Transcriptomes Analysis Reveals the Regulatory Roles of Noncoding RNA in Tanshinones Synthesis Pathway of *Salvia Miltiorrhiza*

Caicai Lin

Shandong Agricultural University

Changhao Zhou

Shandong Agricultural University

Zhongqian Liu

Shandong Agricultural University

Xingfeng Li

Shandong Agricultural University

Zhenqiao Song (✉ szq@sdau.edu.cn)

College of Agronomy, Shandong Agricultural University

Research

Keywords: miRNA, lncRNA, circRNA, *Salvia miltiorrhiza*, tanshinone

Posted Date: August 9th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-736448/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Long noncoding RNAs (lncRNAs), circular RNAs (circRNAs), and microRNAs (miRNAs) have been shown to play fundamental roles in plant development. However, the information of these noncoding RNAs (ncRNAs) in *Salvia miltiorrhiza* remains largely unexplored. In this study, the expression pattern of ncRNAs in six tissues from the same strain of *S. miltiorrhiza* was analyzed to study the biological function of ncRNAs on active ingredients synthesis.

Methods: Analysis of tanshinone content differences of two root samples was carried out on high-performance liquid chromatography (HPLC). RNA sequencing, GO and KEGG enrichment analysis were applied to analyzing the targets of differentially expressed ncRNAs in different organs.

Results: A total of 6,929 lncRNAs, 6,239 circRNAs, and 360 miRNAs were identified. Forty-eight lncRNAs, 70 miRNAs, and 26 circRNAs expressed differentially between red and white root tissues with significantly different tanshinone content. GO and KEGG pathway analysis of target genes of differently expressed ncRNAs indicated that some target genes are involved in the synthesis pathway of terpene, including diterpene and sesquiterpene. We also found many target genes related to secondary metabolites, including 2-C-Methyl-d-erythritol 2,4-cyclodiphosphate Synthase (*SmMCS*) and several CYP450s. Furthermore, most target genes may be related to the resistance of pathogens, such as receptor kinases, disease-resistant proteins, and pentatricopeptide repeat-containing proteins.

Conclusions: The present study exhibited the tissue-specific expression patterns of ncRNAs preliminarily in *S. miltiorrhiza*, which may reflect that the formation of white root or red root is related to regulation by ncRNAs. It would provide a basis for further research about the regulation mechanism in the tanshinone synthesis process.

Background

With the rapid development of high-throughput sequencing technology, whole-transcriptome research sequencing analysis can integrate various RNA information, including mRNA and ncRNA (miRNA, lncRNA, circRNA, etc.), and comprehensively explore the potential network of regulation mechanism at the RNA level [1–3]. Mounting evidence suggests that ncRNAs transcribed from DNA are not useless "junk" while playing a pivotal role in controlling gene expression by binding to proteins or other molecules [4–7].

Recent progress has revealed that miRNAs had an essential regulatory function in the secondary metabolism of the plant [8]. Several target genes regulated by the miR414 family can encode enzyme which play an essential role in artemisinin biosyntheses, such as HMG-CoA reductase (HMGR) and farnesyl pyrophosphate synthase (FPS), which are critical enzymes in artemisinin biosynthesis, and also the amorpho-4,11-diene synthase (ADS) and CYP450, which participate in the synthesis of artemisinic acid, a precursor of artemisinin [9]. High throughput sequencing and degradation group sequencing analysis showed that miR171 and miR164 targeted two genes that lead to paclitaxel biosynthesis, taxane 2 α -O-benzoyltransferase and taxane 13 α hydroxylase, respectively [10]. Another study showed that the

genes involved in essential oil biosynthesis might be regulated by miR156, miR414, and miR5021 [11]. Until now, few studies about the function of miRNAs have been reported in *S. miltiorrhiza*. miR5072 was predicted to regulate acetyl-CoA C-acetyltransferase (AACT), which is the upstream rate-limiting enzyme in the mevalonic acid (MVA) pathway of the terpenoid backbone and tanshinone synthesis [12]. Five miRNAs from different families may be involved in response to the continuous cropping obstacle of *S. miltiorrhiza*, whose target genes including *SPL*, *ARF*, *NAC*, and *GRF*, which were related to root growth and development [13].

Different reports also proved that lncRNAs also play a significant role in a wide range of cellular mechanisms, from almost all aspects of gene expression to protein translation and stability [14–16]. lncRNAs regulate various critical biological pathways, including development, metabolism, disease resistance, stress response, hormone signaling, and maintenance of genome integrity [17, 18]. The regulation and function mechanism of lncRNAs were mainly identified in the model plant, such as *Zea mays* L., *Oryza sativa* L. [19], and *Arabidopsis thaliana* [20] mainly focus on growth and development [21], stress response [22] as well as reproductive development.

Studies have revealed that spatiotemporal expression characters of circRNAs suggested that they might involve in gene regulation and affecting the development of different tissues [23–25]. In addition, circRNAs also participate in regulatory abiotic and biological stress in plants [26, 27]. CircRNA can also regulate the transcriptional expression of source genes and act as a competitive endogenous RNA (ceRNA) to bind intracellular miRNAs and block the inhibition of miRNAs on target genes [28].

Salvia miltiorrhiza bunge, the red root of ‘Danshen’ in Chinese, is a well-known and vital traditional Chinese herbal medicine [29]. It has medicinal value due to its famous blood-activating and stasis-relieving effects. It has been widely applied in treating cardio-cerebrovascular diseases, such as angina, coronary heart disease, myocardial infarction in China and other Asian countries for thousands of years [30, 31]. The main active components of dry, red root in *S. miltiorrhiza* are lipid-soluble tanshinones and water-soluble phenolic acids compounds [32, 33]. Tanshinones are a kind of diterpenoid quinone and also a pigment component. About 40 kinds of tanshinones have been isolated and identified in *S. miltiorrhiza*, such as tanshinone IIA (flaky orange crystal), tanshinone I (reddish-brown columnar crystal), tanshinone IIB (purple acicular crystal), cryptotanshinone (tangerine acicular crystal), dihydrotanshinone I (red columnar crystal), and so on [34–36].

Tanshinones are mainly biosynthesized and accumulated in the periderm of *S. miltiorrhiza* root. The content of tanshinone IIA in the periderm is ~ 185 times higher than that in the phloem and ~ 17 times higher than that in the xylem of *S. miltiorrhiza* root [37]. In general, the color of periderm in *S. miltiorrhiza* root gradually turned red because of the accumulation of tanshinone [38, 39]. Aboveground organs (such as leaves) also contained significantly lower tanshinone, which implies that genes related to tanshinone synthesis may be preferentially expressed in the periderm of roots [40]. Recent studies confirmed that three essential genes (*SmCPS1*, *SmKSL1*, *SmCYP76AH1*) were involved in the downstream pathway of

tanshinone synthesis. Several essential genes were involved in the upstream pathway were mainly expressed in the periderm tissue [41].

In order to interpret the function of ncRNAs in the development and accumulation of active components in *S. miltiorrhiza*, red and white roots, stems, leaves, and flower tissues were used as materials to conduct noncoding RNA sequencing analysis in the present study. The results would benefit from predicting and screening potential ncRNA involved in the synthesis of active components and exploring its function in the growth and development of *S.miltiorrhiza*.

Materials And Methods

Plant materials and growth conditions

The *S. miltiorrhiza* line DSS3 used in this experiment were grown in a glass bottle with diameter of 10 cm and height of 40 cm in the State Key Laboratory of Crop Biology, Shandong Agricultural University, Taian, Shandong province, China. Plantlets were harvested after grown for 6 months, washed with distilled water, divided into 6 types of tissues including leaf, stem, flower, phloem of white root, phloem of red root, and xylem of red root. Those tissues were immediately frozen in liquid nitrogen and stored at -80°C for further experiment.

HPLC analysis of tanshinone content

Analysis of tanshinone content differences of two root samples was carried out on high-performance liquid chromatography (HPLC). The methods of extraction and detection used in this study were established by our laboratory [42] and mainly used for quantitation of known components.

