Plant Growth Regulation (2020) 92:15–28
https://doi.org/10.1007/s10725-020-00616-1

ORIGINAL PAPER

Identification of biosynthetic pathways involved in flavonoid production in licorice by RNA-seq based transcriptome analysis

Ting Hu[1] · Zhi-Qiang Gao[1] · Jia-Ming Hou[1] · Shao-Kai Tian[1] · Zhi-Xin Zhang[1] · Lin Yang[1] · Ying Liu[1]

Received: 16 May 2019 / Accepted: 12 April 2020 / Published online: 20 April 2020

Abstract

Liquiritin, a flavonoid, is a key medicinal ingredient in licorice (*Glycyrrhiza uralensis*), a commonly used herb in traditional Chinese medicine. Biosynthesis of flavonoids is a complex process that involves not only the phenylpropanoid biosynthetic pathway but also many other secondary metabolic pathways. In this study, we tried to identify the key enzymes and pathways for the biosynthesis of flavonoids in *G. uralensis* by analyzing the gene expression patterns in samples containing different levels of flavonoid. *G. uralensis* seeds were mutagenized by X-ray irradiation and samples were selected based on HPLC analysis. RNA-seq was used to examine the gene expression in two samples with high flavonoid content (H1 and H2) and one control sample (L1) with low flavonoid content. 61.37 million, 54.21 million, and 54.22 million clean reads were obtained in sample H1, H2, and L1, respectively. A total of 1875 core differentially expressed genes (DEGs) were identified. The expression patterns of core DEGs were similar in samples H1 and H2 but not in sample L1. Flavonoid metabolic pathway, terpenoid biosynthetic pathway, plant hormone signal transduction pathway, plant circadian rhythm pathway, and starch and sucrose metabolic pathway were found to play significant roles for flavonoid biosynthesis in licorice. Ten co-expressed DEGs on the five metabolic pathways were further verified by qRT-PCR, which confirmed that the RNA-Seq results were accurate and reliable. This study provides a basis for future functional genes mining and molecular regulatory mechanism elucidation of flavonoid biosynthesis in licorice.

Keywords *Glycyrrhiza uralensis* · X-ray · Transcriptome · Flavonoids · Biosynthesis · Liquiritin

Abbreviations

BLAST	Basic local alignment search tool
CNB	Carbon nutrient balance hypothesis
DEGs	Differentially expressed genes
FC	Fold change
FDR	False discovery rate
FPKM	Fragments per kilobase of exon per million mapped reads
GDB	Growth differentiation balance hypothesis
GO	Gene ontology
KEGG	Kyoto encyclopedia of genes and genomes database
OD	Optimum defense hypothesis
qRT-PCR	Quantitative real-time PCR
Q30	Sequencing error rates lower than 0.1%
RA	Resource availability hypothesis
RNA-Seq	RNA-sequence
RIN	RNA integrity number

✉ Ying Liu
liuyliwd@sina.com

[1] School of Life Sciences, Beijing University of Chinese Medicine, Beijing 102488, China

Introduction

Three original plants, *Glycyrrhiza uralensis*, *Glycyrrhiza glabra*, and *Glycyrrhiza inflate*, are prescribed as licorice and widely used in traditional Chinese medicine (TCM) (*Chinese Pharmacopoeia* 2015). Modern pharmacological studies demonstrate that licorice possesses antitumor (Fukuchi et al. 2016; Rehan et al. 2013), anticancer (Lin et al. 2014), anti-inflammatory (Yang et al. 2017), anti-microbial (Wang et al. 2015a, b; Huang et al. 2012), immune-regulatory (Ma et al. 2013; Kim et al. 2013), liver protective (Seo et al. 2014), and neuroprotective activities (Chakravarthi and Avadhani 2014). In addition, it is also an important raw material for cosmetic brighteners, food additives, and tobacco flavoring agents (Hayashi et al. 2016).

Up to date, more than 300 flavonoids have been isolated from licorice (Wang et al. 2013). Licorice flavonoids possess

various pharmacological properties (He et al. 2017; Gao et al. 2017; Wang et al. 2015a, b; Luo et al. 2016; Park et al. 2015; Gong et al. 2015; Zhou and Ho 2014). Liquiritin, a flavonoid, is the marker component for evaluating the quality of licorice. According to *Chinese Pharmacopoeia*, the content of liquiritin in licorice must be at or above 0.5% to be effective. However, our previous investigation showed that about 60% cultivated licorice didn't meet this requirement. Therefore, how to improve the content of liquiritin in cultivated licorice has become a crucial issue.

In order to increase the content of natural active components, many researchers have focused on the radiation mutation breeding. It has become one of the most effective ways to obtain new germplasm resources in recent years (Tanaka et al. 2010). In our previous studies, we also found that X-ray irradiation improved the contents and yields of flavonoids in *G. uralensis* (Hu et al. 2017). However, the molecular mechanism of the flavonoids accumulation in licorice remains unclear so far.

The flavonoids in licorice are biosynthesized by the phenylpropanoid metabolic pathway which is controlled and regulated by many key enzymes (Fig. 1). Several functional genes have been successfully cloned and characterized, such as cinnamate 4-hydroxylase (C4H) (Li et al. 2016), isoflavone synthase (IFS) (Cheng et al. 2013), and chalcone synthase (CHS) (Awasthi et al. 2016). However, biosynthesis of flavonoid is a complex process, many genes critical for the biosynthesis remain to be identified.

With the development of gene sequencing technology, genomic and transcriptome analyses have attracted increasing attention to explore the molecular mechanism of secondary metabolism in medicinal plants (Han et al. 2016; Rhoads and Au 2015). A recent analysis of the transcriptome of *Carthamus tinctorius* has led to the identification of several key genes involved in the biosynthesis of flavonoids (Chen et al. 2018). In addition, transcriptome sequencing of *Scutellaria baicalensis* has uncovered 54 unigenes encoding 12 key enzymes involved in the biosynthetic pathway of flavonoids (Liu et al. 2015). Similarly, a transcriptome analysis of licorice has identified cytochrome P450 enzymes (CYP) and vacuolar saponin transporters involved in glycyrrhizin production (Ramilowski and Daub 2013).

In the present study, X-ray irradiated licorice seeds were cultivated for one year and two resulting plants with high flavonoid content were selected for transcriptome analysis and comparison with a control plant from untreated seeds. A metabolic network of flavonoids was established on the basis of the RNA-sequencing (RNA-Seq) results. The co-expressed differentially expressed genes (DEGs) involved in the metabolic network of flavonoids were determined and partly analyzed by qRT-PCR. This study provides a basis for functional genes mining and molecular regulatory mechanism elucidation of flavonoid biosynthesis in licorice.

Fig. 1 Liquiritin biosynthesis through phenylpropanoid metabolic pathway in *G. uralensis*. *PDH* prephenate dehydratase, *AAT* aromatic aminotransferase, *GT* glycosyltransferase, *PAL* phenylalanine ammonialyase, *C4H* cinnamic acid 4-hydroxylase, 4CL 4-coumarate CoA ligase, *ACC* acetyl CoA carboxylase, *CHS* chalcone synthase, *CHR* chalcone reductase, *I-CHI* I-chalcone isomerase, *II-CHI* II-chalcone isomerase

Materials and methods

Plant material and X-ray treatment

Healthy seeds of *G. uralensis* were irradiated by six gradient doses of X-rays, 5, 10, 15, 20, 30, and 50 Gy, and cultivated for one year in the herb garden at Beijing University of Chinese Medicine. Healthy seeds without irradiation were used as the blank control. The roots of all samples were collected for HPLC analysis of the contents of liquiritin, isoliquiritin, liquiritigenin, and isoliquiritigenin. Based on the HPLC results, two irradiated samples with high flavonoids contents and one blank sample with low flavonoids contents were selected for RNA-Seq analysis, which were showed in Table 1.

RNA extraction, cDNA library construction and RNA-Seq

Total RNA was extracted from the roots of licorice samples (Ding et al. 2010). RNA purity (OD 260/280) and concentration were detected by Nanodrop. RNA integrity was detected by Agilent 2100 Bioanalyzer. The cDNA library of each sample was constructed by an Illumina TruSeq™ RNA Sample Preparation Kit (Illumina, San Diego, CA, USA) and sequenced by Illumina HisSeq™ X10. Raw reads were obtained (oss://nextomics/FTP/BJXWZ-201707001D/) and filtered by NGS QCToolkit (version 2.3.3) to obtain high-quality clean reads (Patel and Mukesh 2012). All clean reads were mapped to the reference genome sequence (https://ngs-data-archive.psc.riken.jp/Gur-genome/download.pl) using Hisat (version 2.0.5) (Kim et al. 2015).

