, Magne Biotechnology Co., Ltd., Guangzhou, China). The NanoDrop2000spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) was used to quantify the RNA concentration. Following the previous studies [27], the cDNA libraries were contracted from each sample. The RNA-seq data's quality was assessed using the FastQC v0.12.1 software [28], and reads with more than 95% bases and a base quality score of 20 were retained. Hisat2 was used to align paired-end clean reads to the reference genome that could be found in the TPIA database (http://tpdb.shengxin.ren/, accessed on 12 February 2024) [29].

2.3. Functional Annotation and Gene Expressions

The sample gene sequences were compared against the Kyoto Encyclopedia of Genes (KEGG) and Genomes and Gene Ontology (GO) databases for the functional annotation of the proteins in *Camellia sinensis* var. Shuchazao and *Camellia ptilophylla*. The Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses of differentially expressed genes (DEGs) sets were carried out using the KOBAS and topG package (www.bioconductor.org/packages/release/bioc/html/topGO.html, accessed on 14 February 2024) [30].

2.4. Real-Time qRT-PCR

For the real-time PCR, the primers were designed by Premier 6 (Supplementary Table S1). BioRad CFX384TM Real-Time System (Bio-Rad, Hercules, CA, USA) was used to perform the qRT-PCR [31]. Relative expression levels were calculated based on the $2^{-\Delta\Delta CT}$ method, and *CsActin1* was used as the internal reference gene [32].

2.5. Statistical Analysis

Graph-pad Prism 9 software was used to create the double-coordinate figure and correlation analysis was made in Origin 2022b R1. The heat maps were constructed with TBtools [33]. To identify significant differences, Duncan's multiple-range test was performed with IBM SPSS statistics 25 at p-value < 0.05 to reveal significant differences. With Cytoscape (3.10.2), the co-expression network was shown. Each dataset is the mean ± SD (n = 3).

3. Results

3.1. RNA-Seq Analysis

RNA-seq analyses were performed to study the mechanism of differential accumulation levels of purine biosynthesis and degradation, as well as phenylpropanoid and flavonoid biosynthesis in two tea species, *Camellia sinensis* var. Shuchazao (SCZ) and *Camellia ptilophylla* (CAF). A total of six libraries were constructed. These libraries collectively produced a substantial total of 44.85 gigabytes (Gb) of clean reads. The average value of clean reads per sample equaled 7.47 Gb, showcasing the depth and thoroughness of the sequencing effort. The quality of the sequencing data was exceptionally high, as evidenced by the base-calling accuracy metrics. The Q20 and Q30 values, which are indicative of the proportion of reads with a base call accuracy of 95.92% and 90.26%, respectively, were both above 90.00%. Furthermore, the average GC content across all samples was found to be 45.74%, as detailed in (Table S7). This value is consistent with the expected GC content for the organism under study, providing additional validation for the accuracy of the sequencing process. In addition to the high-quality sequencing data, Pearson correlation analysis was performed to assess the repeatability and consistency among the samples. The results of this analysis, shown in Figure 1, revealed a high degree of correlation between the samples. This high repeatability and consistency indicate the robustness and reliability of the data generated by the sequencing methodology.

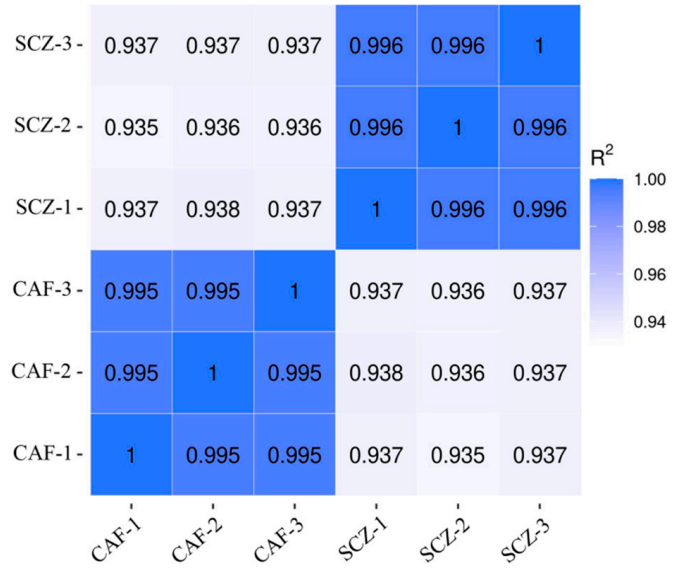


Figure 1. Pearson correlation coefficient of three replicates of CAF and SCZ species.

3.2. Differentially Expressed Genes (DEGs) Analysis

The transcriptome data analysis of *Camellia sinensis* var. Shuchazao (SCZ) and *Camellia ptilophylla* (CAF) revealed 12,728 differentially expressed genes (DEGs), of which 7247 were upregulated while 5481 were noted as downregulated. The proportions of upregulated and downregulated DEGs in CAF vs SCZ were 56.91% vs. 43.05%, respectively. Figure 2 shows that there were more upregulated DEGs than downregulated DEGs.

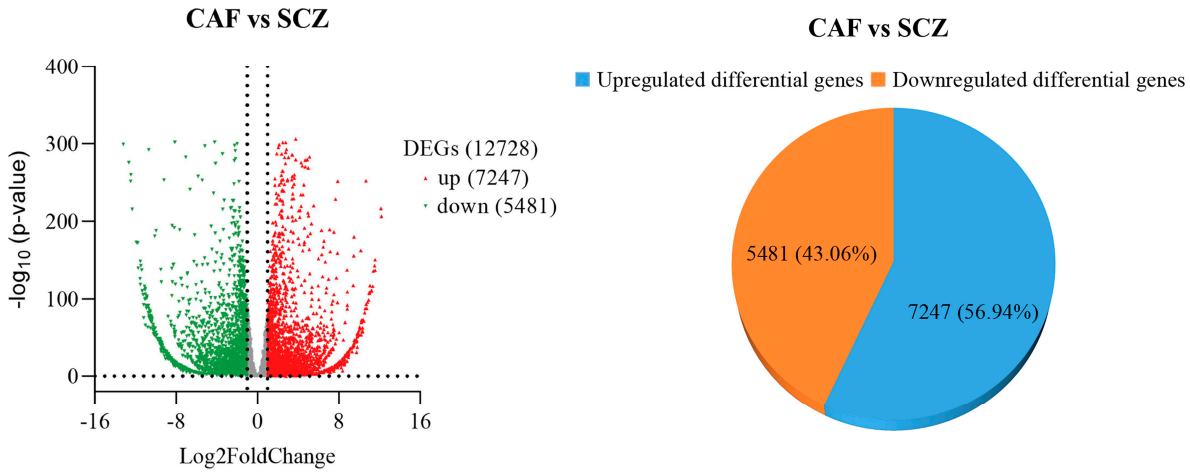


Figure 2. The volcano map of differentially expressed genes (DEGs) in CAF and SCZ species. The percentage distribution of upregulated and downregulated genes is also shown in the figure.

