



Article

# Integrated Analysis of mRNA and microRNA Elucidates the Regulation of Glycyrrhizic Acid Biosynthesis in *Glycyrrhiza uralensis* Fisch

Yanni Li <sup>1</sup>, Caixia Chen <sup>1</sup>, Zhenzhen Xie <sup>1</sup>, Jing Xu <sup>2</sup>, Bin Wu <sup>1,\*</sup> and Wenquan Wang <sup>1,\*</sup>

<sup>1</sup> Key Laboratory of Bioactive Substances and Resources Utilization of Chinese Herbal Medicine, Ministry of Education, Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100193, China; liyanni\_hua1126@126.com (Y.L.); yongtaochen@126.com (C.C.); zhzhxie126@126.com (Z.X.)

<sup>2</sup> Laboratory of Plant Tissue Culture Technology of Haidian District, Beijing 100091, China; xujing-ln@163.com

\* Correspondence: bwu@implad.ac.cn (B.W.); wwq57@126.com (W.W.); Tel.: +86-10-57833201 (B.W.)

Received: 3 March 2020; Accepted: 21 April 2020; Published: 28 April 2020



**Abstract:** Licorice (*Glycyrrhiza*) is a staple Chinese herbal medicine in which the primary bioactive compound is glycyrrhizic acid (GA), which has important pharmacological functions. To date, the structural genes involved in GA biosynthesis have been identified. However, the regulation of these genes in *G. uralensis* has not been elucidated. In this study, we performed a comprehensive analysis based on the transcriptome and small RNAome by high-throughput sequencing. In total, we identified 18 structural GA genes and 3924 transporter genes. We identified genes encoding 2374 transporters, 1040 transcription factors (TFs), 262 transcriptional regulators (TRs) and 689 protein kinases (PKs), which were coexpressed with at least one structural gene. We also identified 50,970 alternative splicing (AS) events, in which 17 structural genes exhibited AS. Finally, we also determined that miRNAs potentially targeted 4 structural genes, and 318, 8, and 218 miRNAs potentially regulated 150 TFs, 34 TRs, and 88 PKs, respectively, related to GA. Overall, the results of this study helped to elucidate the gene expression and regulation of GA biosynthesis in *G. uralensis*, provided a theoretical basis for the synthesis of GA via synthetic biology, and laid a foundation for the cultivation of new varieties of licorice with high GA content.

**Keywords:** *Glycyrrhiza uralensis* Fisch; transcriptome; microRNA; glycyrrhizic acid biosynthesis; regulatory proteins; transporter

## 1. Introduction

Licorice (*Glycyrrhiza*) plants are the most important dual-use plants and have significant value in ecological protection. The roots and rhizomes of *Glycyrrhiza*, Gancao in Chinese, were first recorded in the *Shengnong Bencao Jing* as a Chinese herbal medicine and natural sweetener. In the Chinese Pharmacopoeia, *Glycyrrhiza uralensis* Fisch, *Glycyrrhiza glabra* L., and *Glycyrrhiza inflata* Bat are the source plants of Gancao [1], in which *G. uralensis* is recognized as the most important species according to the findings of ancient books and modern research [2]. The primary bioactive compound in Gancao is glycyrrhizic acid (GA), which has diverse pharmacological actions, such as anti-inflammatory, anti-cancer, anti-allergic, and anti-viral activities [3]. In addition to medicinal applications, Gancao and its extracts are also widely used in food, cosmetics, and other industries. At present, more than 3000 varieties of medicines and health products were developed using *G. uralensis* and its extracts (<http://app1.sfda.gov.cn/datasearchcnda/face3/dir.html>). Therefore, the demand for licorice in domestic and foreign markets is considerable. However, the supply of wild *G. uralensis* is limited. The extraction of GA from artificially cultivated *G. uralensis* is time-consuming. Moreover, in recent years, the quality

of artificially cultivated *G. uralensis* was not determined to be stable because of the impure germplasm in production, hindering efforts to meet the standards of the Chinese Pharmacopoeia. To solve the problems of sustainable development of *G. uralensis* resources and GA shortage, the GA biosynthesis pathway has been heavily researched in recent years [4].

With the release of the *G. uralensis* genome [5], the structural genes involved in GA biosynthesis have gradually been identified. GA belongs to the triterpenoids, and its synthesis is achieved via the mevalonate pathway, which can be classified into two stages: formation of the triterpenoid skeleton and modification of the skeleton [6]. For the first stage, 3-hydroxy-3-methylglutaryl-coenzyme A reductase (HMGR) is the first rate-limiting enzyme that catalyzes the transformation of HMG-CoA to MVA [7]. Squalene synthase (SQS), a key enzyme for the flow of carbon to triterpenes, catalyzes the synthesis of FPP to squalene [8]. Then,  $\beta$ -AS transforms 2,3-oxidosqualene to  $\beta$ -amyrin by two consecutive oxidation reactions [9]. In the second stage, CYP88D6 catalyzes the oxidation of  $\beta$ -amyrin to form 11-oxo- $\beta$ -amyrin, and CYP72A154 oxidizes the 11-oxo- $\beta$ -amyrin at the C-30 position to form glycyrrhetic acid [10]. The cytochrome P450 reductase (GuCPR) is indispensable in the process of C-11 and C-30 oxidation to balance the redox environment in the cell and facilitate the synthesis of glycyrrhetic acid. Finally, *G. uralensis* Fisch UDP-glycosyltransferase (*GuUGT3*) catalyzes the glycosylation of glycyrrhetic acid to form GA [11].

However, the regulation of GA biosynthesis in *G. uralensis* has not been elucidated to date. In recent years, regulatory proteins and microRNAs (miRNAs) have received increasing attention. To the best of our knowledge, there are three types of regulatory proteins. Transcription factors (TFs) are regulatory proteins that play important roles in plant growth and development and the biosynthesis of secondary metabolites [12]. TFs could bind to specific motifs in the promoter region of genes such that these genes are upregulated or downregulated [13]. TFs can simultaneously regulate the expression of multiple structural genes in the same biosynthetic pathway. For example, *GmMYB39* inhibited isoflavone biosynthesis by repressing the transcript levels of phenylalanine ammonia-lyase (*PAL*), cinnamate 4-hydroxylase (*C4H*), chalcone synthase (*CHS*), 4-coumarate: coenzyme A ligase (*4CL*) and chalcone reductase (*CHR*) in soybean [14]. Transcriptional regulators (TRs) are also regulatory proteins that can regulate the expression of their target genes via epigenetic modifications, such as H3K14/K9 acetylation [15], trimethylation of H3K4 [16] and H3K27 [17]. TRs were involved in multiple processes, including core trichome initiation [15], auxin response [18], vernalization [17], chlorophyll biosynthesis [16], leaf senescence [19] and plant immunity [20]. Another regulatory protein, protein kinases (PKs), function in signal transduction pathways and alter the activity of target proteins by phosphorylation [21]. PKs played an important role in plant growth [22], abiotic and biotic stress responses, ABA synthesis [23], activation of the transcription factor [24], and secondary metabolism. For example, in *Arabidopsis*, SNF1-related protein kinase 2 (*SnRK2*) and phosphatase 2C (*PP2C*) are up- and down-regulated after UV-B treatment, respectively, which activated the expression of *CHS* and isoflavone synthase (*IFS*), leading to the accumulation of isoflavone [25]. In soybeans, overexpression of the *Arabidopsis thaliana* calcium-dependent protein kinase 1 (*AtCPK1*) gene upregulated the expression of the *4CL*, *IFS*, hydroxyisoflavanone dehydratase (*HID*), dimethylallyltransferase (*IDMAT*), and coumestrol 4-dimethylallyltransferase (*C4-DMAT*) genes, resulting in the increase in isoflavone aglycones [26].

Plant miRNAs are nonprotein-coding RNAs measuring approximately 21–24 nt, which are usually derived from a primary transcript with a stem-loop structure. After the action of DICER-LIKE 1 (DCL1), SERRATE (SE), HYPONASTIC LEAVES 1 (HYL1), HUA ENHANCER 1 (HEN1), HASTY (HST), ARGONAUTE 1 (AGO1), STABILIZED 1 (STA1) and RNA binding protein. Mature miRNAs are produced [27–29]. These RNAs negatively regulated target genes at the posttranscriptional level by cleavage [30] or translational inhibition [31] or by methylation modification at the transcriptional level [32]. Studies have shown that miRNAs can negatively regulate their target genes to regulate the development of different organs, response to stress [33], and involvement in secondary metabolism [34].

In this study, because roots and leaves in *G. uralensis* have different GA content, we performed transcriptomics and small RNAomics analysis by using these two tissues. We found that TFs, TRs, and PKs genes were coexpressed with multiple structural genes, which were potentially involved in the biosynthesis of GA. We also found that miRNAs potentially regulated the biosynthesis of GA in two aspects. On the one hand, miRNAs potentially targeted structural genes. On the other hand, miRNAs potentially targeted GA-related genes encoding regulatory proteins, including TFs, TRs, and PKs.

Therefore, a study on miRNAs, TFs, TRs, and PKs would help to reveal the regulatory mechanism of GA synthesis, providing a theoretical basis for the synthesis of GA via synthetic biology and laying a foundation for the cultivation of new licorice varieties with a high GA content.

## 2. Results

### 2.1. Construction of RNA-Seq Libraries and Sequence Alignment

Our experiment aimed to elucidate the structural genes involved in the GA biosynthetic pathway and their regulation. Because the content of GA is higher in roots than in leaves in *G. uralensis*, we performed transcriptome sequencing on these two tissues and analyzed the differentially expressed genes (DEGs). First, strand-specific cDNA libraries from the poly(A) mRNAs of leaves and roots samples were constructed and subjected to sequence by the Illumina platform. After sequencing and quality control, a total of 45.65 Gb of clean data were obtained, and the percentage of Q30 base in each sample was more than 93.08% (Table S1).

Then, we used the draft genome of *G. uralensis* as a reference for sequence alignment and subsequent analysis. After the alignment of the clean reads by HISAT2 [35], we used StringTie [36] to compare and assemble the reads to build the longest transcripts and estimate the expression level. The alignment ratio of reads to reference genomes of each sample ranged from 85.72 to 90.58% (Table S2).