Gene annotation and functional enrichment

Gene annotation and functional enrichment were performed using Java Treeview (version 1.1.6), Kyoto Encyclopedia of Genes and Genomes (KEGG) database, and Gene ontology (GO) database (Kanehisa et al. 2008; Ye et al. 2006). The BLASTX algorithm was used to query the assembled sequences against GO and KEGG databases. The IDs of DEGs were submitted to the database for the enrichment analysis of GO biological process terms and KEGG pathway categories. The functional genes influencing the

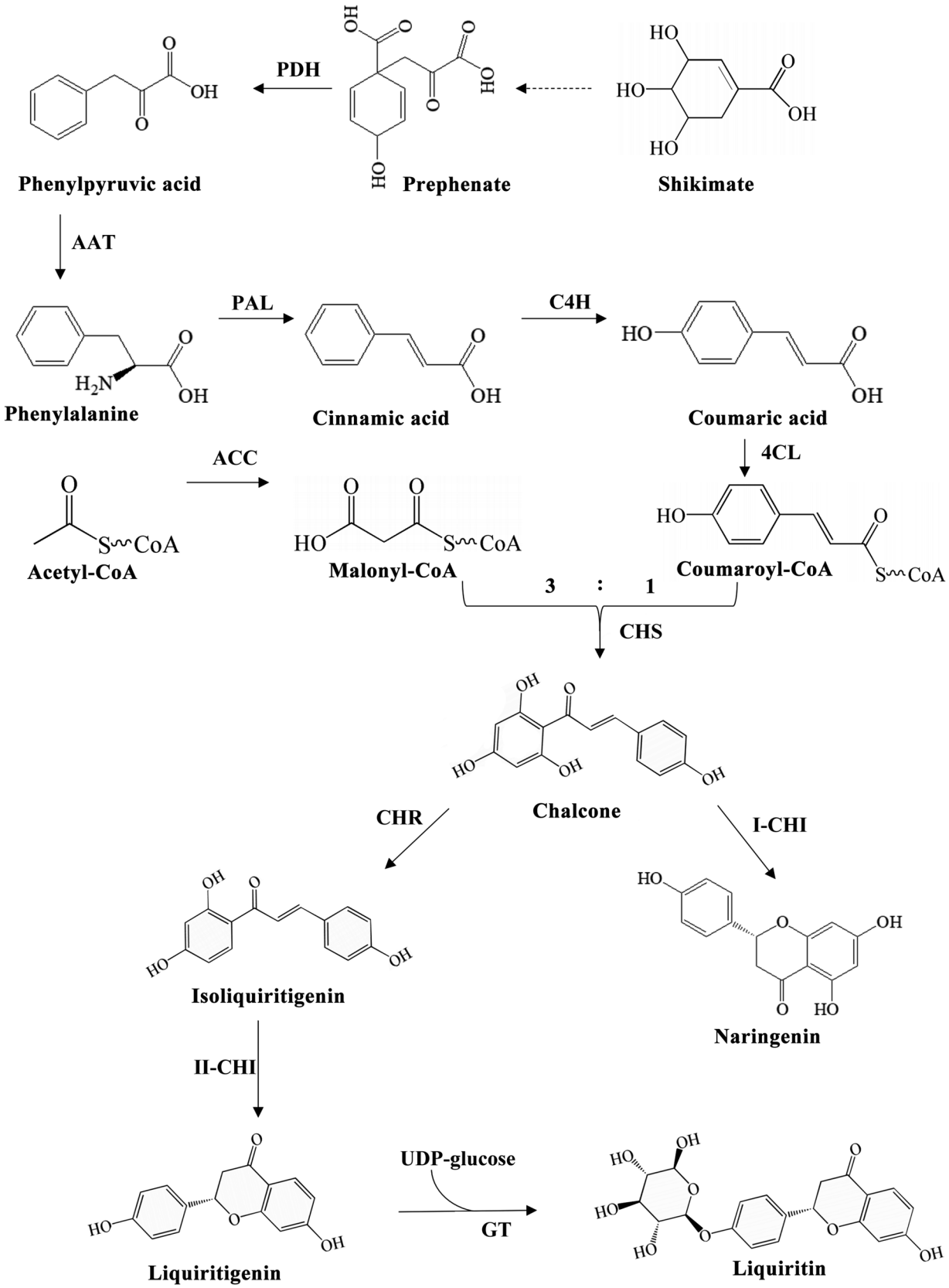
PDH
Phenylpyruvic acid
Prephenate
Shikimate
AAT
PAL
C4H
Phenylalanine
Cinnamic acid
Coumaric acid
4CL
ACC
Acetyl-CoA
Malonyl-CoA
Coumaroyl-CoA
3 : 1
CHS
Chalcone
CHR
I-CHI
Isoliquiritigenin
Naringenin
II-CHI
UDP-glucose
GT
Liquiritigenin
Liquiritin

Table 1 Information of the three *G. uralensis* samples for transcriptome analysis

Sample	X-ray (Gy)	Fresh weight (g)	Dry weight (g)	Content (mg g^{-1})				Yields (mg)			
				Liquiritin	Isoliquiritin	Liquiritigenin	Isoliquiritigenin	Liquiritin	Isoliquiritin	Liquiritigenin	Isoliquiritigenin
L1	0	1.53	0.52	0.0409	0.1050	0.0166	0.0367	0.0213	0.0546	0.0086	0.0191
H1	15	20.29	7.94	10.0178	0.9516	0.2406	0.0691	79.5414	7.5558	1.9106	0.5485
H2	50	42.36	19.47	11.7649	1.8897	0.6036	0.1467	229.0619	36.7929	11.7517	2.8562

Yields = content × dry weight

accumulation of flavonoids in *G. uralensis* were selected according to the gene annotation.

Gene expression quantification and differential expression analysis

The fragments per kilobase of exon per million mapped reads (FPKM) value was used to measure the gene expression level (Mortazavi et al. 2008). Taking the blank sample as a control, DEGs in irradiated samples, H1 and H2, were identified using EdgeR (Robinson et al. 2010). According to the Benjamini–Hochberg method, the significant differences in gene expression were represented by the false discovery rate (FDR). Fold change (FC) indicates the ratio of the gene expression level between two samples. FDR < 0.01 and $\left|\log_2^{(FC)} \geq 1\right|$ were set as the thresholds for gene differential expression. According to the gene expression level, the DEGs were classified into two relative groups, up-regulated and down-regulated genes. The DEGs simultaneously up-regulated or down-regulated in both samples H1 and H2 were co-expressed DEGs. Using the website (https://www.genome.jp/kegg/tool/map_pathway2.html), co-expressed DEGs were mapped to the different pathways. Then, the pathways enriched with DEGs, the biosynthesis map map01060, and various hypotheses about the biosynthetic mechanism of secondary metabolites, including growth differentiation balance hypothesis (GDB), optimum defense hypothesis (OD), carbon nutrient balance hypothesis (CNB), and resource availability hypothesis (RA) (Huang et al. 2010), were combined to construct a metabolic network of flavonoids in licorice.

Relative expression analysis of co-expressed DEGs

To verify the RNA-Seq results, qRT-PCR was performed with the SYBR® Green qPCR Master Mix (High ROX) as the fluorescent dye. The β-actin was used as the internal control gene. All primers of co-expressed DEGs were designed by the Primer Premier 5.0 (Table 2). The qRT-PCR was performed on a Light Cycler480 II (Roche, Switzerland). Gene expression level was calculated by the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen 2001). Correlation of gene expression between RNA-Seq and qRT-PCR data was analyzed by Pearson-Test.

