

Multi-omics Analysis Reveals Whether and how Naphthylacetic Acid Affects the Quality of *Rehmannia Glutinosa*

Jianjun LI

Henan Normal University

Luying SHAO

Henan Normal University

Jialin ZHU

Henan Normal University

Jingxiao MA

Henan Normal University

Yanqing ZHOU (✉ yqzhou@htu.edu.cn)

Henan Normal University

Huimin LI

Henan Normal University

Research article

Keywords: *Rehmannia glutinosa*, Naphthylacetic acid, Omics, quality improvement, differentially expressed transcripts, differentially expressed metabolites

Posted Date: November 19th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-109124/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background: *Rehmannia glutinosa* (*R. glutinosa*) is an important medicinal plant. The tuberous root of *R. glutinosa* is often used as herbal medicine. Naphthylacetic acid (NAA) as expansin can improve its yield, but knowledge about gene regulation and metabolome in its root is limited.

Results: Full-length transcriptome, next generation transcriptome(NGS), small RNA and degradome sequencing and metabolomics were used to elucidate whether and how NAA affected its quality. 30 differential expression metabolites (DEMs) (11 upregulated, 19 downregulated) were identified, but catalpol and Rehmannioside D as quality standards were unchanged in its tuberous roots under control and NAA conditions (CKs and NTs); Their NGS identified 1,113 differentially expressed transcripts (DETs) (596 upregulated, 517 downregulated) verified by RT-qPCR; Small RNA sequencing identified 78 miRNAs (11 known, 67 novel), of which 3 were differentially expressed miRNAs (1 upregulated, 2 downregulated). Among them, 274 differentially expressed miRNAs target transcripts (DEMTs) were predicted found and then validated by degradome sequencing; DETs and DEMTs were mainly related to metabolism. 4 miRNA-mRNA interaction pairs that regulates 4 metabolites (2 negatively correlated, 2 positively correlated) were identified; DETs, DEMs, differentially expressed miRNAs and DEMTs involved in phenylpropanoid biosynthesis regulated metabolites.

Conclusions: The identification of DETs, DEMs, differentially expressed miRNAs and DEMTs could help to elucidate the regulatory networks and molecular mechanisms important for NAA-improving root quality of *R. glutinosa*.

Background

Rehmannia glutinosa (*R. glutinosa*) is a medicinally important species belonging to *Rehmannia* (*Scrophulariaceae*), cultivated all over the home and abroad. In China, it is distributed in Liaoning, Hebei, Henan and other provinces. Its root is rich in bioactive compounds[1], and clinically used as Chinese herbal medicine to treat fever, nervous conditions, diabetes, and hypertension; to strengthen liver function; to enhance the hematopoietic function and immune defense; and serves as an ingredient in tonics[2]. Its yield and quality has been improved by genetic breeding, high yield cultivation techniques, expansins and so on[3–5]. The expansins used in the production of *R. glutinosa* are divided into two types such as water soluble fertilizer and plant growth regulator. The latter includes Liganwei containing Naphthylacetic acid (NAA), N-2-chlorine-4- pyridylphenyl-N-phenylurea (CPPU) and so on, and could increase yield of *R. glutinosa*, of which CPPU increased its catalpol and verbascoside content[3, 6, 7]. NAA is an environment-friendly and easily decomposed plant growth regulator with high efficiency, low toxicity and low residue[8, 9]. Due to its low price and stable properties, now, it is widely used in production practice instead of IAA[10]. It can improve plant photosynthesis, expand and induce adventitious roots, promote cell division, accelerate growth and development, increase yield, enhance plant resistance to drought, cold, salt and heavy metal stress[3, 9–12]. NAA as one expansin could improve the yield of *R. glutinosa* by inhibiting the growth of its shoots and promoting the transfer of nutrients to roots[3], so increasing pharmaceutical farmers and pharmaceutical enterprises or institutes try to improve the yield of *R. glutinosa* by NAA as one expansin. However, whether and how NAA affects its quality remain unknown.

RNA sequencing is an effective and feasible method for transcript quantification, differential gene expression and gene expression profile analysis, transcript reconstruction and the discovery of novel transcripts of plants[13]. Small RNAs are a class of non-coding endogenous RNAs ranging from 20 to 24 nucleotides (nt) that play essential roles in plant growth and development, signal transduction, response to biotic and abiotic stresses and other biological processes[14]. They are divided into nat-siRNA, phasiRNA, miRNAs and so on[15]. Plant miRNA is a type of endogenous non-coding small RNA, about 20-25nt, which regulates plant growth and development, reproduction and stress response by degrading its target gene mRNA and inhibiting its translation[16]. Many miRNAs have been discovered in many medicinal plants and play important regulatory roles mainly by transcript splicing and translation inhibition, many of which are associated with biosyntheses of secondary metabolites in medicinal plants and their root development[17]. NGS of *R. glutinosa* was studied for the identifications of its secondary metabolites-synthesizing genes, lncRNA and miRNAs responding to replanting disease[18–20]. However, the read length, its major limitation would cause the loss of some information from the original full-length transcripts. Now, full-length transcriptome sequencing can considerably increase read lengths compared to NGS. Therefore, it can capture the full-length transcripts, more easily and accurately annotate gene, identify isoform and so on. So far, it has been successively used for whole transcriptome profiling in many different organisms[21]. Nevertheless, to our knowledge, full-length transcriptome sequencing of *R. glutinosa* is not reported until now. The metabolites are usually known as the end products of gene expression in plants. Untargeted metabolomics has been widely used to identify and assess the expression levels of multiple metabolites in plants, including *R. glutinosa*[1, 2], and can quantitatively detect many metabolites and their biochemical states in specific tissues, and study the metabolic phenotype of organisms. Recently, the combination of transcriptome sequencing with metabolomics has increased the power of bioinformatics analysis, and has been used to explore the important genes participating in the regulatory pathways of specific metabolites and to elucidated the molecular mechanisms of melatonin's protective effects in plants[22, 23]; Integration of transcriptome, degradome and smalls sequencing have provides a useful platform for some molecular mechanisms in plants[24]. However, their integration analysis is not used in *R. glutinosa* until now. Therefore, there is a need for comprehensive identification of the DETs, DEMs, differentially expressed miRNAs and their targets regulated by NAA in *R. glutinosa*.

In our current study, based on our NT and CK yield measurement of *R. glutinosa* in advance, we integrated transcriptomic, small RNAs, degradome and metabolomics analyses to compare the gene expression, miRNA and metabolomic profiles between NTs and CKs. The aims of this study were (1) to provide abundant transcriptome resources for a transcript isoform catalog, (2) to detect NAA-regulated metabolites and (3) to systematically identify potential NAA-responsive DETs, DEMs, miRNAs and their targets for providing valuable information regarding NAA-improving quality of *R. glutinosa* Libosh and its the mechanisms.

Results

Full-length transcriptome sequencing data and full length sequence statistics

26.61 GB clean reads from 1 kb to 6 kb were generated via Full-length transcriptome sequencing of Sample F01 on PacBio Sequel platform. Using “full passes ≥ 3 and sequence accuracy > 0.9 ” as criteria, 357,961 circular consensus reads were extracted from these clean reads, which contain 775,062,045 bases and share a mean read length of 2,165 bases. When mean number of passes is 43, these circular consensus reads were divided into full length sequences with 5' primers and 3' primers and poly(A)_n tails, and non-full length sequences without any primers and tails. After those primers and poly(A)_n tails were removed, insertion sequences were obtained for the construction of database. When the database was constructed, synthesis directions of their strands were ascertained based on the differences between their two ends. 307,475 full length non-chimeric clean reads with a percentage of 85.9% were obtained; Among these full length non-chimeric reads, 99,929 consensus isoforms with a read length of 1,969 bases were identified. They were polished for the production of 96,548 polished high-quality consensus isoforms with a percentage of 96.62% and 3,001 polished low-quality consensus isoforms. 61,262 non-redundant high quality consensus isoforms were generated after redundant high quality consensus isoforms were removed using CD-HIT software (identity > 0.99) [25]; In addition, 46,412 coding sequences, 4,241 lncRNAs, 3,070 alternative splicing, 34,013 SSRs and 5,562 transcription factors were predicted. 56,633 full length transcripts were annotated to 8 databases such as COG, GO, KEGG, KOG, Pfam, Swissport, eggNOG and Nr; After removing redundancy, take OrthoDB as the database, transcriptome integrity were assessed by BUSCO [26] as follows: Complete coverage was 1129, of which single copy and duplicated (multi copies) was 532 and 597, respectively. Incomplete coverage was 70. Missing isoforms (no alignment result) was 241. Single copy gene clusters of related species and their genes number was 1440.

The identification and KEGG classification of DETs

42.91 Gb clean data from six samples were presented in Table S1, and aligned to nonredundant transcripts using STAR software and 3rd G nonredundant transcripts as references (Table S2). Based on the positional information of mapped clean reads in 3rd G nonredundant transcripts, the transcript expressions were quantified by RSEM software [27]. 1,113 DET sets between CK and NT were obtained by DESeq software [28] using Fold change ≥ 2 and False discovery rate < 0.01 as criteria, including 596 up-regulated DETs and 517 down-regulated DETs in NT compared to CK (Table S3). The first 25 up-regulated DETs and 25 down-regulated DETs for heatmap were grouped (Fig. 1). To validate our DETs (Table S3), we selected 10 up-regulated DETs and 11 down-regulated DETs for RT-qPCR analyses. Among 10 up-regulated DETs, 9 (90%) were confirmed by RT-qPCR. Among 11 down-regulated DETs, 10 (90.91%) were confirmed by RT-qPCR (Table 1). These results revealed that our sequenced data were highly reliable.

Table 1
qRT-PCR analyses of up- or down-regulated transcripts

ID	Size (bp)	Primer pair(PP)	TM (°C)	Amplicon (bp)	Relative level ± SD	Regulation
10245	2223	PP1	64	188	0.84 ± 0.0759	down
38731	5276	PP2	62	192	3.82 ± 0.5684	up
24916	746	PP3	63	175	66.15 ± 13.2465	up
42226	2886	PP4	64	95	74.62 ± 7.7219	up
37318	5410	PP5	62	138	2.89 ± 0.1497	up
10269	2232	PP6	62	232	2.77 ± 0.2991	up
16137	1703	PP7	63	160	3.94 ± 0.1268	up
37324	1878	PP8	62	96	1.93 ± 0.3363	up
15391	1750	PP9	62	196	2.60 ± 0.1219	up
31933	3060	PP10	61	274	2.11 ± 0.2123	up
31213	2038	PP11	63	102	1.65 ± 0.0544	up
16508	1647	PP12	63	118	0.10 ± 0.0094	down
76643	2129	PP13	61	235	0.91 ± 0.1161	down
60810	3265	PP14	61	191	0.98 ± 0.0356	down
44935	1243	PP15	63	175	0.13 ± 0.0094	down
37583	1269	PP16	63	138	0.73 ± 0.1297	down
84676	2215	PP17	63	144	0.16 ± 0.0047	down
52404	1630	PP18	62	164	0.16 ± 0.0082	down
70202	1634	PP19	60	148	0.41 ± 0.0525	down
34870	2497	PP20	63	96	0.10 ± 0.0125	down
14415	1116	PP21	64	130	0.94 ± 0.0499	down
GAPDH		PP22	65	143	1.00 ± 0.0000	CK