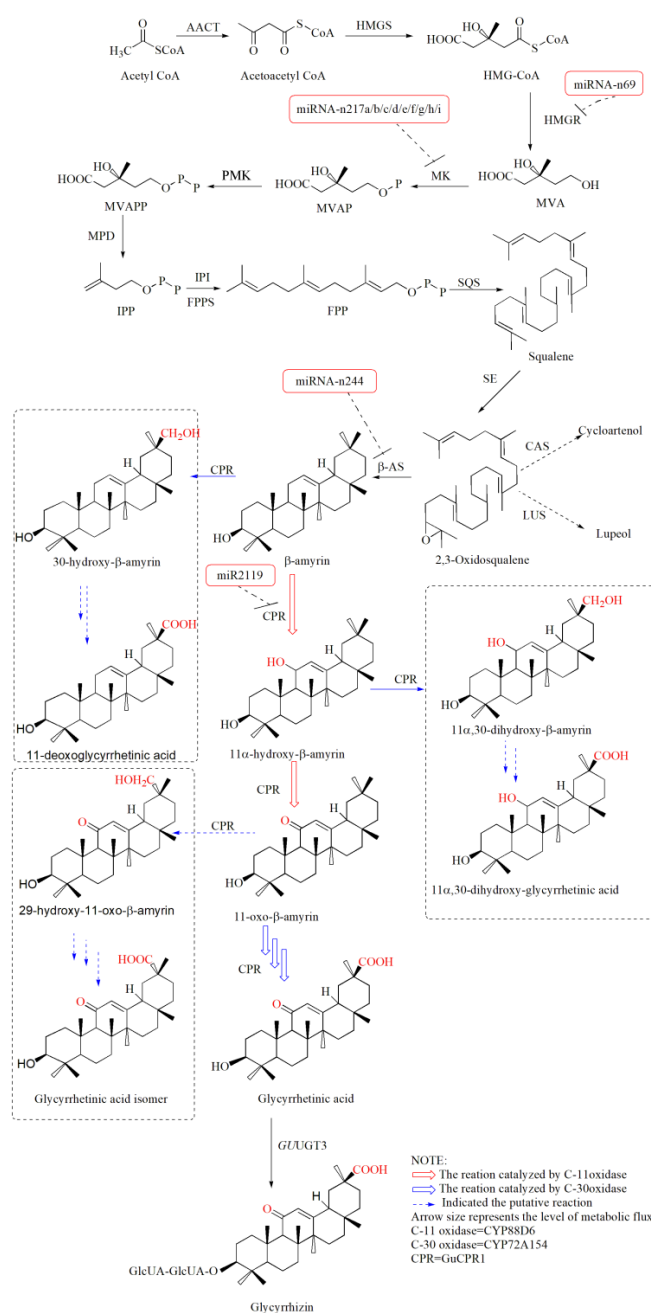
### 2.2. Gene Function Annotation

To determine gene functions, the longest transcripts were compared with the public databases. As a result, 9827, 22,348, 11,768, 16,758, 22,594, 22,762, 29,118 and 31,910 genes were annotated by Cluster of Orthologous Groups of proteins (COG), Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), euKaryotic Ortholog Groups (KOG), Protein family (Pfam), Swiss-Prot protein sequence (Swiss-Prot), Nonsupervised Orthologous Groups (eggNOG), and NCBI nonredundant protein sequences (NR) databases, respectively. We also found a total of 1378 new genes by filtering out sequences encoding peptide chains that were too short (less than 50 amino acid residues) or containing only a single exon (Table S3).

### 2.3. Structural and Transporter Genes Potentially Involved in GA Biosynthesis

To explore the structural genes involved in the GA pathway, we screened the gene annotations based on the NR and KEGG. The results showed that the 3-hydroxy-3-methylglutaryl-coenzyme A reductase (HMGR) gene contains four transcripts; squalene synthase (SQS), CYP88D6 (C-11 oxidase), and CYP72A154 (C-30 oxidase) genes contain two transcripts; acetyl-CoA C-acetyltransferase (AACT), 3-hydroxy-3-methylglutaryl coenzyme A synthase (HMGS), mevalonate kinase (MK), phosphomevalonate kinase (PMK), mevalonate diphosphate decarboxylase (MPD), farnesyl diphosphate synthase (FPPS), beta-amyrin synthase ( $\beta$ -AS), and CPR reductase (GUCPR) genes contain one transcript. The transcript information of these structural genes is shown (Figure 1, Table 1). Unfortunately, the structural genes encoding isopentenyl diphosphate isomerase (IPI), squalene epoxidase (SE), and *Glycyrrhiza uralensis* Fisch UDP-glycosyltransferase (*Gu*UGT 3) were not found in our RNA-Seq data, indicating the low expression level of these genes in our samples. In addition to structural genes, we also focused on the transporter genes because the enzymes, encoded by structural genes, that synthesize the intermediate products of GA are localized in different organelles (Table 1).

Through a comparison with the TCDB transporter database, a total of 3924 transporter candidate genes were identified in *G. uralensis* (Table S4).



**Figure 1.** Proposed GA biosynthetic pathway in *G. uralensis*. Red, blue arrows indicate reactions catalyzed by C-11, C-30 oxidase, respectively. Dotted arrows indicate the proposed reactions. Structural genes potentially targeted by the identified *G. uralensis* miRNAs are represented by red rectangular boxes. The dotted line represents uncertain action in the pathway. AACT: acetyl-CoA C-acetyltransferase; HMGS: 3-hydroxy-3-methylglutaryl coenzyme A synthase; HMGR: 3-hydroxy-3-methylglutaryl-coenzyme A reductase; MK: mevalonate kinase; PMK: phosphomevalonate kinase; MPD: mevalonate diphosphate decarboxylase; IPI: isopentenyl diphosphate isomerase; FPPS: farnesyl diphosphate synthase; SQS: squalene synthase; SE: squalene epoxidase;  $\beta$ -AS: beta-amyrin synthase; C-11 Oxidase: CYP88D6, beta-amyrin 11-oxidase; C-30 Oxidase: CYP72A154, 11-oxo-beta-amyrin 30-oxidase; CPR: cytochrome P450 reductase; GuUGT3: *G. uralensis* Fisch UDP-glycosyltransferase 3.

**Table 1.** Genes potentially involved in the GA biosynthetic pathway.

Annotation	Transcript Number	Transcript ID	Predicted Subcellular Localization	Expression Profiles (FPKM)						Up-Regulated in Roots (Y/N)
				L1	L2	L3	R1	R2	R3	
<i>AACT</i>	1	Glyur000218s00011642.1	Peroxisome	38.32	42.15	47.59	105.87	152.71	143.58	Y
<i>HMGS</i>	1	Glyur000195s00012841.1	Cytoplasm	2.35	2.81	7.56	173.37	280.56	276.05	Y
<i>HMGR</i>	4	Glyur000037s00002618.1	Endoplasmic Reticulum	32.41	57.49	47.67	92.79	193.02	173.28	Y
		Glyur000135s00007303.1	Endoplasmic Reticulum	11.95	15.05	16.18	66.08	35.39	39.91	Y
		Glyur000203s00012900.1	Endoplasmic Reticulum	6.69	12.23	17.49	26.99	42.77	38.50	Y
		Glyur000682s00024324.1	Endoplasmic Reticulum	13.21	16.87	18.17	137.41	182.13	172.25	Y
<i>MK</i>	1	Glyur000069s00004081.1	Extracell	5.15	6.75	6.73	9.61	6.62	6.40	N
<i>PMK</i>	1	Glyur000343s00025703.1	Cytoplasm	16.23	9.45	10.92	5.40	13.19	13.26	N
<i>MPD</i>	1	Glyur000002s00000233.1	Cell Membrane and Extracell	7.42	9.09	9.11	30.96	41.07	42.35	Y
<i>FPPS</i>	1	Glyur000088s00007722.1	Endoplasmic Reticulum and Nucleus	5.14	6.08	11.84	35.58	40.90	40.94	Y
<i>SQS</i>	2	Glyur000017s00002413.1	Extracell	21.40	19.05	24.26	18.33	48.71	46.67	Y
		Glyur000089s00008825.1	Endoplasmic Reticulum.	3.70	3.44	4.82	16.81	43.09	38.43	Y
<i>β-AS</i>	1	Glyur001733s00027628.1	Extracell and Nucleus	0.97	2.28	10.35	45.99	189.49	162.89	Y
<i>CYP88D6</i>	2	Glyur000561s00023443.1	Cytoplasm Endoplasmic Reticulum and Nucleus	0.04	0.39	0.91	12.41	21.66	17.30	Y
		Glyur000561s00023451.1	Endoplasmic Reticulum.	0.25	0.64	2.67	183.15	327.26	297.26	Y
<i>CYP72A154</i>	2	Glyur000890s00019071.1	Endoplasmic Reticulum.	6.92	7.14	6.31	118.31	181.82	175.76	Y
		Glyur001936s00032203.1	Extracell.	10.00	1.51	7.57	103.78	167.71	142.77	Y
<i>GUCPR1</i>	1	Glyur000294s00011848.1	Endoplasmic Reticulum.	38.03	46.50	39.00	16.14	25.88	25.71	Y

<sup>1</sup> “L1, L2, L3” and “R1, R2, R3” represent 3 leaves and roots samples, respectively. “Y” means yes. “N” means no.

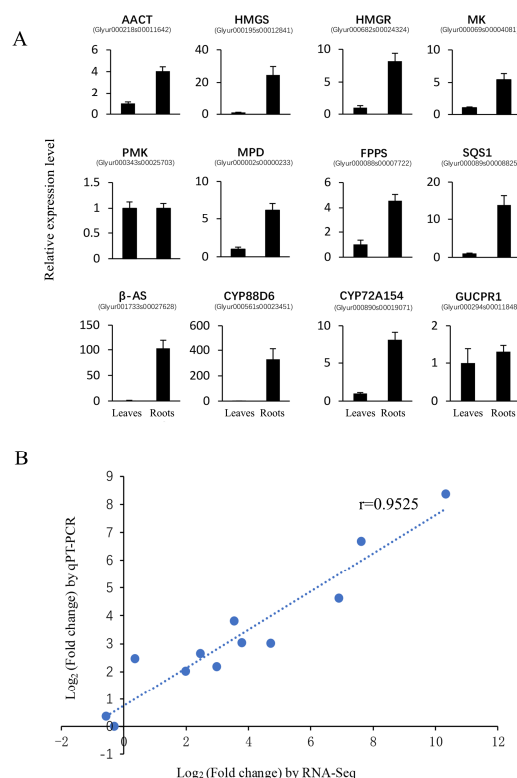
#### 2.4. Identification of Regulatory Proteins in *G. uralensis*

It was reported that regulatory proteins, including TFs, TRs, and PKs, play important roles in plant growth and development and the synthesis of secondary metabolites. To explore the roles of these proteins in *G. uralensis*, we first predicted these regulatory proteins by comparing the nucleotide or coding-protein sequences of the longest transcripts with the iTAK database. In total, we found 1841 TFs belonging to 69 families, in which the number of MYB family TFs was highest (Table S5). We identified a total of 401 TRs belonging to 23 families (Table S6). We also identified 1168 PKs, which belonged to 119 families. Among these PKs, the RLK-Pelle\_DLSV group was the largest family (Table S7).