Results

RNA-Seq sequencing analysis

The RNA integrity number (RIN) of three licorice samples was 8.1, 9.6, and 8.3, respectively, which met the

Table 2 Primer sequences for qRT-PCR analysis

Gene ID	Reference annotation	Gene name	Forward primer sequence (5′–3′)	Reverse primer sequence (5′–3′)
		Actin	CAAAAGGATGCCTATGTGGG	CAGGAGCAACACGCAATTC
Glyur000106s00011717	Phenylalaninammo-nialyase	*PAL*	AAGTGCTTGAATTTGCCTCCT	TTTGCCTACATTGATGACCCT
Glyur000424s00026890	Chalcone synthase	*CHS1*	CTCGTGTTCTGTACCACCTCTG	GTTCTCGGCGATGTCCTTT
Glyur000278s00017280	9-cis-Epoxycarotenoid dioxy-genase	*NCED*	GCGTGTTGTCGGAGATAAGG	TTCACCTCCCCACTCATCAG
Glyur000261s00014360	Gibberellin 2-oxidase	*GA2ox*	TTGGTGTTGAGGAGGAGGTACT	GGCTTGCCTAGAGCTTGGTT
Glyur000231s00022061	1-Deoxy-D-xylulose-5-phosphate synthase	*DXS*	GTTTCCTTCCTCTGTTTCATTCTC	TGTTTCGTAGCGTTTCTCACC
Glyur000158s00011331	Gibberellin receptor	*GID1*	AGGGTGGTTACCGTGAGGAT	ATTACTGCCAGCCATGATGTCT
Glyur002299s00036262	Jasmonate-zim-domain protein 5	*JAZ*	AGTCTGTGAACAAGGGTCCTAAAG	GGGAATGAAGGCTGGCTCT
Glyur000017s00002448	SAUR-like auxin-responsive protein	*SAUR*	GGAGAACGAAGGCACGAAT	CGAGTGGTCCATGGTTACAGA
Glyur000116s00009244	MYB-related transcription factor LHY	*LHY*	TCATTTCGTTGGAATCAGGG	GACAGGGCAAGGAGATATTACACT
Glyur000047s00004005	Beta-amylase	*AMYB*	TAATCGGAGTAGACCTGAAGGG	CGTGGGTGCTGGAAGAAAT

Table 3 Statistics of RNA-seq analysis

Samples	Q30 percentage (%)	Clean reads	Mapped reads	Mapping rate (%)
H1	92.50	61,367,536	54,589,750	88.96
H2	91.50	54,209,318	49,312,496	90.97
L1	91.00	54,223,512	49,264,241	90.85

requirements of cDNA library. The sequencing error rates of more than 91% clean reads were less than 0.1% (Q30) (Table 3). After the data filtering, 61.37 million, 54.21 million, and 54.22 million clean reads were obtained in samples H1, H2, and L1, respectively. More than 88% of clean reads in each library was mapped to the reference genome. Therefore, the transcriptome data of the three samples were obtained with a high correct rate and good genomic coverage.

Gene annotation and pathway enrichment analysis

A total of 16,006 unigenes were annotated to the GO database and classified into three principal categories, including cellular component, molecular function, and biological process, which were further subdivided into 39 categories. The percentage of genes with the function of "binding" was the largest, followed by "catalytic activity" and "metabolic process" (Fig. 2a). A total of 9693 unigenes were annotated to the KEGG database and categorized into 208 KEGG pathways. The percentage of genes with the function of "signal transduction" was the largest, followed by "carbohydrate metabolism" and "translation" (Fig. 2b).

Gene expression analysis and identification of DEGs

3386 up-regulated DEGs and 1141 down-regulated DEGs were identified in the group "H1 vs L1", and 1995 up-regulated DEGs and 2235 down-regulated DEGs in the group "H2 vs L1" (Fig. 3a). A total of 1875 core DEGs was obtained in both group "H1 vs L1" and "H2 vs L1" (Fig. 3b), which was closely related to the accumulation of flavonoids in licorice. The expression patterns of the core DEGs in samples H1 and H2 were similar, but were in contrast with sample L1 (Fig. 3c). In comparison with L1, many genes in H1 and H2 were down-regulated, such as peroxidase gene, beta-glucosidase gene, phenylalanine ammonia-lyase gene, and coniferyl-aldehyde dehydrogenase gene.

Functional enrichment analysis of DEGs

The major GO enrichment terms of the 1875 core DEGs were shown in Fig. 4, including "oxidoreductase activity", "transmembrane transporter activity", and

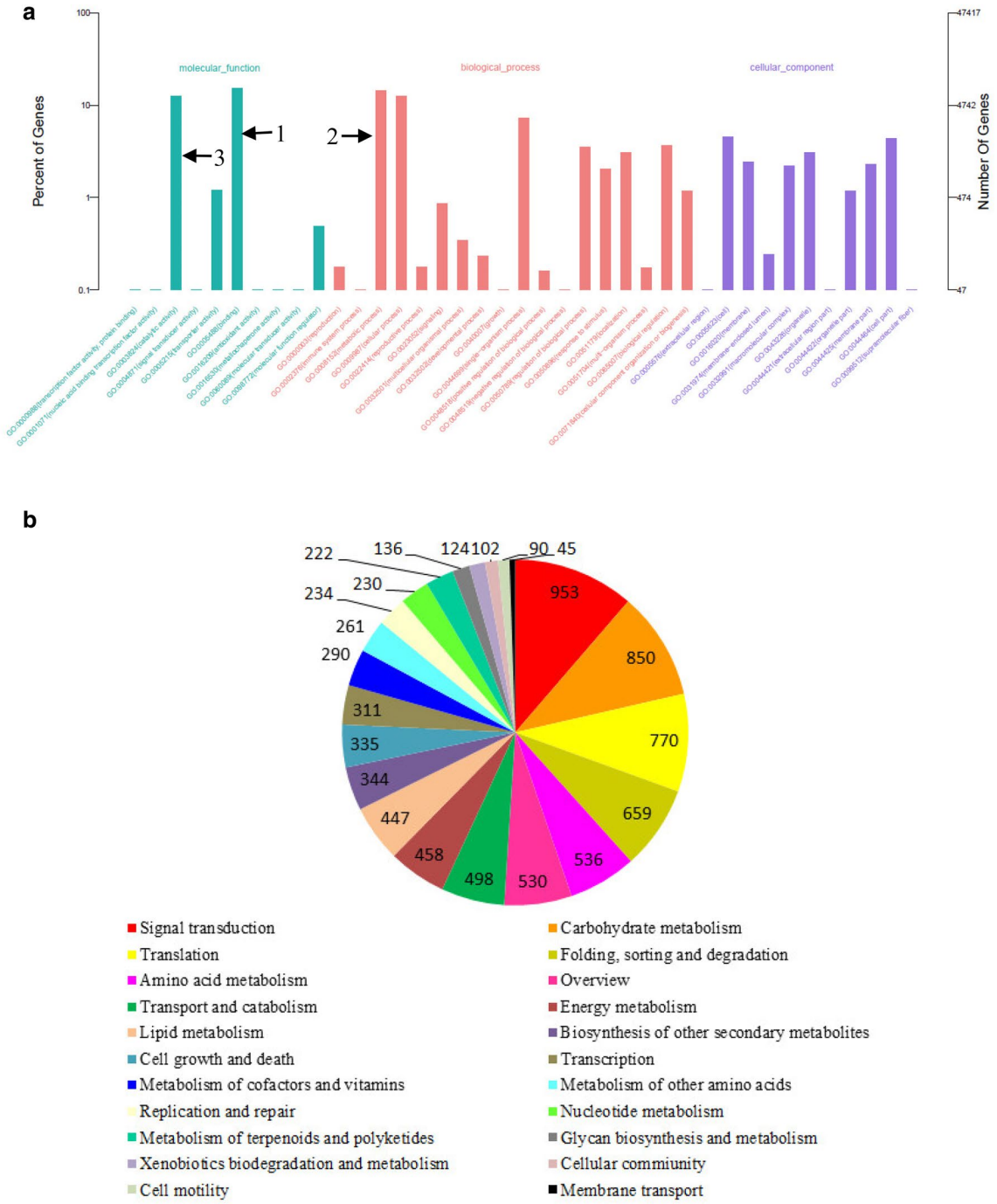

Fig. 2 Unigenes function classification in *G. uralensis*. **a** GO classification of unigenes in *G. uralensis.* Green shows the GO categories of molecular function. Red shows the GO categories of biological process. Purple shows the GO categories of cellular component. Stripes **1**, **2**, **3** show the percentage of genes with the function of protein binding, metabolic process and catalytic activity, respectively. **b** KEGG classification of unigenes in *G. uralensis.* The number of unigenes is marked in pie chart. (Color figure online)