ID: the number in the given transcript, for example, 10245 stands for F01_transcript_10245. Relative level: the expression level in NT, while that is 1 in CK. SD: Standard deviation. Up: Relative level > 1. Down: Relative level < 1. PP1-F: TGCTGCTCTCGTTCTATGTC, R: GTATCCTCCCTGGTGTTC; PP2-F: CAAAGGAGAGGAAAAT GCTG, R: TCACAATCACATACCAACCAA; PP3-F: GAACAGCCTTAGCTCCAATC, R: CTCGTGAACATTACC CATCTC; PP4-F: TGACCTCTTTTCCCTATCT, R: GAACCCCGTGAAGTCTACC; PP5-F: CAAAGACACCC CACAAAAC, R: GCAAAAAGCAAGCGAAACT; PP6-F: CCCATTTCTTTTAGCACAC, R: AGGGTTGTTTTCT TGATTTCC; PP7-F: GTATGTGCTTGGGGAAAGAG, R: TGAGAGTGAAGACGAATAGGAA; PP8-F: TAAACC CGAGCAACAAGAG, R: AATGGAGCCGAGAAGAAG; PP9-F: CATCTTTTGGAGCGTCTT, R: CCACCAT CATCTACCATTCC; PP10-F: GAGAAAGAACAAAAGGGCAA, R: ATTCCGACATCAAACCCA; PP11-F: AGA AGCAAAAAGGATGAGTGG, R: GGGGATAGAGGGAGTGGT; PP12-F: GAGCTAATCAACGAGCCAA, R: CCG ACAACACGACATACG; PP13-F: AAAGGAGCGGAAAATCAAAG, R: TGGGCATCACAAGAAGTCT; PP14-F: CCCTCCTCCCTCTATCTTC, R: GCCTTTGCTACTTCTGGTG; PP15-F: TTGCTATCCTTGGCATCTCT, R: CAC C ACCTTTCACAAACTCA; PP16-F: GATTGTCCTTTATGTGGTGTG, R: ACCTTCTCTTTCTCCTTGTG; PP17-F: TCCAACACTACAAAAGGATGAG, R: AACACAAAACCGACAAAGAGA; PP18-F: ATTGGTCTGGATTGG AG, R: TCTTGTCTTCCCAGTCA; PP19-F: TTCTTGATGGCAGTGGGAGG, R: AAACCTCGGTCCGGCTTT TCT; PP20-F: TACCCTTTTCTCTTTCACACAC, R: CCCTTCTGTAACCTTGGCTT; PP21-F: AGCAATCCTTTA CGACACT, R: TCTCTCACCTCTCAGCAA; PP22-F: TGCCTTGCTCCTTTGGCGAAGG, R: AAGCCGCTC TTCCACCTCTCCA

1,082 of 1,113 DETs were annotated to the following databases such as COG (534DET), GO (859), KEGG (454), KOG (693), Pfam (926), Swiss Prot (867), eggNOG (1,038) and NR (1,076).

The statistical results of KEGG classification of DETs were shown in Fig. 2A. 271 of 454 DETs were annotated to 50 KEGG pathways 387 times, which were classified into five groups such as cellular process with 20 annotated DETs (7.38%), environmental information process (plant hormone transduction) (23.8.49%), genetic information processing (57.21.07%), metabolism (156,57.89%) and organismal systems (14,5017%).

The identification of differential miRNAs and their targets and KEGG classification

110.67 Mb clean reads were generated by small RNA sequencing of 6 samples (Table S4), and mapped to several databases such as Silva, GtRNAdb database, Rfam database and Rfam database by Bowtie software to filter out sequence repeats and ncRNAs, including rRNAs, tRNAs, snRNAs, snoRNAs and unannotated reads with miRNAs (Table S5). The rest were aligned to Reference database for mapped reads (Table S6). 78 miRNAs, including 11 known miRNAs and 67 novel miRNAs, were identified in NTs and CKs (Table S7). These known miRNAs were 20nt to 24nt, while these

novel miRNAs were 18nt or 20nt to 24nt. Based on their expression levels, 3 differential novel miRNAs were screened out, including upregulated novel_miR_57 and downregulated novel_miR_65 and novel_miR_23.

According to the gene sequence information of known miRNAs and novel miRNAs-corresponding species, their DENTs (5077) were predicted by Target Finder software (Table S8), of which 4861 were annotated to the following databases including NR, Swiss-Prot, GO, COG, KEGG, KOG, eggNOG and Pfam by BLAST software. For KEGG annotation (Fig. 2B), 42 DENTs were annotated to 19KEGG pathways, including cellular process with 8 annotated DENTs (18.96%), environmental information process (other than plant hormone transduction) (3, 7.14%), genetic information processing (21, 49.94%), metabolism (8, 19.04%) and organismal systems(2, 4.76%). KEGG pathways enrichment of these DENTs was shown as in Table S9.

miRNA target prediction and validation

To further predict and identify miRNA targets based on miRNA sequences, we performed degradome sequencing, generated 20.67M clean tags. We detected 543 degraded DENTs from these clean tags, which contain 557 cleaved sites. Among them, 247 targets of novel miR23, novel miR57 and novel miR65 were predicted using Target Finder software. Degradome sequencing just identified 12 cleavage DENTs of novel miR23 and 13 of novel miR65 (Table 2). See Table S10 for more details, indicating that 25 of 247 miRNA DENTs were cleaved.

Table 2
List of significantly differential miRNAs, their target genes, regulation and degradation

miRNA ID	miRNA regulated	DEMT	DEMT regulated	Degradation
novel_miR_23	down	F01_transcript_9884,F01_transcript_49425,F01_transcript_49310,F01_transcript_35991 F01_transcript_69027,F01_transcript_52083,F01_transcript_87919,F01_transcript_24294 F01_transcript_77643,F01_transcript_99282,F01_transcript_79157,F01_transcript_57888 F01_transcript_31623,F01_transcript_16601,F01_transcript_822,F01_transcript_75609 F01_transcript_15796,F01_transcript_98227,F01_transcript_53448,F01_transcript_31904 F01_transcript_83403,F01_transcript_98477,F01_transcript_99770,F01_transcript_65741 F01_transcript_71724,F01_transcript_3880,F01_transcript_75904,F01_transcript_90795 F01_transcript_13512,F01_transcript_75138,F01_transcript_96440,F01_transcript_79015 F01_transcript_33176,F01_transcript_99565,F01_transcript_8768,F01_transcript_46928 F01_transcript_32314,F01_transcript_83272,F01_transcript_23341,F01_transcript_84687 F01_transcript_56501,F01_transcript_56997,F01_transcript_57214,F01_transcript_68555 F01_transcript_10641,F01_transcript_57489,F01_transcript_32966,F01_transcript_61126 F01_transcript_76688,F01_transcript_8028,F01_transcript_23078,F01_transcript_14088 F01_transcript_9835,F01_transcript_67250,F01_transcript_50507,F01_transcript_56643 F01_transcript_61,F01_transcript_86562,F01_transcript_55954,F01_transcript_36912 F01_transcript_50525,F01_transcript_72874,F01_transcript_63935,F01_transcript_77339 F01_transcript_6789,F01_transcript_26745,F01_transcript_66299,F01_transcript_7339 F01_transcript_41142,F01_transcript_5378,F01_transcript_1407,F01_transcript_45537 F01_transcript_81751,F01_transcript_61635,F01_transcript_81810,F01_transcript_88940 F01_transcript_8287,F01_transcript_9081,F01_transcript_16747,F01_transcript_10497 F01_transcript_32662,F01_transcript_26254,F01_transcript_17541,F01_transcript_9092 F01_transcript_16533,F01_transcript_71946,F01_transcript_82925,F01_transcript_1272 F01_transcript_48614,F01_transcript_8795,F01_transcript_57499,F01_transcript_69639 F01_transcript_99809,F01_transcript_87819,F01_transcript_45408,F01_transcript_58018 F01_transcript_25596,F01_transcript_88688,F01_transcript_96496,F01_transcript_64707 F01_transcript_14464,F01_transcript_7166,F01_transcript_46191,F01_transcript_8502	Normal	No
novel_miR_23	down	F01_transcript_11039,F01_transcript_25211	up	No
novel_miR_23	down	F01_transcript_12427,F01_transcript_97278,F01_transcript_56882,F01_transcript_54959 F01_transcript_83157,F01_transcript_22329,F01_transcript_98936,F01_transcript_83968 F01_transcript_21523,F01_transcript_33284,F01_transcript_20107,F01_transcript_36246 F01_transcript_39070,F01_transcript_8091,F01_transcript_8952	Normal	Yes
novel_miR_23	down	F01_transcript_37839	down	No
novel_miR_57	up	F01_transcript_49385,F01_transcript_34877	Normal	No

miRNA ID: the ID of miRNA. miRNA regulated: the regulate of miRNA. DEMT: the differentially expressed miRNAs target genes. DEMT regulated: the regulate of differentially expressed miRNAs target genes. Degradation: the status of cleaved sites.

miRNA ID	miRNA regulated	DEMT	DEMT regulated	Degradation
novel_miR_65	down	F01_transcript_64707,F01_transcript_96496,F01_transcript_88688,F01_transcript_8502 F01_transcript_46191,F01_transcript_14464,F01_transcript_7166,F01_transcript_69639 F01_transcript_57499,F01_transcript_8795,F01_transcript_58018,F01_transcript_25596 F01_transcript_45408,F01_transcript_99809,F01_transcript_87819,F01_transcript_17541 F01_transcript_9092,F01_transcript_26254,F01_transcript_1272,F01_transcript_48614 F01_transcript_82925,F01_transcript_16533,F01_transcript_71946,F01_transcript_8287 F01_transcript_81810,F01_transcript_88940,F01_transcript_61635,F01_transcript_81751 F01_transcript_32662,F01_transcript_16747,F01_transcript_10497,F01_transcript_9081 F01_transcript_26745,F01_transcript_66299,F01_transcript_6789,F01_transcript_1407 F01_transcript_45537,F01_transcript_5378,F01_transcript_7339,F01_transcript_41142 F01_transcript_77339,F01_transcript_63935,F01_transcript_50525,F01_transcript_72874 F01_transcript_50507,F01_transcript_67250,F01_transcript_9835,F01_transcript_36912 F01_transcript_55954,F01_transcript_56643,F01_transcript_61,F01_transcript_86562 F01_transcript_76688,F01_transcript_10641,F01_transcript_32966,F01_transcript_57489 F01_transcript_61126,F01_transcript_14088,F01_transcript_23078,F01_transcript_8028 F01_transcript_56501,F01_transcript_84687,F01_transcript_68555,F01_transcript_57214 F01_transcript_56997,F01_transcript_46928,F01_transcript_32314,F01_transcript_23341 F01_transcript_83272,F01_transcript_99565,F01_transcript_79015,F01_transcript_33176 F01_transcript_96440,F01_transcript_8768,F01_transcript_75904,F01_transcript_3880 F01_transcript_71724,F01_transcript_75138,F01_transcript_90795,F01_transcript_13512 F01_transcript_98477,F01_transcript_83403,F01_transcript_65741,F01_transcript_99770 F01_transcript_53448,F01_transcript_31904,F01_transcript_98227,F01_transcript_31623 F01_transcript_16601,F01_transcript_79157,F01_transcript_57888,F01_transcript_99282 F01_transcript_77643,F01_transcript_15796,F01_transcript_822,F01_transcript_75609 F01_transcript_52083,F01_transcript_87919,F01_transcript_35991,F01_transcript_69027 F01_transcript_9884,F01_transcript_49425,F01_transcript_49310,F01_transcript_24294	Normal	No
novel_miR_65	down	F01_transcript_8952,F01_transcript_8091,F01_transcript_36246,F01_transcript_39070 F01_transcript_33284,F01_transcript_20107,F01_transcript_21523,F01_transcript_83968 F01_transcript_98936,F01_transcript_22329,F01_transcript_83157,F01_transcript_54959 F01_transcript_56882,F01_transcript_12427,F01_transcript_97278	Normal	Yes
novel_miR_65	down	F01_transcript_37839	down	No
novel_miR_65	down	F01_transcript_25211,F01_transcript_11039	up	No
miRNA ID: the ID of miRNA. miRNA regulated: the regulate of miRNA. DEMT: the differentially expressed miRNAs target genes. DEMT regulated: the regulate of differentially expressed miRNAs target genes. Degradation: the status of cleaved sites.				