Chromatographic separations were carried out in a reverse-phase C18 column (250 × 4.6 mm, five µm particle size; Thermo, USA), and a 20.0 × 4.6 mm guard column connected to a Waters 600E HPLC System which equipped with an auto-injector, UV detector, and Empowers software (Waters Associates, Milford, MA, USA).

Methods of RNA extraction, detection, and profound sequence of ncRNAs

The RNA samples were extracted with Trizol. RNA integrity was assessed using the RNA Nano 6000 Assay Kit of the Agilent Bioanalyzer 2100 system (Agilent Technologies, CA, USA) to ensure the use of qualified samples for sequencing. RNA degradation and contamination, especially DNA contamination, was monitored on 1.5% agarose gels. RNA concentration and purity were measured using the NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific, Wilmington, DE).

Library preparation for sRNA sequencing

A total of 2.5 ng RNA per sample was used as input material for the RNA sample preparations. Sequencing libraries were generated using NEB Next Ultra-small RNA Sample Library Prep Kit for Illumina

(NEB, USA) following the manufacturer's recommendations. Index codes were added to attribute sequences to each sample. First of all, ligated the 3'SR Adaptor. The mixture system included Mixed 3'SR Adaptor, RNA, and Nuclease-Free Water, which were incubated for 2 minutes at 70 degrees in a preheated thermal cycler. The tube was transferred to ice. Then, add 3'Ligation Reaction Buffer (2X) and 3'Ligation Enzyme Mix ligate the 3'SR Adaptor. They were incubated for 1 hour at 25°C in a thermal cycler. To prevent adaptor-dimer formation, the SR RT Primer hybridizes to the excess of 3'SR Adaptor (that remains free after the 3'ligation reaction). It transforms the single-stranded DNA adaptor into a double-stranded DNA molecule. DsDNAs are not substrates for ligation-mediated. Then, reverse transcription synthetic first chain. After that, PCR amplification and Size Selection. PAGE gel was used for electrophoresis fragment screening purposes, rubber cutting recycling as the pieces get small RNA libraries. At last, PCR products were purified (AMPure XP system), and library quality was assessed on the Agilent Bioanalyzer 2100 system.

Library preparation for lncRNA-Seq and circRNA-Seq

A total amount of 1.5 µg RNA per sample was used as input material for rRNA removal using the Ribo-Zero rRNA Removal Kit (Epicentre, Madison, WI, USA). Sequencing libraries were generated using NEBNext[®] Ultra™ Directional RNA Library Prep Kit for Illumina (NEB, USA). The manufacturer's recommendations and index codes were added to attribute sequences to each sample. Briefly, Fragmentation was carried out using divalent cations under elevated temperature in NEBNext First Strand Synthesis Reaction Buffer (5X). First-strand cDNA was synthesized using a random hexamer primer and Reverse Transcriptase. Second-strand cDNA synthesis was subsequently performed using DNA Polymerase I and RNase H. Remaining overhangs were converted into blunt ends via exonuclease/polymerase activities. After adenylation of 3' ends of DNA fragments, NEBNext Adaptor with hairpin loop structure was ligated to prepare for hybridization. In order to select insert fragments of preferentially 150 ~ 200 bp in length, the library fragments were purified with AMPure XP beads (Beckman Coulter, Beverly, USA). Then three µl USER Enzyme (NEB, USA) was used with size-selected, adaptor-ligated cDNA at 37°C for 15 min before PCR. Then PCR was performed with Phusion High-Fidelity DNA polymerase, Universal PCR primers, and Index(X) Primer. At last, PCR products were purified (AMPure XP system), and library quality was assessed on the Agilent Bioanalyzer 2100 and qPCR.

Clustering, sequencing, and Quality control

According to the manufacturer's instructions, the clustering of the index-coded samples was performed on a cBot Cluster Generation System using TruSeq PE Cluster Kit v4-cBot-HS (Illumina). After cluster generation, the library preparations were sequenced on an Illumina HiSeq Xten platform, and paired-end reads were generated. Raw data (raw reads) of fastq format were firstly processed through in-house Perl scripts. In this step, clean data (clean reads) were obtained by removing reads containing adapter, reads containing ploy-N, and low-quality reads from raw data. The miRNA reads were trimmed and cleaned by removing the sequences smaller than 18 nt or longer than 30 nt. At the same time, Q20, Q30, GC-content, and sequence duplication levels of the clean data were calculated. All the downstream analyses were based on clean data with high quality.

Computational identification of ncRNAs

Identification and Analysis of miRNAs

By using Bowtie (v1.0.0, <http://bowtie-bio.sourceforge.net/index.shtml>) tools, the clean reads were performed sequence alignment respectively with Silva database (<http://www.arb-silva.de/>), GtRNADB database (<http://lowelab.ucsc.edu/GtRNADB/>), Rfam database (<http://rfam.xfam.org/>) and Repbase database (<http://www.girinst.org/repbase/>), then filter ribosomal RNA (rRNA), transfer RNA (tRNA), small nuclear RNA (snRNA), small nucleolar RNA (snoRNA) and other ncRNA and repeats. Finally, unannotated reads containing miRNA were obtained and mapped with the reference genome of *S.miltiarrhiza*

The mapped reads were further used to detect known miRNA by comparing with known miRNAs from miRBase (<http://www.mirbase.org/>) and novel miRNA by miRDeep2 with suitable parameters. Possible precursor sequences were obtained via position information of reads on the genome. Randfold tools soft was used for novel miRNA secondary structure prediction based on Bayesian model algorithm. The 5' end first base composition bias was also analyzed.

Identification and Analysis of lncRNA and CircRNA

The transcriptome was assembled using the StringTie (<https://ccb.jhu.edu/software/stringtie/index.shtml>) based on the reads mapped to the reference genome. Then the assembled transcripts were annotated using the gffcompare program. The unknown transcripts were used to screen for putative lncRNAs. Three computational approaches include CPC (<http://cpc.cbi.pku.edu.cn/>), CNCI (<http://www.ncbi.nlm.nih.gov/pubmed/23892401>), Pfam (<http://pfam.xfam.org/>), CPAT (<http://lilab.research.bcm.edu/cpat/>) were combined to sort non-protein coding RNA candidates from putative protein-coding RNAs in the unknown transcripts. Putative protein-coding RNAs were filtered out using a minimum length and exon number threshold. Transcripts with lengths more than 200 nt and have more than two exons were selected as lncRNA candidates and further screened using CPC/CNCI/Pfam/CPAT that have the power to distinguish the protein-coding genes from the noncoding genes. As well as the different types of lncRNAs include lincRNA, intronic lncRNA, anti-sense lncRNA, sense lncRNA, were selected using cuffcompare. CircRNA were identified by find_circ software.

Differential expression analysis

For the samples with biological replicates, differential expression analysis of two conditions/groups was performed using the DESeq R package (1.10.1, <http://www.bioconductor.org/packages/release/bioc/html/DESeq.html>). DESeq provides statistical routines for determining differential expression in digital miRNA/gene expression data using a model based on the negative binomial distribution. The resulting P values were adjusted using Benjamini and Hochberg's approach for controlling the false discovery rate. MiRNA/genes with an adjusted $p < 0.01$ found by DESeq were assigned as differentially expressed. For the samples without biological replicates,

and about miRNAs, before differential gene expression analysis, for each sequenced library, differential expression analysis of two samples was performed using the IDEG6 [43]. P-value was adjusted using q value [44]. $Q\text{value} < 0.01$ & $|\log_2(\text{Fold Change})| \geq 1$ was set as the threshold for significantly differential expression.

For lncRNA and circRNA, the edgeR program package adjusted the read counts through one scaling normalized factor. Differential expression analysis of two samples was performed using the EBseq (2010, <https://www.biostat.wisc.edu/~kendzior/EBSEQ/>) R package. The resulting FDR (false discovery rate) was adjusted using the PPDE (posterior probability of being DE). The $FDR < 0.05$ & $|\log_2(\text{Fold Change})| \geq 1$ was set as the threshold for significantly differential expression. P-value was used as the key index for screening differentially expressed circRNA.