"carbohydrate metabolic process". The KEGG pathway significantly enriched in 11 biosynthetic pathways, including "starch and sucrose metabolism (map00500)", "terpenoid backbone biosynthesis (map00900)", "diterpenoid biosynthesis (map00904)", "carotenoid biosynthesis (map00906)", "sesquiterpenoid and triterpenoid biosynthesis (map00909)", "phenylpropanoid biosynthesis (map00940)", "flavonoid biosynthesis (map00941)", "flavone and flavonol biosynthesis (map00944)", "plant hormone signal transduction (map04075)", "isoflavonoid

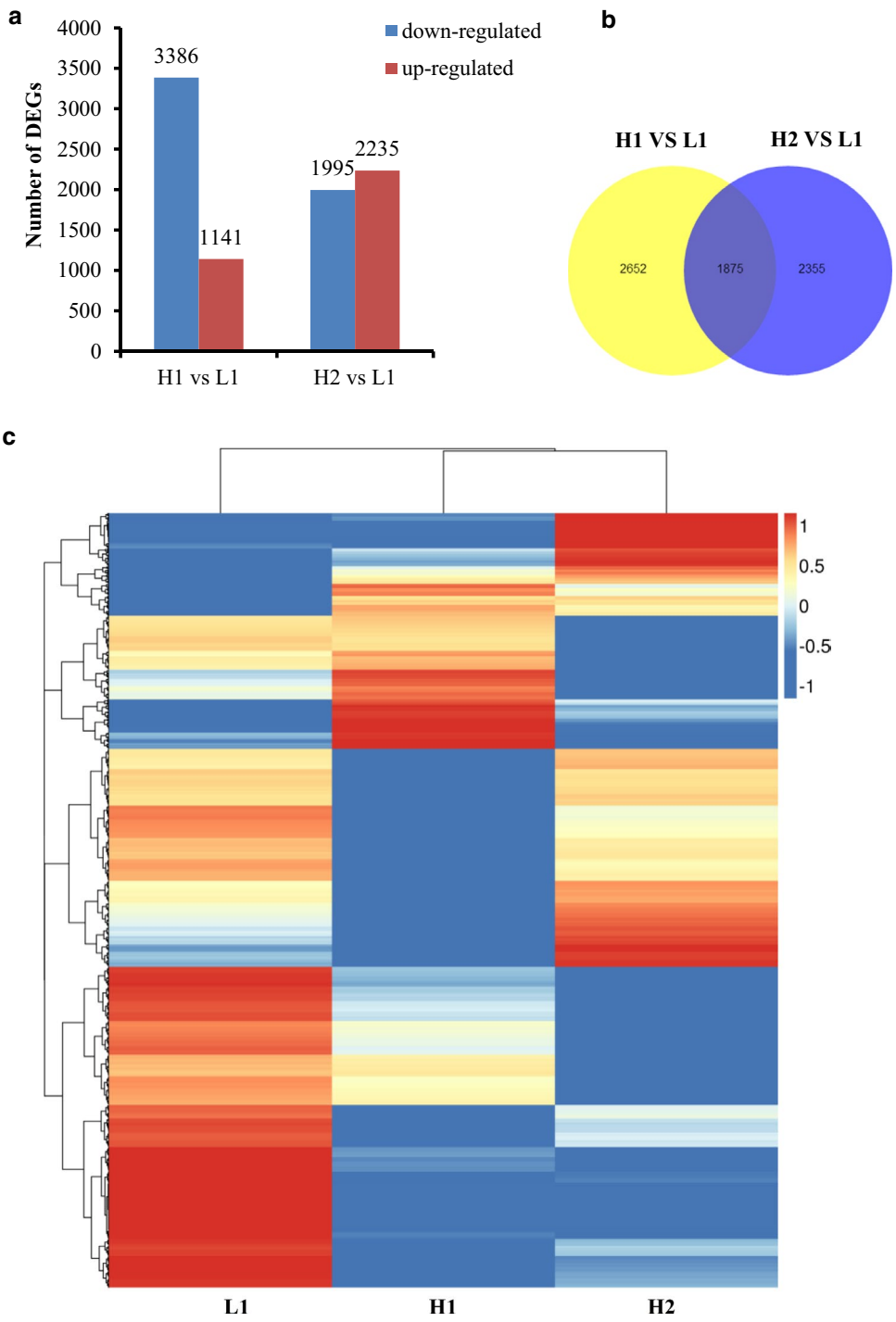

Fig. 3 Expression profiling of core DEGs in the three samples. **a** The numbers of up-regulated and down-regulated DEGs in samples H1 and H2 compared with L1. **b** The core DEGs between H1 and L1, and H2 and L1. **c** Hierarchical clustering graph of the total 1875 core DEGs

biosynthesis (map00943)", and "circadian rhythm-plant (map04712)". A metabolic network of flavonoids in licorice was constructed as shown in Fig. 5. Five KEGG pathways are closely related to the biosynthesis of flavonoids, including the flavonoid metabolic pathway, plant hormone signal transduction pathway, terpenoid biosynthetic pathway, plant circadian rhythm pathway, and starch and sucrose metabolic pathway.

Expression analysis of DEGs in the flavonoid metabolic pathways

Twenty-three DEGs were obtained on the flavonoid metabolic pathway (Table 4), one was up-regulated and

