Transcriptomic and Metabolomic Differences Between Two Saposhnikovia divaricata (Turcz.) Schischk Phenotypes With Single- and Double-Headed Roots

Tao Zhang, Yuqiu Chen, Qinghe Zhang, Peng Yu, Qiong Li, Weichen Qi* and Changbao Chen*

Key Laboratory of Chinese Medicine Planting and Development, Changchun University of Chinese Medicine, Changchun, China

Saposhnikovia divaricata is derived from the dried roots of Saposhnikovia divaricata (Turcz.) Schischk and used as a Chinese herbal medicine for treating respiratory, immune, and nervous system diseases. The continuously increasing market demand for traditional Chinese medicine requires the commercial cultivation of Saposhnikovia divaricata using standardized methods and high yielding genotypes, such as double-headed root plants, for achieving consistent quality and a reliable supply. In this study, we aimed to identify the quantitative differences in chromone, a precursor of flavonoid biosynthesis, between plants with single- and double-headed roots using high-performance liquid chromatography and further explore the two phenotypes at the transcriptomic and metabolomic levels. Our results showed that the chromone content was significantly higher in plants with double-headed roots than in those with single-headed roots. Transcriptomic analysis revealed six significantly differentially expressed genes between the two phenotypes, including five key genes in the flavonoid biosynthesis pathway (4-coumarate-CoA ligase, chalcone synthase 1, vinorine synthase, chalcone-flavonone isomerase 1, and flavanone 3 beta-hydroxylase) and one key gene in the abscisic acid biosynthetic pathway (zeaxanthin epoxidase). Moreover, metabolomic analysis showed that the 126 differentially expressed metabolites were mainly enriched in the biosynthesis of secondary metabolites and phytohormones. Overall, our results suggest that plants with double-headed roots have higher medicinal value than those with single-headed roots, probably due to differences in various biosynthetic pathways. These data might help select the genotypes with superior yield and therapeutic properties.

Keywords: Saposhnikovia divaricata, transcriptome, metabolite, biotechnology, chromone

Abbreviations: ABA, Abscisic acid; DEGs, Differentially expressed genes; DEMs, Differentially expressed metabolites; GAPDH, Glyceraldehyde-3-phosphate dehydrogenase; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; LC-MS, Liquid Chromatography Mass-Spectroscopy; QC, Quality control.
INTRODUCTION

Saposhnikovia divaricata, also known as Fangfeng, is derived from the dried roots of Saposhnikovia divaricata (Turcz.). Schischk and used as a Chinese herbal material owing to its antipyretic, analgesic, antioxidant, antiproliferative, and anti-inflammatory properties (Khan et al., 2011; Okuyama et al., 2019). The chemical composition of S. divaricata is complex and includes more than 100 compounds, such as chromone, which is a precursor of flavonoid biosynthesis, coumarin, and volatile oil. Previous studies have shown that the root has the highest medicinal value during the vegetative stage (Liu et al., 2020), which is significantly reduced in the reproductive stage (Li et al., 2017). Therefore, it is necessary to determine the medicinal substances of S. divaricata and elucidate the underlying formation mechanism (Liu and Zhang, 2007).

Transcriptomics is the study of gene expression at the RNA level under specific physiological conditions and helps identify the differentially expressed genes (DEGs) (Armstrong and Que, 2012; Hamilton and Buell, 2012; Ramaswami et al., 2013) and understand the molecular basis of phenotypic differences (Forrest and Carninci, 2009; Rodriguez-Garcia et al., 2017). Transcriptomics has been used for gene mining, function prediction, and metabolic pathway analysis of various medicinal plants, including Panax ginseng (Hongzhe et al., 2015; Li et al., 2013; Liu et al., 2016; Zhang et al., 2021), Panax quinquefolius (Wu et al., 2013; Zou et al., 2015), Scutellaria baicalensis (Cheng et al., 2018), Salvia miltiorrhiza (Gao et al., 2014), and Bupleurum chinense (Sui et al., 2011). Transcription factors initiate and regulate gene expression by recognizing and binding to cis-acting elements in the promoter region of genes. Studies have shown that WRKY, bHLH, and bZIP transcription factors play important roles in defense responses against various abiotic stresses (Madhunika and Ralf, 2014), in plant growth, stress resistance, and signal transduction (Liu et al., 2015), and in plant development and physiological metabolic processes (Zhang et al., 2011), respectively. Therefore, genome-wide transcriptomics, together with transcription factor analysis, allows us to understand plant responses via transcriptional changes and elucidate the molecular mechanisms underlying the biosynthesis of secondary metabolites in medicinal plants.

Metabolomics is the comprehensive qualitative and quantitative characterization of all endogenous small molecule metabolites in biological systems. It has been widely applied to study metabolites in medicinal plants due to its high sensitivity, resolution, accuracy, and wide dynamic range. Phytohormones are signal molecules that are naturally present in plants in extremely low concentrations, in plants, which is a precursor of flavonoid biosynthesis, coumarin, and volatile oil. Previous studies have shown that the root has the highest medicinal value during the vegetative stage (Liu et al., 2020), which is significantly reduced in the reproductive stage (Li et al., 2017). Therefore, it is necessary to determine the medicinal substances of S. divaricata and elucidate the underlying formation mechanism (Liu and Zhang, 2007).

Transcriptomics is the study of gene expression at the RNA level under specific physiological conditions and helps identify the differentially expressed genes (DEGs) (Armstrong and Que, 2012; Hamilton and Buell, 2012; Ramaswami et al., 2013) and understand the molecular basis of phenotypic differences (Forrest and Carninci, 2009; Rodriguez-Garcia et al., 2017). Transcriptomics has been used for gene mining, function prediction, and metabolic pathway analysis of various medicinal plants, including Panax ginseng (Hongzhe et al., 2015; Li et al., 2013; Liu et al., 2016; Zhang et al., 2021), Panax quinquefolius (Wu et al., 2013; Zou et al., 2015), Scutellaria baicalensis (Cheng et al., 2018), Salvia miltiorrhiza (Gao et al., 2014), and Bupleurum chinense (Sui et al., 2011). Transcription factors initiate and regulate gene expression by recognizing and binding to cis-acting elements in the promoter region of genes. Studies have shown that WRKY, bHLH, and bZIP transcription factors play important roles in defense responses against various abiotic stresses (Madhunika and Ralf, 2014), in plant growth, stress resistance, and signal transduction (Liu et al., 2015), and in plant development and physiological metabolic processes (Zhang et al., 2011), respectively. Therefore, genome-wide transcriptomics, together with transcription factor analysis, allows us to understand plant responses via transcriptional changes and elucidate the molecular mechanisms underlying the biosynthesis of secondary metabolites in medicinal plants.