3.3. Function Analysis and Annotation of Differentially Expressed Genes

Go analysis revealed that the three primary categories of gene annotations were cell fractions (CCs), molecular functions (MFs), and biological processes (BPs). The responses to stress, organic substances, oxygen-containing compounds, endogenous stimuli, hormones, external stimuli, biotic stimuli, other organisms, acidic chemicals, immune response, innate immune response, and fungal responses were the main biological processes for which the differentially expressed genes were enriched, as shown in Figure 3. In the cellular components, the differentially expressed genes were mainly enriched in plastid, cytosol, chloroplast, membrane part, non-membrane-bounded organelle, protein-containing complex, extracellular region, intracellular non-membrane-bounded organelle, vacuole, chloroplast part, and nucleoplasm. In the molecular function, the differentially expressed genes were mainly enriched in transferase activity, small molecule binding, hydrolase activity, nucleoside phosphate binding, catalytic activity, and purine nucleotide binding. According to these data, the three classes of differentially expressed genes displayed distinct GO functions (Figure 3).

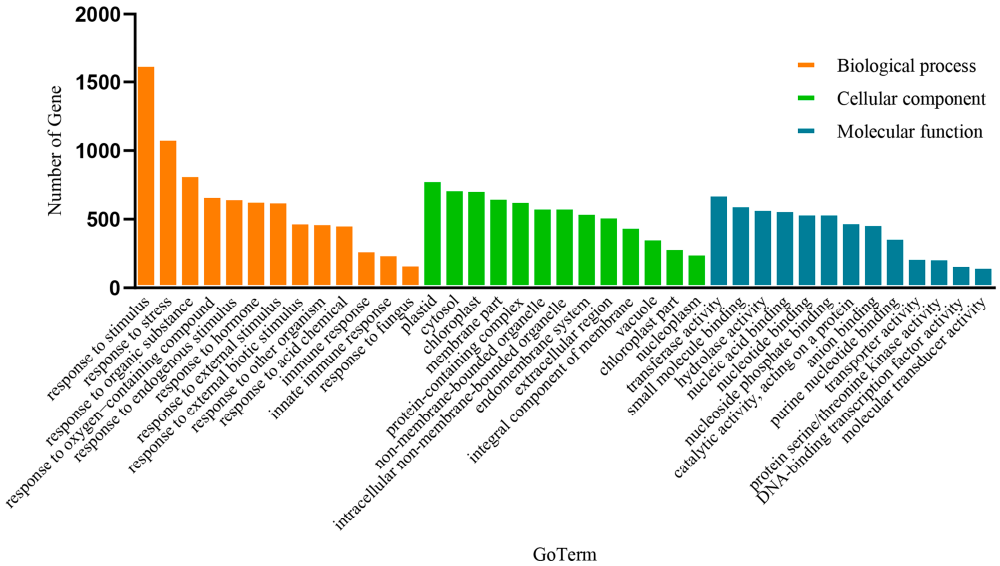


Figure 3. GO analysis of differentially expressed genes (DEGs) in CAF and SCZ species.

3.4. Pathway Analysis of Differentially Expressed Genes

A total of 128 metabolic pathways were identified by KEGG enrichment analysis in the SCZ and CAF. The 128 metabolic pathways included ribosome, phenylpropanoid biosynthesis, plant–pathogen interaction, starch and sucrose metabolism, purine metabolism, plant hormone signal transduction, flavonoid biosynthesis, and glutathione metabolism. Fuhrer analysis of pathways was performed based on the synthesis of caffeine and catechins. In phenylpropanoid biosynthesis, 29 DEGs were enriched; in purine metabolism and flavonoid biosynthesis, 20 DEGs were enriched in each pathway, respectively (Figure 4A). The main focused pathways for further study were purine metabolism, phenylpropanoid, and flavonoid pathways in the present study. The purine metabolism pathway is responsible for the biosynthesis and degradation of caffeine. For catechins synthesis, phenylpropanoid and flavonoid biosynthesis were studied. In addition, 14 genes were chosen, and the RT-qPCR was performed to validate the RNA-Seq data. A list of the primers used in RT-qPCR is given in the Supplementary Materials (Table S1). The RT-qPCR analysis results were in accordance with the RNA-Seq analysis, which shows the reliability of the sequence data (Figure 4B). Genes in the KEGG pathway are listed in the Supplementary Materials (Table S4).

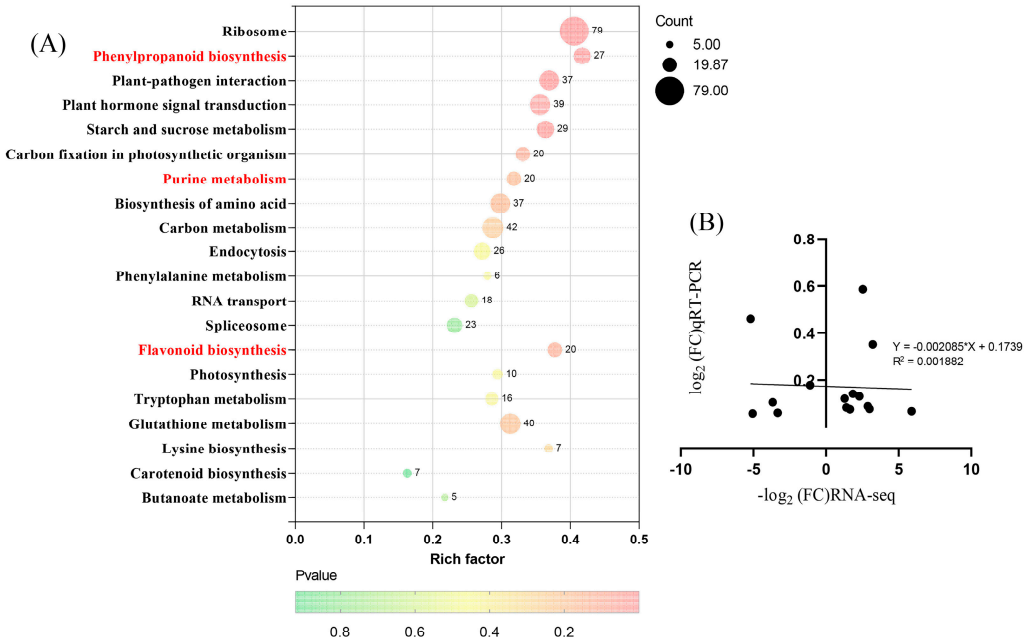


Figure 4. (**A**) The DEGs' KEGG pathways between CAF and SCZ. The three pathways with red color were further analyzed. (**B**) The scatterplot of RNA-seq transcript abundance folds change RNA-seq values (x-axis) against the qRT-PCR expression values (y-axis).