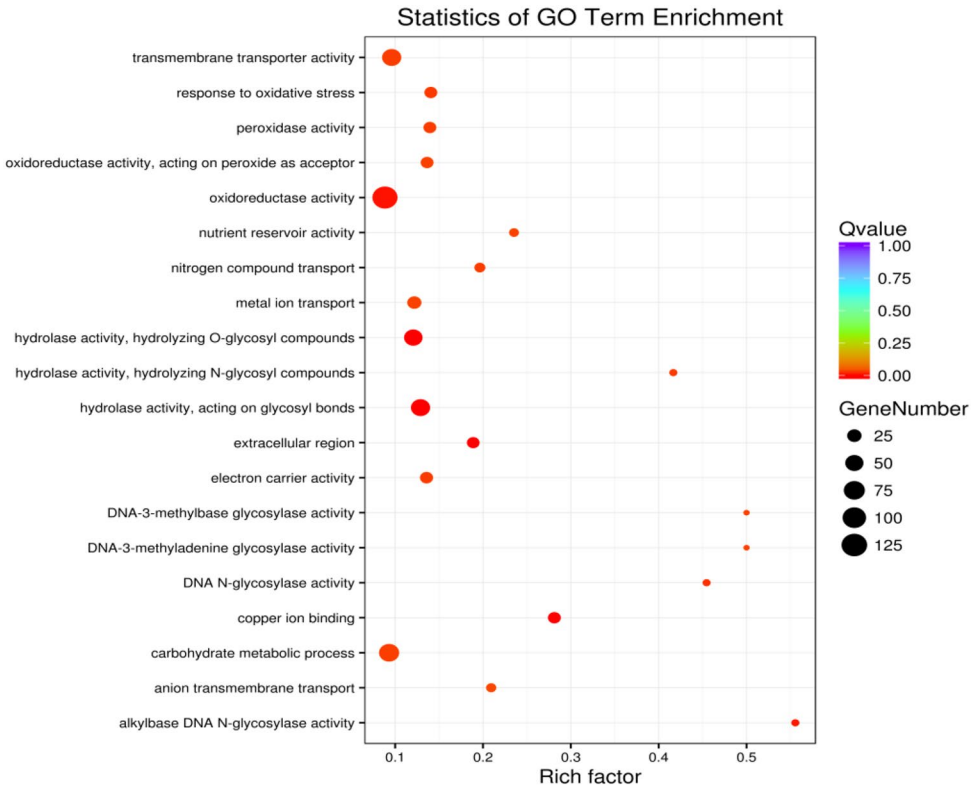

Fig. 4 GO enrichment scatter plot of the total 1875 core DEGs

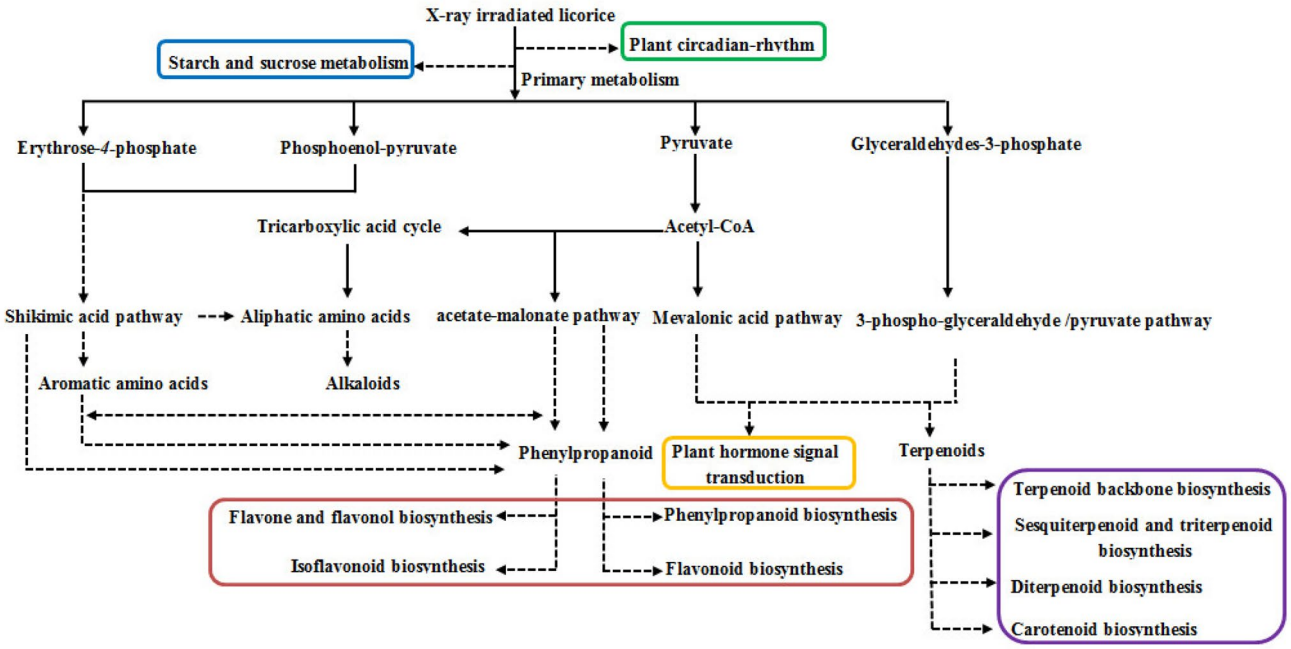

Fig. 5 Network diagram of the flavonoid metabolic pathway and secondary metabolites in licorice. Blue box marks the starch and sucrose metabolic pathway. Green box marks the plant circadian rhythm pathway. Red box marks the flavonoid metabolic pathway. Orange box marks the plant hormone signal transduction pathway. Purple box marks the terpenoid biosynthetic pathway. (Color figure online)

Table 4 FPKM value of DEGs on the five metabolic pathways

Flavonoid metabolic pathway				FPKM			
Gene ID	Name	Abbreviation	Enzyme	H1	H2	L1	Trend
Glyur000120s00006085	Peroxidase	POD	EC:1.11.1.7	0.0969	0.0000	6.7602	Down
Glyur000169s00018357	Feruloyl-CoA ortho-hydroxylase	FCH	EC:1.14.11.–	5.1148	396.0742	105.9758	
Glyur000084s00009470	Peroxidase	POD	EC:1.11.1.7	0.1850	0.4357	12.0702	Down
Glyur002512s00038136	Peroxidase	POD	EC:1.11.1.7	0.0000	0.0000	2.8910	Down
Glyur000244s00014794	Peroxidase	POD	EC:1.11.1.7	0.0000	1.5314	13.6156	Down
Glyur000615s00024015	Peroxidase	POD	EC:1.11.1.7	0.0000	37.0109	10.0119	
Glyur001100s00028602	Beta-glucosidase	GLU	EC:3.2.1.21	0.0518	0.1507	3.0083	Down
Glyur000106s00011717	Phenylalanine ammonia-lyase	PAL	EC:4.3.1.24	1.0363	1.8349	19.9917	Down
Glyur000951s00029021	Peroxidase	POD	EC:1.11.1.7	0.0000	0.1708	3.6665	Down
Glyur000702s00024650	Peroxidase	POD	EC:1.11.1.7	0.0000	0.0000	0.5134	Down
Glyur004683s00047277	Peroxidase	POD	EC:1.11.1.7	0.0000	0.2429	7.9331	Down
Glyur000350s00020051	Coniferyl-aldehyde dehydrogenase	CADH	EC:1.2.1.68	0.0518	0.0703	1.6217	Down
Glyur000097s00008307	Peroxidase	POD	EC:1.11.1.7	0.0248	0.0727	5.2786	Down
Glyur002942s00044115	Peroxidase	POD	EC:1.11.1.7	0.0000	0.1164	2.8671	Down
Glyur007502s00042768	Coniferyl-aldehyde dehydrogenase	CADH	EC:1.2.1.68	0.0353	0.0474	1.0221	Down
Glyur002918s00034906	Peroxidase	POD	EC:1.11.1.7	0.3257	0.0381	16.5382	Down
Glyur002341s00037864	Peroxidase	POD	EC:1.11.1.7	0.0000	0.3444	5.0457	Down
Glyur000058s00006169	Peroxidase	POD	EC:1.11.1.7	0.0532	0.1164	2.6090	Down
Glyur000327s00026461	Phenylalanine ammonia-lyase	PAL	EC:4.3.1.24	2.5772	63.1369	22.6671	
Glyur000006s00001663	Peroxidase	POD	EC:1.11.1.7	0.0000	0.1450	53.8290	Down
Glyur000163s00011138	Phenylalanine ammonia-lyase	PAL	EC:4.3.1.24	0.1937	47.9303	2.6548	
Glyur000257s00022018	Peroxidase	POD	EC:1.11.1.7	0.0234	0.0000	1.1591	Down
Glyur000424s00026890	Chalcone synthase	CHS1	EC:2.3.1.74	288.2868	734.3510	68.7575	Up
Terpenoid biosynthetic pathway				FPKM			
Gene ID	Name	Abbreviation	Enzyme	H1	H2	L1	Trend
Glyur000607s00021618	Gibberellin 3-beta-dioxygenase	GA3OX1	EC:1.14.11.15	6.0104	1.6766	0.0744	Up
Glyur006124s00045728	Beta-amyrin 24-hydroxylase	AH	EC:1.14.99.43	0.0000	41.7190	14.3325	
Glyur006124s00045727	Beta-amyrin 24-hydroxylase	AH	EC:1.14.99.43	0.0000	151.9030	23.1999	
Glyur000231s00022061	1-Deoxy-D-xylulose-5-phosphate synthase	DXS	EC:2.2.1.7	1.0633	15.0166	0.1633	Up
Glyur001733s00027628	Lupeol synthase 2	LUP2	EC:5.4.99.39	1.0245	97.1168	29.3650	
Glyur000278s00017280	9-cis-Epoxycarotenoid dioxygenase	NCED	EC:1.13.11.51	43.5519	13.0738	0.6141	Up
Glyur000261s00014360	Gibberellin 2-oxidase	GA2OX	EC:1.14.11.13	42.4023	4.8963	0.3707	Up
Glyur000844s00023272	Capsanthin/capsorubin synthase	CS1	EC:5.3.99.8	0.3243	0.0999	4.6769	Down
Glyur000882s00021659	Gibberellin 3-beta-dioxygenase	GA3OX1	EC:1.14.11.15	20.4821	9.5546	2.7838	Up
Starch sucrose metabolic pathway				FPKM			
Gene ID	Name	Abbreviation	Enzyme	H1	H2	L1	Trend
Glyur000003s00001132	Sucrose synthase	SS	EC:2.4.1.13	0.0185	0.1945	4.8183	Down
Glyur000047s00004005	Beta-amylase	β-AL	EC:3.2.1.2	1301.5789	286.1003	6.5242	Up
Glyur000075s00007156	1,4-Alpha-glucan branching enzyme	GBE	EC:2.4.1.18	463.1354	378.2404	130.3577	Up
Glyur000164s00014492	Alpha-amylase	α-AL	EC:3.2.1.1	0.4069	0.2846	18.0329	Down
Plant hormone signal transduction pathway				FPKM			
Gene ID	Name	Abbreviation	Enzyme	H1	H2	L1	Trend
Glyur002113s00042685	Auxin-responsive GH3 family protein	GH3		0.2510	0.8678	19.0283	Down
Glyur000012s00000534	Jasmonate-zim-domain protein 5	JAZ		1.2578	56.7972	13.4666	
Glyur000013s00003187	Response regulator 4	RR4		7.4824	2.7850	64.5557	Down

Table 4 (continued)