Metabolomics is the comprehensive qualitative and quantitative characterization of all endogenous small molecule metabolites in biological systems. It has been widely applied to study metabolites in medicinal plants due to its high sensitivity, resolution, accuracy, and wide dynamic range. Phytohormones are signal molecules that are naturally present in plants in extremely low concentrations, influencing plant growth and development (Hong and Xu, 2013). Of these, brassinolide, salicylic acid, and jasmonic acid are involved in plant-pathogen interactions (Jameson et al., 2016); cytokinin (CTK) is a derivative of adenine and, together with indolyl acetic acid, induces plant cell division, inhibits leaf senescence, and defends against abiotic stress (Werner and Schmulling, 2009); zeatin nucleoside, which is the main form of CTK transported in the xylem, controls the differentiation of plant flower buds; and abscisic acid (ABA) regulates plant and environmental signals (Zhang et al., 2010). Plant endogenous hormones play an important role in the process of plant growth and development, which may be the cause of the double-headed roots of S. divaricata.

The market demand for traditional Chinese medicine continues to grow worldwide. Thus, the commercial cultivation of S. divaricata using standardized methods is required for achieving consistent quality and a reliable supply. The conventional cultivation of S. divaricata includes plants with single-headed roots (one aerial part); however, we found that those with double-headed roots (two aerial parts) have a higher yield potential. In this study, we aimed to identify the differences in the chromone content between plants with single- and double-headed roots using high-performance liquid chromatography and further explore the two phenotypes via transcriptomic and metabolomic analyses. Our study provides an important theoretical basis for the quality research of S. divaricata medicinal materials.

MATERIALS AND METHODS

Plant Material

The S. divaricata roots of Umbelliferae plants were selected as the test materials. Samples of S. divaricata were collected from its cultivation base in Baicheng City, Jilin Province (E122.51, N45.37) at the end of September 2020. Single and double-headed root plants were still grown under the well-cultivated conditions. Those were dug out from the soil, all root tissues from every 3 individual plants were mixed as one biological replicate to be stored in liquid nitrogen. Three biological replicates were respectively collected from the treatment and control groups for transcriptome sequencing. Six biological replicates were respectively collected from the treatment and control groups.
for metabolomic sequencing. The remaining samples were dried at 45°C to a constant weight, mixed and crushed, for determining ginsenoside content. The differences between plants with single- and double-headed roots are shown in Figure 1.

**Determination of the Chromone Content**

After drying the samples, these were ground into a powder. The ground powder was filtered through a 60-mesh sieve and then weighed in three replicates, each weighing 1.0 g. Chromogen ketones were extracted ultrasonically (extraction conditions: ultrasonic frequency, 40 kHz; extraction temperature, 30°C; extraction time, 30 min; liquid ratio, 1:15), extracted three times with methanol, filtered in a funnel, and then the filtered liquids were combined. After the filtrate was evaporated in an evaporating dish and transferred to a 5 ml volumetric flask, the volume was adjusted to 5 ml and the solution was filtered using a 0.22 μm filter. High-performance liquid chromatography was performed using an Agilent 1,200 series high-performance liquid chromatography system (Agilent, Palo Alto, CA, United States), equipped with an autosampler and a UV detector with a C18 column (4.6 mm × 250 mm, 5 μm; Agilent). Gradient elution was performed using solvent A (100% methanol) and solvent B (100% water) at 30°C, according to the following gradient program: 0–20 min, 30% A; 20–25 min, 45% A; 25–50 min, 60% A; 50–55 min, 90% A; and 55–60 min, 100% A. The flow rate was maintained at 1.0 ml/min, the sample injection volume was 10 μl, and UV absorption was measured at 254 nm. Quantitative analysis was performed using the one-point curve method using an external standard of authentic chromones (prim-O-glucosylcimifugin, 4′-O-β-D-glucosyl-5-O-methylvisamminol, cimifugin, and sec-O-glucosylhamaudol) and the following equations:

\[
y = 5013.2x - 4988.3 \quad (r^2 = 0.9983)
\]
\[
y = 7810.9x - 5120.6 \quad (r^2 = 0.9990)
\]
\[
y = 6177.9x - 3435.5 \quad (r^2 = 0.9997)
\]
\[
y = 2867.9x - 2658.2 \quad (r^2 = 0.9992)
\]

For **prim-O-glucosylcimifugin**, **4′-O-β-D-glucosyl-5-O-methylvisamminol**, cimifugin, and sec-O-glucosylhamaudol, respectively.
RNA Sequencing

The total RNA was extracted according to the instruction manual of TRIzol reagent (Invitrogen, Waltham, MA, United States). Bioanalyzer 2100 and RNA 1000 Nano LabChip Kit (Agilent) were used to analyze the total RNA quantity and purity. Poly(A) RNA was purified from the total RNA using magnetic beads linked to poly-T oligonucleotides. Next, at a temperature of 95°C, divalent cations were used to fragment the mRNA into small pieces. The RNA-Seq sample preparation kit (Illumina, San Diego, CA, United States) was used to reverse transcribe the cleaved RNA fragments to create the final cDNA library. The average insert size of the double-ended library was 300 ± 50 bp. Paired-end sequencing was performed by LC Sciences (Houston, TX, United States) using an IlluminaHiseq 4,000 sequencer.