3.5. DEGs Involved in Caffeine Biosynthesis in Purine Alkaloids

The most common existing purine alkaloid in tea plants is caffeine; theobromine and theophylline are also present in small amounts; in the unique bitter tea plants, theacrine can also be found. The synthesis and degradation of both pathways were considered in the purine alkaloid metabolism pathway (Figure 5). The synthesis pathway plays a crucial and important role in influencing purine alkaloids. Through N-methyltransferase, the xanthine can be converted to 7-methylxanthine, which uses *TCS* to form theobromine. Through *TCS*, theobromine can be converted to caffeine. Theophylline is the downstream metabolite of caffeine found in typical tea varieties and species, and theacrine is one of the metabolites found in bitter tea varieties and species. The S-adenosylmethionine (SAM), adenosine monophosphate (AMP), guanine monophosphate (GMP), de novo pathway, and

(NAD) nucleic acids degradation route are the five different pathways for the synthesis of xanthosine. Xanthosine is a crucial precursor for the synthesis of caffeine and theobromine.

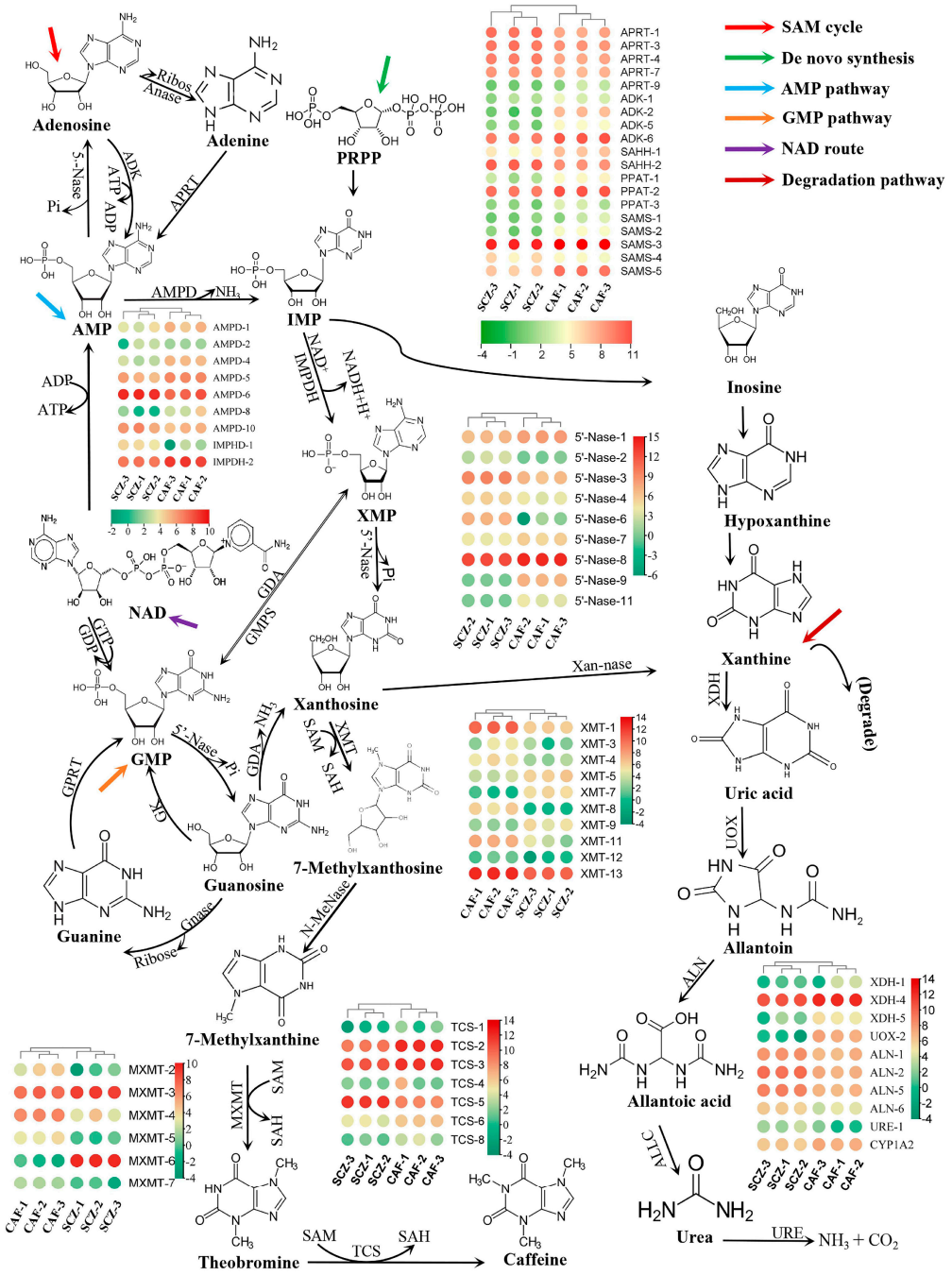


Figure 5. Caffeine biosynthesis pathway and expression profiles of genes in CAF and SCZ species. Five pathways towards caffeine synthesis: (1) de novo route, (2) AMP pathway, (3) SAM cycle, (4) NAD pathway, and (5) GMP pathway were listed. Genes in the respective pathways are shown in the heat map with their expression levels. A list of abbreviations is given in Supplementary Table S6.

The current study examined and identified the gene expression related to synthesis and degradation in caffeine synthesis. The heat maps of the related gene expression in purine alkaloid metabolism are shown in Figure 2. The expression data of the genes in the purine alkaloid metabolism are also available in the Supplementary Materials (Table S2). In the xanthosine biosynthesis pathway, the significantly upregulated genes included SAMS-3 in both species, CAF and SCZ; *SAMS-5* showed significant upregulation in CAF, with low

expression noted in the CSZ. *PPAT-2*, *SAHH-2*, and *ADK-6* were significantly upregulated in CAF and SCZ, while *SAHH-1* and *ADK-2* showed upregulation only in CAF. The expression levels of *APRT-1*, *APRT-3*, *APRT-4*, and *APRT-7* were upregulated in Camellia sinensis var. Shuchazao (SCZ). *AMPD-6* and *IMPDH-2* were upregulated in the AMP pathway in both species. The *5′-Nase-8* and *XMT-13* genes were significantly upregulated in SCZ and CAF. *MXMT-3*, *MXMT-6*, *TCS-5*, and *TCS-3* genes were upregulated considerably in the CSZ species, showing the roles of these genes in the caffeine biosynthesis. The *TCS-1*, *TCS-6*, and *TCS-8* genes showed downregulation in CAF, which indicates the low caffeine content in the *Camellia ptilophylla* (CAF).

In the purine alkaloid metabolism pathway, alkaloid content was affected by the degradation pathway. Generally, the purine alkaloid degradation was less active in SCZ than in CAF. *XDH-4* was upregulated in both SCZ and CAF. *ALN-1*, *ALN-2*, and *ALN-5* (*allantoinase*) were upregulated in SCZ compared to CAF, and the *URE-1* gene was downregulated in CAF. The *Cytochrome P450 family 1 subfamily A polypeptide* gene (*CYP1A2*) was more downregulated in SCZ than CAF, indicating less or no caffeine degradation in SCZ.