Plant hormone signal transduction pathway				FPKM			
Gene ID	Name	Abbreviation	Enzyme	H1	H2	L1	Trend
Glyur000017s00002448	SAUR-like auxin-responsive protein family	SAUR		6.7240	3.9876	293.6897	Down
Glyur000017s00002449	SAUR-like auxin-responsive protein family	SAUR		0.0000	0.0785	16.8546	Down
Glyur000061s00006533	Indoleacetic acid-induced protein 10	IAA		0.5928	6.5050	63.4823	Down
Glyur000069s00004052	Histidine-containing phosphotransfer factor 5	AHP5		0.0000	0.0000	1.4751	Down
Glyur000086s00011528	Abscisic acid receptor PYR	PYR		926.2463	14.0138	2.6005	Up
Glyur000101s00010319	SAUR-like auxin-responsive protein family	SAUR		0.0854	0.0361	1.8945	Down
Glyur000140s00011477	SAUR-like auxin-responsive protein family	SAUR		0.0000	0.0000	13.5484	Down
Glyur000156s00011998	Protein transport inhibitor response 1	TIR1		6.4809	6.8081	1.7057	Up
Glyur000158s00011331	Gibberellin receptor GID1	GID1		57.7521	83.9343	15.2811	Up
Glyur000357s00018138	Indoleacetic acid-induced protein 10	IAA10		3.2700	15.5026	163.3004	Down
Glyur001068s00035968	SAUR-like auxin-responsive protein family	SAUR		0.0000	0.0000	1.6627	Down
Glyur001825s00027392	indoleacetic acid-induced protein 10	IAA10		0.2642	0.6174	5.9161	Down
Glyur002299s00036262	Jasmonate-zim-domain protein 5	JAZ		130.9788	335.2809	36.9156	Up
Glyur003343s00045195	Transcription factor MYC2	MYC2		2.5779	173.9687	20.3530	
Glyur007908s00046209	Regulatory component of ABA receptor 1	PYR		2.7065	84.8236	19.0443	
Plant circadian rhythm pathway				FPKM			
Gene ID	Name	Abbreviation	Enzyme	H1	H2	L1	Trend
Glyur000116s00009253	Pseudo-response regulator 5	PRR5		190.3406	54.5714	10.0149	Up
Glyur000015s00000787	Pseudo-response regulator 5	PRR5		15.0716	6.3635	0.9138	Up
Glyur000072s00008110	Pseudo-response regulator 5	PRR5		25.0265	18.0544	3.9343	Up
Glyur001652s00038197	Protein flowering locus T	FT		1.5170	2.0001	0.0000	Up
Glyur002560s00044392	Protein flowering locus T	FT		26.3535	32.2115	0.0000	Up
Glyur000116s00009244	MYB-related transcription factor LHY	LHY		64.0074	15.8056	3.9804	Up
Glyur000424s00026890	Chalcone synthase	CHS1		288.2868	734.3510	68.7575	Up

Red-labeled genes were verified by qRT-PCR

eighteen were down-regulated in both samples H1 and H2, while the left four genes were up-regulated in sample H1 but down-regulated in sample H2. Eighteen DEGs were obtained on the plant hormone signal transduction pathway, four were up-regulated and eleven were down-regulated in both samples H1 and H2, while the left three genes were up-regulated in sample H1 but down-regulated in sample H2. Nine DEGs were obtained on the terpenoid biosynthetic pathway, five were up-regulated and one was down-regulated in both samples H1 and H2, while the left three genes were up-regulated in sample H1 but down-regulated in sample H2. Seven up-regulated DEGs were obtained on the plant circadian rhythm pathway in both samples H1 and H2. Four DEGs were obtained on the starch sucrose metabolic pathway, two were up-regulated and two were down-regulated in both samples H1 and H2. The further analysis of the above DEGs on the five metabolic pathways are showed in "Discussion".

Verification of gene expression by qRT-PCR

The gene expression level of ten co-expressed DEGs, *PAL* (Glyur000106s00011717), *CHS1* (Glyur000424s00026890), *NCED* (Glyur000278s00017280), *GA2ox* (Glyur000261s00014360), *DXS* (Glyur000231s00022061), *GID1* (Glyur000158s00011331), *JAZ* (Glyur002299s00036262), *SAUR* (Glyur000017s00002448), *LHY* (Glyur000116s00009244), and *AMYB* (Glyur000047s00004005), were verified by qRT-PCR. The gene expression levels of the ten co-expressed DEGs were measured in RNA-Seq (Fig. 6a). The correlation coefficient between the $\log_2^{FRKM}$ value and $\log_2^{average(2^{-\Delta\Delta CT})}$ value was 0.6324 ($P < 0.0001$) (Fig. 6b). Except *DXS*, the expression level of the other nine genes was consistent with the RNA-Seq results (Fig. 6c). The expression levels of *NCED*, *GA2ox*, *LHY*, and *AMYB* were highest in sample H1, and those of *CHS1* and *JAZ* were highest in sample H2 (Fig. 6c).

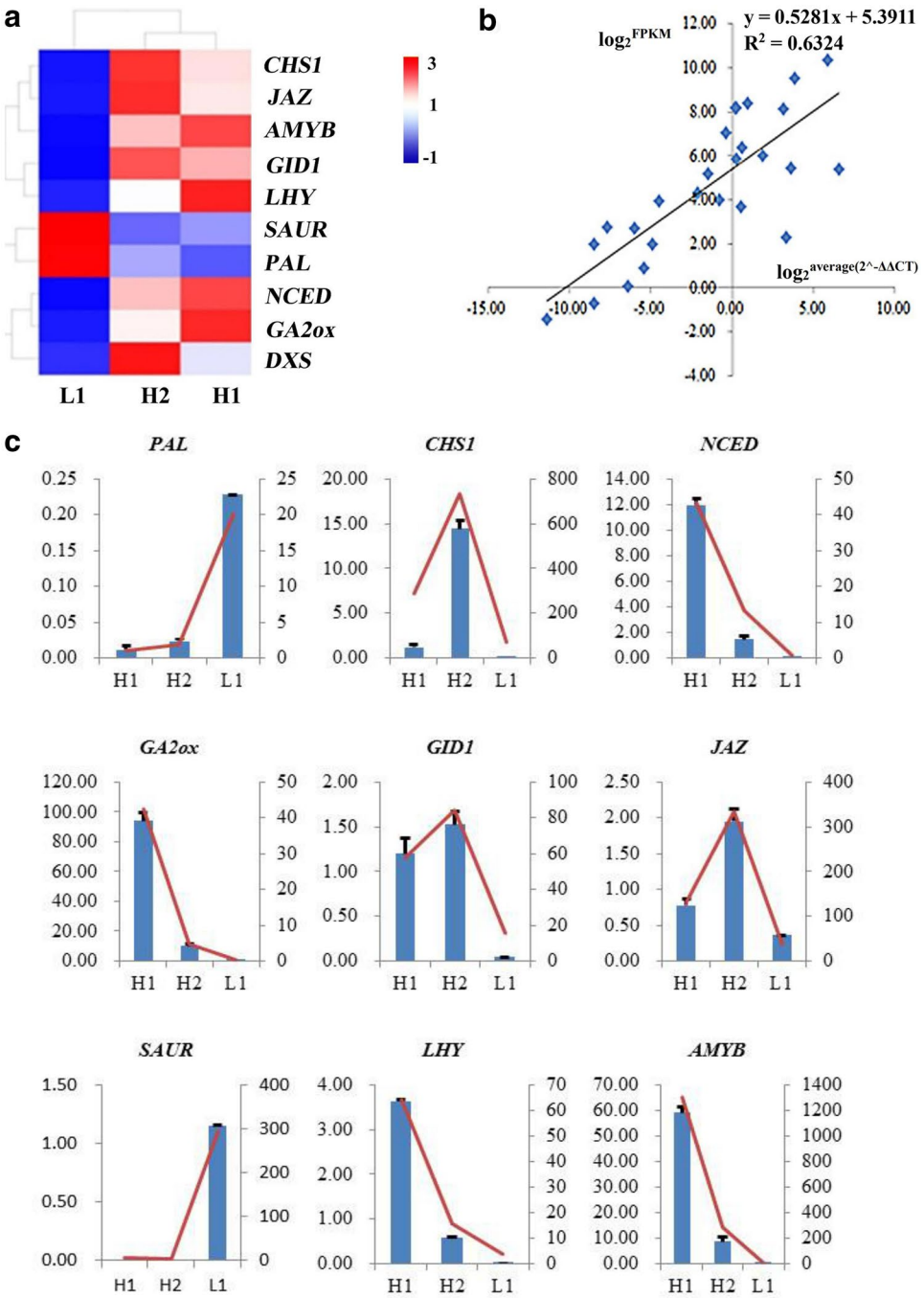

Fig. 6 Gene expression levels of the co-expressed DEGs in RNA-Seq and qRT-PCR. **a** Gene expression levels of the ten co-expressed DEGs measured in RNA-Seq. **b** Correlation scatter plot between $\log_2^{\text{average}(2^{-\Delta\Delta CT})}$ and $\log_2^{\text{FPKM}}$, which shows the relationship between RNA-Seq and qRT-PCR. **c** Gene expression levels of the nine co-expressed DEGs in qRT-PCR. The blue columns corresponding to the ordinate axis on the left show the expression level of co-expressed DEGs in qRT-PCR. The red polyline corresponding to the ordinate axis on the right show the expression of co-expressed DEGs in RNA-Seq. (Color figure online)