De Novo Assembly, Unigene Annotation, and Functional Classification

The workflow of the bioinformatics analysis is shown in Figure 2. Cutadapt and Perl scripts were used to delete the reads containing adapter contamination, low-quality bases, and undetermined bases (Martin, 2011). FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) was used to verify sequence quality and calculate the Q20, Q30, and GC content of clean data. All downstream analyses were based on high-quality, clean data. The de novo assembly of the transcriptome data was performed using Trinity 2.4.0. The transcripts were grouped into clusters according to the shared sequence content, and the transcript with the largest expansion in each cluster was selected as the unigene (Grabherr et al., 2011). All assembled unigenes were linked to the non-redundant protein database (http://www.ncbi.nlm.nih.gov/), Gene Ontology (GO; http://www.geneontology.org), SwissProt for comparison (http://www.expasy.ch/sprot/), Kyoto Encyclopedia of Genes and Genomes (KEGG; http://www.genome.jp/kegg/), and eggNOG (http://eggnodb.embl.de/). A DIAMOND threshold E value of <0.00001 was used (Buchfink et al., 2015).

Identification of the DEGs

Salmon (https://combine-lab.github.io/salmon/) was used to determine the expression level of unigenes by calculating the transcripts per million (Patro et al., 2017). The “edgeR” package in R was used to select the DEGs in the sample with p < 0.05 when the log2 (multiple change) > 1 or log2 (multiple change) < −1 (Robinson et al., 2010).

Metabolite Extraction and Liquid Chromatography Mass-Spectroscopy

The collected samples were thawed on ice, and 120 µl of pre-chilled 50% methanol buffer was used to extract the metabolites from 20 µl of each sample. The metabolite mixture was vortexed for 1 min, incubated at 20°C–25°C for 10 min, and stored at −20°C overnight. The mixture was centrifuged at 4,000 × g for 20 min, and the supernatant was transferred to a 96-well plate. The samples were stored at −80°C before the LC-MS analysis. Mixed quality control (QC) samples were prepared by combining 10 µl of each extraction mixture.

LC-MS Analysis

All test samples were analyzed using a TripleTOF 5,600 Plus high-resolution tandem mass spectrometer (SCIEX, Warrington, United Kingdom) in the positive and negative ion modes, and an ultra-high-performance liquid chromatography system equipped with Acuity ultra-high-performance liquid T3 columns (SCIEX) for separation (100 mm × 2.1 mm, 1.8 µm; Waters, Milford, MA, United States). The mobile phase consisted of solvent A (water, 0.1% formic acid) and solvent B (acetonitrile, 0.1% formic acid). The gradient elution conditions were as follows: flow rate, 0.4 ml/min: 5% solvent B, for 0–0.5 min; 5%–100% solvent B for 0.5–7 min; 100% solvent B for 7–8 min; 100%–5% solvent B for 8–8.1 min; and 5% solvent B for 8.1–10 min. The column temperature was maintained at 35°C. The TripleTOF 5,600 Plus system was used to detect the metabolites eluted from the column. The curtain gas pressure was set to 30 psi, and the pressure of the ion source gas 1 and 2 was set to 60 psi. The temperature of the interface heater was set to 650°C. For the positive and negative ion modes, the ion spray float voltage values used were 5 and −4.5 kV, respectively. MS data were collected in IDA mode. The time of flight mass range used was 60–1,200 Da. Survey scans were collected every 150 ms, and up to 12 production scans were collected if the threshold of 100 counts exceeded the 1+ charge state. The total cycle time was set to 0.56 s. The four-time periods of each scan were summed at a pulse frequency of 11 kHz by monitoring using a 40 GHz multi-channel thermal conductivity detector with four anodes/channel detection. Dynamic exclusion was set to 4 s. During the entire collection period, every 20 samples, the system was calibrated for mass accuracy, and every ten samples, one QC sample was analyzed to evaluate the stability of the LC-MS.
Metabolomic Data Processing

XCMS was used to preprocess the obtained LC-MS data. The original data file was converted to the mzXML format and then processed using the “XCMS,” “CAMERA,” and “metal” packages in R. Each ion was identified based on comprehensive information about the retention time and m/z. The intensity of each peak was determined, and a three-dimensional matrix containing any specified peak index (retention time-m/z pair), the sample name (observation result), and ion intensity information (variable)

### TABLE 1 | The RNA-seq data for these 6 samples.

| Sample | Raw reads (G) | Raw bases (G) | Valid reads | Valid bases (G) | Valid % | Q20% | Q30% | GC % |
|--------|---------------|---------------|-------------|----------------|--------|------|------|------|
| D_R_1  | 41,486,368    | 6.22          | 40,820,710  | 5.72           | 98.40  | 98.18| 93.77| 43.69|
| D_R_2  | 52,422,016    | 7.86          | 51,496,090  | 7.21           | 98.23  | 98.21| 93.91| 43.59|
| D_R_3  | 47,555,580    | 7.13          | 46,691,176  | 6.54           | 98.18  | 98.29| 94.12| 43.92|
| S_R_1  | 47,065,328    | 7.06          | 46,114,706  | 6.46           | 97.96  | 98.18| 93.86| 43.83|
| S_R_2  | 48,422,758    | 7.26          | 47,510,030  | 6.65           | 98.12  | 98.22| 93.97| 43.79|
| S_R_3  | 48,893,016    | 7.33          | 48,060,970  | 6.73           | 98.30  | 98.27| 94.05| 43.71|

### TABLE 2 | Summary of assembly results of SD.

| Index   | All GC% | Min length | Median length | Max length | Total assembled bases | N50 |
|---------|---------|------------|---------------|------------|-----------------------|-----|
| Transcript | 40.10   | 201        | 755           | 15,586     | 108,266,339           | 1,599|
| Gene    | 40.18   | 201        | 610.00        | 15,586     | 43,027,640            | 1,617|

### TABLE 3 | Summary of function annotation of SD.

| DB           | Num | Ratio (%) |
|--------------|-----|-----------|
| All          | 44,258 | 100.00    |
| GO           | 23,417 | 52.91     |
| KEGG         | 18,487 | 41.77     |
| Pfam         | 20,999 | 47.45     |
| SwissProt    | 19,125 | 43.21     |
| eggNOG       | 25,735 | 58.15     |
| NR           | 28,726 | 64.91     |

Metabolomic Data Processing

XCMS was used to preprocess the obtained LC-MS data. The original data file was converted to the mzXML format and then processed using the “XCMS,” “CAMERA,” and “metal” packages in R. Each ion was identified based on comprehensive information about the retention time and m/z. The intensity of each peak was determined, and a three-dimensional matrix containing any specified peak index (retention time-m/z pair), the sample name (observation result), and ion intensity information (variable)
was generated. Then, the information was matched with internal and public databases. KEGG and the human metabolome database were used to annotate the metabolites by matching their molecular weight data (m/z) within the 10-ppm threshold. The “metal” R package was used to further process the peak intensity data. The features detected in <50% of the QC samples or 80% of the test samples were deleted, and the k-nearest neighbor algorithm was used to extrapolate the

FIGURE 5 | Number of DEGs in the single-head root and double-head root SD.