3.6. Phenylpropanoid and Flavonoid Biosynthesis Pathway

The phenylpropanoid route's first three steps combine to form the so-called general phenylpropanoid pathway, which is linked to specialized and core metabolism and provides precursors for all metabolites that occur downstream. *PAL* (*phenylalanine ammonia lyase*), *C4H* (*cinnamic acid 4- hydroxylase*), and *4CL* (*4-coumarate-CoA ligase*) analyses were performed to catalyze the reactions in the phenylpropanoid pathway [34]. *PAL* and *C4H* did not show a noticeable expression in SCZ and CAF in the phenylpropanoid biosynthesis pathway. Overall, the *PAL* and *C4H* genes did not show a noticeable expression in SCZ and CAF in the mentioned pathway. The *4-coumarate-CoA ligase* (*4CL-1*) was significantly expressed in the CAZ samples, whereas *4CL-4* showed slight expression in CAF samples, and significant expression of *4CL-4* was noted in SCZ. *Chalcone synthase* (*CHS-2* and *CHS-6*) were noted with significantly greater expression in SCZ than CAF and *cinnamate 3-hydroxylase C3H-2* had slight expression in SCZ and CAF, while (*C3H-3*) had high expression in SCZ, as shown in Figure 6.

The key component of the phenylpropanoid metabolism is flavonoid metabolism. A diphenylpropane (C6-C3-C6) skeleton with two aromatic rings connected by a three-carbon chain makes up the molecular structure of flavonoids. The synthesis of a particular flavonoid, which starts with the formation of chalcones, is signified by the inclusion of p-coumaroyl-CoA in the flavonoid biosynthesis pathway [34]. In the flavonoid pathway, the chalcone isomerase genes *CHI-1*, *CHI-2*, *CHI-3*, *CHI-6*, and *CHI-7* were downregulated in SCZ and CAF, while *CHI-4* showed significant upregulation in both species. *Flavonoid 3′-hydroxylase-1* (*F3′H-1*) showed upregulation, while *F3′H-2* was downregulated. *Flavonoid 3′5′-hydroxylase* (*F3′5′H-2*) showed upregulation in SCZ. *Flavonoid 3-hydroxylase* (*F3H-3* and *F3H-4*) was significantly upregulated in SCZ.

The *flavonol synthase* (*FLS-1* and *FLS-8*) were downregulated in SCZ while *FLS-2*, *FLS-3*, *FLS-4*, and *dihydroflavonol 4-reductase* (*DFR-3* and *DFR-4*) were significantly upregulated in SCZ. *Leucoanthocyanidin reductase* (*LAR*) showed significant upregulation in SCZ and CAF, except for *LAR-2* in CAF. The significant upregulation of *glucosyltransferases* and *serine carboxypeptidase-like acyltransferases 1 subfamily A* (*SCPL1A*) and *SCPL13-1* were noted in SCZ and CAF, which resulted in the accumulation of GCG in SCZ. *Anthocyanidin reductase* (*ANR-1*) was upregulated, which was consistent with the epicatechin and epigallocatechin. These key genes play a major role in the phenylpropanoid and flavonoid concentration in

Camellia sinensis var. Shuchazao (SCZ) and *Camellia ptilophylla* (CAF). The gene expression data are listed in Supplementary Materials (Table S3).

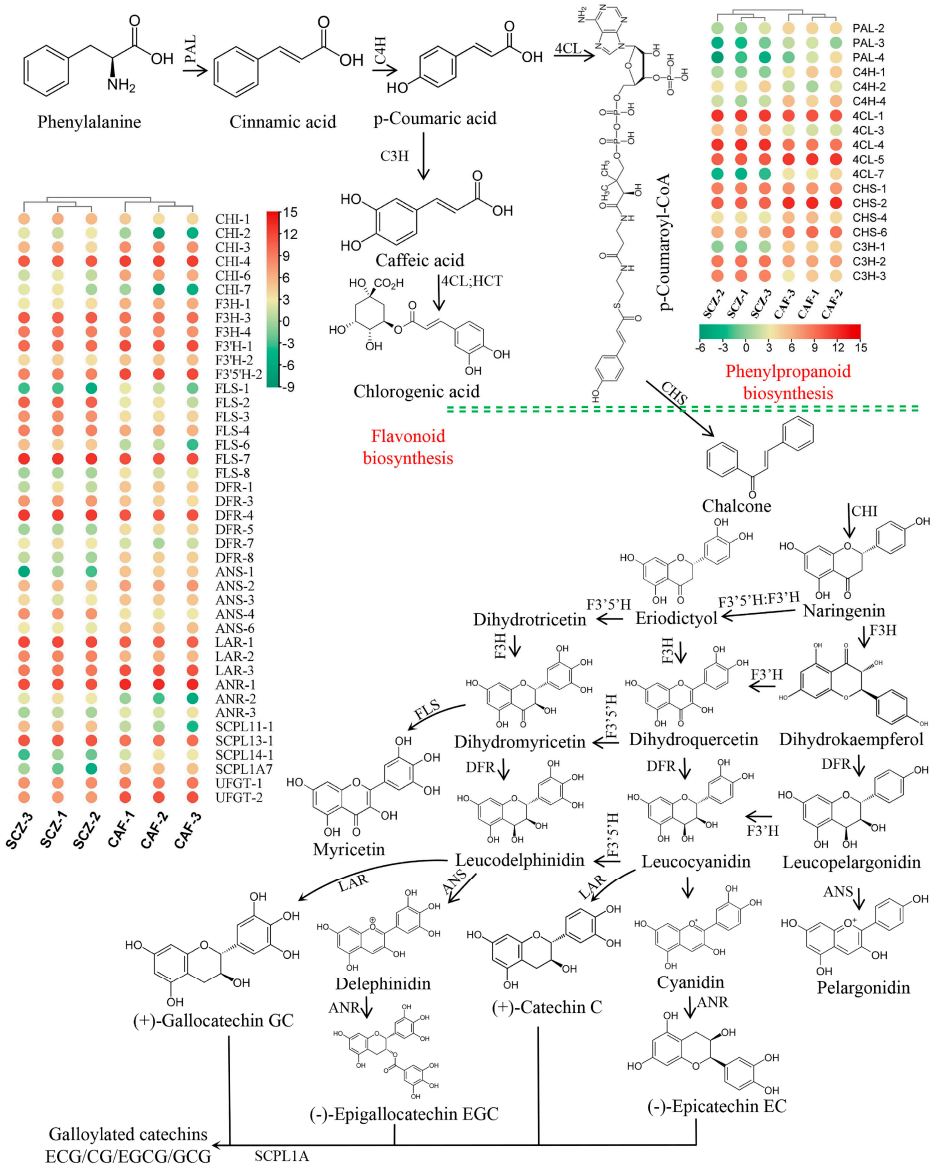


Figure 6. Phenylpropanoid and flavonoid biosynthesis pathways and gene expression in CAF and SCZ species. Genes involved in the respective pathways are shown in the heat map with their expression levels. A list of abbreviations is given in the Supplementary Materials, Table S6.

3.7. Coexpression Network

From the RNA-seq data, a total of 394 transcription factors (TFs) were identified with their different distribution rates. The AP2/ERF family had the most significant number (20%) of overall TFs, followed by WRKY (12%) and NAC transcription factors (TFs) with 11% of the 394. The details of the distribution pie chart are given in Supplementary Materials (Figure S1). Additionally, in the purine alkaloid, phenylpropanoid, and flavonoid biosynthesis pathways, the coexpression networks were constructed to identify putative regulatory mechanisms for differential gene expression.