Validation by quantitative real-time polymerase chain reaction (qRT-PCR)

Total RNA was obtained and purified from phloem and periderm of red and white roots using RNAprep Pure Plant Plus Kit (Polysaccharides&Polyphenolics-rich) (Tiangen Biotech, Beijing, China). No less than 1 µg total RNA was used to synthesize cDNA by HiScript[®] 1st Strand cDNA Synthesis Kit (+ gDNA wiper) (Vazyme Biotech Co., Ltd., Nanjing, China) according to the manufacturer's instructions. The microRNA (miRNAs) of the two groups mentioned above was isolated using miRcute Plant miRNA Isolation Kit (Tiangen Biotech, Beijing, China). High efficient poly(A) tail addition and first-strand cDNA synthesis are performed by TransScript[®] miRNA RT Enzyme Mix (containing tailing enzyme and RT enzyme) and 2×TS miRNA Reaction Mix (Transgen Biotech, Beijing, China). Quantitative reverse transcription-polymerase chain reaction (qRT-PCR) was performed using QuantStudio 6 Flex Real-Time PCR System (Applied Biosystems, Foster City, USA) and TransStart[®] Top Green qPCR SuperMix (Transgen Biotech, Beijing, China) with gene-specific primer pairs. The reaction procedure of qPCR was performed as follows: 94°C for 30 s, followed by 40 cycles of 94°C for 5 s and 60°C for 30 s, and with a final cycle of 95°C for 10 s and 60°C lasting for 5 s. The relative expression level of the gene was calculated with the reference gene *SmActin* by the comparative CT method ($2^{-\Delta\Delta C_t}$).

Target prediction and functional annotation

Target prediction

We used TargetFinder software to predict the target of known miRNAs and novel miRNAs. The nearest mRNAs at the range of 100kb upstream and downstream of lncRNAs were identified as *cis*-target genes by Perl. We investigated the complementary sequence between lncRNA and mRNA by the LncTar tool and then calculated and standardized the free energy of the matching sites. Those mRNA with standardized free energy threshold < -0.1 were considered as the target genes of lncRNAs. The mRNA with the highest matching degree with circRNA was selected as the parental genes, and the circRNA was classified according to their position on the genome. Miranda (v3.3a) software was used to predict the target miRNA of circRNA, and TargetFinder software was used to predict the target mRNA of target miRNA.

Target functional annotation

Gene function was annotated based on the following databases: Nr (NCBI non-redundant protein sequences); Nt (NCBI non-redundant nucleotide sequences) Pfam (Protein family) KOG/COG (Clusters of Orthologous Groups of proteins) Swiss-Prot (A manually annotated and reviewed protein sequence database) KO (KEGG Ortholog database) GO (Gene Ontology).

GO enrichment analysis

Gene Ontology (GO) enrichment analysis of the differentially expressed genes (DEGs) was implemented by the Goseq R packages based on Wallenius non-central hypergeometric distribution for targets of differential expression miRNAs. For targets of differential expression lncRNAs and circRNAs, Gene Ontology (GO) enrichment analysis of the differentially expressed genes (DEGs) was implemented by the topGO R packages.

KEGG pathway enrichment analysis

We used KOBAS software to test the statistical enrichment of differential expression targets of differential expression ncRNAs in KEGG pathways.

Results

Effective constituents contents in white root and red root

S. miltiorrhiza is an important medicinal plant, and the tanshinone and salvianolic acid in its roots have a remarkable curative effect on cardiovascular and cerebrovascular diseases [45, 46]. In the designed experiment, these red roots and white roots are located in different parts of the same root. The roots close to the base of stem and near the ground surface show white roots, while the roots far away from the ground surface show red roots. This experiment is specially designed to produce the difference between white roots and red roots, resulting in the difference of dark and dim light (This experiment will be specially reported). The HPLC results showed that the four main tanshinone components in the phloem and periderm tissues of red roots, including tanshinone IIA and tanshinone I, dihydrotanshinone I and cryptotanshinone, were significantly higher than those of white roots (Table 1). In order to explore the molecular mechanism of tanshinone synthesis, differentially expressed profiles of miRNAs, lncRNAs, and circRNAs of these two kinds of roots were analyzed in this study.

High-throughput ncRNAs sequencing data from *S. miltiorrhiza*

Six different organs (periderm and phloem in the red root, xylem in the red root, periderm and phloem in white root, leaf, stem, flower) were collected and sequenced by next-generation sequencing to acquire the spatial distribution pattern of ncRNAs.

One hundred thirty-two known miRNAs were identified, and 228 novel miRNAs were forecasted by applying miRDeep2 core algorithms with modifications for plant miRNAs (**Table S1**). The length of the novel mature miRNA was mainly in the range of 20 nt to 24 nt (Fig. 1a). We further analyzed the 5' end first base of novel miRNAs and showed the AU content of the precursor miRNA was higher than the GC content, and Uracil was the predominant base in the mature miRNA sequences (Fig. 1b). Based on the sequence similarity, these miRNAs could be grouped into 72 miRNA families (**Table S1**).

The cDNA library was sequenced through the Illumina HiSeq Xten platform, and 222.53 Gb clean data were obtained for bioinformatics analysis. Reads mapped to the reference genomes of *S. miltiorrhiza* were assembled and spliced by StringTie software to identify new transcripts from previously unannotated transcriptional regions. Afterward, transcripts with characteristics (length ≥ 200 bp, exon count ≥ 2 , and FPKM ≥ 0.1) were initially selected and further analyzed by four computational approaches include CPC, CNCI, Pfam, and CPAT. Eventually, 6,929 candidate lncRNAs without protein-coding potential were identified in all the samples (Fig. 2a). Most of the lncRNAs were lincRNAs (5503, 79.4%), followed by antisense-lncRNAs (617, 8.9%), sense lncRNAs (546, 7.9%), and intronic-lncRNAs (263, 3.8%) (Fig. 2b). The distribution of lncRNAs on chromosomes was shown in the circular diagram, which describes the distribution of different kinds of lncRNAs on chromosomes (Fig. 2c).

Furthermore, 6,239 circRNAs were detected via find_circ software, of which 5,633 (90.3%) were exonic circRNAs (ecircRNAs) generated from exons of a single protein-coding gene. The majority of circRNAs were between 200 to 1000nt and derived from exons, while the other circRNAs were over 3000 nt and from intergenic regions. We compared the locations of circRNAs and their source genes on the reference genome and then named circRNA ID after their source genes (Fig. 2d and 2e).

Comparison of mRNAs and lncRNAs features

In order to understand the differences in features between mRNAs and lncRNAs, the terms of length, exon number, ORF of mRNA, and lncRNA were compared. The results showed that the number of mRNAs with a length of 400nt was the highest, followed by those with a length $\geq$ of 3000nt. For lncRNAs, the number of lncRNAs with a length of 400nt was the largest, while the number of lncRNAs showed a decreasing trend with the increase of length (Fig. 3a). The number of exons of lncRNAs ranges from 2 to 8, and most lncRNAs have two exons. In contrast, most mRNAs contain one or two exons. The number of mRNA decreases gradually with the increase of exon (Fig. 3b). We predicted that the length of ORF of mRNA was mainly in the range of 100 ~ 500, while that of lncRNAs was mainly in the range of 50 ~ 200nt (Fig. 3c). By comparing the number of variable shear isomers of mRNA and screened lncRNAs, it was concluded that most mRNA and lncRNAs had one variable shear isomer (Fig. 3d).