The identification and heat map of DEMs

83637 precursor molecules were obtained from CKs and NTs by LC-MS technique, which included 46364 in the positive ion mode and 37273 in the negative ion mode; Based on the screening standards such as $VIP \geq 1$, $FC \geq 1.5$ or ≤ 0.667 and $p\text{-value} \leq 0.05$, 6125 differential metabolites between CKs and NTs were screened out of these precursor molecules, including 2,690 up-regulated and 3,435 down-regulated in NTs; After the exact masses of metabolites were confirmed according to mass error < 15 ppm, 30 DEMs were analyzed (Fig. 3A) and identified (Fig. 3B)(Table S11) including 11 up-regulated ($Z\text{-score} > 0$) and 19 down-regulated ($Z\text{-score} < 0$) in NTs (Fig. 3C); In their heat map constructed with R (v3.1.3) (Fig. 3D), the closer to 1 correlation coefficient, the more positively significant the correlation between any two metabolites. the closer to -1 correlation coefficient, the more negatively significant the correlation between any two metabolites; KEGG pathways of differential metabolites were presented in Table S12.

Integrated analysis of NGS and small RNA Sequencing based on KEGG pathway classification

KEGG pathway classifications of both DETs and DEMTs were all divided into Cellular process, Environmental information process, Genetic information processing, Metabolism and Organismal systems (Fig. 2). Moreover, both had some common KEGG pathways such as endocytosis in Cellular process, spliceosome and RNA degradation in Genetic information processing, biosynthesis of amino acids, glycerolipids, starch and sucrose metabolism and phenylpropanoid biosynthesis in Metabolism, plant-pathogen interaction in Organismal systems. These results indicated that DETs and DEMTs regulated common KEGG pathways to improve *R. glutinosa*.

Integrated analysis of NGS and metabolomics based on KEGG pathway

Correlation of DEMs and DETs was presented in Fig. 4. It was seen from it that 22 DEMs were correlated to 3846 DETs (Table S13). Among them, 1,357 transcripts were negatively correlated to DEMs, while 2,309 transcripts were positively correlated to DEMs. According to KEGG pathways, the relations of DEMs and DETs screened by correlation were mapped. One DEM was correlated to 1 ~ 16 DETs in one KEGG pathway. On the other hand, according to the results of differential metabolite analysis and differential transcripts analysis, the differential transcripts and differential metabolites in the same group were mapped to the same KEGG pathways at the same time for us to better understand the relationship between genes and metabolites. There were 24 KEGG pathways simultaneously mapped by DETs and DEMs (Table S14). According to $p\text{-value} < 0.05$, 6 KEGG pathways were commonly enriched, which included flavone and flavonol biosynthesis, flavonoid biosynthesis, phenylpropanoid biosynthesis, biosynthesis of unsaturated fatty acids, phenylalanine metabolism and purine metabolism (Fig. 4). Among them, 3 KEGG pathways, such as flavone and flavonol biosynthesis, flavonoid biosynthesis and phenylpropanoid biosynthesis, were extremely significant ($p\text{-value} < 0.01$). These results suggested that these DEMs were correlated to a core set of DETs.

Discussions

Previous studies indicated that NAA as one expansin could improve the yield of *R. glutinosa* by inhibiting the growth of its shoots and promote the transfer of nutrients to roots[3]. However, the effect of NAA on its quality and the genetic mechanism underlying its effect remain unknown. *R. glutinosa* is a non-model plant species without a reference of known genome, so its genetic basis is still narrow. In the current study, both were investigated using *LC-MS technology*, the full-length transcriptome *sequencing technology*, *small RNA Illumina sequencing and degradome sequencing*.

The effect of NAA on the metabolites in the mature tuberous roots of *R. glutinosa*

Chemical composition of *R. glutinosa* is an important part of its quality, of which catalpol, acteoside (ever) and Rehmannioside D are its index components for quality control. In this study, untargeted metabolomics was used to evaluate the effect of NAA on its metabolites. 6,125 differential metabolites between CK and NT were identified. Among them, 2,690 were up-regulated, while 3,435 were down-regulated in NTs, of which 30 DETs were matched in the databases and further identified; According to Z-score, 30 DETs were divided into 11 up-regulated DETs ($Z\text{-score} > 0$), such as melibiose, thromboxane B2, nicotinate D-ribonucleoside and so on, and 19 down-regulated DETs ($Z\text{-score} < 0$), including rosmarinic acid, lipoxin A4, linoleic acid and so on in NTs. Therefore, NAA treatment did not change the kinds of metabolites, but change the contents of some metabolites in the mature tuberous roots of *R. glutinosa*. However, Catalpol and acteoside were not differential metabolites between CK and NT. This result of ours is not consistent with the previous report, indicating the increased catalpol and acteoside contents in CPPU-treated *R. glutinosa* cultivar Wen 85-5[7], even though both are a type of expansins. The inconsistency could be caused by different cultivars of *R. glutinosa* or /and different chemical composition between NAA and CPPU. What is worse, Rehmannioside D was not identified. Anyway, NAA treatment could affect the contents of some metabolites in *R. glutinosa*, suggesting some change in its quality. On the other hand, these result of our current study also indicated that LC-MS technology had a potential to identify many metabolites in *R. glutinosa* as stated by previous reports[1,2], but could not identify any metabolite in *R. glutinosa*.

Full-length transcriptome sequencing and NGS-based analyses

RNA-Sequencing has been widely used in plant gene identification, expression and regulation. One of its most popular applications is for differential expression analysis where one identifies genes expressed at different levels between two classes of samples[29,30]. So far, the NGS technology has produced most of the plant transcriptomes, but the short reads from NGS sequencers cause incompletely assembled transcripts to be lack of some important information, for example, alternative splicing, limiting better understanding of transcriptome data. Based on the single-molecule real-time (SMRT) sequencing technology, the PacBio platform does not need to break mRNA randomly, can sequence the transcript from 5' end to 3' UTR region at once, generate longer and even full-length transcripts from observations of single molecules without assembly. The full-length transcripts can be used to investigate alternative splicing, alternative polyadenylation, novel genes, non-coding RNAs and fusion transcripts[31]. Based on these advantages of full-length transcriptome sequencing, we conducted the full-length transcriptome sequencing of *R. glutinosa*, a non-model plant, and constructed its full-length transcriptome database for the first time, which would be used as a reference database for other studies. Its database capacity of 26.61Gb clean data with cDNA size of 1-6kb is bigger than its NGS transcriptome databases such as 42.91Gb clean data in this study and one of its previous studies[20]. 46,412 coding sequences with complete ORF, obtained based on this full-length transcriptome database, made its gene cloning faster and more efficient and easier than homology cloning and electronic cloning[32,33]. By using the full-length transcriptome as reference, we performed Illumina sequencing to analyze the differential gene expression between NTs and CKs, found 596 upregulated DETs and 517 downregulated DETs in NT, and validated some of them by qRT-PCR. Based on their annotation results to four databases such as GO, KEGG, COG and eggNOG, a large number of transcripts related to metabolism, signal transduction, posttranslational modification and response to stimulus were significantly upregulated. On the other hand, according to GO annotation analysis, some transcripts related to cellular component organization, reproduction, growth and structural molecule activity were significantly downregulated. These findings showed that NAA improved the metabolites contents, the quality *R. glutinosa* by regulating some DETs mentioned above.

Identification and functions of miRNAs and their target genes

After 110.67 Mb clean data from 6 samples generated by small RNA sequencing were mapped to the full length transcriptome reference, 11 known miRNAs and 67 novel miRNAs were obtained, while no nat-siRNA and phasiRNA were identified. Among them, only novel_miR_57 was upregulated, while novel_miR_65 and novel_miR_23 were downregulated in NT. 18nt miRNAs were also discovered in other plants[34]. These miRNAs enlarged its known miRNA database and their categories. Before our work, Yang *et al.*[35] identified 25 conserved miRNA families from its cultivar Wen85-5 subjected to continuous cropping; Li *et al.*[36] identified 598 conserved miRNA families from the same cultivar again. However, when our identified miRNAs were aligned to their miRNAs, it was found that many miRNAs between theirs and ours were incomparable. The possible reasons leading to this incomparability are the disadvantage of their confirmation by bioinformatics method, the loss of some potential miRNAs during data analysis, the low annotation rate percent of 18.09-30.94, different cultivar and stress, of which the first three reasons were in accordance with the report [37]. 5,077DEMTs, including 2,285 of 11 known miRNAs and 3,179 of novel miRNAs, were predicted, of which 4,861 were annotated to 8databases. Among 4,861, 159 differential expressed genes were annotated to 8databases. In terms of KEGG annotation, 63DEMTs were annotated, of which 42 were annotated into 19types of KEGG pathways, dealing with cellular processes, environmental and genetic information processing, metabolism and organismal systems. 21(50%) DEMTs were related to genetic information processing such as RNA degradation and transport, spliceosome, ribosome and protein processing in endoplasmic reticulum and base excision repair; 8(19.05%) associated with metabolism were phenylpropanoid biosynthes, biosynthesis of amino acids, starch and sucrose metabolism, carotenoid biosynthesis glycerolipid metabolism and so on. These miRNAs down-regulated their target genes in order to regulate metabolism and metabolites accumulation. In addition, 247 DEMTs of novel_miR_23, novel_miR_57 and novel_miR_65 were further verified by degradome sequencing. We found that 25 of 247 DEMTs (10.12%) were cleaved, and that one miRNA could regulating many target genes. This phenomenon of one miRNA relative to multiple targets was also observed in other plant species.