The key genes involved in purine alkaloid biosynthesis, which showed significantly upregulated expressions, were *ADK-6*, *APRT-1*, *APRT-3*, *APRT-4*, *IMPDH-2*, *5′-Nase-8*, *XMT-13*, *MXMT-3*, *MXMT-6*, *TCS-5*, *TCS-3*, *XDH-4*, *ALN-1*, *ALN-2*, and *ALN-5*. The key gene in caffeine synthesis, *TCS-2*, showed a significant correlation with zf-HD

(TEA019835.1) and two MYB (TEA007284.1 and TEA021527.1). *TCS-5* was correlated with MYB (TEA001604), and AP2/ERF (TEA010818.1) also showed a correlation with MXMT-3 [27]. MYB (TEA031358.1), B3 (TEA022127.1), two RWP-RK (TEA022777.1 and TEA033593.1), and WKRY (TEA028963.1) were significantly related with MXMT-4 the key gene in theobromine synthesis (Figure 7).

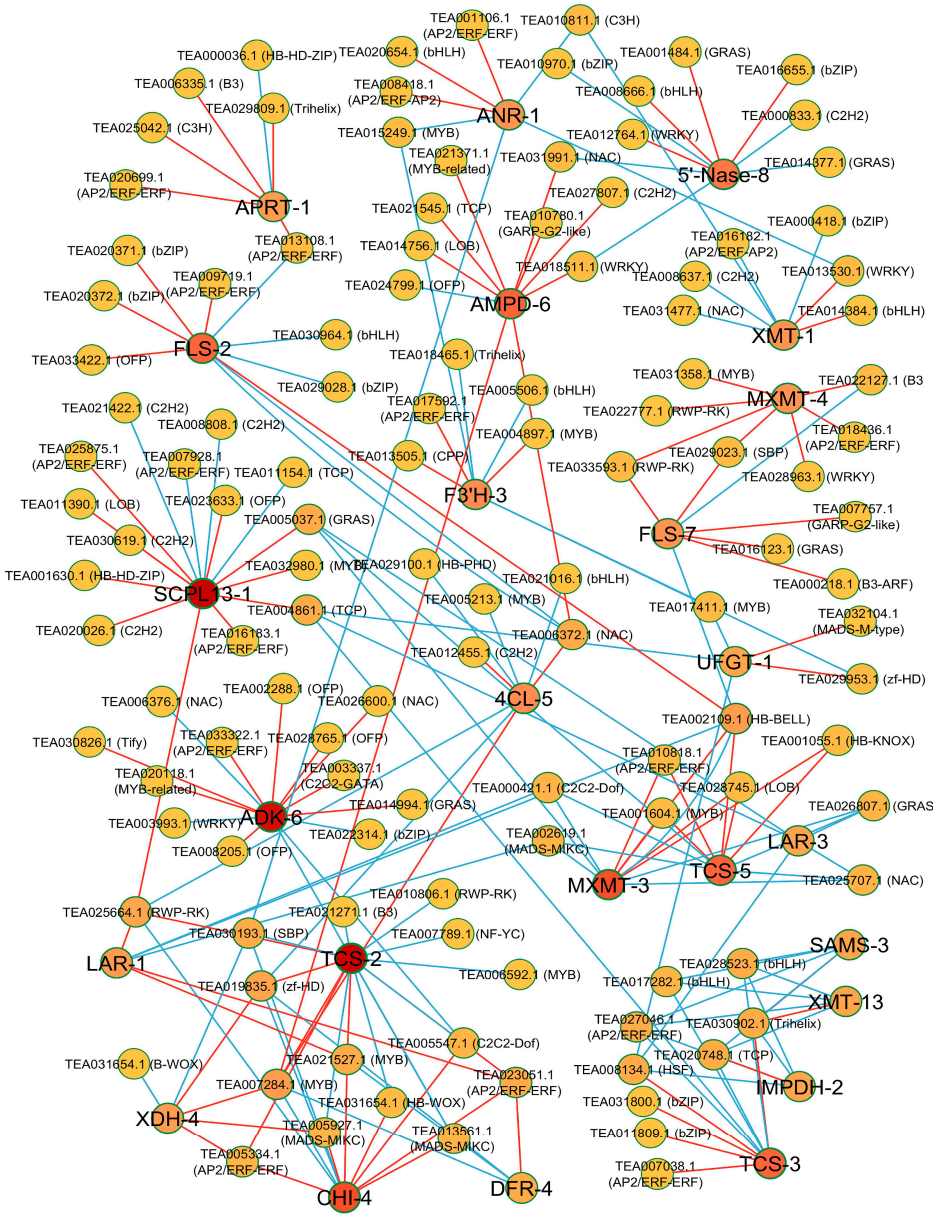


Figure 7. Coexpression network of TFs and purine alkaloid and phenylpropanoid biosynthesis genes, shown with colored lines (Pearson's correlation test, $p \leq 0.01$). The gene correlation expression networks were visualized using Cytoscape software. The size and color of nodes depend on the degree of coexpression.

In phenylpropanoid and flavonoid biosynthesis pathways, the key genes that showed significantly upregulated expressions were *4CL-1*, *CHS-2*, *CHS-6*, *C3H-3*, *FLS-2*, *FLS-3*, *FLS-4*, *DFR-3*, *DFR-4*, *SCPL13-1*, and *ANR-1*. The gene in phenylpropanoid biosynthesis *4CL-5* had significant expression in the pathway, which was correlated with C2H2 (TEA012455.1) and NAC (TEA006372.1). The *ANR-1* was significantly correlated with two AP2/ERF (TEA001106.0 and TEA008418.1) and bHLH (TEA020654.1). The key gene *SCPL13-1* in flavonoid biosynthesis was significantly upregulated and had a significant correlation

with two C2H2 (TEA020026.1 and TEA030619.1), MYB (TEA032980.1), and AP2/ERF (TEA025875.1 and TEA016183.1). These results indicate these TFs play a key role in the regulation of the purine alkaloid, phenylpropanoid, and flavonoid biosynthesis pathways. The annotated genes as TFs are listed in Supplementary Materials (Table S5).