#### 2.5. Identification of DEGs and qRT-PCR Validation

Compared with roots, those genes that had an expression level of [fold change] > 2 and FDR < 0.01 in leaves found by StringTie were assigned as DEGs. In total, we obtained 4284 DEGs, including the GA pathway structural genes *HMG5*(Glyur000195s00012841), *HMGR*(Glyur000037s00002618)/(Glyur000203s00012900)/(Glyur000682s00024324) and *SQS* (Glyur000089s00008825), 680 transporters, 321 TFs, 43 TRs, and 213 PKs (Table S8).

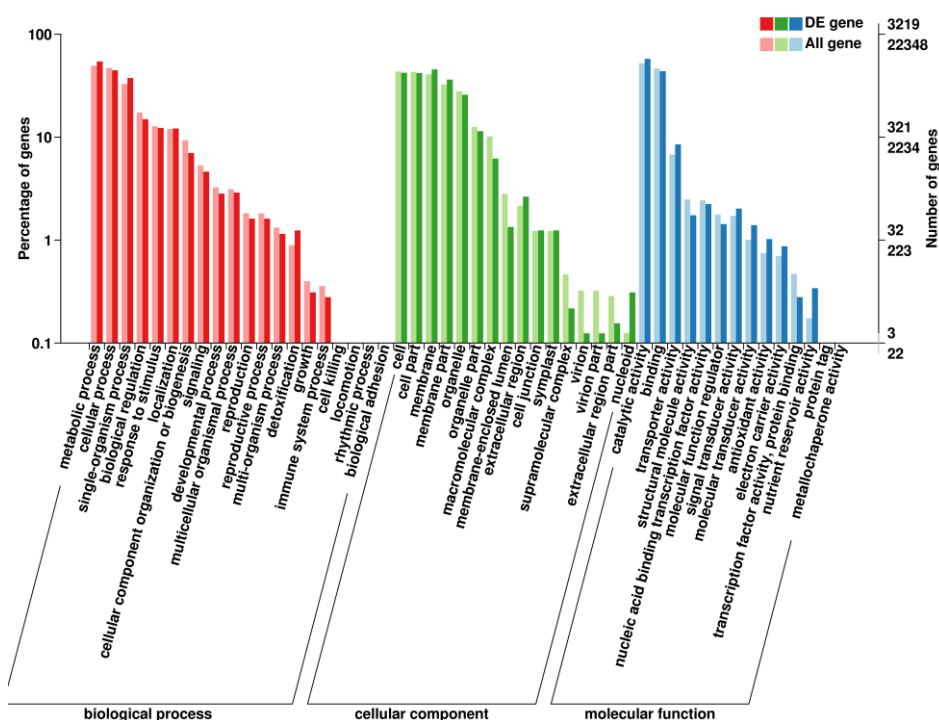
To verify the reliability of RNA-Seq results, 12 structural genes, including *AACT* (Glyur000218s00011642.1), *HMG5*(Glyur000195s00012841.1), *HMGR*(Glyur000682s00024324.1), *MK*(Glyur000069s00004081.1), *PMK*(Glyur000343s00025703.1), *MPD*(Glyur000002s00000233.1), *FPPS*, (Glyur000088s00007722.1), *SQS1*(Glyur000089s00008825.1),  $\beta$ -AS (Glyur001733s00027628.1), *CYP88D6*(Glyur000561s00023451.1), *CYP72A154*(Glyur000890s00019071.1), and *GUCPR1* (Glyur000294s00011848.1) were selected for qRT-PCR validation. The results of qRT-PCR were highly consistent with the RNA-Seq data, and the correlation coefficient was 0.9525 (Figure 2).



**Figure 2.** qRT-PCR validation of structural genes (A) and the correlation between qRT-PCR and RNA-Seq results (B). In Figure 2 (A), the expression level of the structural gene in the leaves was set to 1, the relative expression of genes in roots was calculated, and the error bar is shown. In Figure 2 (B), the X-axis represents the log<sub>2</sub> (fold change) of the RNA-Seq results, and the Y-axis is the log<sub>2</sub> (fold change) of the qRT-PCR results.

## 2.6. GO Enrichment Analysis of DEGs

The GO database describes gene function across species. The results showed that a total of 3219 DEGs were annotated by GO (Table S9). The top 50 secondary subclassifications are listed in Figure 3, including biological process (20), cellular component (16) and molecular function (14). The top five of the DEGs are involved in catalytic activity (1852), metabolic process (1752), cellular process (1436), binding (1411) and cell (1360), which showed the same tendency with all the GO annotated genes (Figure 3).

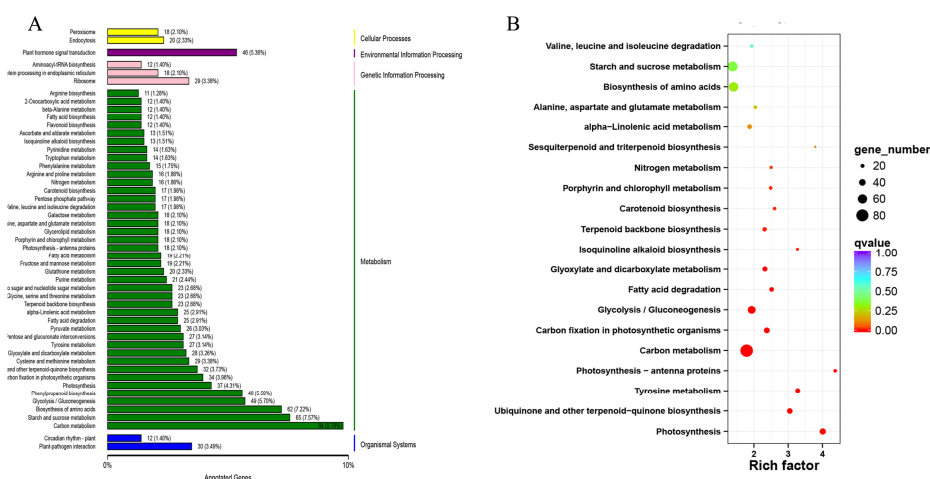


**Figure 3.** Results of GO enrichment. The percentage of genes assigned to partial subclassifications is shown. Red, green and blue columns represent the biological, cellular, and molecular functions, respectively.

## 2.7. KEGG Enrichment Analysis of DEGs

KEGG allows for functional annotations of gene products according to various metabolic pathways. Figure 4A shows the top 50 subcategories, including cellular processes (38), environmental information processing (46), and genetic information processing (59). Interestingly, many DEGs are involved in metabolism (1047) and organismal systems (42) (Figure 4A, Table S10).

The results of the KEGG pathway enrichment analysis of DEGs are shown in Figure 4B, which shows the top 20 pathways with the Q value. We found that some genes are enriched in terpenoid synthesis pathways and metabolism processes, in which 32, 23, and 7 genes are enriched in ubiquinone and other terpenoid-quinone biosynthesis (7 transporters), terpenoid backbone biosynthesis (4 transporters), and sesquiterpenoid and triterpenoid biosynthesis, respectively (Figure 4B).



**Figure 4.** (A) KEGG classification of DEGs. The X-axis indicates the percentage of the annotation genes, and the Y-axis represents the KEGG pathway. (B) KEGG pathway enrichment of DEGs for the comparison of roots vs. leaves. The X-axis indicates the enrichment factor and  $\log_{10}$  of the Q-value. The Y-axis represents the different KEGG pathways.

## 2.8. Correlation Analysis between Transporter and Structural Genes

Of the 3924 transporter genes, 3556 have expression levels (Table S11). Then, we calculated the Pearson correlation coefficients ( $r$ ) of the expression profiles between transporter genes and the GA structural genes. With the threshold of  $|r| > 0.8$ , a total of 2374 transporter genes were identified. Among these genes, 713, 435, 161, 66, 41, 41, 136, 311, and 1077 transporter genes were found to be highly correlated with 10, 9, 8, 7, 6, 5, 4, 3, 2, and 1 structural gene(s), respectively (Table S12). Subsequently, subcellular localization prediction of the 2374 transporters was carried out by ngLOC (<http://genome.unmc.edu/ngLOC/index.html>) (Table S13), in which 53 transporters predicted to be localized in the vacuole and 23 transporters to be localized in the endoplasmic reticulum (ER) (Table S14), consistent with the location of most of the GA metabolic enzymes. This information strongly enhances our knowledge of transporters involved in GA metabolism.

## 2.9. Correlation Analysis between Regulatory Protein and the Structural Genes

Of the 1841 TFs, 1552 had expression levels (Table S15). Of the 401 TRs, 368 had expression levels (Table S16). Of the 1168 PKs, 1034 had expression levels (Table S17).

To decipher the role of regulatory protein genes in GA metabolism, coexpression analysis was performed to characterize the expression patterns of TF, TR and PK genes and the structural genes. Using a threshold of  $|r| > 0.8$ , a total of 1040 TFs were associated with at least one structural gene in the GA metabolism pathway, in which 289, 204, 80, 43, 27, 13, 18, 62, 148, and 445 TFs were highly correlated with 10, 9, 8, 7, 6, 5, 4, 3, 2, and 1 enzyme gene(s), respectively. Among the 1040 TFs, BHLH (82), MYB (74), C2H2 (64), AP2/ERF (60), WRKY (59), bZIP (47), MYB-related (45), NAC (44), C3H (41), HB-HD-ZIP (35) were the top 10 families. To date, six TF families in plants have been reported to be associated with terpenoid metabolism, including AP2/ERF, BHLH, MYB [37], NAC [38], WRKY [39], and bZIP [40].