Expression profiles of ncRNAs in different tissues

Tissue-specific expression patterns of ncRNAs would explain its spatiotemporal distribution and its role in gene regulation of different tissues. FPKM value was used as an indicator to measure the expression level of the transcript. A total of 321 miRNAs, 1969 lncRNA, and 270 circRNAs were found to be

differentially expressed with the $FDR < 0.05$ (False Discovery Rate, FDR) and $FC \geq 2$ (Fold Change) (**Table S2**). Hierarchical clustering analyses were conducted on the differentially expressed ncRNAs; the results of differentially expressed ncRNA are exhibited in **Fig. S1**. Given the significant difference of tanshinones content between red and white roots, we focused on the expression profiles of ncRNAs between the phloem and periderm of the red and white root. After that, 70 miRNAs, 48 lncRNAs, and 26 circRNAs were discovered to be differentially expressed. Among them, 40 miRNAs, 22 lncRNAs, and 14 circRNAs were down-regulated in the phloem and periderm of red root than that of white root, 30 miRNAs, 26 lncRNAs, and 12 circRNAs were up-regulated, respectively (Fig. 4; Table 2).

Functional annotation of target genes of different expression ncRNAs

In total, 1232 target genes of DE-miRNAs, 304 *cis*-target genes and 11 *trans*-target genes of DE-lncRNAs, and 12 source genes DE-circRNAs between the phloem and periderm of white and red root were annotated. To explore the potential functions of these targets, GO and KEGG Pathway analysis was performed. The function annotation was integrated based on multiple databases, such as NR, Swiss-Prot, COG, KOG, Pfam (**Table 3**; **Table S3**).

GO enrichment analysis revealed that 473 target genes of DE-miRNAs could be classified into 19 biological processes, 15 molecular functions, and 11 cellular component terms. For biological processes, the “lignin catabolic process,” “regulation of transcription, DNA-dependent,” and “auxin-activated signaling pathway” were the three most dominant GO categories. As for molecular functions, “hydroquinone: oxygen oxidoreductase activity” and “copper ion binding” were the two most significant GO terms. The three most dominant GO terms in cellular components were “apoplast,” “chloroplast stroma,” and “coated vesicle membrane” (Fig. 5a). Based on the KEGG pathway analysis, 162 DE-miRNAs targets gene are mainly involved in 58 metabolic pathways. Among them, the “Plant-pathogen interaction pathway,” “Biosynthesis of amino acids,” and “diterpenoid biosynthesis” pathway contain the most DEGs. These results indicate that miRNA plays an essential regulatory role in the growth, development, and metabolism in *S. miltiorrhiza* (Fig. 5b; **Table S4**). For the terpenoid biosynthesis pathways, seven targets matched the “Diterpenoid biosynthesis” pathway, such as the sly-miR164b-3p target Ent-kaurenoic acid oxidase (EVM0005756 and EVM0001334), aly-miR166g-5p target leucoanthocyanidin dioxygenase-like (Salvia_newGene_57657), unconservative_Lachesis_group0_16803 target kaurene synthase-like (Salvia_newGene_54337 and EVM0017060), unconservative_Lachesis_group5_24210 target ent-kaurene oxidase (EVM0019420 and EVM0013257). Besides, one target gene also matches the “Sesquiterpenoid and triterpenoid biosynthesis” pathway, aly-miR172c-5p target premnaspirodiene oxygenase (EVM0008863). These target genes are likely to play a significant role in the biosynthesis of terpenoids in *S. miltiorrhiza*.

In addition, the Nr, Swiss-Prot, and other database annotation results also provide other clues. The results indicated that lots of target genes encoded receptor kinases, disease-resistant proteins, proteins containing BTB/POZ domains, Pentatricopeptide repeat-containing proteins, F-box proteins, chaperone

dnaJ, GDSL esterase/lipase, β -glucanases, and so on. These genes may be relevant to the plant-pathogen interaction. Some target genes were also associated with epigenetic modifications, such as DNA (cytosine – 5) methyltransferase 1. Some miRNAs were predicted to regulate transcription factor families, such as NAC, WRKY, ERF, MYB, TCP, SPL, bHLH, MADS-Box, and AFR. Some targets were deemed to possibly involve secondary metabolite syntheses, such as the 2-C-methyl-D-erythritol 2, 4-cyclodiphosphate synthase (*SmMCS*), Cinnamate 4-Hydroxylase (*SmC4H*), laccase, and CYP450s (CYP76AK5, CYP72A219, CYP749A22).

Both the *cis* and *trans* targets of the DE-lncRNAs were analyzed. GO enrichment analysis displayed that 165 *cis*-regulated targets were significantly enriched in biological processes, cellular components, and molecular function. The GO terms associated with biological processes that were most abundant contained “protein retention in ER lumen,” “regulation of G2/M transition of the mitotic cell cycle,” and “lipid metabolic process.” The most relevant GO terms related to molecular functions were “chromatin binding,” “oxidoreductase activity,” and “monodehydroascorbate reductase (NADH) activity.” In the cellular component group, targets were enriched in terms of “plant-type vacuole,” “peroxisome,” and “anchored component of membrane” (Fig. 6a). The six *trans*-targets were correlated with 34 GO terms in three main classifications, contained “membrane,” “organelle,” “binding and cellular process,” and so on (Fig. 6c). KEGG analysis revealed the enrichment pathway of 111 *cis*-regulated targets mainly included “ribosome, carbon metabolism, plant hormone signal transduction, glycerophospholipid metabolism, protein processing in the endoplasmic reticulum,” and so on. The Ribosome pathway is the most significant enrichment pathway (Fig. 6b; **Table S4**). There were only two KEGG pathways for two *trans*-regulated targets of DE-lncRNAs, “Proteasome” and “Ubiquinone and another terpenoid-quinone biosynthesis” (Fig. 6d; **Table S4**). Annotation results for other databases also shown that target genes of DE-lncRNAs encoded some receptor protein kinase, pentatricopeptide repeat-containing protein, peroxidase, membrane protein, transporter protein, zinc finger protein, hormone synthesis, and signal transduction related protein.

The functions associated with the parent's gene of DE-circRNAs were dissected by GO and KEGG analysis. The significant functions of source genes were involved in “acetyl-CoA metabolic process,” “purine ribonucleoside monophosphate metabolic process,” “glycosyl compound biosynthetic process,” “ATPase coupled ion transmembrane transporter activity,” “cation-transporting ATPase activity,” “FMN binding,” “bounding membrane of organelle,” “myosin complex,” “actin cytoskeleton,” and so on (Fig. 7a). As a result of KEGG analysis, the target gene of DEcircRNA was significantly enriched in only one pathway. The unigene in the alpha-Linolenic acid metabolism pathway was analyzed between white phloem and red phloem samples (Fig. 7b; **Table S4**).

In addition, the 201 targets of 59 miRNAs combined with DE-circRNAs also were Analysis. Target genes were significantly enriched in biological process contained “defense response to the bacterium,” “cellular biogenic amine metabolic process,” “organic cyclic compound biosynthetic process,” “purine nucleotide biosynthetic process” “heterocycle biosynthetic process.” The enriched term related to molecular function included “sequence-specific DNA binding transcript,” “ligase activity,” “symporter activity,” “oxidoreductase

activity,” and “oxidoreductase” (Fig. 7c). The function of the DE targets gene is mainly involved in 15 metabolic pathways, such as “Plant-pathogen interaction, Endocytosis, Phenylpropanoid biosynthesis, Porphyrin, and chlorophyll metabolism, Galactose metabolism.” Thereinto, the “Sesquiterpenoid and triterpenoid biosynthesis” pathway had been paid close attention. Only premnaspirodiene oxygenase (EVM0021656) was enriched in the pathway targeted by unconservative_Lachesis_group2_8821 (Fig. 7d; Table S4).

Validation of differential expression ncRNAs

To verify the validity of RNA-Seq data, we picked six DE-mRNAs, five DE-lncRNAs at random and validated their expression by quantitative RT-PCR (Fig. 8). Experiment results indicated that qRT-PCR expression profiles of these ncRNAs were consistent with the results of RNA-Seq, as that confirmed the effectiveness of RNA-Seq data. Therefore, present results might serve as an efficient reference for researching the molecular mechanism of these ncRNAs in the future.