FIGURE 6 | Analysis of D_R vs S_R differentially expressed gene GO enrichment.
values of the missing peaks to further improve data quality. Principal component analysis was performed to use the processed data set to detect any outliers and batch effects. QC data about the injection sequence were fitted to Q C - b a s e d local regressions to minimize the signal intensity drift over time. In addition, the relative standard deviations of the metabolic characteristics were calculated for all QC samples, and those with standard deviations >30% were eliminated. Before performing QC-robust spline batch correction, the data were normalized using a probability quotient normalization algorithm. Student’s t-test was used to determine the significant differences, and then the Benjamini-Hochberg procedure was used to adjust for multiple tests. Furthermore, the “metal” R package was used to perform a supervised partial least squares discriminant analysis using a variable importance cutoff value of 1.0 to identify the differentially expressed metabolites (DEMs) between the two phenotypes.

Quantitative Reverse Transcription-PCR
The instructions of the TaKaRa MiniBEST Universal RNA Extraction Kit were followed for total RNA extraction (TaKaRa, Kusatsu, Shiga, Japan), and the PrimeScriptTM RT Master Mix Kit (TaKaRa) was used for reverse transcription. RT-qPCR was performed on a 96-well plate using an Agilent Technologies Stratagene Mx3000P thermal cycler and a SYBR Green-based PCR kit. Each reaction involved 1 μl of cDNA template (1 mg/ml), 10 μl of SYBR Green Mix (TaKaRa), 1 μl of forward primer (1 mM), 1 μl of reverse primer (1 mM), and 7 μl of ddH2O. The thermal conditions were as follows: 95°C for 3 min, followed by 40 cycles at 95°C for 5 s, 55°C for 32 s, and 72°C for 20 s. The melting curve was obtained by gradually increasing the temperature from 55°C to 95°C at a heating rate of 0.1°C/s. RT-qPCR analysis was performed using three biological replicates. GAPDH (glyceraldehyde-3-phosphate dehydrogenase) was used as a housekeeping gene to estimate the relative gene expression using the 2−ΔΔCt method. The primer sequences used in this study are shown in Supplementary Table S1.

Statistical Analysis
The original data were compiled using the MS Excel 2016 software (Microsoft, Redmond, WA, United States), and SPSS 19.0 was used for data analysis (IBM, Armonk, NY, United States). GraphPad Prism 6.0 (GraphPad Software, San Diego, CA, United States) and Origin 9.0 (OriginLab, Northampton, MA, United States) were used to create the graphic illustrations.

RESULTS
Chromone Content
The chromone content in the single- and double-headed roots is shown in Figure 3. The prim-O-glucosylimcifugin, cimifugin, and total chromone contents in the double-headed roots were significantly higher (p < 0.01) than those in the single-headed roots, with 2.36, 1.76, and 1.84-fold differences, respectively. On the basis of the detection indicators stipulated in the 2020 edition of the Chinese Pharmacopoeia, the measurement indicators were enriched (Commission, 2020). The measurement results showed
that the double-headed roots were superior to the single-headed roots in terms of the effective ingredients. Prim-O-glucosylcimifugin is colling and relieves pain, which are good pharmacological effects; therefore, it can be used as the first choice medicine in clinical practice.

**Transcriptomic Data Analysis**

Six cDNA libraries from each phenotype generated 285,845,066 raw reads. After removing the adapters and filtering out the low-quality sequences, 280,693,682 clean reads were obtained from the six libraries (Table 1). After de novo assembly, mapping to contigs, and redundancy removal, 44,258 unigenes were obtained with an N50 length of 1,599 nucleotides (Table 2, Supplementary Figure S1). Therefore, the quality of the experimental data obtained via sequencing was high, meeting the conditions for the subsequent experimental analysis. Of the 44,258 unique sequences, 90.67% showed homology to known genes of existing models (88.51% to *Daucus carota*, 0.87% to *Quercus suber*, 0.57% to *Vitis vinifera*, 0.36% to *Actinidia chinensis*, and 0.36% to *Helianthus annuus*; Table 3 and Figure 4). Salmon was used to determine the expression levels of the unigenes. A total of 5,208 DEGs were screened in the D_R vs. S_R comparison group, of which 2,829 were upregulated and 2,379 were downregulated. The number of upregulated genes was higher than that of downregulated genes, indicating that the DEGs upregulated in the double-headed roots can be used as important candidate genes to identify the quality differences between both phenotypes (Figure 5).

Furthermore, GO function enrichment analysis revealed that the DEGs were enriched in biological processes (i.e., biological process, regulation of transcription, DNA-templated transcription, DNA-templated defense response, and protein phosphorylation), cellular components (i.e., nucleus, plasma membrane, cytoplasm, integral components of membrane, and chloroplast), and molecular functions (i.e., molecular function, protein binding, DNA binding transcription factor activity, ATP binding, and DNA binding) (Figure 6). To further determine the internal cause of the difference in quality observed between the