NAA regulated genes and metabolites involved in phenylpropanoid biosynthesis to improve the quality of *R. glutinosa*

We found some gene and metabolites were regulated by NAA. In terms of phenylpropanoid pathway in which they took part, the general phenylpropanoid metabolism generates an enormous array of secondary metabolites[38]. The phenylpropanoid pathway is responsible for the synthesis of numerous compounds important for plant growth and responses to all aspects of plant responses towards biotic and abiotic stimuli[38,39] and might have potential implications in manipulating the tea quality in tea plants[40], which is consistent with ours. Phenylpropanoids a group of plant secondary metabolites derived from phenylalanine, exert beneficial pharmacological effects on human health such as anticancer, anti-inflammatory, antiviral, and antibacterial activity, and aid in wound healing[41], as well as have a wide variety of functions both as structural and signaling molecules[42]. We identified 14DEMTs taking part in Phenylpropanoid biosynthesis. Among them, 10genes such as F01_transcript_45042,14367,57187,57503,79209.83451,86382,88550, 45952,48883 encode beta-glucosidase; F01_transcript_54233 and 88383encode caffeoyl-CoA O-methyltransferase (CCoAOMT); F01_transcript_94508 and F01_transcript_33286 encode peroxidase and 4-coumarate-CoA ligase (4CL). (1) For beta-glucosidase, it responds to drought stress[43], dehydration and NaCl stress[44], freezing stress[45], induces the accumulation of higher abscisic acid (ABA) levels, which results in the formation of dwarf creeping plant[46], has high glycosyl hydrolase activity and plays a role in fruit ripening via the rapid production of free ABA[47], increases biomass yields and convert lignocellulose[48] and deals with stomatal development[49]; (2) For 4CL, phenylpropanoid pathway provides precursors for numerous secondary metabolites in plants. In this pathway, 4CL is the main branch point enzyme which generates activated thioesters. We identified one F01_transcript_33286 (4CL). Products of 4CL are subsequently

used by various oxygenases, reductases and transferases for biosynthesis of lignin, flavonoids, anthocyanins, aurones, stilbenes, coumarins, suberin, cutin, sporopollenin, etc[38]. 4CL plays multiple important roles in plant growth and development by catalyzing the formation of CoA ester[50], regulates the biosynthesis of lignin, flavonoids and wall-bound phenolics[51,52] plays a compensatory role in normal development[35], mediates the rice blast resistance, floret development and lignin biosynthesis[53] and affects plant phenotype and resulted in dwarfed plants with a "bonsai tree-like" appearance[54]; (3) For peroxidase, peroxidase possesses salt tolerance and antioxidant reaction[55], tolerance to chilling[56] and drought resistance[57], and promotes stomatal closure[42], and related to ROS and NO[58] and modulates ascorbic acid metabolism[59]; (4) For CCoAOMT, it is a key enzyme in the lignin biosynthesis pathway[60], and involved in lignin biosynthesis[61] and vanillin biosynthesis[62], and affects the accumulation of phenolic acids[14], and increases biomass production, Cd uptake and translocation[63]. In a word, these four genes were associated with phenylpropanoid biosynthesis, lignin biosynthesis, biomass, stress tolerance, growth, development and so on.

Integrated analysis of multi-omics

A combined transcriptomic and metabolic analysis can better explain the transcriptional regulation of metabolic pathways. To identify the main pathways involved in the roots of naphthylacetic acid-improving *R. glutinosa*, we simultaneously mapped DETs and DEMs from the same treatment groups to KEGG pathways to clarify the relationships between DETs and DEMs (Fig. 4). As summarized in results, 30 significantly differentially expressed metabolites were assigned to 12 KEGG pathways, including phenylpropanoid biosynthesis (ko00940), biosynthesis of unsaturated fatty acids (ko01040), flavone and flavonol biosynthesis (ko00944), and cysteine and methionine metabolism (ko00270), glyoxylate and dicarboxylate metabolism (ko00630), serine and threonine metabolism (ko00260), fatty acid biosynthesis (ko00061), alpha-Linolenic acid metabolism (ko00592), nicotinate and nicotinamide metabolism (ko00760), arginine and proline metabolism (ko00330), steroid biosynthesis (ko00100), galactose metabolism (ko00052). Among these pathways, phenylpropanoid biosynthesis (ko00940) included seven genes (F01_transcript_45042, F01_transcript_79209, F01_transcript_88383, F01_transcript_48883, F01_transcript_57503, F01_transcript_54233, F01_transcript_83451) and two metabolites (Sinapoyl aldehyde and (E)-3-(4-Hydroxyphenyl)-2-propenal), the seven genes were also enriched for alkaloid biosynthetic process (GO:0009821), lignin biosynthetic process (GO:0009809), carbohydrate metabolic process (GO:0005975), response to hormone (GO:0009725) and glucosinolate catabolic process (GO:0019762); steroid biosynthesis (ko00100) included one genes (F01_transcript_18023) and one metabolite (beta-Sitosterol), the genes was enriched for lipid catabolic process (GO:0016042); In addition, cysteine and methionine metabolism (ko00270) included one genes (F01_transcript_65625) and one metabolite (S-Adenosylmethionine), the genes was enriched for methionine biosynthetic process (GO:0009086) and homocysteine metabolic process (GO:0050667). For further study, A total of 78 miRNAs were identified by sRNA sequencing, including 11 known miRNAs and 67 new miRNAs, and 3 miRNAs were differentially expressed under NAA treatment. Through the joint analysis of miRNA and its target genes, 4 miRNA-mRNA pairs were identified, which were novel miRNA23, novel miRNA65, F01_transcript_81522 and F01_transcript_11999 annotated to phenylpropanoid biosynthesis (ko00940), respectively. Taken together, a coexpression regulatory network was constructed based on the profiles of the differentially expressed miRNA-mRNA-metabolite pairs, suggesting that NAA could improve the quality of *R. glutinosa* through multiple mechanisms (Fig. 5)

Conclusion

In our study, we first constructed a full length transcriptome database (26.61 Gb) with 46,412 coding sequences of *R. glutinosa*, 42.91 Gb transcriptome database, 110.67 Mb small RNA database and 19.72 Mb degradome database, identified 596 upregulated DETs and 517 downregulated DETs, 78 miRNAs, including 3 differentially expressed novel miRNAs, 11 up-regulated DEMs and 19 down-regulated DEMs, and predicted 5,077 DEMTs in NTs. NAA as expansin did not change the contents of catalpol and Rehmannioside D as quality standards but that of 30 metabolites, suggesting that it should not change the quality of *R. glutinosa*. Multi-omics analysis showed that NAA may regulate some metabolites contents of *R. glutinosa* via coexpression of DETs, DEMs, differentially expressed miRNAs and DEMTs linked to phenylpropanoid biosynthesis. These results enlarge *R. glutinosa* transcriptome and miRNA databases, provide novel insights into NAA improvement in its quality. The candidate genes and metabolites presented here will lay the foundation for some metabolite accumulation and biosynthesis mechanism elucidation, for example, acteoside, in *R. glutinosa*.

Methods

Plant materials

A total of 12 samples of *Rehmannia glutinosa*, divided into CK and NT groups because of the different treatment conditions. Materials were allowed and provided by Wen County Agricultural Science institute and were identified as *Rehmannia glutinosa* by Professor Yanqing ZHOU of Henan Normal University and Agronomist Cuihong Lu of Wen County Agricultural Science Institute. Group1 was not treated without any NAA (CK), while Group2 was treated three times with NAA with foliar spraying method: (1) A mixed solution of 50ml 5% naphthalene acetic acid with 15 liter water was sprayed into the young leaves of young *R. glutinosa* plants on June 20th. (2) A mixed solution of 5g 99% naphthalene acetic acid powder with 15 liter water was sprayed into the young leaves of young *R. glutinosa* plants on June 30th. (3) the same as (1) on July 10th. When its plants were mature, their tuberous roots were collected and stored in liquid nitrogen on Nov.16, 2019. The roots from three plants were mixed as one sample. Samples CK1-CK6 and NT1-NT6 were used for Illumina HiSeq2500 sequencing, metabolomics and qRT-PCR. A mixture of CK and NT as well as the leaves from both plants was

called Sample F01 for PacBio Full length transcriptome sequencing using sequencing by synthesis. Moreover, the yield of NT is 10% higher than that of CK.

Full-length transcriptome sequencing

Pacific Biosciences Long Read Processing

Raw reads were processed into error corrected reads of insert (ROIs) using Iso-seq pipeline with minFullPass=3 and minPredicted Accuracy=0.9. Next, full-length and non-chemeric (FLNC) transcripts were determined by searching for the polyA tail signal and the 5' and 3' cDNA primers in ROIs. ICE (Iterative Clustering for Error Correction) was used to obtain consensus isoforms and full-length (FL) consensus sequences from ICE was polished using Quiver[64]. High quality FL transcripts were classified with the criteria post-correction accuracy above 99%.

Remove redundant and Structure analysis

Iso-Seq high quality FL transcripts for removing redundancy using CD-HIT(identity > 0.99). Simple sequence repeats (SSRs) of the transcriptome were identified using MISA (<http://pgrc.ipk-gatersleben.de/misa/>); TransDecoder (<https://github.com/TransDecoder/TransDecoder/releases>) identifies candidate coding regions(CDS) within transcript sequences. The AS Gap should be larger than 100bp and at least 100bp away from the 3'/5' end[65].

Next generation transcriptome sequencing

Data Processing

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 poly-N and low quality reads from raw data. At the same time, Q20, Q30, GC-content and sequence duplication level of the clean data were calculated. All the downstream analyses were based on clean data with high quality[14]. These clean reads were then mapped to the Pacbio reference genome sequence. Only reads with a perfect match or one mismatch were further analyzed and annotated based on the reference genome. Hisat2 tools software were used to map with reference genome.

Quantification of gene expression levels and differential expression analysis

Gene expression levels were estimated by fragments per kilobase of transcript per million fragments mapped. For the samples with biological replicates: Prior to differential gene expression analysis, for each sequenced library, the read counts were adjusted by edgeR program package through one scaling normalized factor. Differential expression analysis of two samples was performed using the EBSeq R package[2]. The resulting FDR (false discovery rate) were adjusted using the PPDE(posterior probability of being DE). The FDR < 0.05 and $|\log_2(\text{foldchange})| \geq 1$ was set as the threshold for significantly differential expression.

Validation of gene expression using RT-qPCR

The RT-qPCR was performed on RNA extracted from CK and NT using the *GAPDH* gene as the internal reference. Specific primer pairs of 22 selected genes were designed using the Primer Premier 5.0 (Table 1). Data were presented as relative transcript level based on the $2^{-\Delta\Delta Ct}$ method [23].

Functional enrichment analysis

Gene Ontology (GO) enrichment analysis of the DETs was implemented by the GOseq R packages based Wallenius non-central hyper-geometric distribution[66], which can adjust for gene length bias in DETs; KEGG[67] is a database resource for understanding high-level functions and utilities of the biological system, such as the cell, the organism and the ecosystem, from molecular-level information, especially large-scale molecular datasets generated by genome sequencing and other high-throughput experimental technologies (<http://www.genome.jp/kegg/>). We used KOBAS software[68] to test the statistical enrichment of differential expression genes in KEGG pathways.

Small RNA sequencing and data analyses

The small RNAs from CK and NT were sequenced and their data were analyzed as stated by Wang *et al.*[14].

Metabolomics detection

Metabolites extraction[69-71] Accurately weigh 50mg ($\pm 1\%$) of each sample in 2 mL EP tube, and add 0.6 mL 2-chlorophenylalanine (4 ppm) methanol ($-20\text{ }^{\circ}\text{C}$), vortex for 30 seconds; Add 100 mg glass beads and grind the samples by a high-throughput tissue grinder for 90 seconds at 60 Hz; Ultrasonic treatment for 15 minutes at room temperature; Centrifuge at $4\text{ }^{\circ}\text{C}$ for 10 min at 14000 rpm, and the supernatant was filtered through $0.22\text{ }\mu\text{m}$ membrane to obtain the prepared samples for LC-MS; Take $30\text{ }\mu\text{L}$ from each sample to the quality control (QC) samples (These QC samples were used to monitor deviations of the analytical results from these pool mixtures and compare them to the errors caused by the analytical instrument itself)[72]; Use the rest of the samples for LC-MS detection.