Discussion

A large number of ncRNAs were identified in Danshen

Most ncRNAs in plants, which have preliminarily obtained functional clues, regulate multiple aspects of development and response to environmental stress-specific in tissue cells and exist in complex regulatory networks [47–49]. Previous studies on ncRNA are still preliminary. In this study, six tissues, including stem, leaf, flower, red and white roots derived from the same plant, were collected for deep sequencing of non-coding RNA and preliminarily explored this regulatory mechanism in tanshinone biosynthesis at the transcriptional and post-transcriptional levels. A total of 360 miRNAs, 6,929 lncRNAs, and 6,239 circRNAs were identified in all the samples. Differential expression analysis showed that a total of 321 miRNAs, 1,969 lncRNAs, and 270 circRNAs were screened, indicating the tissue expression characteristics of noncoding RNAs. Based on the difference in tanshinone content, we focused on analyzing the differentially expressed ncRNAs between the red and white roots. A total of 70 miRNAs, 48 lncRNAs, and 26 circRNAs were found to be differentially expressed, which indicated their potential regulatory roles in biosynthesis of diterpenoid tanshinone. The number of differentially expressed ncRNAs were less than found in previous studies [50, 51], which might result from the similar period of development and same genetic background between red and white root.

NcRNAs expressed differentially in root tissues and their target genes

In order to explore the potential regulatory functions of these differentially expressed ncRNAs, we carried out GO classification and KEGG enrichment analysis of their target genes. We mined the target genes related to terpenoid synthesis. The results showed that seven miRNAs were enriched in the “diterpene biosynthesis pathway,” and their target genes were mainly annotated as terpenes synthase and oxidase

including *ent*-kaurenoic acid oxidase, kaurene synthase. A few miRNAs were enriched in “sesquiterpene and triterpene biosynthesis pathways”. Their target genes are also oxygenases. Combining with previous studies [10–12] and the test results, suggestion was proposed that the miRNAs generally involved in secondary metabolic pathways, and the study of regulatory function has important significance in the future.

Studies have shown that the expression patterns of mRNA-like noncoding RNAs (mlncRNAs) exhibited tissue specificity, suggesting that mlncRNAs may be involved in regulating the growth and development of *S.miltiorrhiza*. In addition, many mlncRNAs were responsive to the treatment of elicitors (yeast extract, Ag⁺ and MeJA) which lead to the production of bioactive compounds in *S. miltiorrhiza*, suggesting that lncRNAs might be involved in the production of active substances [52]. In this study, analysis of *cis*- and *trans*-target genes of differentially expressed lncRNAs revealed that most of the target genes were involved in the primary metabolism, such as the ribosome, carbon metabolism, plant hormone signal transduction. Only one target gene was involved in the biosynthesis of ubiquinone and other terpene quinones. Our results indicated that lncRNAs may not directly target some key enzyme genes for secondary metabolite synthesis, but indirectly influence the synthesis of active substances by responding to the external environment.

The source genes of the circRNA and the interacting miRNAs were also analyzed. CircRNA source genes may be involved in some primary metabolic processes, such as the Acetyl-CoA Metabolic process; target genes of miRNAs interacting with them participate in defense response to bacterium process and plant interaction pathway. At present, the functions of circRNAs in most plants are still unknown, especially in the regulation of secondary metabolites. Our results provide some references for future studies.

Combined with previous research results [12, 52], tanshinone synthesis may be generally regulated by non-coding RNA. In addition, based on the expression patterns of the three non-coding RNAs in different tissues, miRNA is more commonly and directly involved in the regulation of secondary metabolite synthesis than lncRNA and circRNA.

In addition, the integration of annotations results from Nr, Swiss-Prot, and other databases showed that some target genes of differentially expressed ncRNAs were involved in secondary metabolism. The enzyme 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase(*SmMCS*) was a key enzyme in the 2-C-methyl-D-erythritol-4-phosphate (MEP) pathway that was one of upstream terpene biosynthetic pathways [53, 54]. Cytochrome P450 enzymes (CYPs) play significant roles in generating highly functionalized terpenoids [55, 56]. Besides, differentially expressed ncRNAs also regulate transcription factors such as ERFs, MYB, bHLH, and WRKY. Studies have shown that several transcription factor families regulate the tanshinone synthesis mechanism [57, 58].

In addition, there were still most target genes of these DE ncRNAs encoded some receptor protein kinase, zinc finger protein, disease-resistant proteins, F-box Protein, hormone synthesis, and signal transduction related protein. This means that these genes may be involved in plant and pathogenic interactions, which

may be related to the abundance of microbial flora in the soil [59, 60]. The pericarp of the root approached the soil directly and closely is stimulated by microorganisms, which initiates the defense system of the root system [61]. It is worth mentioning that lots of genes are differentially expressed in the red and white roots, which indicates that the production of secondary metabolites, primarily diterpenes, may be related to the interaction of plant pathogens. Some studies suggest that secondary metabolites produced by plants are natural antimicrobial substances [62, 63]. The potential regulatory relationships between the production of secondary metabolites and the interactions among plants and pathogens would focus on future study. The study also provides new insight into the regulation mechanisms in the terpene synthesis process.

Conclusion

In this study, six tissues of the same *S.miltiorrhiza* plant, including stem, leaf, flower, red and white root, were collected for deep sequencing of non-coding RNA and preliminarily explored the regulatory mechanism of tanshinone biosynthesis at the transcriptional level. In view of the difference in tanshinone content between red and white roots, we conducted an in-depth analysis of the target genes of differentially expressed ncRNAs between them. The results showed that seven miRNAs targeted terpene synthase and oxidase in the diterpenoid synthesis pathway, suggesting that miRNAs may be directly involved in the regulation of terpenoid synthesis. The target genes of lncRNAs and circRNAs are mainly involved in the primary metabolic process, suggesting that they indirectly affect the synthesis of active substances by responding to the environment. In addition, we found that a large number of target genes of differentially expressed ncRNAs are involved in the plant-pathogen interaction pathway, suggesting that the production of secondary metabolites (mainly diterpenoids) may be related to the interaction of plant pathogens. In conclusion, the data provided in this study will provide a reference for the study on the regulation of tanshinone synthesis.

Abbreviations

lncRNAs: Long noncoding RNAs; CircRNAs: Circular RNAs; MiRNAs: MicroRNAs; NcRNAs: Noncoding RNAs; *S.miltiorrhiza*: *Salvia miltiorrhiza*; HPLC: High-performance liquid chromatography; KEGG: Kyoto Encyclopedia of Genes and Genomes; GO: Gene Ontology; MCS: 2-C-Methyl-d-erythritol 2,4-cyclodiphosphate Synthase; CYP450: Cytochrome P450; HMGR: HMG-CoA reductase; FPS: Farnesyl pyrophosphate synthase; ADS: Amorpha-4,11-diene synthase; AACT: Acetyl-CoA C-acetyltransferase; MVA pathway: The mevalonate pathway; MEP pathway: The 2-C-methyl-d-erythritol 4-phosphate pathway; SPL: Squamosa promoter-binding protein like; ARF: Auxin response factor; NAC: (NAM-ATAF1/2-CUC1/2) transcription factor; GRF: Growth-regulating factor; CeRNA: Competing endogenous RNAs; CPS: Copalyl diphosphate synthase; KSL: Kaurene synthase-like; ERF: Ethylene-responsive factor; MYB: (v-MYB avian myeloblastosis viral oncogene homolog) transcription factor; TCP: TEOSINTE BRANCHED1/CYCLOIDEA/PROLIFERATING CELL FACTORS; bHLH: Basic helix-loop-helix transcription factor; C4H: Cinnamate 4-Hydroxylase; MlncRNAs: MRNA-like noncoding RNAs

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The datasets supporting the conclusions of this article are included within the article and its additional files.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by funds from the National Natural Science Foundation of China (81872949), the Natural Science Foundation of Shandong province of China (ZR2019HM081), and the double first-class construction project of Shandong Agricultural University (Z. Song and X. Li).

Authors' contributions

ZS and XL conceived and designed the research, coordinated the study and wrote the manuscript. CL and CZ performed RNA-seq, RT-PCR. CZ performed downstream analysis of the data and generation of additional files. ZL and CL planted the samples for sequencing and conducted the extraction and determination of tanshinone content for all samples. All authors read and approved the final manuscript.