In total, 262 TRs were associated with at least one structural gene in the GA metabolism pathway, in which 52, 29, 15, 5, 6, 6, 5, 35, 68 and 93 TRs were related to 10, 9, 8, 7, 6, 5, 4, 3, 2, and 1 structural gene(s), respectively. Among the 262 TRs, others (43), GNAT (27), SET (26), SNF2 (26), AUX/IAA (25), mTERF (24), PHD (19), TRAF (11), ARID (11), and SWI/SNF-BAF60b (11) were the top 10 family members.

A total of 689 PKs were associated with at least one structural gene in the GA metabolism pathway, in which 188, 132, 59, 26, 15, 8, 14, 38, 98, and 299 PKs were related to 10, 9, 8, 7, 6, 5, 4, 3, 2, and 1



structural gene(s), respectively. Among the 689 PKs, RLK-Pelle\_DLSV (50), RLK-Pelle\_LRR-XI-1 (45), RLK-Pelle\_RLCK-VIIa-2 (29), RLK-Pelle\_LRR-III (25), CAMK\_CDPK (23), RLK-Pelle\_LRK10L-2 (21), CAMKL-CHK1(20), RLK-Pelle\_LysM (20), TKL-Pl-4(19), RLK-Pelle\_L-LEC (19) were the top 10 family members in PKs (Table S18).

All 1040 TFs, 262 TRs, and 689 PKs supplied useful candidate components to enhance GA biosynthesis in *G. uralensis*.

#### 2.10. Binding Site Analysis of R2R3-MYB in the Promoter Region of the Structural Gene

The recognition and binding of TF to the target gene promoter is the most critical link in gene expression regulation. In plants, the MYB factor is one of the largest TF families [41]. MYB could be divided into three subfamilies according to the number of adjacent repeats in the MYB domain, in which the R2R3-MYB subfamily, which mainly regulates plant-specific processes [12], was further studied. The current research has verified that the core sequence of the R2R3-MYB binding site is TAACTG [42]. Therefore, we searched it at the promoter regions (2000 bp upstream of the initiation codon) of the structural genes (Table 2). As shown in Table S19, the promoter regions of *AACT* (Glyur000218s00011642.1), *HMGR* (Glyur000682s00024324.1), *MK* (Glyur000069s00004081.1), *MPD* (Glyur000002s00000233.1), *SQS1* (Glyur000089s00008825.1), and *CYP88D6* (Glyur000561s00023451.1) had binding sites. There are 119 TFs of the MYB family found in *G. uralensis*, of which three R2R3-MYBs (Glyur000618s00030214, Glyur000049s00003134 and Glyur000266s00012631) were related to 10, 10 and 9 structural genes, respectively [37].

**Table 2.** Analysis of the R2R3-MYB binding site.

Gene ID	Structural Gene Name	Predicted Binding Site Upstream ATG (bp)
Glyur000218s00011642.1	<i>AACT</i>	−1934
Glyur000682s00024324.1	<i>HMGR</i>	−1544
Glyur000069s00004081.1	<i>MK</i>	−1875
Glyur000002s00000233.1	<i>MPD</i>	−1748
Glyur000089s00008825.1	<i>SQS1</i>	−696
Glyur000561s00023451.1	<i>CYP88D6</i>	−340 and −347

#### 2.11. AS Analysis

In total, 50,970 AS events were identified by ASprofile, including 2168 exon skipping (ES), 20,126 Alt 3' (alternative 3' last exon, TTS), 20,700 Alt 5' (alternative 5' first exon, TSS) and 3671 intron retention (IR), respectively [43]. ES could be classified into SKIP (skipped exon), XSKIP (approximate SKIP), MSKIP (multiexon SKIP) and XMSKIP (approximate MSKIP). IR included IR (intron retention), XIR (approximate IR), MIR (multi-IR) and XMIR (approximate MIR). Then, we focused on the AS in the GA biosynthetic structural genes. The results are shown in Table 3. More than one AS event occurred in the 17 structural genes of GA. All the structural genes were AS of the TSS and TTS types. We also found a total of 2694 transporters with AS, including 89 ES, 2655 Alt 3', 2616 Alt 5' and 171 IR. A total of 1073 TFs occurred as AS events, including 41 ES, 1065 Alt 3', 1047 Alt 5' and 50 IR. A total of 297 TRs occurred with AS, including 12 ES, 294 Alt 3', 293 Alt 5', and 35 IR. A total of 775 PK occurred AS, including 42 ES, 763 Alt 3', 752 Alt 5', and 74 IR. Most of the transporter, TF, TR and PK genes occurred in the TSS and TTS types of AS (Table S20).

**Table 3.** AS occurred in structural genes of GA.

Gene ID	Gene Name	Event Type	Chrom	Event Pattern
Glyur000294s00011848	CPR	TTS	Scaffold00294	18154
		TSS	Scaffold00294	10567
		TTS	Scaffold00890	8048
Glyur000890s00019071	CYP 72A154	XIR_ON	Scaffold00890	8315–8798
		TSS	Scaffold00890	10471
		TTS	Scaffold01936	49804
Glyur001936s00032203	CYP 72A154	TSS	Scaffold01936	52193
		TTS	Scaffold00218	225653
Glyur000218s00011642	AACT	TSS	Scaffold00218	229590
		TTS	Scaffold00195	66634
Glyur000195s00012841	HMGS	TSS	Scaffold00195	71499
		TTS	Scaffold00682	96974
Glyur000682s00024324	HMGR	TSS	Scaffold00682	95210
		TTS	Scaffold00037	147378
Glyur000037s00002618		TSS	Scaffold00037	149004
		TTS	Scaffold00135	140846
Glyur000135s00007303		TSS	Scaffold00135	142733
		TTS	Scaffold00203	184607
Glyur000203s00012900		TSS	Scaffold00203	186414
		TTS	Scaffold00069	351209
Glyur000069s00004081	MK	TSS	Scaffold00069	345493
		IR_OFF	Scaffold00069	347728–347912,348276–348369
		IR_ON	Scaffold00069	347728–348369
Glyur000343s00025703	PMK	TTS	Scaffold00343	178004
		TSS	Scaffold00343	171667
Glyur000002s00000233	MPD	TTS	Scaffold00002	176640
		TSS	Scaffold00002	173112
Glyur000088s00007722	FPPS	TTS	Scaffold00088	174563
		TSS	Scaffold00088	178062
		TTS	Scaffold00017	277461
Glyur000017s00002413	SQS2	TTS	Scaffold00017	277670
		TTS	Scaffold00017	277888
		TSS	Scaffold00017	271392
		SKIP_OFF	Scaffold00089	317631–318581
Glyur000089s00008825	SQS1	SKIP_ON	Scaffold00089	317631,318018–318093,318581
		TTS	Scaffold00089	314669
		TSS	Scaffold00089	319057
		IR_OFF	Scaffold00089	315038–315129,315471–315559
		IR_ON	Scaffold00089	315038–315559
Glyur001733s00027628	$\beta$ -AS	TTS	Scaffold01733	36913
		TSS	Scaffold01733	44011
Glyur000561s00023451	CYP88D6	TTS	Scaffold00561	103653
		TTS	Scaffold00561	139473
		TSS	Scaffold00561	143458

### 2.12. Features of *G. uralensis* Small RNA Population

To identify miRNAs in *G. uralensis*, we first constructed small RNA libraries of roots and leaves (each tissue with three replicates). Then, these libraries were sequenced by the HiSeq NovaSeq platform. After sequencing, to collect the clean data, the raw reads were filtered by removing adapters, low-quality sequences, sequences less than 18 nt and sequences larger than 30 nt. Then, the clean data of the six libraries were pooled together to analyze the overall features of the *G. uralensis* sRNA population. In total, 116,807,113 clean reads represented by 20,048,413 unique sRNA sequences were obtained. Among these reads, the group of 24 nt clean reads was the highest, accounting for 19.24% of the total clean reads (Figure S1A), and the group of 24 nt unique sRNA sequences was the largest, accounting for 40.43% of the total unique sequences (Figure S1B), which was consistent with the findings of previous reports in some plants, such as *Arabidopsis* [44], *Oryza sativa* [45], and *Zea mays* [46].

As shown in Figure S1C, the groups of 21 nt and 24 nt sRNAs had the highest proportions of U and A nucleotides at the first base, respectively, indicating the different proportions of sRNAs. It was reported that miRNAs and trans-acting small interfering RNAs prefer beginning with a U base, and antisense transcript-derived siRNAs and repeat-associated siRNAs tended to begin with an A base (Figure S1C).

### 2.13. Identification and Validation of MiRNAs in *G. uralensis*

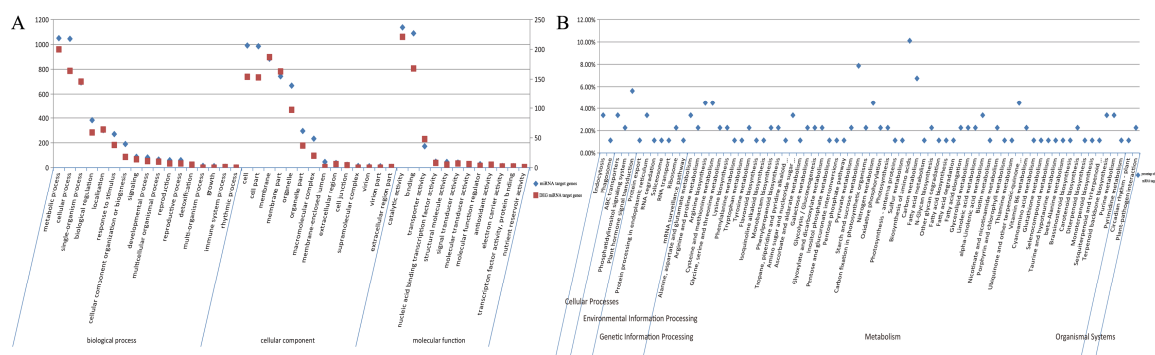
The psRobot software was used to predict the miRNAs in *G. uralensis*. Then, these predicted miRNAs were checked manually according to the criteria of miRNA, as described by Meyer et al. In total, 1894 miRNAs were identified in *G. uralensis*. After comparison with miRbase, it was found that 198 miRNAs were conserved and could be divided into 29 families, of which miR4389 was the most heavily represented family, with 73 members. The remaining 1696 miRNAs were identified as *G. uralensis*-specific miRNAs, which could be divided into 518 families according to the different precursors. Among these miRNAs, the largest family (miR-n4) had 294 members (Table S21).