**TABLE 4 | Candidate genes differentially expressed in single-head root and double-head root SD.**

| Name | Gene ID          | Gene expression (log2 ratio) | up/down |
|------|------------------|------------------------------|---------|
| 4CLL9 | TRINITY_DN17215_c0_g5 | ≈1.30                               | down    |
| 4CLL7 | TRINITY_DN15920_c0_g4 | ≈1.53                               | down    |
| CHS3 | TRINITY_DN13271_c0_g7 | ≈1.42                               | down    |
| 4CL  | TRINITY_DN12193_c0_g5 | 1.06                                   | up      |
| CHS2 | TRINITY_DN12770_c0_g2 | 1.20                                   | up      |
| ZEP  | TRINITY_DN6974_c0_g1 | 4.15                                   | up      |
single- and double-headed roots (that is, the metabolic pathways associated with the quality of the medicinal material), we mapped the DEGs to the KEGG database. KEGG pathway analysis showed that the DEGs were highly associated with several pathways, including plant-pathogen interaction, plant hormone signal transduction, MAPK signaling pathway-plant, phenylpropanoid biosynthesis, and galactose metabolism (Figure 7). The DEGs were enriched in plant hormone signal transduction pathways, indicating that hormones may be essential for the quality of the medicinal materials. In addition, many genes are involved in the MAPK signaling pathway and phenylpropanoid biosynthesis-related pathways. Chromone is the most important active ingredient in S. divaricata and the basic framework for flavonoid synthesis. Therefore, this research focused on the synthesis pathways of plant hormones and flavonoids. Cluster analysis demonstrated that the number of upregulated genes was significantly higher than that of downregulated genes Figure 8. The test results indicated that the higher number of upregulated genes in double-headed roots may be related to the accumulation of active ingredients. This experimental result was mutually corroborated by the previous differential gene expression analysis.

A transcription factor is a protein with a special structure that regulates gene expression. Transcription factors initiate and regulate gene expression by recognizing and binding to cis-acting elements in the promoter region of genes. We also performed a statistical analysis of the differential transcription factors in the D_R vs. S_R comparison group and screened 4,166 differential transcription factors. These factors mainly belonged to the MYB (290 genes), bHLH (278 genes), NAC (241 genes), ERF (186 genes), C2H2 (174 genes), WRKY (123 genes), and bZIP (71 genes) families, as shown in Figure 9. This shows that the transcription factors of these families play an important role in the formation of double-headed root medicinal materials.

Furthermore, through the functional analysis of the differential genes and further screening, we obtained six key differential genes related to the synthesis of flavonoids and plant hormones, including five key genes in the flavonoid synthesis pathway (4-coumarate-CoA ligase, chalcone synthase 1, vinorine synthase, chalcone-flavonone isomerase 1, and flavanone 3 beta-hydroxylase) and one key gene in the ABA biosynthetic pathway (zeaxanthin epoxidase) (Table 4). Therefore, the above six genes can be used as key candidate genes to study the quality differences of double-headed root medicinal materials.
FIGURE 12 | Results of hierarchical clustering analysis of altered metabolite pools in the single-headed roots and double-headed roots (p < 0.01). Heatmap color indicate the abundance of each metabolite in the single-headed roots and double-headed roots.
Validation of the Transcriptome Data
To verify the reliability of the transcriptome data, qRT-PCR was performed for ten randomly selected DEGs. qRT-PCR was used to compare and analyze the expression of these genes between samples and the results were consistent with the transcriptome analysis results for all the targets, except for TRINITY_DN12193_c0_g5. These results are summarized in Figure 10 and support our assertion that the transcriptome sequencing data were reliable.

Metabolomic Data and Analysis
The analysis of the metabolomic data revealed 126 DEMs (30 organic acids and derivatives, 30 lipids and lipid-like molecules, 20 phenylpropanoids and polyketides, 11 organoheterocyclic compounds, eight organic oxygen compounds, six benzenoids, two nucleosides, nucleotides, and analogs, one alkaloid and derivatives, one hydrocarbon, one organooxygen compound, and 16 other compounds; Supplementary Table S2). According to principal component analysis for the positive electrospray ionization (ESI) mode, the first PC (PC1) and the second PC (PC2) explained 38.54 and 7.24% of the total measured metabolite variation, respectively, in the direction of the treatments. Similar results were obtained for the negative ESI mode (Figure 11). Using heat map analysis, we observed that the abundance of primary metabolites was significantly different between phenotypes (Figure 12). The major metabolites in the double-headed root phenotype were phenylpropanoids, polyketides, lipids, lipid-like molecules, organic acids, and derivatives. Moreover, the cudraflavone A and coumarin contents were 6.38 and 3.27-fold higher, respectively, in the double-headed roots than in the single-headed roots. KEGG analysis of the DEMs revealed that they were mainly enriched in phenylalanine, tyrosine, tryptophan, and glucosinolate biosynthesis, in the biosynthesis of plant secondary metabolites and plant hormones, 2-oxocarboxylic acid metabolism, and in the biosynthesis of amino acids (Figure 13).

Correlation analysis between the DEGs and the DEMs showed that the genes associated with flavonoid synthesis, such as flavanone isomerase (EC 5.5.1.6), chalcone synthase (EC 2.3.1.74), and vinorine synthase (EC 2.3.1.133), were upregulated, whereas flavanone 3 beta-hydroxylase (EC 1.14.11.9) was downregulated in the double-headed root phenotype (Figure 14). A network diagram of the relevant genes and metabolites was constructed to analyze the correlation between the DEGs and DEMs in the two phenotypes (Figure 15). The results showed that the changes in 5,7-Dimethoxycoumarin, 5,7-Dimethoxy-4-methylcoumarin, 5,7-Dihydroxy-4-methylcoumarin, 7-Allyloxycoumarin, and trans-zeatin-riboside-O-glucoside were correlated with several genes.

DISCUSSION
Here, we assessed the morphology of S. divaricata with single- and double-headed roots, and tested samples for four kinds of medicinal ingredients. Simultaneously, we used transcriptomic and metabolomic analyses to identify the differences between the two phenotypes. We found that plants with double-headed roots had a superior yield potential and efficiency in accumulating organic matter. Moreover, the contents of prim-O-glucosylcimifugin, cimifugin, and total chromone in the double-headed root phenotype were significantly higher than those in the single-headed root phenotype. Flavonoids are a major class of secondary metabolites (Pourcel et al., 2007; Takao et al., 2021), and the chromone content is considered
an important index for evaluating the medicinal value of a plant (Lv et al., 2019; Kuang et al., 2020). The active ingredients in \( S. \text{ divaricata} \), prim-O-glucosylcimifugin and 4′-O-β-D-glucosyl-5-O-methylvisamminol, can significantly reduce the body temperature of rats with yeast-induced fever and simultaneously effectively inhibit the pain caused by various stimuli in mice (Xue et al., 2000). Furthermore, research findings show that chromone has a significant antipyretic effect (Yang D.-s. et al., 2017). Double-headed roots have high contents of medicinal ingredients; thus, this phenotype might have better therapeutic properties. Therefore, it is of great significance to elucidate the internal mechanism of the quality difference between single- and double-headed root plants.