4. Discussion

Tea plants, including the Shuchazao variety, contain purine alkaloids such as caffeine, theobromine, and theophylline [35]. These alkaloids are important metabolites found in tea. Specifically, the shoot tips of the Shuchazao tea cultivar contain about 35.87 mg/g of caffeine on a dry mass basis [27]. The alkaloid content and composition can vary across different tea species and cultivars. A study reported the differential metabolites, including alkaloids, in seven major tea cultivars grown in China, including the Shuchazao variety [36]. In early shoots of tea plants and allied species, caffeine is the purine alkaloid that accumulates the most; its amount ranges from 6.2 to 27.7 mg/g of dry-leaf [37,38]. Typically utilized to promote central nervous system stimulation, caffeine provides an almost bitter taste to tea beverages [11]. Excessive consumption of caffeine intake may cause problems in sleep and also can cause caffeine poisoning, which shows symptoms like anxiety, agitation, gastrointestinal disorders, and insomnia in some customers [39,40]. Higher expression levels of multiple genes related to caffeine biosynthesis, flavonoid and catechin biosynthesis, and other factors that affect tea quality have been clearly defined based on the genome assembly of *Camellia sinensis* and comparative research among other *Camellia* varieties and species [41]. Flavonoids and catechins in tea plants are significantly influenced by genetic factors, different abiotic stressors, developmental stages, and the processing of many tea types. The phenylpropanoid pathway is the source of flavonoids and catechins. According to molecular research, they share basic enzymatic genes such as *chalcone synthase* (*CHS*), *chalcone isomerase* (*CHI*), and *dihydroflavonol reductase* (*DFR*) [3,42]. Catechin production is aided by the later genes *leucoanthocyanidin reductase* (*LAR*) and *anthocyanidin reductase* (*ANR*) [11].

Purine alkaloids, phenylpropanoid, and flavonoid biosynthesis pathways were comparatively investigated in the present search research work based on the differential concentrations of the metabolite in various varieties and species of tea plants. In the previous study, it was shown that *C. ptilophylla* (CAF) is caffeine-free, and the phenylpropanoid and flavonoid biosynthesis pathway produce abundant GCG, and EGCG was also noted as a significant catechin in CAF. RNA-Seq analysis of CAF and SCZ was used to further explore the mechanism behind the different concentrations of caffeine and flavonoid content in tea plants. In the SAM cycle, the *SAMS-3* gene showed significant upregulation in SCZ and CAF while downregulation of *SAMS-5* was observed. The diverse regulation of SAM cycle genes was also reported in previous studies [43,44].

S-adenosyl-L-homocysteine hydrolases (*SAHH-2*) showed greater upregulation in SCZ than CAF in xanthosine biosynthesis. The ADK transcript abundance increased in association with changes in *SAHH* (*S-adenosyl-L-homocysteine hydrolase*) gene transcript levels, most probably as a result of changes in the SAM cycle's fluctuation, it has been thought that transcription of ADK increases in association with SAM-dependent methylation activities [45]. The three methylation steps from xanthosine to caffeine [27] are followed by one nucleoside removal step, *methyltransferase* (*XMT*), *7-methylxanthine methyltransferase* (*MXMT*), and *tea caffeine synthase* (*TCS*) [3]. The *MXMT-3*, *MXMT-6*, and *TCS-5* genes had upregulation in SCZ than CAF, while *TCS-2* and *TCS-3* had upregulation in both SCZ and CAF species. Transcription factor MYB is a key transcription factor reported as an important TF for caffeine biosynthesis regulation in tea plants [3]. In this detailed work,

MYB (TEA001604.1) and AP2/ERF (TEA010818.1) showed a positive correlation with *TCS-5* and *MXMT-3*, which were significantly upregulated in SCZ.

In the phenylpropanoid pathways, *4CL-1*, *4CL-4*, *C3H-3*, and *C3H-2* showed upregulation in SCZ and *CHS-2* and *CHS-6* were noted with upregulation in CAF, while *4CL-5* was upregulated in both species and in the network analysis transcription factor C2H2 (TEA012455.1), and NAC (TEA006372.1) correlated with *4CL-5*. *4-coumarate-CoA ligase* (*4CL*) genes play an essential role in the biosynthesis of phenylpropanoid and flavonoid [46,47]. A total of 42 genes were involved in the flavonoid biosynthesis pathway, as shown in Figure 6. Among them, *F3′H-1* was significantly upregulated in SCZ and CAF, which reported elevated GCG content in flavonoid biosynthesis [36]; *F3′5′H-2* is considered an exclusive enzyme that catalyzes the 5′ position hydroxylation of the B ring in flavonoid molecules, and higher expression of *F3′5′H-2* is closely related to the higher accumulation of catechins [36]. Similarly, *DFR-4*, *LAR-1*, *LAR-3*, and *ANR-1* were significantly upregulated in SCZ, which are the key genes in flavonoid biosynthesis. The transcription factors AP2/ERF (TEA023051.1) were positively correlated with *DFR-4*, and two AP2/ERF (TEA008418.1 and TEA001106.1) and bHLH (TEA020654.1) were positively correlated with *ANR-1*. *SCPL13-1* showed upregulation in SCZ and CAF, which are the key galloylated cis-catechins. Three transcription factors, AP2/ERF (TEA025875.1, TEA007928.1, and TEA016183.1), GRAS (TEA005037.1), and MYB (TEA032980.1), which were significantly related to *SCPIL13-1* in coexpression network analysis, may play an essential role in regulating the expression of *SCPIL* genes.

5. Conclusions

The present study investigated the species characteristics of the tea plants and their potential molecular mechanisms leading to purine alkaloid, phenylpropanoid, and flavonoid biosynthesis pathways by performing a thorough transcriptome analysis of two tea plant species, *Camellia sinensis* var. Shuchazao (SCZ) and *Camellia ptilophylla* (CAF). The low-caffeine content CAF had more active xanthosine synthesis pathways than SCZ and the high expression of tea caffeine synthase genes and 7-methylxanthine methyltransferase levels were noted. In the degradation pathway, the expression of *XDH-4* and *CYPIA2* was higher in CAF than in SCZ. The transcription factors AP2/ERF (20%) and WRKY (12%) with higher numbers were related to the purine alkaloid and phenylpropanoid biosynthesis. Furthermore, the higher expression levels of *4CL-4*, *F3H-3*, *FLS-2*, *FLS-3*, and *SCPL13-1* indicated that these genes are crucial for the production of flavonoids and phenylpropanoids in SCZ. In summary, the additional details that these data reveal about the distinctive metabolites and the essential genes that regulate both tea species provide insight for future research on tea plants and the selection of cultivars.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/horticulturae11010008/s1, Figure S1: Distribution of transcription factors.; Table S1: List of primers; Table S2: Genes involved in purine metabolism.; Table S3: Genes involved in phenylpropanoid and flavonoid biosynthesis pathway.; Table S4: KEGG pathway and genes involved; Table S5: Genes annotated as TF; Table S6: List of abbreviations; Table S7: Data output and quality statistics.

Author Contributions: W.K.: formal analysis, methodology, writing—original draft. B.S.: conceptualization, formal analysis, review. S.L.: conceptualization, resources. P.Z.: resources, review. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the Guangdong Province Natural Science Foundation (Grant Nos. 2023A1515010786 and 2022A1515111141).