LC-MS test

Chromatographic separation was accomplished in an Shimadzu LC-30A system equipped with an ACQUITY UPLC[®] HSS T3 ($150\times 2.1\text{ mm}$, $1.8\text{ }\mu\text{m}$, Waters) column maintained at $40\text{ }^{\circ}\text{C}$. The temperature of the autosampler was $4\text{ }^{\circ}\text{C}$. Gradient elution of analytes was carried out with 0.1% formic acid in water (A) and acetonitrile (B) at a flow rate of 0.3 mL/min . Injection of $5\text{ }\mu\text{L}$ of each sample was done after equilibration. An increasing linear gradient of solvent B (v/v) was used as follows: $0\sim 0.5\text{ min}$, 2% B; $0.5\sim 9\text{ min}$; 2%~50% B; $9\sim 12\text{ min}$, 50%~98% B; $12\sim 13\text{ min}$, 98% B; $13\sim 14\text{ min}$, 98%~2% B; $14\sim 15\text{ min}$, 2% B[73].

The ESI-MSⁿ experiments were executed on the AB 5600+ mass spectrometer with the spray voltage of 5.50 kV and -4.50 kV in positive and negative modes, respectively. Gas 1 and gas 2 were both set at 50 psi. Curtain gas was 35 psi. The source temperature was $500\text{ }^{\circ}\text{C}$. The mass analyzer scanned over a mass range of $m/z\ 100\text{--}500$ for full scan at the collision energy of 45 eV. Dynamic exclusion was implemented to remove some unnecessary information in MS/MS spectra[73].

Sample data preprocessing

Through proteowizard software (v3.0.8789), the original data is converted into mzXML format (xcms input file format)[74]; Using xcms package of R (v3.3.2) to identify, filter and align peaks. The main parameters are $\text{BW} = 5$, $\text{PPM} = 15$, $\text{peakwidth} = \text{C}(5,30)$, $\text{mzwidth} = 0.015$, $\text{mzdiff} = 0.01$, $\text{method} = \text{"centWave"}$; Data matrix, including mass to charge ratio (M/z), retention time and peak area, are obtained, in which 46364 precursor molecules were obtained in the positive ion mode, 37273 precursor molecules were obtained in the negative ion mode. The data were exported to Excel for follow-up analysis; In order to compare the data of different magnitude, the batch normalization of peak area is carried out.

Multivariate statistical analyses

The methods of multivariate statistical analysis (software package Simca-p (v13.0) and R language ropls package)[75] are as follows: Principal component analysis (PCA), Partial least squares discriminant analysis (PLS-DA) and Orthogonal partial least squares discriminant analysis (OPLS-DA).

Screening and identification of differential metabolites

Based on the screening criteria for differential metabolites : $p\text{-value} \leq 0.05$ and $\text{VIP} \geq 1$ [76,77], final differential metabolites were screened out. The exact masses of metabolites were confirmed according to mass error $< 15\text{ ppm}$. Then, Fragment information from MS / MS mode in data matrix were further matched in the databases such as HumanMetabolome Database (HMDB) (<http://www.hmdb.ca>), Metlin (<http://metlin.scripps.edu>), massbank (<http://www.massbank.jp/>), LipidMaps (<http://www.lipidmaps.org>), mzcloud (<https://www.mzcloud.org>) and Self-built standard product database[72].

Relative contents statistical analyses of differential metabolites

Z-score is a value based on the relative content conversion of metabolites, which is used to measure the relative content of metabolites on the same level[78]. The calculation of Z-score is based on the mean and standard deviation of reference data set (control group). The specific formula is $Z = (x - \mu) / \sigma$, in which x is a specific fraction, μ is the average, σ is the standard deviation[72].

Correlation heat map of differential metabolites

When the linear relationship between the two metabolites is enhanced, the correlation coefficient tends to 1 or -1: positive correlation tends to 1, negative correlation tends to -1. R (v3.1.3) cor.test function was used to test the correlation analysis of metabolites, $p\text{-value} \leq 0.05$ had significant correlation[79].

KEGG pathways of differential metabolites

MetPA database is a part of metaanalyst (www.metaboanalyst.ca). Its analyses are mainly based on the metabolic pathway of KEGG[80]. MetPA database can be used to analyze the metabolic pathways related to two groups of different metabolites. MetPA database identifies the metabolic pathways that may be disturbed by organisms through metabolic pathway concentration and topological analysis, and then analyzes the metabolic pathways of metabolites[72].

Combined analysis of transcriptome and metabolome

Correlation of differential genes and metabolites was analyzed as stated by Liu *et al.*[81]. Both KEGG pathways were analyzed as mentioned above. Both common enriched KEGG pathways were analyzed based on $p\text{-value} < 0.05$.

Abbreviations

Rehmannia glutinosa; *R. glutinos*; NAA: Naphthylacetic acid; NT: NAA-treated tuberous roots; CK: Non-treated tuberous roots; CPPU: N-2-chlorine-4-pyridylphenyl-N-phenylurea; DETs: differentially expressed transcripts; DETs: Differentially expressed metabolites; DEMTs: Differentially expressed miRNAs target genes; NGS: Next generation transcriptome; SMRT: Single-molecule real-time; KEGG: Kyoto Encyclopedia of Genes and Genomes; GO: Gene Ontology; NR: NCBI non-redundant protein sequences; Pfam: Protein family; COG: Clusters of Orthologous Groups of proteins.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The datasets used and analyzed in the current study are available from the Supplemental Materials.

Competing interests

No conflicts of interest declared.

Funding

This work was supported by National Key Research and Development Project “modernization research on traditional Chinese medicine” (No. 2017YFC1702800), Henan Provincial Major Science and Technology Projects (No. 2017AB002), Natural Science Foundation of Henan province (No. 182300410018) and 2016 National Project Training Funds of Henan Normal University (No. 2016PL11). There is no role of the funding body in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Authors' Contributions

LJJ funded this study via National Key Research and Development Project “modernization research on traditional Chinese medicine” and Henan Provincial Major Science and Technology Projects. ZYQ and LJJ conceived and designed this study. ZYQ supervised its experiments and wrote manuscript. SLY and ZJL collected samples, analyzed the data and helped ZYQ to write the manuscript. MJX and LHM detected the traits of CKs and NTs with SLY. All Authors read and approved the final manuscript.

Acknowledgments

Rehmannia glutinosa materials of this work were provided by Agronomist cuihong Lu (Wen County Agricultural Science institute). At the point of finishing this paper, I'd like to express my sincere thanks to all those who have lent me hands in the course of my writing this paper.

Author details

¹College of Life Sciences, Henan Normal University, Xinxiang 453007, Henan, P.R.China.

References

1. Yang K. Metabolite accumulation and metabolic network in developing roots of *Rehmannia glutinosa* reveal its root developmental mechanism, quality and verbascoside biosynthesis pathway, Master Thesis, Henan Normal University, Xinxiang. 2019.
2. Zhou YQ, Yang K, Zhang DD, Duan HY, Liu YK and Guo MM. Metabolite accumulation and metabolic network in developing roots of *Rehmannia glutinosa* reveals its root developmental mechanism and quality. *Scientific Reports*. 2018;8:14127
3. Lu CH, Dong J and Xu YX. A preliminary report on the application of expansin (CPPU) in the production of *Rehmannia glutinosa*. *Anhui Agri Sci Bull*. 2019;25(14):20-27.
4. Zhou YQ, Yao HL, Zhou CE, Duan HY and Lu SX. Review on *Rehmannia glutinosa* breeding. *Guihaia*. 2010;30(3): 373-376.
5. Zhang Y. High yield cultivation methods of *Rehmannia glutinosa*. China, CN201811084848.1.2020.
6. Chen MX, Zhao XT, Li JJ, Zhao XL, Wang W, Wang GF, Liu XX and Li MJ. Yield and quality of virus-free *Rehmannia glutinosa* 85-5 applied with CPPU, *Guizhou Agricultural Sciences*. 2014;42(12):109-111.
7. Chen MX, Li JJ, Li MJ, Li XW, Bai YH, Zhang XL. Effects of combined application of paclobutrazol and for chlorfenuron on the growth and quality of virus-free *Rehmannia glutinosa* Hueichingensis (Chan et Seih) Hsiao. *Journal of Southern Agriculture*. 2014;45(11):1947-1950.
8. Lin HB. α -naphthylacetic acid: a regulator of plant growth. *Chemical education*. 2009;10:6-7.
9. Huang Y. Effects of root application sodium naphthalene acetate and polyaspartic acid on growth, physiological characteristics and yield of *Cucumber*. Master Thesis. Chinese academy of agricultural sciences. 2017.
10. Sun J, Xiong SM, Pi XL, Le MW, Rao YL, Yan TX, Yan XW and Zhou HY. Effects of foliar spraying naphthylacetic acid at different periods on yield and seed quality of autumn *Sesame* variety Ganzhi No.13. *Journal of Southern Agriculture*. 2018;49(7):1318-1323
11. Liang YP. Effects of culture substrates and expansin on metabolism of carbohydrates during dulbs development of *Tulip*. Master Thesis, Qinghai university. 2013.
12. Xia CB, Xiao Y, Lin YH, Peng YF, Wu MC, Wu YJ, Wu Q, Liu M and Zeng JF. Experimental report on the effect of 5% NAA on *Tomato* growth and yield. *Modern horticulture*. 2018;7: 10-11.
13. Wang Z, Gerstein M and Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet*. 2009;10(1):57-63.
14. Wang Y, Wang Q, Gao L, Zhu B, Ju Z, Luo Y and Zuo J. Parsing the Regulatory Network between Small RNAs and Target Genes in Ethylene Pathway in *Tomato*. *Front Plant Sci*. 2017;8:527
15. Banerjee S, Sirohi A, Ansari AA and Gill SS. Role of small RNAs in abiotic stress responses in plants. *Plant Gene*. 2017;11:180-189.
16. Huang JJ, Liu WW, Guo YR, Jiang TH, Ren Q, Wang HH and Liang WH. Research progress of microRNA in plant development. *Biotechnology Bulletin*. 2019;35(11):141-149.
17. Shi L, Guo YB and Shen Y. Regulation of growth, development and secondary metabolism of medicinal plants by miRNA. *Chinese Journal of Biochemistry and Molecular Biology*. 2019;35(04):361-370.
18. Yang YH, Li MJ, Yi YJ, Li RF and Dong C. Identification of 12 lncRNAs as potential miRNA Targets in roots of *Rehmannia glutinosa* responding to replanting disease. *International journal of agriculture and biology*. 2018;20:1201-1210.
19. Yang YH, Chen XJ, Chen JY, Xu HX, Li J and Zhang ZY. Identification of Novel and Conserved microRNAs in *Rehmannia glutinosa* L. by Solexa Sequencing. *Plant molecular biology reporter*. 2011;29(4): 986-996.
20. Zhou YQ, Wang XN, Wang WS and Duan HY. De novo transcriptome sequencing-based discovery and expression analyses of verbascoside biosynthesis-associated genes in *Rehmannia glutinosa* tuberous roots. *Molecular Breeding*. 2016;36(10):139.
21. Cui JW, Shen N, Lu ZG, Xu GL, Wang YY and Biao JB. Analysis and comprehensive comparison of PacBio and nanopore-based RNA sequencing of the *Arabidopsis* transcriptome. *Plant Methods*. 2020;16:85.
22. Hu ZC, Fu QS, Zheng J, Zhang A and Wang HS. Transcriptomic and metabolomic analyses reveal that melatonin promotes melon root development under copper stress by inhibiting jasmonic acid biosynthesis, *Horticulture Research*. 2020;7;1-15.
23. Zhang ZL, Liu ZH, Song HN, Chen MH and Cheng SP. Protective role of leaf variegation in *Pittosporum tobira* under low temperature: insights into the physio-biochemical and molecular mechanisms. *International Journal of Molecular Sciences*. 2019;20(19):4857.
24. Liang GT, Guo J, Zhang SY and Zhang GC. Integration of small RNAs, degradome, and transcriptome sequencing in *Populus x euramericana* "Neva" provides insights into the allelopathic interference of para-hydroxybenzoic acid. *Canadian Journal of Forest Research*. 2019;50(4):422-437.
25. Li WZ and Godzik A. CD-HIT: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*. 2006;22(13):1658-1659.
26. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV and Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*. 2015;31(19): 3210-3212.