Prim-O-glucosylcimifugin and 4′-O-β-D-glucosyl-5-O-methylvisamminol are quality evaluation indicators, as stipulated in the 2020 edition of the Chinese Pharmacopoeia. Both belong to the chromone group of ingredients. The differential expression of the key enzyme genes (i.e., 4-coumarate-CoA ligase, chalcone synthase, and chalcone-flavonone isomerase) in the synthesis pathway leads to differences in the contents of the medicinal components (Cheng et al., 2018). Thus, the analysis of its synthesis pathway is vital to explain the difference in the quality of medicinal materials. In the present study, transcriptome sequencing analysis results showed significant differences regarding 4-coumarate-CoA ligase, chalcone synthase, and chalcone-flavonone isomerase in the flavonoid synthesis pathway. Different copies of 4-coumarate-CoA ligase showed inconsistent expression patterns; chalcone synthase 1 was upregulated, whereas chalcone-flavonone isomerase 1 was downregulated in the double-headed root phenotype, suggesting modifications in the flavonoid biosynthesis pathway. Thus, the accumulation of active ingredients in the double-headed root phenotype might be closely related to the high expression of flavonoid synthesis genes. Phytohormones refer to trace organic matter that is synthesized to regulate various plant processes. ABA affects plant growth and development by regulating the plant and environmental signals and phytohormones, which are small molecular signal substances that transmit exogenous signals to the internal biosynthesis process of the plant (Hong and Xu, 2013). Here, the key enzyme gene of the ABA synthesis pathway, zeaxanthin epoxidase, was found in the double-headed root.

FIGURE 14 | Pathway analysis of flavonoid biosynthesis pathways in single-headed roots and double-headed roots. The proposed metabolic pathways are based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) database of metabolic pathways. The metabolites, written in circle, were detected in this study. Omitted metabolic processes are represented by dashed lines. The rectangle represent the genes. The red colors indicates upregulated expression and green colors indicates downregulated expression.
phenotype, with a 4.15-fold difference compared to the single-headed root phenotype ($p < 0.01$), leading to changes in ABA accumulation, and then transmitting the signal to flavonoids. In the synthetic pathway, the differential accumulation of the secondary metabolites of Fangfeng eventually leads to differences in the quality of medicinal materials (Zhang et al., 2010). These six genes can be used as important candidate genes to reveal the differences in the quality of the single- and double-headed root *S. divaricata* medicinal materials.

Metabolomics analysis helps identify the changes in metabolites (Hall et al., 2002), which are the final products of cell regulation processes and are regarded as responses to genetic and environmental cues (Fiehn, 2002). Therefore, metabolomic analysis enables the investigation of the relationship between biological processes and phenotypes. Furthermore, some intuitive changes can also be observed at the metabolic level (Zhang et al., 2017). Previous studies have examined the metabolic responses of different plant species, such as *Fagopyrum esculentum* (Kim et al., 2013) and *Camellia sinensis* (Yang JM. et al., 2017), and have suggested that they affect flavonoid biosynthesis. Therefore, here, we conducted a metabonomic analysis of samples based on transcriptomic sequencing. Here, we identified 126 DEMs between the two phenotypes, of which cudraflavone A and coumarin showed the biggest differences (6.38-fold and 3.27-fold, respectively). Moreover, the DEMs were mainly enriched in the biosynthesis of secondary metabolites and phytohormones. Thus, metabolomics might help elucidate the internal mechanism leading to the medicinal property differences between plants with single- and double-headed roots.

We used the transcriptome and metabolome data to identify the genes related to flavonoid biosynthesis and explain any differences between the single- and double-headed root phenotypes. Plants with double-headed roots probably have superior medicinal value than those with single-headed roots. Six DEGs were identified, including five key genes in the flavonoid synthesis pathway and one in the ABA synthesis pathway. Correlation analysis revealed that most of the DEGs were significantly related to chromone and coumarin metabolites. Moreover, *trans*-zeatin-riboside-O-glucoside, an intermediate metabolite in ABA synthesis, was also correlated with the DEGs. Thus, the accumulation of active ingredients in double-headed root plants is closely related to hormone synthesis, which is affected by flavonoid synthesis. ABA regulation in *S. divaricata* might also affect its medicinal properties.

**CONCLUSION**

Our study showed significant differences in the quality of medicinal materials between single- and double-headed roots. The prim-O-glicosylcimifugin, cimifugin, and total chromone contents in double-headed roots were significantly higher than those in the single-headed roots. Using metabolome and transcriptome analyses, we identified six DEGs and 126 DEMs between the single- and double-headed root phenotypes, indicating that flavonoids and ABA synthesis pathways may play an important role in the synthesis of the active ingredients of *S. divaricata*. Furthermore, the six genes can be
used as important candidate genes for analyzing the quality differences in wind-breaking medicinal materials. In general, our research provides important information on the quality differences of the *S. divaricata* medicinal materials. The double-headed root phenotype can be used as a candidate phenotype for clinical drugs, providing theoretical support for selecting and breeding *S. divaricata* medicinal materials.

**DATA AVAILABILITY STATEMENT**

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding authors.

**AUTHOR CONTRIBUTIONS**

TZ, WQ, and CC conceived and designed the experiments. TZ performed most of the experiments, analyzed the data, and completed the first draft. YC, QZ, PY, and QL worked together with TZ to perform the experiments. All authors read and approved the manuscript.