Acknowledgments

Not applicable.

References

1. Kim YH, Liang H, Liu X, Lee JS, Cho JY, Cheong JH, Kim H, Li M, Downey TJ, Dyer MD, et al. AMPK α modulation in cancer progression: multilayer integrative analysis of the whole transcriptome in Asian gastric cancer. *Cancer Res.* 2012;72(10):2512-21. <https://doi.org/10.1158/0008-5472.CAN-11-3870>.
2. Guttman M, Rinn JL. Modular regulatory principles of large non-coding RNAs. *Nature.* 2012;482(7385):339-46.<https://doi.org/10.1038/nature10887> .

3. Xue YC, Chen RS, Qu LH, Cao XF. Noncoding RNA: from dark matter to bright star. *Sci. China Life Sci.* 2020;63(4):463-8. <https://doi.org/10.1007/s11427-020-1676-5>.
4. Hou JN, Lu DD, Mason AS, Li BQ, Xiao ML, An SF, Fu DH. Non-coding RNAs and transposable elements in plant genomes: emergence, regulatory mechanisms and roles in plant development and stress responses. *Planta.* 2019;250(1):23-40. <https://doi.org/10.1007/s00425-019-03166-7>.
5. Wang JJ, Meng XW, Dobrovolskaya OB, Orlov YL, Chen M. Non-coding RNAs and their roles in stress response in plants. *Genomics Proteomics Bioinformatics.* 2017;15(5):301-12. <https://doi.org/10.1016/j.gpb.2017.01.007>.
6. Cech TR, Steitz JA. The noncoding RNA revolution-trashing old rules to forge new ones. *Cell.* 2014;157(1):77-94. <https://doi.org/10.1016/j.cell.2014.03.008>.
7. Waititu JK, Zhang C, Liu J, Wang H. Plant non-coding RNAs: origin, biogenesis, mode of action and their roles in abiotic stress. *Int J Mol Sci.* 2020;21(21):8401. <https://doi.org/10.3390/ijms21218401>.
8. Qin JP, Tang ZH, Ma XX, Meng YJ. Investigating the regulatory roles of the microRNAs and the Argonaute 1-enriched small RNAs in plant metabolism. *Gene.* 2017;628:180-189. <https://doi.org/10.1016/j.gene.2017.07.016>.
9. Pani A, Mahapatra RK, Behera N, Naik PK. Computational identification of sweet wormwood (*Artemisia annua*) microRNA and their mRNA targets. *Genomics Proteomics Bioinformatics.* 2011;9(6):200-10. [https://doi.org/10.1016/S1672-0229\(11\)60023-5](https://doi.org/10.1016/S1672-0229(11)60023-5).
10. Hao DC, Yang L, Xiao PG, Liu M. Identification of *Taxus* microRNAs and their targets with high-throughput sequencing and degradome analysis. *Physiol Plant.* 2012;146(4):388-403. <https://doi.org/10.1111/j.1399-3054.2012.01668.x>.
11. Singh N, Srivastava S, Shasany AK, Sharma A. Identification of miRNAs and their targets involved in the secondary metabolic pathways of *Mentha spp.* *Comput Biol Chem.* 2016;64:154-62. <https://doi.org/10.1016/j.compbiolchem.2016.06.004>.
12. Xu XB, Jiang QH, Ma XY, Ying QC, Shen B, Qian YS, Song HM, Wang HZ. Deep sequencing identifies tissue-specific microRNAs and their target genes involving in the biosynthesis of tanshinones in *Salvia miltiorrhiza*. *PLoS One.* 2014;9(11):e111679. <https://doi.org/10.1371/journal.pone.0111679>.
13. Zhang HH, Jin WB, Zhu XL, Liu L, He ZG, Yang SS, Liang ZS, Yan XJ, He YF, Liu Y. Identification and characterization of *Salvia miltiorrhizain* miRNAs in response to replanting disease. *PLoS One.* 2016;11(8):e0159905. <https://doi.org/10.1371/journal.pone.0159905>.
14. Schmitz SU, Grote P, Herrmann BG. Mechanisms of long noncoding RNA function in development and disease. *Cell Mol Life Sci.* 2016;73(13):2491-509. <https://doi.org/10.1007/s00018-016-2174-5>.
15. Wu T, Du Y. LncRNAs: from basic research to medical application. *Int J Biol Sci.* 2017;13(3):295-307. <https://doi.org/10.7150/ijbs.16968>.
16. Delás MJ, Gregory JH. LncRNAs in development and disease: from functions to mechanisms. *Open Biol.* 2017;7(7):170121. <https://doi.org/10.1098/rsob.170121>.
17. Zhang XP, Wang W, Zhu WD, Dong J, Cheng YY, Yin ZJ, Shen FF. Mechanisms and functions of long non-coding RNAs at multiple regulatory levels. *Int J Mol Sci.* 2019;20(22):5573.

<https://doi.org/10.3390/ijms20225573>.

18. Sun X, Zheng HX, Sui N. Regulation mechanism of long non-coding RNA in plant response to stress. *Biochem Biophys Res Commun*. 2018;503(2):402-7. <https://doi.org/10.1016/j.bbrc.2018.07.072>.
19. Wang H, Niu QW, Wu HW, Liu J, Ye J, Yu N, Chua NH. Analysis of non-coding transcriptome in rice and maize uncovers roles of conserved lncRNAs associated with agriculture traits. *Plant J*. 2015;84(2):404-16. <https://doi.org/10.1111/tpj.13018>.
20. Zhu QH, Stephen S, Taylor J, Helliwell CA, Wang MB. Long noncoding RNAs responsive to *Fusarium oxysporum* infection in *Arabidopsis thaliana*. *New Phytol*. 2014;201(2):574-84. <https://doi.org/10.1111/nph.12537>.
21. Huang L, Dong H, Zhou D, Li M, Liu Y, Zhang F, Feng YY, Yu DL, Lin S, Cao JS. Systematic identification of long non-coding RNAs during pollen development and fertilization in *Brassica rapa*. *Plant J*. 2018;96(1):203-22. <https://doi.org/10.1111/tpj.14016>.
22. Huanca-Mamani W, Arias-Carrasco R, Cárdenas-Ninasivincha S, Rojas-Herrera M, Sepúlveda-Hermosilla G, Caris-Maldonado JC, Bastías E, Maracaja-Coutinho V. Long Non-coding RNAs responsive to salt and boron stress in the hyper-arid Lluteño maize from Atacama Desert. *Genes*. 2018;9(3):170. <https://doi.org/10.3390/genes9030170>.
23. Qu SB, Yang XS, Li XL, Wang JL, Gao Y, Shang RZ, Sun W, Dou KF, Li HM. Circular RNA: A new star of noncoding RNAs. *Cancer Lett*. 2015;365(2):141-8. <https://doi.org/10.1016/j.canlet.2015.06.003>.
24. Panda AC, Grammatikakis I, Munk R, Gorospe M, Abdelmohsen K. Emerging roles and context of circular RNAs. *Wiley Interdiscip Rev RNA*. 2017;8(2):10.1002/wrna.1386. <https://doi.org/10.1002/wrna.1386>.
25. Tan JJ, Zhou ZJ, Niu YJ, Sun XY, Deng ZP. Identification and functional characterization of tomato circRNAs derived from genes involved in fruit pigment accumulation. *Sci Rep*. 2017;7(1):8594. <https://doi.org/10.1038/s41598-017-08806-0>.
26. Tong W, Yu J, Hou Y, Li FD, Zhou QY, Wei CL, Bennetzen JL. Circular RNA architecture and differentiation during leaf bud to young leaf development in tea (*Camellia sinensis*). *Planta*. 2018; 248(6):1417-29. <https://doi.org/10.1007/s00425-018-2983-x>.
27. Wang JX, Lin J, Wang H, Li XG, Yang QS, Li H, Chang YH. Identification and characterization of circRNAs in *Pyrus betulifolia* Bunge under drought stress. *PLoS One*. 2018;13(7):e0200692. <https://doi.org/10.1371/journal.pone.0200692>.
28. Pan T, Sun XQ, Liu YX, Li H, Deng GB, Lin HH, Wang SH. Heat stress alters genome-wide profiles of circular RNAs in *Arabidopsis*. *Plant Mol Biol*. 2018;96(3):217-29. <https://doi.org/10.1007/s11103-017-0684-7>.
29. Song ZQ, Wang JH, Li XF. Expression profiles of genes involved in tanshinone biosynthesis of two *Salvia miltiorrhiza* genotypes with different tanshinone contents. *Journal of Genetics*. 2016;95(2):433-9. <https://doi.org/10.1007/s12041-016-0621-6>.
30. Song ZQ, Guo LL, Liu T, Lin CC, Wang JH, Li XF. comparative rna-sequence transcriptome analysis of phenolic acid metabolism in *Salvia miltiorrhiza*, a traditional chinese medicine model plant. *Int J*