27. Li B and Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC bioinformatics*. 2011;12:323.
28. Anders S and Huber W. Differential expression analysis for sequence count data. *Genome Biology*. 2010;11(10):R106.
29. Rao GD, Zhang JG, Liu XX and Luo Y. Identification of putative genes for polyphenol biosynthesis in olive fruits and leaves using full-length transcriptome sequencing. *Food chemistry*. 2019;300:125246.
30. Zhang XK and Jonassen I. RASflow: an RNA-Seq analysis workflow with Snakemake. *BMC Bioinformatics*. 2020;21(1):110.
31. Ren YP, Zhang JQ, Sun Y, Wu ZF, Ruan JS, He BJ, Liu GQ, Gao S, Bu WJ. Full-length transcriptome sequencing on PacBio platform. *Chinese Science Bulletin*. 2016;61(11):1250-1254.
32. Guo MM. SNP markers and RgURT,RgTAT,Rg4CL,RgC4H and RgC3H1 genes of *Rehmannia glutinosa* Master Thesis, Henan Normal University, Xinxiang. 2020:29-33.
33. Zhang DD. Discovery and expressions of key verbascoside biosynthesis and glycolysis-associated enzyme genes in *Rehmannia glutinosa*, Master Thesis, Henan Normal University, Xinxiang. 2018.
34. Zhang Y, Rahmani RS, Yang XY, Chen JM and Shi T. Integrative expression network analysis of microRNA and gene isoforms in sacred lotus. *BMC Genomics*. 2020;21(1):429.
35. Yang, JM, Chen, F, Yu, O and Beachy RN. Controlled silencing of 4-coumarate:CoA ligase alters lignocellulose composition without affecting stem growth. *Plant Physiology and Biochemistry*. 2011;49(1):103-109.
36. Li MJ, Yang YH, Chen XJ, Wang FQ, Lin WX, Yi YJ, Zeng L, Yang SY, Zhang ZY. Transcriptome/ Degradome-Wide Identification of *R. glutinosa* miRNAs and Their Targets: The Role of miRNA Activity in the Replanting Disease. *Plos One*. 2013;8(7): e68531.
37. Sun P, Cheng C, Lin Y, Zhu Q, Lin J and Lai Z. Combined small RNA and degradome sequencing reveals complex microRNA regulation of catechin biosynthesis in tea (*Camellia sinensis*). *PLoS ONE*. 2017;12(2): e0171173
38. Vogt T. Phenylpropanoid biosynthesis. *Molecular Plant*. 2010;3(1): 2–20.
39. Kurepa J, Shull TE, Karunadasa SS and Smalle JA. Modulation of auxin and cytokinin responses by early steps of the phenylpropanoid pathway. *BMC Plant Biology*. 2018;18(1): 278.
40. Li X, Zhang LP, Zhang L, Yan P, Ahammed G and Han WY. Methyl salicylate enhances flavonoid biosynthesis in tea leaves by stimulating the phenylpropanoid pathway. *Molecules*, 2019;24(2): 362.
41. Cuong DM, Ha TW, Park CH, Kim NS, Yeo HJ, Chun SW, Kin C, Park S. Effects of LED lights on expression of genes involved in phenylpropanoid biosynthesis and accumulation of phenylpropanoids in wheat sprout, *Agronomy*. 2019;9(6): 307.
42. Liu FR, Xie LF, Yao ZY, Zhou YL and Gong CM. Caragana korshinskii phenylalanine ammonialyase is up-regulated in the phenylpropanoid biosynthesis pathway in response to drought stress. *Biotechnology & Biotechnological Equipment*. 2019;33(1):842-854.
43. Wang PT, Liu H, Hua HJ, Wang L and Song CP. A vacuole localized beta-glucosidase contributes to drought tolerance in *Arabidopsis*. *Chinese Science Bulletin*. 2012;56(33): 3538-3546.
44. Xu ZY, Lee KH, Dong T, Jeong JC, Jin JB, Kanno Y, Kim DH, Kim SY, Seo M, Bressan R, Yun DJ, Hwang I. A vacuolar beta-glucosidase homolog that possesses glucose-conjugated abscisic acid hydrolyzing activity plays an important role in osmotic stress responses in *Arabidopsis*. *Plant Cell*. 2012;24(5): 2184-2199.
45. Hashempour A, Ghasemnezhad M, Sohani MM, Ghazvini RF and Abedi A. Effects of freezing stress on the expression of fatty acid desaturase (FAD2, FAD6 and FAD7) and beta-glucosidase (BGLC) genes in tolerant and sensitive *olive* *Russian Journal of Plant Physiology*. 2019;66(2): 214-222.
46. Han YJ, Cho KC, Hwang OJ, Choi YS, Shin AY, Hwang I and Kim JI. Overexpression of an *Arabidopsis* beta-glucosidase gene enhances drought resistance with dwarf phenotype in creeping bentgrass. *Plant Cell Reports*. 2012;31(9): 1677-1686.
47. Zhang SH, Sun JH, Dong YH, Shen YY, Li CL, Li YZ and Guo JX. Enzymatic and functional analysis of beta-glucosidase FaBG1 during strawberry fruit ripening. *Journal of Horticultural Science & Biotechnology*. 2014;89(6), 733-739.
48. Cho EJ, Nguyen QA, Lee YG, Song Y, Park BJ and Bae HJ. Enhanced Biomass Yield of and Saccharification in transgenic *Tobacco* over-Expressing beta-Glucosidase. *Biomolecules*. 2020;10(5):806.
49. Jiao QS, Chen TS, Niu GT, Zhang HC, Zhou CF and Hong Z. N-glycosylation is involved in stomatal development partially by modulating the release of active abscisic acid and auxin by beta-glucosidase in *Arabidopsis thaliana*. *Journal of experimental botany*. 2020;71(19):5865-5879.
50. Chen XH, Wang HT, Li XY, Ma K, Zhan YG and Zeng FS. Molecular cloning and functional analysis of 4-coumarate-CoA ligase (4CL-like 1) from *Fraxinus mandshurica* and its role in abiotic stress tolerance and cell wall synthesis, *BMC Plant Biology*. 2019;19(1): 231.
51. Li SS, Chang Y, Li B, Shao SL and Zhang ZZ. Functional analysis of 4-coumarate: CoA ligase from *Dryopteris fragrans* in transgenic tobacco enhances lignin and flavonoids, *Genetics and Molecular Biology*. 2020;43(2): e20180355.
52. Liu XY, Wang PP, Wu YF, Cheng AX and Lou HX. Cloning and functional characterization of two 4-coumarate: CoA ligase genes from *Selaginella moellendorffii*. *Molecules*. 2018;23(3):595.
53. Liu H, Guo ZH, Gu FW, Ke SW, Sun DY, Dong SY, Liu W, Huang M, Xiao WM, Yang GL, Liu YZ, Guo T, Wang H, Wang JF, Chen ZQ. 4-Coumarate-CoA ligase-like gene OsAAE3 negatively mediates the rice blast resistance, floret development and lignin biosynthesis. *Front. Plant Sci*. 2017;7:2041.