To verify the authenticity of miRNA, 6 miRNAs, including miR172c-3P, miR156a, miR164a-5P, miR156c-3P, miR166-3P and miR2119, which targeted TFs or the *GUCPR1* genes, were verified by stem-loop PCR. The results showed that the melting curves of the six miRNA amplification products were all single peaks, which validated the authenticity of these miRNAs (Figure S2).

### 2.14. Target Prediction of MiRNAs in *G. uralensis*

MiRNAs could bind and negatively regulate their target genes in a not completely complementary pairing manner [27]. To decipher the function of miRNAs in *G. uralensis*, targets of the miRNAs were predicted by psRNATarget software using an expectation penalty score of  $\leq 3$ . As a result, we predicted 553 and 2474 targets for the conserved and species-specific miRNAs in *G. uralensis*, respectively. In total, 2866 target genes were obtained, of which 2171 had GO annotations and 1081 had KEGG annotations. A total of 448 target genes were DEGs, of which 374 had GO annotations and 166 genes had KEGG annotations (Table S22). GO enrichment of target genes showed that the top five classifications were catalytic activity (1137), binding (1089), metabolic process (1050), cellular process (1045), and cell (991). GO terms of the top five classifications of the 374 DEGs were catalytic activity (221), metabolic process (200), membrane (187), binding (168), and cellular process (164) (Figure 5A, Table S22).

KEGG enrichment analysis of target genes is shown in Figure 5B. We found that some genes were enriched in terpenoid synthesis pathways and metabolism processes, in which 1, 1, 1, and 3 genes were enriched in diterpenoid biosynthesis, monoterpenoid biosynthesis, sesquiterpenoid and triterpenoid biosynthesis, and terpenoid backbone biosynthesis, respectively (Figure 5B, Table S22).



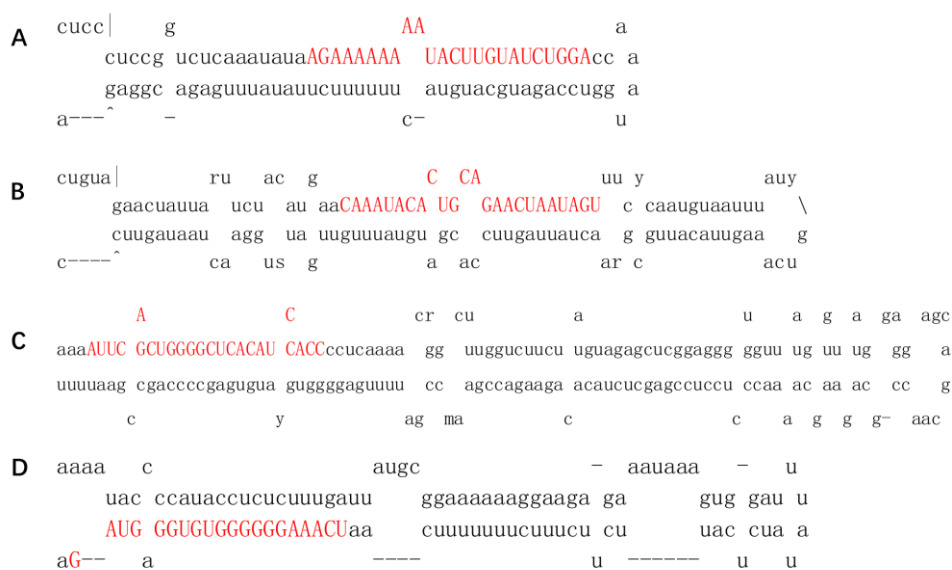
**Figure 5.** (A) GO subclassification of miRNA target genes. (B) KEGG subclassification of miRNA target genes.

One of our concerns was to determine if the structural genes of GA were targeted by miRNAs. As a result, four structural genes were found to be potentially targeted by 12 miRNAs belonging to

four families (Table 4). Among them, *MK*(Glyur000069s00004081.1) was potentially targeted by the nine miRNAs of the miRNA-n217 family. miRNA-n244, miRNA-n69 and miR2119 potentially targeted  $\beta$ -AS(Glyur001733s00027628.1), *HMGR*(Glyur000682s00024324.1), and *CPR1*(Glyur000294s00011848.1) genes, respectively. The structures of miRNA-n217g, miRNA-n69, miRNA-n244, and miR2119 precursors are shown in Figure 6.

**Table 4.** miRNAs target structural genes.

miRNA Sequence	miRNA Name	Target Gene ID	Target Annotation	Expectation
AGCAAAAAAGUACAUGUAUCUGGA	miRNA-n217c	Glyur000069s00004081.1	<i>MK</i>	3
AGCAAAAAAGUACAUGUAUCUGGA	miRNA-n217b	Glyur000069s00004081.1	<i>MK</i>	3
AGCAAAAAAGUACAUGUAUCUGGA	miRNA-n217a	Glyur000069s00004081.1	<i>MK</i>	3
AGCAAAAAAGUACAUGUAUCUGGA	miRNA-n217e	Glyur000069s00004081.1	<i>MK</i>	3
AGCAAAAAAGUACAUGUAUCUGGA	miRNA-n217d	Glyur000069s00004081.1	<i>MK</i>	3
AAGAAAAAAGUACAUGCAUCUGA	miRNA-n217f	Glyur000069s00004081.1	<i>MK</i>	2
AGAAAAAAAUAACAUGUAUCUGGA	miRNA-n217i	Glyur000069s00004081.1	<i>MK</i>	1
AGAAAAAAAUAACAUGUAUCUGGA	miRNA-n217h	Glyur000069s00004081.1	<i>MK</i>	1
AGAAAAAAAUAACAUGUAUCUGGA	miRNA-n217g	Glyur000069s00004081.1	<i>MK</i>	1
AUUCAGCUGGGGCACAUCCACC	miRNA-n244	Glyur001733s00027628.1	$\beta$ -AS	3
CAAAUACACUGCAGAACUAAUGU	miRNA-n69	Glyur000682s00024324.1	<i>HMGR</i>	3
UCAAGGGGGGUGUGGAGUAG	miR2119	Glyur000294s00011848.1	<i>CPR1</i>	2.5



**Figure 6.** Precursor structure of miRNA-n217g (A), miRNA-n244 (B), miRNA-n69 (C), and miR2119 (D). The red capitalized bases represent the sequences of mature miRNA.

We also investigated GA-related transporter genes which are regulated by miRNAs. The results are presented in Table S23. We found that 83 conserved and 884 *G. uralensis*-specific miRNAs targeted 53 and 269 transporter genes, respectively, in which 21 transporters were potentially regulated by conserved and *G. uralensis*-specific miRNAs simultaneously. Of these 322 transporters, 9 transporters were localized in the vacuole, and 1 transporter was localized in the ER.

miRNAs might regulate GA synthesis by acting with GA-related regulatory protein genes. The results showed that 223 conserved and 195 *G. uralensis*-specific miRNAs targeted 60 and 90 TF genes, respectively, in which 22 TFs were potentially regulated by conserved and *G. uralensis*-specific miRNAs simultaneously (Table S24).

It was reported that overexpression of MicroRNA156 could improve drought stress tolerance and biomass yield in *Medicago sativa* L [47,48]. However, under drought conditions, the synthesis of GA was enhanced. Therefore, we further analyzed the target genes of miRNA156 in *G. uralensis*. In total,

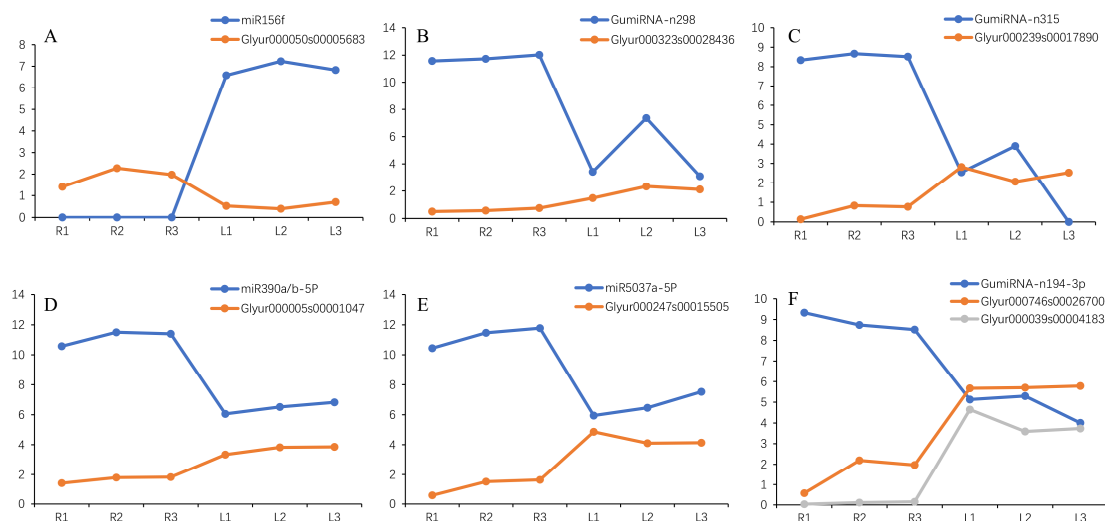
10 members of miRNA156 were identified, potentially targeting 15 potential GA-related TFs, including 12 *SPLs* (Table S24), in which *SPL3*, *SPL 6*, *SPL12* and *SPL13* cleavage by MicroRNA156 was reported in *Salvia miltiorrhiza* and *Medicago sativa* L [48,49].

For TRs, we found 6 conserved and 74 *G. uralensis*-specific miRNAs acted on 60 and 90 TR genes, respectively, in which 2 TRs were potentially regulated by conserved and *G. uralensis*-specific miRNAs simultaneously (Table S25). We also found 31 conserved and 187 *G. uralensis*-specific miRNAs potentially acted on 16 and 83 PKs genes, respectively, in which 5 PKs were potentially regulated by conserved and *G. uralensis*-specific miRNAs at the same time (Table S26).