Discussion

In this study, a total of 1875 core DEGs involved in the secondary metabolic pathways were obtained in three licorice samples by RNA-Seq analysis, and a metabolic network for the biosynthesis of flavonoids was established. Five metabolic pathways were identified that play important roles for flavonoid accumulation in licorice and analyzed as follows.

The flavonoid metabolic pathway: the up-regulated expression of chalcone synthase gene (*CHS*) in samples H1 and H2 is expected to influence the biosynthesis of flavonoids (Wang et al. 2018). The down-regulated expression of beta-glucosidase gene (*GLU*), coniferyl-aldehyde dehydrogenase gene (*CADH*), and peroxidase gene (*POD*) may result in attenuation of the downstream alternative pathways, and hence shunt biosynthetic substrates towards the flavonoid synthesis pathways.

The terpenoid biosynthetic pathway: two up-regulated genes, gibberellin 2-oxidase gene (*GA2ox*) and gibberellin 3-β-dioxygenase gene (*GA3ox*), are involved in

photosynthesis and plant primary metabolism (Zhou et al. 2011). The up-regulated expression of 1-deoxy-D-xylulose-5-phosphate synthase gene (*DXS*) is expected to promote the biosynthesis of monoterpene, diterpene and carotenoids. The down-regulated expression of capsanthin synthase gene (*CS1*) may favor biosynthesis of carotenoid and promote accumulation of flavonoids (Zhou et al. 2017).

The starch sucrose metabolic pathway: the up-regulated genes, the 1,4-α-glucan branching enzyme gene (*GBE*) and β-amylase gene (*AMYB*), are involved in conversion of starch to dextrin and maltose. The down-regulated expression of the α-amylase gene and sucrose synthase gene may lead to inhibition of sucrose synthesis, which may be related to the increased biomass of *G. uralensis* observed in our previous studies (Hu et al. 2017).

The plant hormone signal transduction pathway: the up-regulated expression of the protein transport inhibitor response 1 gene (*TIR1*) and gibberellin receptor gene (*GID1*) is expected to promote the protein ubiquitination (Maraschin et al. 2010) and diterpene biosynthesis. The up-regulated expression of the jasmonate methyl-domain protein 5 gene (*JAZ*) is expected to enhance stem maturation and stress resistance. The down-regulated expression of the SAUR-like auxin-responsive protein gene (*SAUR*), auxin-reactive GH3 family protein gene (*GH3*), and indoleacetic acid-induced protein 10 gene (*IAA10*) indicates a down-regulation of auxin metabolism (Luo et al. 2018), while the down-regulated expression of the histidine-containing phosphotransfer factor 5 gene (*AHP*) and response regulator 4 gene (*RR4*) suggests a down-regulation of cytokinin metabolism (Verma et al. 2015). The down-regulation of auxin and cytokinin metabolism may cause the slow growth rate of the two samples in the premature state, which, according to the growth-differentiation balance hypothesis, the optimum defense hypothesis, and the resource availability hypothesis, leads to the accumulation of secondary metabolites and promotes the biosynthesis of flavonoids.

The plant circadian rhythm pathway: the up-regulated expression of the chalcone synthase gene (*CHS*) and pseudo-responsive regulator gene 5 (*PRR5*) may protect licorice against the injury caused by the X-ray irradiation to increase the biomass of *G. uralensis*. The up-regulated expression of the flowering locus T gene (*FT*) and MYB-related transcription factor gene (*LHY*) may influence the flowering period and anthocyanid production, resulting in an increased accumulation of flavonoids.

Many genes were identified critical for biosynthesis of bioactive components in medicinal plants in the past several years (Wei et al. 2015; Wang et al. 2016). In this study, sixty-one genes (thirty different kinds) involved in five pathways were mined (Table 4) that play an important role for flavonoids biosynthesis in licorice. Among them, *AH*, *DXS*, *LUP*, *CHS*, and *SQS* have already been identified from licorice, and *POD*, *FCH*, *GLU*, *PAL*, *CADH*, *GA3OX1*, *NCED*, *CS1*, *β-AL*, *GBE*, *α-AL*, *GH3*, *JAZ*, *RR4*, *SAUR*, *PYR*, *TIR1*, *GID1*, *MYC2*, *FT*, and *LHY* have been identified from other Leguminosae plants. However, for most of these genes the roles of them in flavonoids production in high plants were unclear. Among the ten genes we selected for qRT-PCR analysis, *CHS* has been reported to be related to a high level of flavonoid accumulation (Wang et al. 2018). In our current studies we also find that *CHS* and *DXS* are able to regulate flavonoid biosynthesis. These studies provide important insights regarding the role of the identified genes in production of various bioactive components.

With the development of 2nd generation sequence technology, the genomic and transcriptome analyses have become powerful tools to analyze biosynthetic pathways in plants (Han et al. 2016; Rhoads and Au 2015). For example, to clarify the mechanism of gibberellin-regulated flowering in *Jatropha curcas*, a lot of genes involved in gibberellin metabolism and signaling pathways were identified by the genomic and transcriptome analyses (Gao et al. 2015). The present study utilizes a genowide based analysis to uncover changes in transcription associated with high content of flavonoids. The results show that five pathways are involved in the biosynthesis of flavonoids in licorice. Changes in the five pathways are able to affect production of flavonoids, which make them potential targets for genetic manipulation in order to increase flavonoid production in licorice. Our transcriptome study will provide a frame-work for further determining the role of each identified gene in the process.

Acknowledgements This work was supported by the National Natural Science Foundation of China (Grant No. 81503181). We thank Dr. Yu Jiang at University of Pittsburgh for critical reading the manuscript.

References

Awasthi P, Mahajan V, Jamwal VL (2016) Cloning and expression analysis of chalcone synthase gene from *coleus forskohlii*. J Genet 95(3):647–657

Chakravarthi KK, Avadhani R (2014) Enhancement of hippocampal CA3 neuronal dendritic arborization by *glycyrrhiza glabra* root extract treatment in wistar albino rats. J Nat Sci Biol Med 5(1):25–29

Chen J, Tang XH, Ren CX et al (2018) Full-length transcriptome sequences and the identification of putative genes for flavonoid biosynthesis in safflower. BMC Genom 19(1):548. https://doi.org/10.1186/s12864-018-4946-9

Cheng H, Wang J, Chu S et al (2013) Diversifying selection on flavanone 3-hydroxylase and isoflavone synthase genes in cultivated soybean and its wild progenitors. PLoS ONE 8(1):e54154. https://doi.org/10.1371/journal.pone.0054154

Chinese Pharmacopoeia Commission (2015) Pharmacopoeia of the People's Republic of China. Part 1. China Medical Science Press, Beijing87

Ding Y, Liu XZ, Li YH et al (2010) Comparison of extraction methods of total RNA from *Eucommia Ulmoides*. Oliv Med Plant 1(11):14–17

Fukuchi K, Okudaira N, Adachi K et al (2016) Antiviral and antitumor activity of licorice root extracts. Vivo (Athens, Greece) 30(6):777–786

Gao CC, Jun N, Chen MS et al (2015) Characterization of genes involved in gibberellin metabolism and signaling pathway in the biofuel plant jatropha curcas. Plant Divers Resour 37(2):157–167

Gao YX, Cheng BF, Lian JJ et al (2017) Liquiritin, a flavone compound from licorice, inhibits IL-1β-induced inflammatory responses in SW982 human synovial cells. J Funct Foods 33:142–148