- Genomics. 2017;2017:9364594. <https://doi.org/10.1155/2017/9364594>.
31. Deng CP, Hao XL, Shi M, Fu R, Wang Y, Zhang Y, Zhou W, Feng Y, Makunga NP, Kai GY. Tanshinone production could be increased by the expression of *SmWRKY2* in *Salvia miltiorrhiza* hairy roots. Plant Sci. 2019;284:1-8. <https://doi.org/10.1016/j.plantsci.2019.03.007>.
 32. Liu L, Yang DF, Xing BC, Zhang HH, Liang ZS. *Salvia castanea* hairy roots are more tolerant to phosphate deficiency than *salvia miltiorrhiza* hairy roots based on the secondary metabolism and antioxidant defenses. Molecules. 2018;10:23(5):1132. <https://doi.org/10.3390/molecules23051132>.
 33. Feng YY, Guo LL, Jin H, Lin CC, Zhou CH, Fang XS, Wang JH, Song ZQ. Quantitative trait loci analysis of phenolic acids contents in *Salvia miltiorrhiza* based on genomic simple sequence repeat markers. Industrial crops and products. 2019;133:365-372. <https://doi.org/10.1016/j.indcrop.2019.01.063>.
 34. Zhang JH, Lv HZ, Liu WJ, Ji AJ, Zhang X, Song JY, Luo HM, Chen SL. bHLH transcription factor SmbHLH92 negatively regulates biosynthesis of phenolic acids and tanshinones in *Salvia miltiorrhiza*. Chinese Herbal Medicines. 2020;12(3):237-46. <https://doi.org/10.1016/j.chmed.2020.04.001>.
 35. Wang HT, Wang ZD, Chen HT, Yang Y. Study on Tanshen red pigment (i)—extraction and chemical components. Food Science. 2004; 25(5):86-91.
 36. Wang LL, Ma RF, Liu CY, Liu HX, Zhu RY, Guo SZ, Tang MK, Li Y, Niu JZ, Fu M, et al. *Salvia miltiorrhiza*: a potential red light to the development of cardiovascular diseases. Curr Pharm Des. 2017;23(7):1077-97. <https://doi.org/10.2174/1381612822666161010105242>.
 37. Xu ZC, Peters RJ, Weirather J, Luo HM, Liao BS, Zhang X, Zhu YJ, Ji Aj, Zhang B, Hu SN, et al. Full-length transcriptome sequences and splice variants obtained by a combination of sequencing platforms applied to different root tissues of *Salvia miltiorrhiza* and tanshinone biosynthesis. Plant J. 2015;82(6):951-61. <https://doi.org/10.1111/tpj.12865>.
 38. Wang QR, Li XY, Sun Y, Mu DX, Yan ZY, Chen X. Research on distribution of tanshinones in different parts of danshen's root by raman spectroscopy. The journal of light scattering. 2018;30(04):351-6. <https://doi.org/10.13883/j.issn1004-5929.201804009>.
 39. Song ZQ, Lin CC, Xing PY, Feng YY, Jin H, Zhou CH, Gu YQ, Wang JH, Li XF. A high-quality reference genome sequence of *Salvia miltiorrhiza* provides insights into tanshinone synthesis in its red rhizomes. The Plant Genome. 2020;13(1):e20041. <https://doi.org/10.1002/tpg2.20041>.
 40. Li SL, Zhu NL, Tang CP, Duan HN, Wang YW, Zhao GR, Liu J, Ye Y. Differential distribution of characteristic constituents in root, stem and leaf tissues of *Salvia miltiorrhiza* using MALDI mass spectrometry imaging. Fitoterapia. 2020;146:104679. <https://doi.org/10.1016/j.fitote.2020.104679>.
 41. Guo J, Zhou YJ, Hillwig ML, Shen Y, Yang L, Wang YJ, Zhang XN, Liu WJ, Peters RJ, Chen XY, et al. CYP76AH1 catalyzes turnover of miltiradiene in tanshinones biosynthesis and enables heterologous production of ferruginol in yeasts. Proc Natl Acad Sci U S A. 2013;110(29):12108-13. <https://doi.org/10.1073/pnas.1218061110>.
 42. Fang XS, Wang JH, Zhang S, Zhao QQ, Zheng ZK, Song ZQ. Simultaneous extraction of hydrosoluble phenolic acids and liposoluble tanshinones from *Salviae miltiorrhizae radix* by an optimized

- microwave-assisted extraction method. *Separation & Purification Technology*. 2012;86:149-156
<https://doi.org/10.1016/j.seppur.2011.10.039>.
43. Romualdi C, Bortoluzzi S, D'Alessi F, Danieli GA. IDEG6: a web tool for detection of differentially expressed genes in multiple tag sampling experiments. *Physiol Genomics*. 2003;12(2):159-162
<https://doi.org/10.1152/physiolgenomics>.
44. Storey JD. The positive false discovery rate: a bayesian interpretation and the q-value. *The Annals of Statistics*. 2003;31(6):2013-2035.
45. Li HQ, Li CL, Deng YX, Jiang XW, Lu SF. The pentatricopeptide repeat gene family in *Salvia miltiorrhiza*: genome-wide characterization and expression analysis. *Molecules*. 2018;23(6):1364.
<https://doi.org/10.3390/molecules23061364>.
46. Liu T, Guo LL, Pan YL, Zhao Q, Wang JH, Song ZQ. Construction of the first high-density genetic linkage map of *Salvia miltiorrhiza* using specific length amplified fragment (SLAF) sequencing. *Sci Rep*. 2016;6:24070. <https://doi.org/10.1038/srep24070>.
47. Ahmed W, Xia YS, Li RH, Bai GH, Siddique KHM, Guo PG. Non-coding RNAs: Functional roles in the regulation of stress response in *Brassica* crops. *Genomics*. 2020;112(2):1419-24.
<https://doi.org/10.1016/j.ygeno.2019.08.011>.
48. Xu S, Jiang YL, Cui WT, Jin QJ, Zhang YH, Bu D, Fu JY, Wang R, Zhou F, Shen WB. Hydrogen enhances adaptation of rice seedlings to cold stress via the reestablishment of redox homeostasis mediated by miRNA expression. *Plant Soil*. 2017;414:53-67. <https://doi.org/10.1007/s11104-016-3106-8>.
49. Zhang JS, Zhang H, Srivastava AK, Pan YJ, Bai JJ, Fang JJ, Shi HZ, Zhu JK. Knockdown of rice microRNA166 confers drought resistance by causing leaf rolling and altering stem xylem development. *Plant Physiol*. 2018;176(3):2082-94. <https://doi.org/10.1104/pp.17.01432>.
50. Zuo JH, Wang YX, Zhu BZ, Luo YB, Wang Q, Gao LP (2018) Analysis of the coding and non-coding RNA transcriptomes in response to *Bell Pepper* chilling. *Int J Mol Sci* 19(7):2001
<https://doi.org/10.3390/ijms19072001>
51. Ou LJ, Liu ZB, Zhang ZQ, Wei G, Zhang YP, Kang LY, Yang BZ, Yang S, Lv JH, Liu YH, et al. Noncoding and coding transcriptome analysis reveals the regulation roles of long noncoding RNAs in fruit development of hot pepper (*Capsicum annuum* L.). *Plant Growth Regul*. 2017;83:141-156.
<https://doi.org/10.1007/s10725-017-0290-3>.
52. Li DQ, Shao FJ, Lu SF. Identification and characterization of mRNA-like noncoding RNAs in *Salvia miltiorrhiza*. *Planta*. 2015; 241(5):1131-1143. <https://doi.org/10.1007/s00425-015-2246-z>.
53. Calisto BM, Perez-Gil J, Bergua M, Querol-Audi J, Fita I, Imperial S. Biosynthesis of isoprenoids in plants: structure of the 2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase from *Arabidopsis thaliana*. Comparison with the bacterial enzymes. *Protein Sci*. 2007;16(9):2082-8.
<https://doi.org/10.1110/ps.072972807>.
54. Ma XH, Ma Y, Tang JF, He YL, Liu YC, Ma XJ, Shen Y, Cui GH, Lin HX, Rong QX, et al. The biosynthetic pathways of tanshinones and phenolic acids in *Salvia miltiorrhiza*. *Molecules*. 2015;20(9):16235-54.