### 2.15. Negative Correlation Expression between MiRNA and Target Transcripts

We identified a total of 1894 miRNAs in *G. uralensis*, of which 1726 had expression levels (Table S27). Using the threshold of  $|\log_2 \text{fold change}| \geq 2$  and  $\text{FDR} < 0.05$  found by DESeq 2 package, we identified a total of 35 DE miRNAs. Among these miRNAs, 29 were downregulated, and 6 were upregulated (Table S28). Then, we focused on the expression patterns of negative correlations between miRNAs and their target genes. As a result, six miRNA/protein-coding gene pairs were identified (Figure 7). This group included miR156f/*SPL6* (Glyur000050s00005683), GumiRNA-n298/Putative transporter arsB (Glyur000323s00028436), GumiRNA-n315/phyloquinone biosynthesis protein (Glyur000239s00017890), miR390a/b-5P/hypothetical protein LR48\_Vigan09g172300 (Glyur000005s00001047), miR5037a-5P/pentatricopeptide repeat-containing protein and chloroplast (Glyur000247s00015505), and GumiRNA-n194-3p/hypothetical protein KK1\_020744 (Glyur000039s00004183)/chalcone reductase genes (Glyur000746s00026700) (Table S29, Figure 7).

Then, we focused on the miRNAs/GA-related TFs. For miR156f, one of its target genes (Glyur000050s00005683) was determined to encode *SPL6*. The expression profiles of Glyur000050s00005683 were negatively correlated with nine structural genes (Table S18).



**Figure 7.** Expression patterns of six miRNAs and target genes. L, R stand for leaves and roots, respectively. Every tissue has three biological replicates.

### 3. Discussion

In this study, we performed transcriptional analysis to reveal the expression and regulation of GA biosynthesis in *G. uralensis*. For the structural genes of GA, three aspects should be considered. First, some family genes contain more than one locus in the genome. To understand the gene function comprehensively, it is necessary to study all members because of the possibility of functional diversity for different members. For example, diterpene syntheses genes have different functions in *Salvia miltiorrhiza* [50]. Taking the cellulose synthase (*CESA*) genes as another example, there are 10 *CESA*

genes in *Arabidopsis* [51]. The *irx1*, *irx3*, and *irx5* mutants of *Arabidopsis* are caused by mutations in the *AtCESA7*, *AtCESA8*, and *AtCESA4* genes, respectively; the encoding proteins of these three genes also interact with one another and participate in the cellulose synthesis of the secondary cell wall [52]. Second, AS could enhance the protein diversity of genes [53]. In medicinal plants, such as *Digitalis purpurea* and *Salvia miltiorrhiza*, AS occurred in genes involved in the biosynthesis of cardiac glycosides, and tanshinone was deciphered [53,54]. Moreover, it is of significance to study the polymorphisms of structural genes, such as SNPs or inserts/indels, in different cultivars, subspecies or mutants, which could lead to the change of enzyme activity. For example, in *Oryza Sativa*, the *gh2* mutant is a lignin-deficient mutant, and gold hull and internode2 (*GH2*) encodes a cinnamyl-alcohol dehydrogenase [55].

Transporters function in the transportation of various substrates, such as anions, sugars/maltose, amino acids, proteins, mRNAs, electrons, and hormones, which play an important role in biological processes [56]. In this study, we found that 2374 transporter genes were coexpressed with at least structural genes, of which 53 transporters were predicted to be in the vacuole and 23 transporters to be localized in the ER. Because most of the enzyme of GA is localized in the ER, our results indicate that the 23 transporters are potentially involved in the transportation metabolic intermediate of GA. Further experiments may confirm this possibility.

To date, regulatory protein genes involved in the biosynthesis of GA have not been reported. In this study, we found that 1040 TFs, 262 TRs, and 689 PKs genes were correlated with the structural genes of GA for the first time. Previously, TFs have been shown to be involved in terpenoid metabolism. For example, MsMYB, a R2R3-MYB TF, repressed monoterpene production [37]. We also found that TFs of AP2/ERF, bHLH, MYB, NAC, WRKY, and bZIP were potentially associated with terpenoid metabolism, consistent with the findings of a previous report. Among these TFs, R2R3-MYBs, including Glyur000618s00030214, Glyur000049s00003134, and Glyur000266s00 012631, were coexpressed with 10, 10, and 9 structural genes, respectively. Then, we further analyzed the binding site (TAACTG) of R2R3-MYB in the promoter region of the structural genes of GA. Fortunately, this site was found in the promoter regions of six structural genes (Table 2). Therefore, the three R2R3-MYBs were highly likely to interact with these structural genes. To date, no TRs have been reported to be involved in secondary metabolism. For PKs, it is reported that PKs are involved in the biosynthesis of isoflavone [25,26]. The discovery of GA-related TFs, TRs, and PKs supply candidate components for synthetic biology.

miRNAs may regulate the biosynthesis of GA in two aspects. On the one hand, some miRNAs potentially target structural genes, that is, genes encoding regulatory proteins positively related to GA synthesis (Figure 1). To cultivate a new variety of *G. uralensis* with higher GA content, it is necessary to knock out these miRNAs. On the other hand, other miRNAs may target genes encoding regulatory proteins negatively related to GA synthesis. Overexpression of these miRNAs might enhance GA biosynthesis in *G. uralensis*. For the complex regulation of miR156, this miRNA potentially targets 15 TFs, 7 and 5 of which were negatively and positively related to the structural genes of GA, respectively, indicating that GA synthesis may be enhanced by the overexpression of miR156. Further experiments may validate this possibility.

Taken together, these results helped to characterize gene expression and regulation in *G. uralensis*, provided a theoretical basis for the synthesis of GA via synthetic biology, and laid a foundation for cultivating new licorice varieties with high GA content.

## 4. Materials and Methods

### 4.1. Plant Materials and RNA Extraction

Two-year-old *G. uralensis* “B11”, wild accession (unpublished) plants were grown in the experimental field of Beijing Medicinal Plant Garden of the Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences and Peking Union Medical College (Beijing, China) in natural growing conditions. Leaves and roots were collected from three plants (one plant for a biological

repeat) on September 3, 2018, and frozen in liquid nitrogen. At that time, the temperature ranged from 20 to 32 °C. Total RNA from leaves and roots was extracted using the RNAprep Pure Plant Kit (DP171221) (TIANGEN Biotech (Beijing, China) Co., Ltd.) and TRIzol reagent (Invitrogen, Carlsbad, CA, USA), respectively, according to the manufacturer's instructions. Then, RNAs were assessed and quantified as described previously [57].

#### 4.2. Illumina Transcriptome Library Construction and Sequencing

The poly(A) mRNAs of leaf and root samples (every organ with three replicates) were separated from the total RNA using oligo(dT) magnetic beads (Dyna, Oslo). Then, the mRNAs were used to construct strand-specific transcriptome libraries, as in our previous work [57]. Finally, these libraries were sequenced using the Illumina platform.

#### 4.3. Illumina Data Mapping and Quantification of Gene Expression Levels

Raw reads in FASTQ format were first processed using in-house Perl scripts. Clean reads were obtained by removing reads containing adapters, reads containing poly(N) and low-quality reads by Trimmomatic software (v0.30, Aachen, Germany) [58]. Then, these clean reads were mapped to the reference genome of *G. uralensis* using HISAT (v2, Baltimore, MD, USA) [35] with the default parameter. Subsequently, the mapping reads were assembled, and the expression levels were calculated using StringTie (v1, Baltimore, MD, USA) [36]. Compared with roots, those genes that had a  $|\text{fold change}| \geq 2$  and  $\text{FDR} < 0.01$  found by StringTie were assigned as differentially expressed genes (DEGs).

#### 4.4. Gene Function Annotation

To obtain annotation information, the longest transcripts were annotated by comparison with the data in the Nr [59], Swiss-Prot [60], GO [61], COG [62], KOG [63], Pfam [64], and KEGG [65] databases, respectively, by BLAST (v2.2.17) [66]. Compared with previous annotations in the draft genome of *G. uralensis*, new genes were acquired. Then, these new genes were compared with the Pfam database (<http://pfam.sanger.ac.uk/>) using the HMMER software (v3.1b2) [67].

#### 4.5. Identification of TF, TR, PK, and Transporter

To predict candidate TFs, TRs, and PKs in *G. uralensis*, nucleotide sequences for TFs and TRs and translated protein sequences for PKs of the longest transcripts were compared to the sequences in the iTAK database (<http://bioinfo.bti.cornell.edu/cgi-bin/itak/index.cgi>) using the default parameters [21]. To analyze transporters, the protein sequence of the longest transcripts was compared with the sequence in the Transporter Classification Database (TCDB) [68] by BLAST using the default parameter.

#### 4.6. Detection of AS Events

AS events were identified as follows. First, the alignment results of HISAT (v2) [35] were spliced using StringTie software [36], and the alternative splicing type and corresponding expression amount of each sample were obtained using ASprofile software (v2, Baltimore, MD, USA) [43].

#### 4.7. Validation of the Expression Profile of Transcripts by qRT-PCR

The RNA samples, as described for RNA-Seq library construction, were used for qRT-PCR analyses. First, the RNAs were digested with RNase-free DNase I to remove residual genomic DNA during the RNA extraction process. Then, reverse transcription, qRT-PCR and calculation of gene relative expression levels were performed as previously described in He [57]. The primers are listed in Table S30. The Pearson correlation coefficient of the expression profiles was also calculated.

#### 4.8. Small RNA Library Construction and High-Throughput Sequencing

A small RNA library was prepared as described [69]. Briefly, 10–30 nt small RNAs were purified by a 15% denaturing polyacrylamide gel and then ligated to the adapters. After reverse transcription by M-MLV (TaKaRa, Dalian, China), these small RNAs were amplified by PCR. Finally, high-throughput sequencing was performed using the HiSeq NovaSeq platform.