**REFERENCES**

Armstrong, F., and Que, L. (2012). Current Opinion in Chemical Biology. Curr. Opin. Chem. Biol. 16 (1-2), 1–2. doi:10.1016/j.cbpa.2012.03.011

Buchfink, B., Xie, C., and Huson, D. H. (2015). Fast and Sensitive Protein Alignment Using DIAMOND. Nat. Methods 12 (1), 59–60. doi:10.1038/nmeth.3176

Cheng, L., Han, M., Yang, L.-m., Li, Y., Sun, Z., and Zhang, T. (2018). Changes in the Physiological Characteristics and Baicalin Biosynthesis Metabolism of Scutellaria Baicalensis Georgii under Drought Stress. Ind. Crops Prod. 122 (122), 473–482. doi:10.1016/j.indcrop.2018.06.030

Commission, C. P. (2020). Pharmacopoeia of the People’s Republic of China. China: The Medicine Science and Technology Press of China.

Fiehn, O. (2002). Metabolomics - the Link between Genotypes and Phenotypes. Plant Mol. Biol. 48, 155–171. doi:10.1023/A:1015852112314

Forrest, A. R. R., and Carninci, P. (2009). Whole Genome Transcriptome Analysis. RNA Biol. 6 (2), 107–112. doi:10.4161/rna.6.2.7931

Gao, W., Sun, H. X., Xiao, H., Cui, G., Hillwig, M. L., Jackson, A., et al. (2014). Combining Metabolomics and Transcriptomics to Characterize Tanshinone Biosynthesis in Salvia Miltiorrhiza. Bmc Genomics 15 (1), 73–14. doi:10.1186/1471-2164-15-73

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2011). Full-length Transcriptome Assembly from RNA-Seq Data without a Reference Genome. Nat. Biotechnol. 29 (7), 644–652. doi:10.1038/nbt.1883

Hall, R., Beale, M., Fiehn, O., Hardy, N., Summer, L., and Bino, R. (2002). Plant Metabolomics. Plant Cell 14, 1437–1440. doi:10.1105/tpc.140720

Hamilton, J. P., and Robin Buell, C. (2012). Advances in Plant Genome Sequencing. Plant J. 70 (1), 177–190. doi:10.1111/j.1365-313X.2012.04894.x

Hong, J. H., Seah, S. W., and Xu, J. (2013). The Root of ABA Action in Environmental Stress Response. Plant Cell Rep 32 (7), 971–983. doi:10.1007/s00299-013-1439-9

Hongqe, C., Mohammed, N., Hao, X., Jingjia, H., Kunlu, W., Xianghui, C., et al. (2015). Transcriptome Analysis of Methyl Jasmonate-Elicited Panax ginseng Adventitious Roots to Discover Putative Ginsenoside Biosynthesis and Transport Genes. Int. J. Mol. Sci. 16 (2), 3035–3057. doi:10.3390/ijms16023035

**FUNDING**

This work was financially supported by the National Key Research and Development Plan of China (2019YFC1710704), the Natural Science Foundation of Jilin Province (YDZJ202101ZYTS012), and the fifth batch of Jilin Province young scientific and technological talent support project (QT202117).

**ACKNOWLEDGMENTS**

We would like to thank WQ and CC, Changchun University of Chinese Medicine, for providing financial support for part of this work.

**SUPPLEMENTARY MATERIAL**

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2021.764093/full#supplementary-material

**Supplementary Figure 1** | The length distribution and GC content.

Jameson, P. E., Dhandapani, P., Novak, O., and Song, J. (2016). Cytokinins and Expression of SWEET, SUT, CWINV and AAP Genes Increase as Pea Seeds Germinate. Int. J. Mol. Sci. 17 (12), 1–13. doi:10.3390/ijms17122013

Khan, S., Shin, E. M., Choi, R. J., Jung, Y. H., Kim, J., Tosun, A., et al. (2011). Suppression of LPS-Induced Inflammatory and NF-κB Responses by Anomalin in RAW 264.7 Macrophages. J. Cell. Biochem. 112 (8), 2179–2188. doi:10.1002/jcb.23137

Kim, Y. B., Park, S.-Y., Thwe, A. A., Seo, J. M., Suzuki, T., Kim, S.-J., et al. (2013). Metabolomic Analysis and Differential Expression of Anthocyanin Biosynthetic Genes in White- and Red-Flowered Buckwheat Cultivars (Fagopyrum Esculentum). J. Agric. Food Chem. 61 (44), 10525–10533. doi:10.1021/jf402238f

Kuang, L., Li, C., Cao, S., Li, H., Wang, Z., Wu, S., et al. (2020). Pharmacopoeia of People’s Republic of China. Beijing: China Medical Science and Technology Press.

Li, C., Zhu, Y., Guo, X., Sun, C., Luo, H., Song, J., et al. (2013). Transcriptome Analysis Reveals Ginsenosides Biosynthetic Genes, microRNAs and Simple Sequence Repeats in Panax ginsengC. A. Meyer. Bmc Genomics 14 (1), 245. doi:10.1186/1471-2164-14-245

Li, Y., Wang, X., Li, Z., Zhang, J., Meng, H., Lv, Q., et al. (2017). Separation and Identification of Chemical Constituents of Saposhnikovia Divaricata. Chin. J. Exp. Formulas 23 (15), 60–64. doi:10.16166/j.jchrbmb.2010.09.027

Liu, L., Zhang, N., Jin, H., and Ma, S. (2020). Comparative Study on Chromones in Wild Boidet and Unboidet Saposhnikovia Divaricata. Chin. Pharm. J. 55 (08), 637–642. doi:10.11669/cpj.2020.08.012

Liu, S., Wang, S., Liu, M., Yang, F., Zhang, H., Liu, S., et al. (2016). De Novo sequencing and Analysis of the Transcriptome of Panax Ginseng in the Leaf-Expansion Period. Mol. Med. Rep. 14 (2), 1404–1412. doi:10.3892/mmr.2016.5376

Liu, S., and Zhang, L. (2007). Quality Analysis of Different Medicinal Parts of Cultivated Saposhnikovia Divaricata. Ginseng Res. (1), 8–11.