54. Wagner A, Donaldson L, Kim H, Phillips L, Flint H, Steward D, Torr K, Gerald K, Uwe S, John R. Suppression of 4-coumarate-CoA ligase in the coniferous gymnosperm *Pinus* Plant Physiology. 2009;149(1): 370-383.
55. Jin T, Sun YY, Zhao RR, Shan Z, Gai JY and Li Y. Overexpression of Peroxidase Gene GsPRX9 Confers Salt Tolerance in *Soybean*. International Journal of Molecular Sciences. 2019;20(15):3745.
56. Chen ZH, Lu HH, Hua SM, Lin KH, Chen ND, Zhang YW, You ZY, Kuo YW, Chen SP. Cloning and overexpression of the ascorbate peroxidase gene from the yam (*Dioscorea alata*) enhances chilling and flood tolerance in transgenic *Arabidopsis*. Journal of Plant Research. 2019;132(6): 857-866.
57. Liu JX, Feng K, Duan AQ, Li H, Yang QQ, Xu ZS and Xiong AS. Isolation, purification and characterization of an ascorbate peroxidase from celery and overexpression of the AgAPX1 gene enhanced ascorbate content and drought tolerance in *Arabidopsis*. BMC Plant Biology. 2019;19(1):488.
58. Kim YH, Yun BW and Kwak SS. Expression of the sweet potato peroxidase gene swpa4 in *Arabidopsis* activates defense genes mediated by reactive oxygen species and nitric oxide. Plant Biotechnology Reports. 2019;13(4):329-336.
59. Li H, Liu H, Wang Y, Teng RM, Liu JY, Lin SJ and Zhuang J. Cytosolic ascorbate peroxidase 1 modulates ascorbic acid metabolism through cooperating with nitrogen regulatory protein P-II in tea plant under nitrogen deficiency stress. Genomics. 2020;112(5):3497-3503.
60. Li XY, Chen WJ, Zhao Y, Xiang Y, Jiang HY, Zhu SW and Cheng BJ. Downregulation of caffeoyl-CoA O-methyltransferase (CCoAOMT) by RNA interference leads to reduced lignin production in maize straw. Genetics and Molecular Biology. 2013;36(4):540-546.
61. Wang ZJ, Ge Q, Chen C, Jin XX, Cao XY and Wang ZZ. Function Analysis of Caffeoyl-CoA O-Methyltransferase for Biosynthesis of Lignin and Phenolic Acid in *Salvia miltiorrhiza*. Applied Biochemistry and Biotechnology. 2017;181:562-572.
62. Widiez T, Hartman TG, Dudai N, Yan Q, Lawton M, Havkin-Frenkel D and Belanger FC. Functional characterization of two new members of the caffeoyl CoA O-methyltransferase-like gene family from *Vanilla planifolia* reveals a new class of plastid-localized O-methyltransferases. Plant Molecular Biology. 2011;76:475-488.
63. Xia Y, Liu J, Wang Y, Zhang XX, Shen ZG and Hu ZB. Ectopic expression of *Vicia sativa* Caffeoyl-CoA O-methyltransferase (VsCCoAOMT) increases the uptake and tolerance of cadmium in *Arabidopsis*. Environmental And Experimental Botany. 2018;145:47-53.
64. Gordon SP, Tseng E, Salamov A, Zhang JW, Meng XD, Zhao ZY, Kang DW, Underwood J, V.Grigroriev I, Figueroa M, S.Schilling J, Chen F and Wang Z. Widespread Polycistronic Transcripts in Fungi Revealed by Single-Molecule mRNA Sequencing. PLoS One. 2015;10:e0132628.
65. Deng Y, Zheng H, Yan ZC, Liao DY, Li CL, Zhou JY and Liao H. Full-Length transcriptome survey and expression analysis of *Cassia obtusifolia* to discover putative genes related to Aurantio-Obtusin Biosynthesis, Seed Formation and Development, and Stress Response. International Journal of Molecular Sciences. 2018;19:2476.
66. Young MD, Wakefield MJ, Smyth GK and Oshlack A. Gene ontology analysis for RNA-seq: accounting for selection bias. Genome Biol. 2010;11:R14.
67. Kanehisa M, Goto S, Kawashima S, Okuno Y and Hattori M. The KEGG resource for deciphering the genome. Nucleic Acids Research. 2004;32:D277-D280.
68. Mao X, Cai T, Olyarchuk JG and Wei L. Automated genome annotation and pathway identification using the KEGG Orthology(KO) as a controlled vocabulary. Bioinformatics. 2005;21(19):3787-3793.
69. Ponnusamy K, Choi JN, Kim J, Lee SY and Lee CH. Microbial community and metabolomic comparison of irritable bowel syndrome faeces. J Med Microbiol. 2011;60:817-827.
70. Ng JS, Ryan U, Trengove RD and Maker GL. Development of an untargeted metabolomics method for the analysis of human faecal samples using *Cryptosporidium*-infected samples. Mol Biochem Parasitol. 2012;185:145-150.
71. Sangster T, Major H, Plumb R, Wilson AJ and Wilson ID. A pragmatic and readily implemented quality control strategy for HPLC-MS and GC-MS-based metabolomic analysis. Analyst. 2006;131:1075-1078.
72. Zhang YW, Chen WX, Chen HM, Zhong QP, Yun YH and Chen WJ. Metabolomics analysis of the deterioration mechanism and storage time limit of tender coconut water during storage. Foods. 2020;9:46.
73. Sun H, Zhang AH, Song Q, Fang H, Liu XY, Su J, Yang L, Yu MD and Wang XJ. Functional metabolomics discover pentose and glucuronate interconversion pathways as promising targets for Yang Huang syndrome treatment with Yinchenhao Tang. RSC Advances, 2018;8:36831-36839.
74. Smith CA, Want EJ, O'Maille G, Abagyan R and Siuzdak G. XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. Analytical chemistry. 2006;78:779-787.
75. Thevenot EA, Roux A, Xu Y, Ezan E and Junot C. Analysis of the human adult urinary metabolome variations with age, body mass index and gender by implementing a comprehensive workflow for univariate and OPLS statistical analyses. Journal of Proteome Research. 2015;14:3322-3335.
76. Trygg J and Wold S. Orthogonal projections to latent structures (O-PLS). Journal of chemometrics. 2002;16:119- 128.
77. Wang JB, Pu SB, Sun Y, Li ZF, Niu M, Yan XZ, Zhao YL, Wang LF, Qin XM, Ma ZJ, Zhang YM, Li BS, Luo SQ, Gong M, Sun YQ, Zou ZS and Xiao XH. Metabolomic profiling of autoimmune hepatitis: the diagnostic utility of nuclear magnetic resonance spectroscopy. Journal of proteome research. 2014;13:3792-3801.
78. Chen X, Xie C, Sun L, Ding JH and Cai HB. Longitudinal Metabolomics Profiling of Parkinson's Disease Related α -Synuclein A53T Transgenic Mice. PLoS one. 2015;10(8), e0136612.

79. Rao G, Sui J and Zhang J. Metabolomics reveals significant variations in metabolites and correlations regarding the maturation of walnuts (*Juglans regia* L.). *Biology open*. 2016;5:829-836.
80. Kanehisa M and Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research*. 2000;28(1):27-30.
81. Liu XX, Zhang JG, Luo Y and Rao GD. Metabolome and Transcriptome Analyses Reveal Tissue-Specific Variations in Gene Expression and Metabolites of *Olive*. *Journal of plant biology*. 2020;63:73-82.

Figures

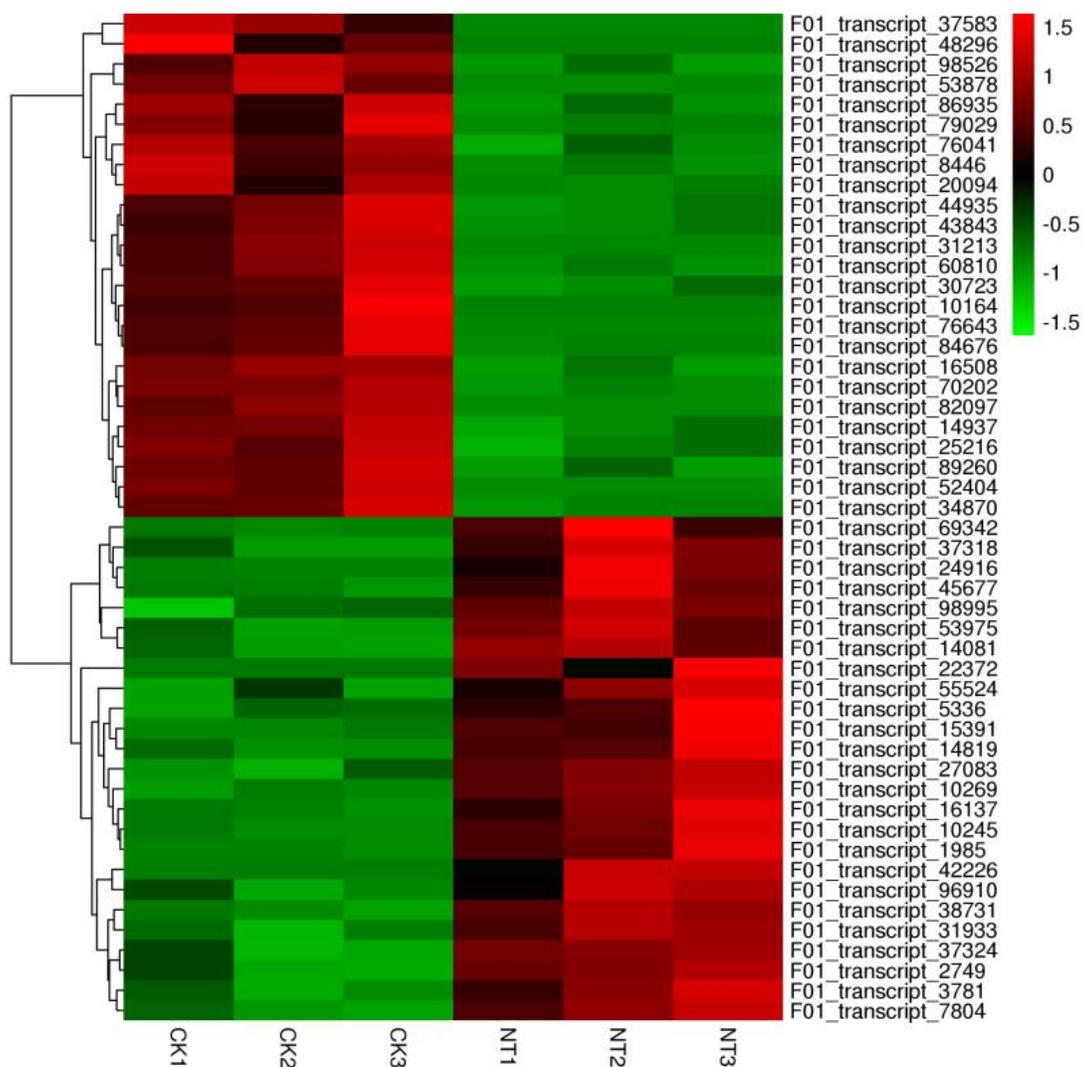


Figure 1

DETs cluster. Ordinate: Abscissa: the names and cluster of samples. Ordinate: DETs cluster. Different columns represent different samples. Different rows represent different transcripts. Color represents the expression levels (log₂FPKMs) of DETs in the samples. F01_transcript_37583: unknown, F01_transcript_98526: unknown, F01_transcript_53878: unknown, F01_transcript_86935: TT12, F01_transcript_79029: PP2C 27, F01_transcript_76041: ENODL2, F01_transcript_8446: CSLB3, F01_transcript_20094: Lipase_GDSL, F01_transcript_44935: ELIP1, F01_transcript_43843: ELIP1, F01_transcript_31213: U2AF65B, F01_transcript_60810: VALRS, F01_transcript_30723: unknown, F01_transcript_10164: PP2C6, F01_transcript_76643: NIPA8, F01_transcript_84676: VTC2, F01_transcript_16508: unknown, F01_transcript_70202: VTC2, F01_transcript_82097: PP2C39, F01_transcript_14937: PP2C27, F01_transcript_25216: HBPL1, F01_transcript_89260: TPRL3, F01_transcript_52404: KAT2, F01_transcript_34870: ELIP1, F01_transcript_69342: unknown, F01_transcript_37318: DMS3, F01_transcript_24916: unknown, F01_transcript_45677: FERONIA, F01_transcript_98995: CKA, F01_transcript_53975: CAX3, F01_transcript_14081: CAX3, F01_transcript_22372: APT1, F01_transcript_55524: TBL11, F01_transcript_5336: BGAL, F01_transcript_15391: ALN, F01_transcript_14819: ASIL2, F01_transcript_27083: CXXS1, F01_transcript_10269: unknown, F01_transcript_16137: CYP716A52v2, F01_transcript_10245: RH56, F01_transcript_1985: POT11, F01_transcript_42226: RNHX1, F01_transcript_96910: RNHX1, F01_transcript_38731: DMS3, F01_transcript_31933: FRS5, F01_transcript_37324: MES17, F01_transcript_2749: unknown, F01_transcript_3781: CBSX3, F01_transcript_7804: HSP70-6.