Gong H, Zhang BK, Yan M et al (2015) A protective mechanism of licorice (*glycyrrhiza uralensis*): isoliquiritigenin stimulates detoxification system via Nrf2 activation. J Ethnopharmacol 162:134–139

Han R, Rai A, Nakamura M, Suzuki H et al (2016) De novo deep transcriptome analysis of medicinal plants for gene discovery in biosynthesis of plant natural products. Method Enzymol 576:19–45

Hayashi H, Tamura S, Chiba R et al (2016) Field survey of glycyrrhiza plants in central Asia (4). Characterization of *G. glabra* and *G. bucharica* collected in Tajikistan. Biol Pharm Bull 39(11):1781–1786

He SH, Liu HG, Zhou YF et al (2017) Liquiritin (LT) exhibits suppressive effects against the growth of human cervical cancer cells through activating caspase-3 in vitro and xenograft mice in vivo. Biomed Pharmacother 92:215–228

Hu T, Gao ZQ, Ma YS et al (2017) Effect of x-ray irradiation treatment on contents of triterpenoids and flavonoids in glycyrrhizae radix et rhizoma. Chin J Exp Tradit Med Formulae 23(17):17–21

Huang W, Chen X, Li Q et al (2012) Inhibition of intercellular adhesion in herpex simplex virus infection by glycyrrhizin. Cell Biochem Biophys 62(1):137–140

Huang LQ, Gao W, Zhou J et al (2010) Systems biology applications to explore secondary metabolites in medicinal plants. Chin J Chin Mater Med 35(1):8–12

Kanehisa M, Araki M, Goto S et al (2008) KEGG for linking genomes to life and the environment. Nucleic Acids Res 36(Database):D480–D484

Kim ME, Kim HK, Kim DH et al (2013) 18*β*-glycyrrhetinic acid from licorice root impairs dendritic cells maturation and Th1 immune responses. Immunopharmacol Immunotoxicol 35(3):329–335

Kim D, Langmead B, Salzberg SL (2015) Hisat: a fast spliced aligner with low memory requirements. Nat Methods 12(4):357–360

Li W, Yang LX, Jiang LZ et al (2016) Molecular cloning and functional characterization of a cinnamate 4-hydroxylase-encoding gene from *Camptotheca acuminata*. Acta Physiol Plant 38(11):256. https://doi.org/10.1007/s11738-016-2275-7

Lin D, Zhong W, Li J et al (2014) Involvement of bid translocation in glycyrrhetinic acid and 11-deoxy glycyrrhetinic acid-induced attenuation of gastric cancer growth. Nutr Cancer 66(3):463–473

Liu JX, Hou JY, Jiang C et al (2015) Deep sequencing of the *scutellaria baicalensis* georgi transcriptome reveals flavonoid biosynthetic profiling and organ-specific gene expression. PLoS ONE 10(8):e0136397. https://doi.org/10.1371/journal.pone.0136397

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods 25(4):402–408

Luo JJ, Li ZB, Wang JJ et al (2016) Antifungal activity of isoliquiritin and its inhibitory effect against *peronophythora litchi* chen through a membrane damage mechanism. Molecules 21(2):237–248

Luo J, Zhou JJ, Zhang JA (2018) Aux/IAA gene family in plants: molecular structure, regulation, and function. Int J Mol Sci 19(1):259. https://doi.org/10.3390/ijms19010259

Ma C, Ma Z, Liao X et al (2013) Immunoregulatory effects of glycyrrhizic acid exerts anti-asthmatic effects via modulation of Th1/Th2 cytokines and enhancement of cd4(+)cd25(+)Foxp3+ regulatory t cells in ovalbumin-sensitized mice. J Ethnopharmacol 148(3):755–762

Maraschin FDS, Memelink J, Offringa R (2010) Auxin-induced, SCFTIR1-mediated poly-ubiquitination marks AUX/IAA proteins for degradation. Plant J 59(1):100–109

Mortazavi A, Williams B, Schaeffer L et al (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 5(7):621–628

Park SM, Ki SH, Han NR et al (2015) Tacrine, an oral acetylcholinesterase inhibitor, induced hepatic oxidative damage, which was blocked by liquiritigenin through GSK3-beta inhibition. Biol Pharm Bull 38(2):184–192

Patel RK, Mukesh J (2012) NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. PLoS ONE 7(2):e30619. https://doi.org/10.1371/journal.pone.0030619

Ramilowski JA, Daub CO (2013) *Glycyrrhiza uralensis* transcriptome landscape and study of phytochemicals. Plant Cell Physiol 54(5):697–710

Rehan K, Quaiyoom KA, Abdul L et al (2013) Glycyrrhizic acid suppresses the development of precancerous lesions via regulating the hyperproliferation, inflammation, angiogenesis and apoptosis in the colon of wistar rats. PLoS ONE 8(2):e56020. https://doi.org/10.1371/journal.pone.0056020

Rhoads A, Au KF (2015) PacBio sequencing and its applications. GPB 13:278–289

Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26(1):139–140

Seo JY, Han JH, Kim YJ et al (2014) Protective effects of dehydroglyasperin c against carbon tetrachloride-induced liver damage in mice. Food Sci Biotechnol 23(2):547–553

Tanaka A, Shikazono N, Hase Y (2010) Studies on biological effects of ion beams on lethality, molecular nature of mutation, mutation rate, and spectrum of mutation phenotype for mutation breeding in higher plants. J Radiat Res 51(3):223–233

Verma V, Sivaraman J, Srivastava AK et al (2015) Destabilization of interaction between cytokinin signaling intermediates AHP1 and ARR4 modulates *Arabidopsis* development. New Phytol 206(2):726–737

Wang FB, Ren GL, Li FS et al (2018) A chalcone synthase gene *AeCHS* from *Abelmoschus esculentus* regulates flavonoid accumulation and abiotic stress tolerance in transgenic *Arabidopsis*. Acta Physiol Plant 40(5):97. https://doi.org/10.1007/s11738-018-2680-1

Wang LQ, Yang R, Yuan BC et al (2015a) The antiviral and antimicrobial activities of licorice, a widely-used Chinese herb. Acta Pharm Sin B 5(4):310–315

Wang QJ, Lei XY, Zheng LP et al (2016) Molecular characterization of an elicitor-responsive 3-hydroxy-3-methylglutaryl coenzyme A reductase gene involved in oleanolic acid production in cell cultures of *Achyranthes bidentata*. Plant Growth Regul 81:335–343

Wang R, Zhang CY, Bai LP et al (2015b) Flavonoids derived from liquorice suppress murine macrophage activation by up-regulating heme oxygenase-1 independent of Nrf2 activation. Int Immunopharmacol 28(2):917–924

Wang X, Zhang H, Chen L et al (2013) Liquorice, a unique "guide drug" of traditional Chinese medicine: a review of its role in drug interactions. J Ethnopharmacol 150(3):781–790

Wei W, Wang PP, Wei YJ et al (2015) Characterization of *panax ginseng* UDP-glycosyltransferases catalyzing protopanaxatriol and biosyntheses of bioactive ginsenosides F1 and Rh1 in metabolically engineered yeasts. Mol Plant. https://doi.org/10.1016/j.molp.2015.05.010

Yang BW, Hahm YT (2018) Transcriptome analysis using de novo RNA-seq to compare ginseng roots cultivated in different environments. Plant Growth Regul 84:149–157

Yang R, Yuan BC, Ma YS et al (2017) The anti-inflammatory activity of licorice, a widely used Chinese herb. Pharm Biol 55(1):1–14

Ye J, Fang L, Zheng HK et al (2006) WEGO: a web tool for plotting GO annotations. Nucleic Acids Res 34(Web Server Issue):W293–W297

Zhou B, Peng D, Lin JZ et al (2011) Heterologous expression of a gibberellin 2-oxidase gene from *Arabidopsis thaliana* enhanced the photosynthesis capacity in *Brassica napus* L. J Plant Biol 54(1):23–32

Zhou XW, Li JY, Zhu YL et al (2017) De novo assembly of the *Camellia nitidissima* transcriptome reveals key genes of flower pigment biosynthesis. Front Plant Sci 8:1545. https://doi.org/10.3389/fpls.2017.01545

Zhou Y, Ho WS (2014) Combination of liquiritin, isoliquiritin and isoliquirigenin induce apoptotic cell death through upregulating p53 and p21 in the A549 non-small cell lung cancer cells. Oncol Rep 31(1):298–304