Data Availability Statement: All the relevant data are available within the manuscript and Supplementary Materials.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Zhang, X.; Chen, S.; Shi, L.; Gong, D.; Zhang, S.; Zhao, Q.; Zhan, D.; Vasseur, L.; Wang, Y.; Yu, J. Haplotype-resolved genome assembly provides insights into evolutionary history of the tea plant *Camellia sinensis*. *Nat. Genet.* **2021**, *53*, 1250–1259. [CrossRef] [PubMed]
2. Chaturvedula, V.S.P.; Prakash, I. The aroma, taste, color and bioactive constituents of tea. *J. Med. Plants Res.* **2011**, *5*, 2110–2124.
3. Li, P.; Ye, Z.; Fu, J.; Xu, Y.; Shen, Y.; Zhang, Y.; Tang, D.; Li, P.; Zuo, H.; Tong, W. CsMYB184 regulates caffeine biosynthesis in tea plants. *Plant Biotechnol. J.* **2022**, *20*, 1012. [CrossRef] [PubMed]
4. Changchun, Y.; Suhua, S.; Chuangxing, Y. Phylogentic relationship of *Camellia ptilophylla* Chang with its allied species and disintegration of its population. *Zhongshan Da Xue Xue Bao. Zi Ran Ke Xue Ban= Acta Sci. Nat. Univ. Sunyatseni* **1999**, *38*, 72–76.
5. Yang, C.S.; Hong, J. Prevention of chronic diseases by tea: Possible mechanisms and human relevance. *Annu. Rev. Nutr.* **2013**, *33*, 161–181. [CrossRef] [PubMed]
6. Zhang, S.; Jin, J.; Chen, J.; Ercisli, S.; Chen, L. Purine alkaloids in tea plants: Component, biosynthetic mechanism and genetic variation. *Beverage Plant Res.* **2022**, *2*, 13. [CrossRef]
7. Chacko, S.M.; Thambi, P.T.; Kuttan, R.; Nishigaki, I. Beneficial effects of green tea: A literature review. *Chin. Med.* **2010**, *5*, 13. [CrossRef]
8. Pinto, G.; Illiano, A.; Carpentieri, A.; Spinelli, M.; Melchiorre, C.; Fontanarosa, C.; di Serio, M.; Amoresano, A. Quantification of polyphenols and metals in Chinese tea infusions by mass spectrometry. *Foods* **2020**, *9*, 835. [CrossRef] [PubMed]
9. Dong, C.; Li, F.; Yang, T.; Feng, L.; Zhang, S.; Li, F.; Li, W.; Xu, G.; Bao, S.; Wan, X. Theanine transporters identified in tea plants (*Camellia sinensis* L.). *Plant J.* **2020**, *101*, 57–70. [CrossRef] [PubMed]
10. Jin, J.-Q.; Chai, Y.-F.; Liu, Y.-F.; Zhang, J.; Yao, M.-Z.; Chen, L. Hongyacha, a naturally caffeine-free tea plant from Fujian, China. *J. Agric. Food Chem.* **2018**, *66*, 11311–11319. [CrossRef] [PubMed]
11. Ashihara, H.; Sano, H.; Crozier, A. Caffeine and related purine alkaloids: Biosynthesis, catabolism, function and genetic engineering. *Phytochemistry* **2008**, *69*, 841–856. [CrossRef] [PubMed]
12. Juliano, L.M.; Griffiths, R.R. A critical review of caffeine withdrawal: Empirical validation of symptoms and signs, incidence, severity, and associated features. *Psychopharmacology* **2004**, *176*, 1–29. [CrossRef]
13. Li, Z. Comparative analysis of Fenghuang Dancong, Tieguanyin, and Dahongpao teas using headspace solid-phase microextraction coupled with gas chromatography-mass spectrometry and chemometric methods. *PLoS ONE* **2022**, *17*, e0276044. [CrossRef] [PubMed]
14. Jin, J.; Zhou, C.; Ma, C.; Yao, M.; Ma, J.; Chen, L. Identification on purine alkaloids of representative tea germplasms in China. *J. Plant Genet. Resour.* **2014**, *15*, 279–285.
15. Chen, L.; Zhou, Z.-X. Variations of main quality components of tea genetic resources [*Camellia sinensis* (L.) O. Kuntze] preserved in the China National Germplasm Tea Repository. *Plant Foods Hum. Nutr.* **2005**, *60*, 31–35. [CrossRef]
16. Tritsch, N.; Steger, M.C.; Segatz, V.; Blumenthal, P.; Rigling, M.; Schwarz, S.; Zhang, Y.; Franke, H.; Lachenmeier, D.W. Risk assessment of caffeine and epigallocatechin gallate in coffee leaf tea. *Foods* **2022**, *11*, 263. [CrossRef] [PubMed]
17. Song, X.; Shi, X.; Li, Y.; Peng, L.; Li, C.; Yang, X.; Ye, C.; Li, J.; He, Y.; Huang, H. Studies of cocoa tea, a wild tea tree containing theobromine. *Front. Biol. China* **2009**, *4*, 460–468. [CrossRef]
18. Suzuki, T. The participation of S-adenosylmethionine in the biosynthesis of caffeine in the tea plant. *FEBS Lett.* **1972**, *24*, 18–20. [CrossRef]
19. Koshiishi, C.; Kato, A.; Yama, S.; Crozier, A.; Ashihara, H. A new caffeine biosynthetic pathway in tea leaves: Utilisation of adenosine released from the S-adenosyl-L-methionine cycle. *FEBS Lett.* **2001**, *499*, 50–54. [CrossRef]
20. Hamberger, B.; Ellis, M.; Friedmann, M.; de Azevedo Souza, C.; Barbazuk, B.; Douglas, C.J. Genome-wide analyses of phenylpropanoid-related genes in *Populus trichocarpa*, *Arabidopsis thaliana*, and *Oryza sativa*: The Populus lignin toolbox and conservation and diversification of angiosperm gene families. *Botany* **2007**, *85*, 1182–1201. [CrossRef]
21. Stankovic, M.S.; Niciforovic, N.; Topuzovic, M.; Solujic, S. Total phenolic content, flavonoid concentrations and antioxidant activity, of the whole plant and plant parts extracts from *Teucrium Montanum* L. Var. *Montanum*, F. *Supinum* (L.) Reichenb. *Biotechnol. Biotechnol. Equip.* **2011**, *25*, 2222–2227. [CrossRef]
22. Tohge, T.; Watanabe, M.; Hoefgen, R.; Fernie, A.R. The evolution of phenylpropanoid metabolism in the green lineage. *Crit. Rev. Biochem. Mol. Biol.* **2013**, *48*, 123–152. [CrossRef] [PubMed]
23. Nabavi, S.M.; Šamec, D.; Tomczyk, M.; Milella, L.; Russo, D.; Habtemariam, S.; Suntar, I.; Rastrelli, L.; Daglia, M.; Xiao, J. Flavonoid biosynthetic pathways in plants: Versatile targets for metabolic engineering. *Biotechnol. Adv.* **2020**, *38*, 107316. [CrossRef]