Liu, Y. M., Zhang, L., Zhou, J. Y., and Cao, M. J. (2015). Research Progress of the bHLH Transcription Factors Involved in Genic Male Sterility in Plants. Yi Chuan 37 (12), 1194–1203. doi:10.16288/j.yczs.15-229

Lv, X., Chen, S., and Wang, Y. (2019). Advances in Understanding the Physiological and Molecular Responses of Sugar Beet to Salt Stress. Front. Plant Sci. 10, 1431. doi:10.3389/fpls.2019.01431
Madhunita, B., and Ralf, O. (2014). WRKY Transcription Factors: Jack of many Trades in Plants. *Plant Signal. Behav.* 9 (2), e27700. doi:10.4161/psb.27700

Martin, M. (2011). Cutadapt Removes Adapter Sequences from High-Throughput Sequencing Reads. *Embrret J.* 17 (1), 1–3. doi:10.14806/ej.17.1.200

Okuyama, E., Hasegawa, T., Matsushita, T., Fujimoto, H., Ishibashi, M., and Yamazaki, M. (2019). Analgetic Components of Saposhnikovia Root (Saposhnikovia Divaricata). *Chem. Pharm. Bull. (Tokyo)* 49 (2), 154–160. doi:10.1248/cpb.49.154

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., and Kingsford, C. (2017). Salmon Provides Fast and Bias-Aware Quantification of Transcript Expression. *Nat. Methods* 14 (4), 417–419. doi:10.1038/nmeth.4197

Pourcel, L., Routaboul, J., Cheynier, V., Lepiniec, L., and Debeaujon, I. (2007). Flavonoid Oxidation in Plants: from Biochemical Properties to Physiological Functions. *Trends Plant Sci.* 12 (1), 29–36. doi:10.1016/j.tplants.2006.11.006

Ramaseswami, G., Zhang, R., Piskol, R., Keegan, L. P., Deng, P., O’Connell, M. A., et al. (2013). Identifying RNA Editing Sites Using RNA Sequencing Data Alone. *Nat. Methods* 10 (2), 128–132. doi:10.1038/nmeth.2330

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: A Bioconductor Package for Differential Expression Analysis of Digital Gene Expression Data. *Bioinformatics* 26 (1), 139–140. doi:10.1093/bioinformatics/btp616

Rodriguez-Garcia, A., Sola-Landa, A., and Barreiro, C. (2017). RNA-Seq-Based Comparative Transcriptomics: RNA Preparation and Bioinformatics. *Methods Mol. Biol.* 1645, 59–72. doi:10.1007/978-1-4939-7183-1_5

Sui, C., Zhang, J., Wei, J., Chen, S., Li, Y., Xu, J., et al. (2011). Transcriptome Analysis of Bupleurum Chinense Focusing on Genes Involved in the Biosynthesis of Saposhnikoviae. *Bmc Genomics* 12 (1), 539. doi:10.1186/1471-2164-12-539

Takao, K., Nagaishi, Y., Nagai, J., Hoshii, K., Nakamura, S., et al. (2021). Synthesis and Biological Evaluation of 3-styrylchromone Derivatives as Selective Monoamine Oxidase B Inhibitors. *Bioorg. Med. Chem.* 42, 116253. doi:10.1016/j.bmc.2021.116253

Werner, T., and Schmülling, T. (2009). Cytokinin Action in Plant Development. *Curr. Opin. Plant Biol.* 12, 527–538. doi:10.1016/j.pbi.2009.07.002

Wu, D., Austin, R. S., Zhou, S., and Brown, D. (2013). The Root Transcriptome for North American Ginseng Assembled and Profiled across Seasonal Development. *Bmc Genomics* 14 (1), 564. doi:10.1186/1471-2164-14-564

Xue, B., Li, W., Li, L., and Xiao, Y. (2000). Study on the Pharmacological Activity of the Chromone Glycosides of Saposhnikovia Divaricata. *China J. Chin. Materia Med.* 25 (5), 41–45. doi:10.3321/j.issn:1001-5302.2000.05.016

Yang, D.-s., Zhang, J., Li, M.-x., and Shi, L.-x. (2017a). Metabolomics Analysis Reveals the Salt-Tolerant Mechanism in Glycine Soja. *J. Plant Growth Regul.* 36 (2), 460–471. doi:10.1007/s10344-016-9654-6

Yang, J. M., Jiang, H., Dai, H. L., Wang, Z. W., Jia, G. Z., and Meng, X. C. (2017b). Fleeble Antipyretic, Analgesic, and Anti-inflammatory Activities Were Found with Regular Dose 4‘-O-β-D-Glucosyl-5-O-Methylvisamminol, One of the Conventional Marker Compounds for Quality Evaluation of Radix Saposhnikoviae. *Pharmacogn Mag.* 13 (49), 168–174. doi:10.4103/0973-1296.197637

Zhang, H., Han, W., De Smet, I., Talboys, P., Loya, R., Hassan, A., et al. (2010). ABA Promotes Quiescence of the Quiescent centre and Suppresses Stem Cell Differentiation in the Arabidopsis Primary Root Meristem. *Plant J.* 64 (5), 764–774. doi:10.1111/j.1365-313x.2010.04367.x

Zhang, J. Y., Shen-Chun, Q. U., Guo, Z. R., Xiao-Li, D. U., Bei-Bei, D. U., and Zhang, Z. (2011). Biology Function of bZIP Transcription Factors in Plants. *Acta Botanica Boreali-Occidentalia Sinica* 31 (3), 1066–1075.

Zhang, Q., Liu, M., and Ruan, J. (2017). Metabolomics Analysis Reveals the Metabolic and Functional Roles of Flavonoids in Light-Sensitive tea Leaves. *BMC Plant Biol.* 17 (1), 64. doi:10.1186/s12870-017-1012-8

Zhang, T., Gao, Y., Han, M., and Yang, L. (2021). Changes in the Physiological Characteristics of Panax Ginseng Embryogenic Calli and Molecular Mechanism of Ginsenoside Biosynthesis under Cold Stress. *Planta* 253 (79), 1–23. doi:10.1007/s00425-020-03535-7

Zou, L.-q., Kuang, X.-j., and Sun, C. (2015). Advances in Transcriptomic Studies and Ginsenoside Biosynthesis of American Ginseng. *Chin. Herbal Medicines* 7 (2), 116–122. doi:10.1674/1674-6384(15)06028-4

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher’s Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

**Copyright © 2021 Zhang, Chen, Zhang, Yu, Li, Q and Chen.** This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.