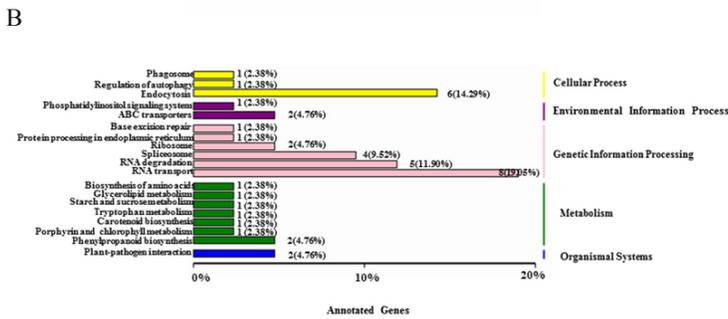
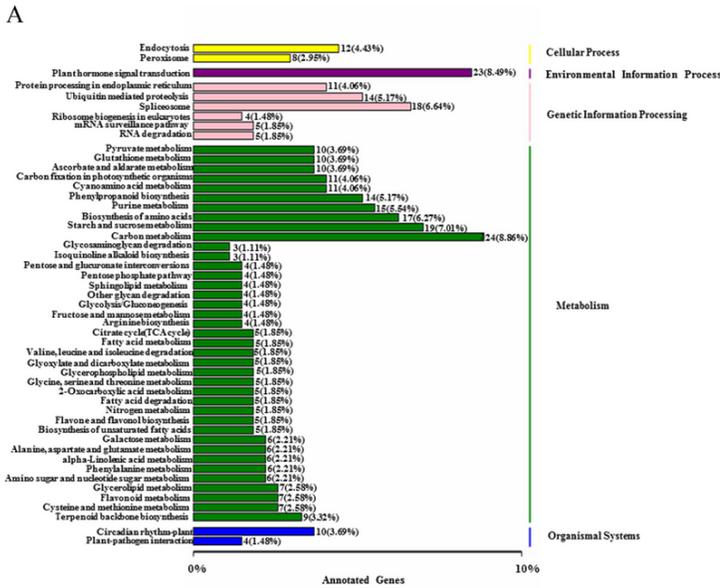


Figure 2

KEGG pathway classification of annotated DETs and DEMTs. (A) KEGG classification of annotated DETs. Abscissa: Number of DETs annotated to a KEGG pathway and percentage of the number to all transcripts annotated to all KEGG pathways. Ordinate: Names of KEGG pathways. 280 differentially expressed transcripts were annotated 387 times in 50 KEGG pathways, classified into five groups such as cellular process, environmental information process, genetic information processing, metabolism and organismal systems. (B) KEGG pathway classification of annotated DEMTs. Abscissa: the number of DEMT(s) annotated to one KEGG pathway, the number: total number of DEMT(s) annotated to all of KEGG pathways.

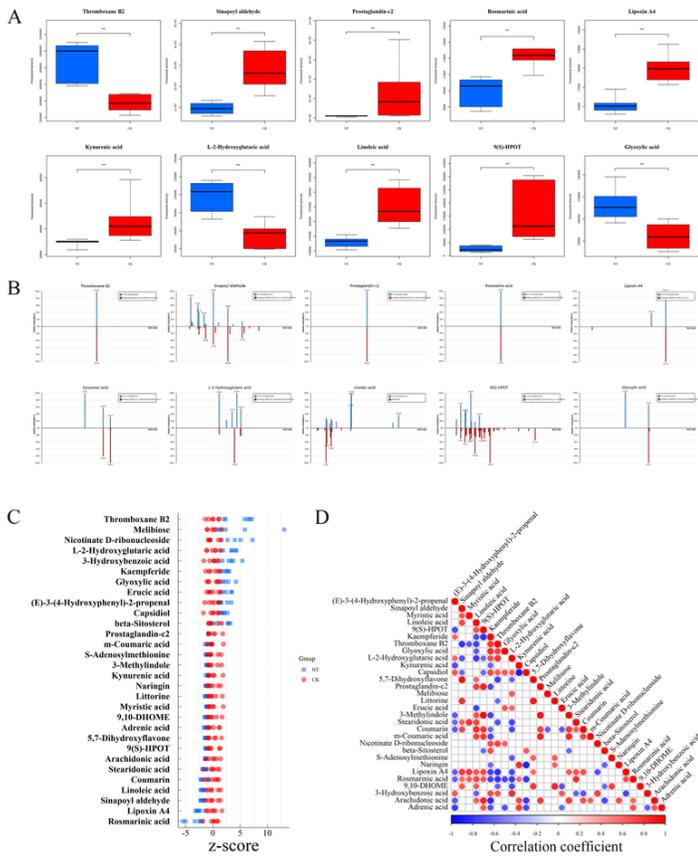


Figure 3

Identification and analysis of DEMs. (A) Box plot of DEMs. (B) Identification of DEMs. (C) Z-score plot (NT vs CK). Abscissa: Z-score of 30 differential metabolites. Ordinate: 30 differential metabolites matched and further identified in the database. (D) Correlation heat map of differential metabolites. Abscissa: correlation coefficient of 30 differential metabolites. Ordinate: 30 differential metabolites matched and further identified in the database. Red ball: CK. Blue ball: NT.

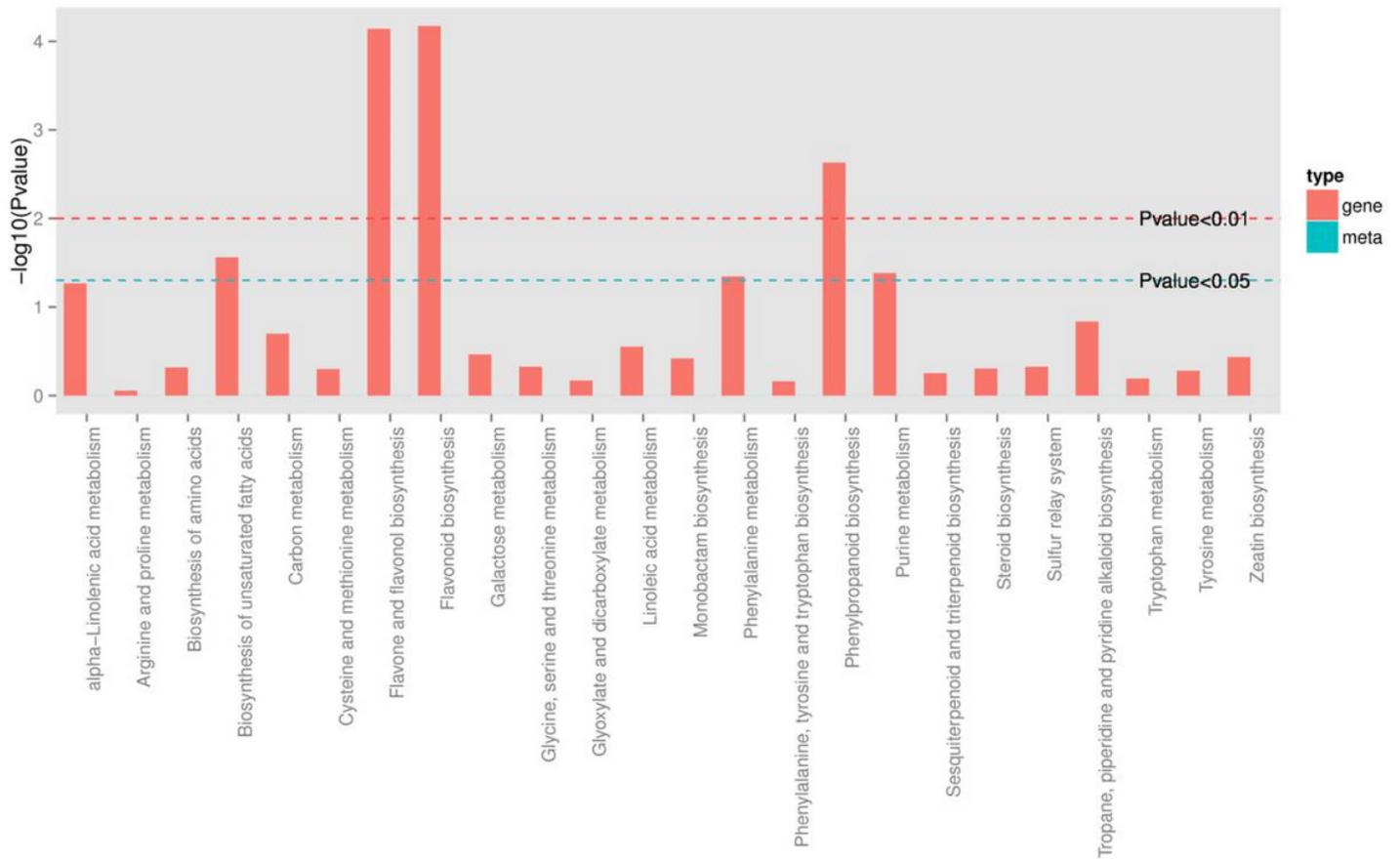


Figure 4

Commonly enriched KEGG pathways of differential metabolites and differential genes. Abscissa: All KEGG pathways enriched. Ordinate: Pvalue of DEMs and DETs.

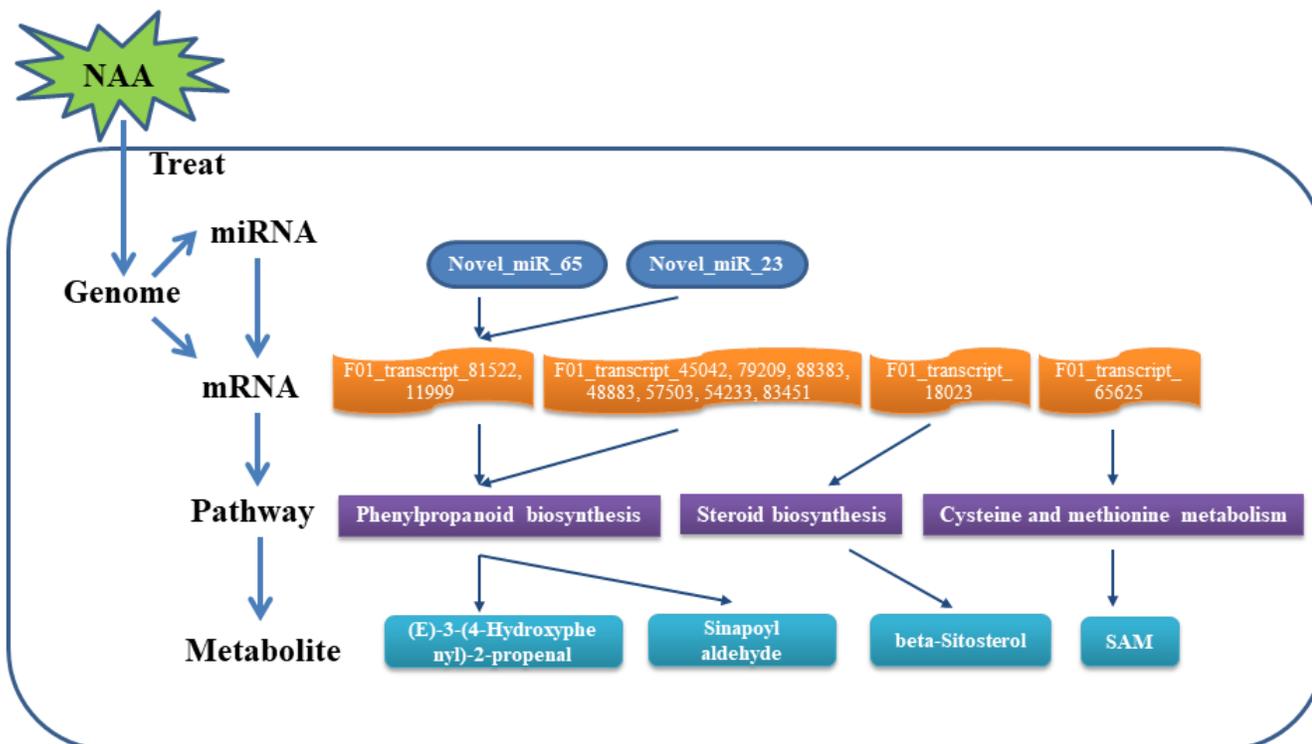


Figure 5

Schematic diagram of the mechanism by which NAA improves quality of *R. glutinosa* Libosh. A coexpression regulatory network was constructed based on the profiles of the differentially expressed miRNA-mRNA-metabolite pairs, suggesting that NAA could improve the quality of *R. glutinosa* through multiple mechanisms.

Supplementary Files

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

- [Supplementaryinformation.pdf](#)