24. Wang, J.; Li, G.; Li, C.; Zhang, C.; Cui, L.; Ai, G.; Wang, X.; Zheng, F.; Zhang, D.; Larkin, R.M. NF-Y plays essential roles in flavonoid biosynthesis by modulating histone modifications in tomato. *New Phytol.* **2021**, *229*, 3237–3252. [CrossRef] [PubMed]
25. Deng, X.; Bashandy, H.; Ainasoja, M.; Kontturi, J.; Pietiäinen, M.; Laitinen, R.A.; Albert, V.A.; Valkonen, J.P.; Elomaa, P.; Teeri, T.H. Functional diversification of duplicated chalcone synthase genes in anthocyanin biosynthesis of *Gerbera hybrida*. *New Phytol.* **2014**, *201*, 1469–1483. [CrossRef]
26. Zhang, X.; Abrahan, C.; Colquhoun, T.A.; Liu, C.-J. A proteolytic regulator controlling chalcone synthase stability and flavonoid biosynthesis in *Arabidopsis*. *Plant Cell* **2017**, *29*, 1157–1174. [CrossRef] [PubMed]
27. Deng, C.; Ku, X.; Cheng, L.-L.; Pan, S.-a.; Fan, L.; Deng, W.-W.; Zhao, J.; Zhang, Z.-Z. Metabolite and transcriptome profiling on xanthine alkaloids-fed tea plant (*Camellia sinensis*) shoot tips and roots reveal the complex metabolic network for caffeine biosynthesis and degradation. *Front. Plant Sci.* **2020**, *11*, 551288. [CrossRef]
28. Ogino, A.; Taniguchi, F.; Yoshida, K.; Matsumoto, S.; Fukuoka, H.; Nesumi, A. A new DNA marker CafLess-TCS1 for selection of caffeine-less tea plants. *Breed. Sci.* **2019**, *69*, 393–400. [CrossRef]
29. Kim, D.; Langmead, B.; Salzberg, S.L. HISAT: A fast spliced aligner with low memory requirements. *Nat. Methods* **2015**, *12*, 357–360. [CrossRef] [PubMed]
30. Bu, D.; Luo, H.; Huo, P.; Wang, Z.; Zhang, S.; He, Z.; Wu, Y.; Zhao, L.; Liu, J.; Guo, J. KOBAS-i: Intelligent prioritization and exploratory visualization of biological functions for gene enrichment analysis. *Nucleic Acids Res.* **2021**, *49*, W317–W325. [CrossRef]
31. Li, J.; Liu, S.; Chen, P.; Cai, J.; Tang, S.; Yang, W.; Cao, F.; Zheng, P.; Sun, B. Systematic analysis of the R2R3-MYB family in camellia sinensis: Evidence for galloylated catechins biosynthesis regulation. *Front. Plant Sci.* **2022**, *12*, 782220. [CrossRef] [PubMed]
32. Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the $2-\Delta\Delta CT$ method. *Methods* **2001**, *25*, 402–408. [CrossRef] [PubMed]
33. Chen, C.; Chen, H.; Zhang, Y.; Thomas, H.R.; Frank, M.H.; He, Y.; Xia, R. TBtools: An integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* **2020**, *13*, 1194–1202. [CrossRef]
34. Stobiecki, M.; Kachlicki, P. Isolation and identification of flavonoids. In *The Science of Flavonoids*; Springer: Cham, Switzerland, 2006; pp. 47–69.
35. Wang, L.; Di, T.; Peng, J.; Li, Y.; Li, N.; Hao, X.; Ding, C.; Huang, J.; Zeng, J.; Yang, Y. Comparative metabolomic analysis reveals the involvement of catechins in adaptation mechanism to cold stress in tea plant (*Camellia sinensis* var. sinensis). *Environ. Exp. Bot.* **2022**, *201*, 104978. [CrossRef]
36. Ting, G.; Hou, B.-H.; Shao, S.-X.; Xu, M.-t.; Zheng, Y.-c.; Shan, J.; Wang, P.-j.; Ye, N.-X. Differential metabolites and their transcriptional regulation in seven major tea cultivars (*Camellia sinensis*) in China. *J. Integr. Agric.* **2023**, *22*, 3346–3363. [CrossRef]
37. Jin, J.-Q.; Yao, M.-Z.; Ma, C.-L.; Ma, J.-Q.; Chen, L. Association mapping of caffeine content with TCS1 in tea plant and its related species. *Plant Physiol. Biochem.* **2016**, *105*, 251–259. [CrossRef] [PubMed]
38. Cai, J.; Qiu, Z.; Liao, J.; Li, A.; Chen, J.; Wu, Z.; Khan, W.; Sun, B.; Liu, S.; Zheng, P. Comprehensive Analysis of the Yield and Leaf Quality of Fresh Tea (*Camellia sinensis* cv. Jin Xuan) under Different Nitrogen Fertilization Levels. *Foods* **2024**, *13*, 2091. [CrossRef]
39. Negishi, O.; Ozawa, T.; Imagawa, H. Biosynthesis of caffeine from purine nucleotides in tea plant. *Biosci. Biotechnol. Biochem.* **1992**, *56*, 499–503. [CrossRef]
40. Ashihara, H.; Kato, M.; Chuang-xing, Y. Biosynthesis and metabolism of purine alkaloids in leaves of cocoa tea (*Camellia ptilophylla*). *J. Plant Res.* **1998**, *111*, 599–604. [CrossRef]
41. Mei, S.; Yu, Z.; Chen, J.; Zheng, P.; Sun, B.; Guo, J.; Liu, S. The physiology of postharvest tea (*Camellia sinensis*) leaves, according to metabolic phenotypes and gene expression analysis. *Molecules* **2022**, *27*, 1708. [CrossRef] [PubMed]
42. Qiu, Z.; Li, A.; Huang, W.; Chen, J.; Lin, X.; Yao, J.; Pan, L.; Sun, B.; Liu, S.; Zheng, P. Metabolic profiling reveal changes in shoots and roots of nitrogen-deficient tea plants (*Camellia sinensis* cv. Jinxuan). *Sci. Hortic.* **2024**, *337*, 113528. [CrossRef]
43. Ying, C.; Chen, J.; Chen, J.; Zheng, P.; Zhou, C.; Sun, B.; Liu, S. Differential accumulation mechanisms of purine alkaloids and catechins in *Camellia ptilophylla*, a natural theobromine-rich tea. *Beverage Plant Res.* **2023**, *3*, 15. [CrossRef]
44. Wang, W.; Zhou, Y.; Wu, Y.; Dai, X.; Liu, Y.; Qian, Y.; Li, M.; Jiang, X.; Wang, Y.; Gao, L. Insight into catechins metabolic pathways of *Camellia sinensis* based on genome and transcriptome analysis. *J. Agric. Food Chem.* **2018**, *66*, 4281–4293. [CrossRef]
45. Schoor, S.; Moffatt, B.A. Applying high throughput techniques in the study of adenosine kinase in plant metabolism and development. *Front. Biosci.* **2004**, *9*, 1771–1781. [CrossRef]
46. Wang, Y.; Guo, L.; Zhao, Y.; Zhao, X.; Yuan, Z. Systematic analysis and expression profiles of the 4-coumarate: *CoA ligase (4CL)* gene family in pomegranate (*Punica granatum* L.). *Int. J. Mol. Sci.* **2022**, *23*, 3509. [CrossRef]
47. Li, Y.; Kim, J.I.; Pysh, L.; Chapple, C. Four isoforms of *Arabidopsis* 4-coumarate: CoA ligase have overlapping yet distinct roles in phenylpropanoid metabolism. *Plant Physiol.* **2015**, *169*, 2409–2421. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.