#### 4.9. MiRNA Identification and Target Prediction

After removing adapters and filtering low-quantity sequences by Trimmomatic software [58], the 18–30 nt clean reads of small RNAs in leaves and roots were pooled together, and small RNAs with counts  $\geq 5$  were further compared with the draft genome sequences of *G. uralensis* using the precursor predicted module “psRobot-mir” of the psRobot software (v1.2) [70]. The criteria described by Meyers and Zhang were applied to check *G. uralensis* microRNAs manually [71,72]. The minimal folding free energy index (MFEI) was calculated as described previously [73]. To categorize these miRNAs, we first distinguished and named them according to the precursor. Then, for the analysis of conserved miRNAs, the sequences were compared with the sequences in miRbase, allowing three bases of mismatch. To compare the expression level of miRNAs, the count of each miRNA was normalized to transcripts per million (TPM) following the statistical method as described previously. Compared with the expression level in roots, those miRNAs in leaves that had a  $|\log_2 \text{fold change}| \geq 2$  and  $\text{FDR} < 0.05$  found by DESeq (v2) [74] were assigned as differentially expressed miRNAs (DEmiRNAs). Targets of the miRNAs were predicted by the psRNATarget Server using a penalty score  $\leq 3$  [75].

#### 4.10. Validation of the Authenticity of MiRNAs by Stem-Loop PCR

Approximately 200 ng of DNA-free total RNA was hybridized with a miRNA-specific stem-loop RT primer. Then, the hybridized miRNA molecules were reverse transcribed into cDNA as described [76]. The resulting cDNA was diluted and then PCR-amplified in triplicate using the Bio-Rad CFX96 Real-Time PCR System (Bio-Rad, USA). Each PCR analysis was performed in a final volume of 20  $\mu\text{L}$  containing 10  $\mu\text{L}$  2x SYBR Premix Ex Taq II (TaKaRa, Dalian, China), 0.25  $\mu\text{M}$  each of the miRNA-specific forward primer and the universal reverse primer (Table S31), and cDNA reverse-transcribed from 10 ng total RNA was obtained with the following conditions: 95  $^\circ\text{C}$  for 20 s, 40 cycles of 95  $^\circ\text{C}$  for 10 s, 60  $^\circ\text{C}$  for 10 s and 72  $^\circ\text{C}$  for 6 s. To verify the specificity of PCR amplification, melting curves were analyzed.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/1422-0067/21/9/3101/s1>.

**Author Contributions:** Conceptualization, C.C. and B.W.; Data curation, Y.L., Z.X. and B.W.; Formal analysis, Y.L.; Methodology, Y.L., J.X. and B.W.; Project administration, C.C., B.W. and W.W.; Software, Y.L. and B.W.; Supervision, B.W. and W.W.; Validation, Y.L. and Z.X.; Writing—original draft, Y.L. and B.W.; Writing—review & editing, Y.L., J.X. and B.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** The work was supported by the National Key Research and Development Project (Grant No. 2018YFC1706500, 2017YFC1701400) and Natural Science Foundation of China (81274015). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

**Data Availability:** All the raw data of the transcriptome and small RNAs are deposited in the NCBI short read archive (SRA) under the accession numbers PRJNA555398 and PRJNA555703, respectively.

## References

1. Chinese Pharmacopoeia Commission. *Chinese Pharmacopoeia*; The Chemical Industry Publishing House: Beijing, China, 2005; p. 203. [CrossRef]
2. Montoro, P.; Maldini, M.; Russo, M.; Postorino, S.; Piacente, S.; Pizza, C. Metabolic profiling of roots of liquorice (*Glycyrrhiza glabra*) from different geographical areas by ESI/MS/MS and determination of major metabolites by LC-ESI/MS and LC-ESI/MS/MS. *J. Pharm. Biomed. Anal.* **2011**, *54*, 535–544. [CrossRef] [PubMed]



3. Tang, Z.H.; Li, T.; Tong, Y.G.; Chen, X.J.; Chen, X.P.; Wang, Y.T.; Lu, J.J. A Systematic Review of the Anticancer Properties of Compounds Isolated from Licorice (Gancao). *Planta Med.* **2015**, *81*, 1670–1687. [[CrossRef](#)] [[PubMed](#)]
4. Ramilowski, J.A.; Sawai, S.; Seki, H.; Mochida, K.; Yoshida, T.; Sakurai, T.; Muranaka, T.; Saito, K.; Daub, C.O. *Glycyrrhiza uralensis* transcriptome landscape and study of phytochemicals. *Plant Cell Physiol.* **2013**, *54*, 697–710. [[CrossRef](#)] [[PubMed](#)]
5. Mochida, K.; Sakurai, T.; Seki, H.; Yoshida, T.; Takahagi, K.; Sawai, S.; Uchiyama, H.; Muranaka, T.; Saito, K. Draft genome assembly and annotation of *Glycyrrhiza uralensis*, a medicinal legume. *Plant J.* **2017**, *89*, 181–194. [[CrossRef](#)] [[PubMed](#)]
6. Yang, R.; Yuan, B.C.; Ma, Y.S.; Wang, L.Q.; Liu, C.S.; Liu, Y. HMGR, SQS,  $\beta$ -AS, and Cytochrome P450 Monooxygenase Genes in *Glycyrrhiza uralensis*. *Chin. Herb. Med.* **2015**, *7*, 290–295. [[CrossRef](#)]
7. Peacock, R.B.; Hicks, C.W.; Walker, A.M.; Dewing, S.M.; Lewis, K.M.; Abboud, J.C.; Stewart, S.W.A.; Kang, C.; Watson, J.M. Structural and Functional Characterization of Dynamic Oligomerization in Burkholderia cenocepacia HMG-CoA Reductase. *Biochemistry* **2019**, *58*, 3960–3970. [[CrossRef](#)]
8. Qian, J.; Liu, Y.; Ma, C.; Chao, N.; Chen, Q.; Zhang, Y.; Luo, Y.; Cai, D.; Wu, Y. Positive Selection of Squalene Synthase in *Cucurbitaceae* Plants. *Int. J. Genom.* **2019**, *2019*, 5913491. [[CrossRef](#)]
9. Liu, Y.; Chen, H.H.; Wen, H.; Gao, Y.; Wang, L.Q.; Liu, C.S. Enhancing the accumulation of beta-amyrin in *Saccharomyces cerevisiae* by co-expression of *Glycyrrhiza uralensis* squalene synthase 1 and beta-amyrin synthase genes. *Acta Pharm. Sin.* **2014**, *49*, 734–741.
10. Seki, H.; Sawai, S.; Ohyama, K.; Mizutani, M.; Ohnishi, T.; Sudo, H.; Fukushima, E.O.; Akashi, T.; Aoki, T.; Saito, K.; et al. Triterpene functional genomics in licorice for identification of CYP72A154 involved in the biosynthesis of glycyrrhizin. *Plant Cell* **2011**, *23*, 4112–4123. [[CrossRef](#)]
11. Xu, G.; Cai, W.; Gao, W.; Liu, C. A novel glucuronosyltransferase has an unprecedented ability to catalyse continuous two-step glucuronosylation of glycyrrhetic acid to yield glycyrrhizin. *New Phytol.* **2016**, *212*, 123–135. [[CrossRef](#)]
12. Chu, S.; Wang, J.; Zhu, Y.; Liu, S.; Zhou, X.; Zhang, H.; Wang, C.E.; Yang, W.; Tian, Z.; Cheng, H.; et al. An R2R3-type MYB transcription factor, *GmMYB29*, regulates isoflavone biosynthesis in soybean. *PLoS Genet.* **2017**, *13*, e1006770. [[CrossRef](#)] [[PubMed](#)]
13. Rushton, P.J.; Somssich, I.E.; Ringler, P.; Shen, Q.J. WRKY transcription factors. *Trends Plant Sci.* **2010**, *15*, 247–258. [[CrossRef](#)]
14. Liu, X.; Yuan, L.; Xu, L.; Xu, Z. Over-expression of *GmMYB39* leads to an inhibition of the isoflavonoid biosynthesis in soybean (*Glycine max.* L). *Plant Biotechnol. Rep.* **2013**, *7*, 445. [[CrossRef](#)]
15. Wang, T.; Jia, Q.; Wang, W.; Hussain, S.; Ahmed, S.; Zhou, D.X.; Ni, Z.; Wang, S. GCN5 modulates trichome initiation in *Arabidopsis* by manipulating histone acetylation of core trichome initiation regulator genes. *Plant Cell Rep.* **2019**, *38*, 755–765. [[CrossRef](#)] [[PubMed](#)]
16. Zhang, D.; Li, Y.; Zhang, X.; Zha, P.; Lin, R. The SWI2/SNF2 Chromatin-Remodeling ATPase BRAHMA Regulates Chlorophyll Biosynthesis in *Arabidopsis*. *Mol. Plant* **2017**, *10*, 155–167. [[CrossRef](#)] [[PubMed](#)]
17. Kim, D.H.; Sung, S. Accelerated vernalization response by an altered PHD-finger protein in *Arabidopsis*. *Plant Signal. Behav.* **2017**, *12*, e1308619. [[CrossRef](#)] [[PubMed](#)]
18. Chandler, J.W. Auxin response factors. *Plant Cell Environ.* **2016**, *39*, 1014–1028. [[CrossRef](#)] [[PubMed](#)]
19. Cho, E.J.; Choi, S.H.; Kim, J.H.; Kim, J.E.; Lee, M.H.; Chung, B.Y.; Woo, H.R. A Mutation in Plant-Specific SWI2/SNF2-Like Chromatin-Remodeling Proteins, DRD1 and DDM1, Delays Leaf Senescence in *Arabidopsis thaliana*. *PLoS ONE* **2016**, *11*, e0146826. [[CrossRef](#)]
20. Chen, W.; Zhu, Q.; Liu, Y.; Zhang, Q. Chromatin Remodeling and Plant Immunity. *Adv. Protein Chem. Struct. Biol.* **2017**, *106*, 243–260.
21. Zheng, Y.; Jiao, C.; Sun, H.; Rosli, H.G.; Pombo, M.A.; Zhang, P.; Banf, M.; Dai, X.; Martin, G.B.; Giovannoni, J.J.; et al. iTAK: A Program for Genome-wide Prediction and Classification of Plant Transcription Factors, Transcriptional Regulators, and Protein Kinases. *Mol. Plant* **2016**, *9*, 1667–1670. [[CrossRef](#)]
22. Garcia, A.V.; Al-Yousif, M.; Hirt, H. Role of AGC kinases in plant growth and stress responses. *Cell Mol. Life Sci.* **2012**, *69*, 3259–3267. [[CrossRef](#)] [[PubMed](#)]
23. Tan, W.; Zhang, D.; Zhou, H.; Zheng, T.; Yin, Y.; Lin, H. Transcription factor HAT1 is a substrate of SnRK2.3 kinase and negatively regulates ABA synthesis and signaling in *Arabidopsis* responding to drought. *PLoS Genet.* **2018**, *14*, e1007336. [[CrossRef](#)] [[PubMed](#)]