<https://doi.org/10.3390/molecules200916235>.

55. Guo J, Ma XH, Cai Y, Ma Y, Zhan ZL, Zhou YJJ, Liu WJ, Guan MX, Yang J, Cui GH, et al. Cytochrome P450 promiscuity leads to a bifurcating biosynthetic pathway for tanshinones. *New Phytol.* 2016;210(2):525-34. <https://doi.org/10.1111/nph.13790>.
56. Ma Y, Cui GH, Chen T, Ma XH, Wang RS, Jin BL, Yang J, Kang LP, Tang JF, Lai CJS, et al. Expansion within the CYP71D subfamily drives the heterocyclization of tanshinones synthesis in *Salvia miltiorrhiza*. *Nat Commun.* 2021;12(1):685. <https://doi.org/10.1038/s41467-021-20959-1>.
57. Huang Q, Sun MH, Yuan TP, Wang Y, Shi M, Lu SJ, Tang BP, Pan JX, Wang Y, Kai GY. The AP2/ERF transcription factor *SmERF1L1* regulates the biosynthesis of tanshinones and phenolic acids in *Salvia miltiorrhiza*. *Food Chem.* 2019;274:368-75. <https://doi.org/10.1016/j.foodchem.2018.08.119>.
58. Xing BC, Yang DF, Yu HZ, Zhang BX, Yan KJ, Zhang XM, Han RL, Liang ZS. Overexpression of SmbHLH10 enhances tanshinones biosynthesis in *Salvia miltiorrhiza* hairy roots. *Plant Sci.* 2018;276:229-38. <https://doi.org/10.1016/j.plantsci.2018.07.016>.
59. Curtis RH, Pankaj, Powers SJ, Napier J, Matthes MC. The *Arabidopsis* F-box/Kelch-repeat protein At2g44130 is upregulated in giant cells and promotes nematode susceptibility. *Mol Plant Microbe Interact.* 2013;26(1):36-43. <https://doi.org/10.1094/MPMI-05-12-0135-FI>.
60. Cao XB, Zhai XQ, Xu EK, Zhao ZL, Fan GQ. Genome-wide identification of candidate genes related to disease resistance and high biomass in tetraploid *Paulownia*. *Acta Physiol Plant.* 2020;42:171. <https://doi.org/10.1007/s11738-020-03160-7>.
61. Ahuja I, Kissen R, Bones AM. Phytoalexins in defense against pathogens. *Trends Plant Sci.* 2012;17(2):73-90. <https://doi.org/10.1016/j.tplants.2011.11.002>.
62. Zhan CS, Lei L, Liu ZX, Zhou S, Yang CK, Zhu XT, Guo H, Zhang F, Peng M, Zhang M, et al. Selection of a subspecies-specific diterpene gene cluster implicated in rice disease resistance. *Nat Plants.* 2020;6(12):1447-54. <https://doi.org/10.1038/s41477-020-00816-7>.
63. Lu X, Zhang J, Brown B, Li RQ, Rodríguez-Romero J, Berasategui A, Liu B, Xu MM, Luo DP, Pan ZQ, et al. Inferring roles in defense from metabolic allocation of rice diterpenoids. *Plant Cell.* 2018;30(5):1119-31. <https://doi.org/10.1105/tpc.18.00205>.

Tables

Table 1
The effective constituents content in white root and red root in same plantlet

Content (mg/g)	Phloem in red root	Xylem in red root	Phloem in white root	Xylem in white root	Wavelength Detected
Tanshinone IIA	4.34	0.001	0	0	269
Dihydrotanshinone	0.5	0.001	0	0	269
Cryptotanshinone	1.51	0.002	0.001	0	269
Tanshinone IA	0.968	0.001	0	0	269

Table 2
Differentially expressed ncRNAs between two root tissues

ncRNAs	DEG_Number	up_regulated	down_regulated
miRNAs	70	30	40
lncRNAs	48	26	22
circRNAs	26	12	14

Table 3
Statistical table of the number of differentially expressed ncRNA target genes and source genes

Type	anannotated	COG	GO	KEGG	KOG	Swissprot	eggNOG	nr
Targets of DE-miRNAs	1232	375	473	315	677	894	1116	143
<i>cis</i> -targets of DE-LncRNAs	304	125	165	111	184	210	0	0
<i>tran</i> -targets of DE-LncRNAs	11	2	6	3	9	5	0	0
source genes of DE-circRNAs	12	5	9	3	5	9	12	0

Figures

Figure 1

Characterization of identified miRNAs of *S.miltiorrhiza*. (a) The distribution of reads along with mature miRNA length. (b) The presence of the first nucleotide of miRNAs along with the mature miRNA length.

Figure 2

(a) Venn diagrams of coding potential analysis results from four software. (b) The lncRNAs classification. (c) The distribution on the chromosomes, sense_lncRNA ring (green), intergenic lncRNA (red), and intronic lncRNA (blue). (d) CircRNA Length Distribution of All_circRNA. (e) The number of junction reads across chromosomes LG1 to LG8.

(red), anti-sense lncRNA (gray), intron lncRNA (blue), and the outermost layer is the chromosome. (d) CircRNA length distribution. (e) The distribution of circRNAs on the chromosomes.

Figure 3

Comparative Analysis of mRNA and lncRNA. (a) The length distribution of mRNA and lncRNA. (b) The number of exon corresponding to mRNA and lncRNA. (c) The ORF length corresponding to mRNA and lncRNA. (d) Comparison of lncRNA and mRNA variable-shear isoforms.

The gene ontology (GO, a) and the Kyoto encyclopedia of genes and genomes (KEGG, b) analysis of DE miRNAs targets in two root tissues of *S. miltiorrhiza*

Figure 6

The gene ontology (GO, a) and the Kyoto encyclopedia of genes and genomes (KEGG, b) analysis of DE lncRNAs cis-targets, the GO (c) and KEGG (d) analysis of DE lncRNAs trans-targets in two root tissues of *S. miltiorrhiza*

Figure 7

The gene ontology (GO, a) and the Kyoto encyclopedia of genes and genomes (KEGG, b) analysis of DE circRNAs targets, the GO (c) and KEGG (d) analysis of targets of miRNAs combined with DE circRNAs in two root tissues of *S. miltiorrhiza*

Figure 8

Expression of lncRNAs (Right) and miRNAs (Left) in phloem and periderm of red roots (RZP, black) and white roots (WZP, white) of a *S. miltiorrhiza* plants. Expression levels were quantified by qRT-PCR. The level of transcripts in RZP was arbitrarily set to 1, and the level in WZP was given relative to this. Smactin was used as the internal control gene, and three biological replicates were used.

Supplementary Files

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

- [FigureS1.tif](#)
- [SupplementaryTableS1.xls](#)
- [SupplementaryTableS2.xls](#)
- [SupplementaryTableS3.xls](#)

- [SupplementaryTableS4.xls](#)
- [supplymentaryfiletitlepage.docx](#)