24. Luang, S.; Sornaraj, P.; Bazanova, N.; Jia, W.; Eini, O.; Hussain, S.S.; Kovalchuk, N.; Agarwal, P.K.; Hrmova, M.; Lopato, S. The wheat *TabZIP2* transcription factor is activated by the nutrient starvation-responsive SnRK3/CIPK protein kinase. *Plant Mol. Biol.* **2018**, *96*, 543–561. [[CrossRef](#)] [[PubMed](#)]
25. Jiao, C.; Gu, Z. Cyclic GMP mediates abscisic acid-stimulated isoflavone synthesis in soybean sprouts. *Food Chem.* **2019**, *275*, 439–445. [[CrossRef](#)]
26. Veremeichik, G.N.; Grigorchuk, V.P.; Silanteva, S.A.; Shkryl, Y.N.; Bulgakov, D.V.; Brodovskaya, E.V.; Bulgakov, V.P. Increase in isoflavonoid content in *Glycine max* cells transformed by the constitutively active Ca(2+) independent form of the *AtCPK1* gene. *Phytochemistry* **2019**, *157*, 111–120. [[CrossRef](#)]
27. Bartel, D.P. MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* **2004**, *116*, 281–297. [[CrossRef](#)]
28. Ben Chaabane, S.; Liu, R.; Chinnusamy, V.; Kwon, Y.; Park, J.H.; Kim, S.Y.; Zhu, J.K.; Yang, S.W.; Lee, B.H. STA1, an *Arabidopsis* pre-mRNA processing factor 6 homolog, is a new player involved in miRNA biogenesis. *Nucleic Acids Res.* **2013**, *41*, 1984–1997. [[CrossRef](#)]
29. Ren, G.; Yu, B. Critical roles of RNA-binding proteins in miRNA biogenesis in *Arabidopsis*. *RNA Biol.* **2012**, *9*, 1424–1428. [[CrossRef](#)]
30. Mallory, A.C.; Vaucheret, H. Functions of microRNAs and related small RNAs in plants. *Nat. Genet.* **2006**, *38*, S31–S36. [[CrossRef](#)]
31. Brodersen, P.; Sakvarelidze-Achard, L.; Bruun-Rasmussen, M.; Dunoyer, P.; Yamamoto, Y.Y.; Sieburth, L.; Voinnet, O. Widespread translational inhibition by plant miRNAs and siRNAs. *Science* **2008**, *320*, 1185–1190. [[CrossRef](#)]
32. Wu, L.; Zhou, H.; Zhang, Q.; Zhang, J.; Ni, F.; Liu, C.; Qi, Y. DNA methylation mediated by a microRNA pathway. *Mol. Cell* **2010**, *38*, 465–475. [[CrossRef](#)] [[PubMed](#)]
33. Sunkar, R.; Zhu, J.K. Novel and stress-regulated microRNAs and other small RNAs from *Arabidopsis*. *Plant Cell* **2004**, *16*, 2001–2019. [[CrossRef](#)] [[PubMed](#)]
34. Ng, D.W.; Zhang, C.; Miller, M.; Palmer, G.; Whiteley, M.; Tholl, D.; Chen, Z.J. cis- and trans-Regulation of miR163 and target genes confers natural variation of secondary metabolites in two *Arabidopsis* species and their allopolyploids. *Plant Cell* **2011**, *23*, 1729–1740. [[CrossRef](#)] [[PubMed](#)]
35. Kim, T.; Seo, H.D.; Hennighausen, L.; Lee, D.; Kang, K. Octopus-toolkit: A workflow to automate mining of public epigenomic and transcriptomic next-generation sequencing data. *Nucleic Acids Res.* **2018**, *46*, e53. [[CrossRef](#)] [[PubMed](#)]
36. Perteua, M.; Perteua, G.M.; Antonescu, C.M.; Chang, T.C.; Mendell, J.T.; Salzberg, S.L. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **2015**, *33*, 290–295. [[CrossRef](#)]
37. Reddy, V.A.; Wang, Q.; Dhar, N.; Kumar, N.; Venkatesh, P.N.; Rajan, C.; Panicker, D.; Sridhar, V.; Mao, H.Z.; Sarojam, R. Spearmint R2R3-MYB transcription factor *MsMYB* negatively regulates monoterpene production and suppresses the expression of geranyl diphosphate synthase large subunit (*MsGPPS.LSU*). *Plant Biotechnol. J.* **2017**, *15*, 1105–1119. [[CrossRef](#)] [[PubMed](#)]
38. Nieuwenhuizen, N.J.; Chen, X.; Wang, M.Y.; Matich, A.J.; Perez, R.L.; Allan, A.C.; Green, S.A.; Atkinson, R.G. Natural variation in monoterpene synthesis in kiwifruit: Transcriptional regulation of terpene synthases by NAC and ETHYLENE-INSENSITIVE3-like transcription factors. *Plant Physiol.* **2015**, *167*, 1243–1258. [[CrossRef](#)]
39. Yan-Hua, X.; Jia-Wei, W.; Shui, W.; Jian-Ying, W.; Xiao-Ya, C. Characterization of *GaWRKY1*, a cotton transcription factor that regulates the sesquiterpene synthase gene (+)-delta-cadinene synthase-A. *Plant Physiol.* **2004**, *135*, 507–515.
40. Miyamoto, K.; Nishizawa, Y.; Minami, E.; Nojiri, H.; Yamane, H.; Okada, K. Overexpression of the bZIP transcription factor *OsbZIP79* suppresses the production of diterpenoid phytoalexin in rice cells. *J. Plant Physiol.* **2015**, *173*, 19–27. [[CrossRef](#)]
41. Riechmann, J.L.; Heard, J.; Martin, G.; Reuber, L.; Jiang, C.; Keddie, J.; Adam, L.; Pineda, O.; Ratcliffe, O.J.; Samaha, R.R.; et al. *Arabidopsis* transcription factors: Genome-wide comparative analysis among eukaryotes. *Science* **2000**, *290*, 2105–2110. [[CrossRef](#)]
42. Stracke, R.; Werber, M.; Weisshaar, B. The R2R3-MYB gene family in *Arabidopsis thaliana*. *Curr. Opin. Plant Biol.* **2001**, *4*, 447–456. [[CrossRef](#)]
43. Florea, L.; Song, L.; Salzberg, S.L. Thousands of exon skipping events differentiate among splicing patterns in sixteen human tissues. *F1000Research* **2013**, *2*, 188. [[CrossRef](#)] [[PubMed](#)]

44. Lu, C.; Tej, S.S.; Luo, S.; Haudenschild, C.D.; Meyers, B.C.; Green, P.J. Elucidation of the small RNA component of the transcriptome. *Science* **2005**, *309*, 1567–1569. [[CrossRef](#)] [[PubMed](#)]
45. Morin, R.D.; Aksay, G.; Dolgosheina, E.; Ebhardt, H.A.; Magrini, V.; Mardis, E.R.; Sahinalp, S.C.; Unrau, P.J. Comparative analysis of the small RNA transcriptomes of *Pinus contorta* and *Oryza sativa*. *Genome Res.* **2008**, *18*, 571–584. [[CrossRef](#)] [[PubMed](#)]
46. Wang, L. Identification and characterization of maize microRNAs involved in the very early stage of seed germination. *BMC Genom.* **2011**, *12*, 154. [[CrossRef](#)] [[PubMed](#)]
47. Arshad, M.; Feyissa, B.A.; Amyot, L.; Aung, B.; Hannoufa, A. MicroRNA156 improves drought stress tolerance in alfalfa (*Medicago sativa*) by silencing SPL13. *Plant Sci.* **2017**, *258*, 122–136. [[CrossRef](#)]
48. Aung, B.; Gruber, M.Y.; Amyot, L.; Omari, K.; Bertrand, A.; Hannoufa, A. MicroRNA156 as a promising tool for alfalfa improvement. *Plant Biotechnol. J.* **2015**, *13*, 779–790. [[CrossRef](#)]
49. Zhang, L.; Wu, B.; Zhao, D.; Li, C.; Shao, F.; Lu, S. Genome-wide analysis and molecular dissection of the SPL gene family in *Salvia miltiorrhiza*. *J. Integr. Plant Biol.* **2014**, *56*, 38–50. [[CrossRef](#)]
50. Cui, G.; Duan, L.; Jin, B.; Qian, J.; Xue, Z.; Shen, G.; Snyder, J.H.; Song, J.; Chen, S.; Huang, L.; et al. Functional Divergence of Diterpene Syntheses in the Medicinal Plant *Salvia miltiorrhiza*. *Plant Physiol.* **2015**, *169*, 1607–1618. [[CrossRef](#)]
51. Holland, N.; Holland, D.; Helentjaris, T.; Dhugga, K.S.; Xoconostle-Cazares, B.; Delmer, D.P. A comparative analysis of the plant cellulose synthase (CesA) gene family. *Plant Physiol.* **2000**, *123*, 1313–1324. [[CrossRef](#)]
52. Taylor, N.G.; Howells, R.M.; Huttly, A.K.; Vickers, K.; Turner, S.R. Interactions among three distinct CesA proteins essential for cellulose synthesis. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 1450–1455. [[CrossRef](#)] [[PubMed](#)]
53. Wu, B.; Suo, F.; Lei, W.; Gu, L. Comprehensive analysis of alternative splicing in *Digitalis purpurea* by strand-specific RNA-Seq. *PLoS ONE* **2014**, *9*, e106001. [[CrossRef](#)] [[PubMed](#)]
54. Xu, Z.; Peters, R.J.; Weirather, J.; Luo, H.; Liao, B.; Zhang, X.; Zhu, Y.; Ji, A.; Zhang, B.; Hu, S.; et al. Full-length transcriptome sequences and splice variants obtained by a combination of sequencing platforms applied to different root tissues of *Salvia miltiorrhiza* and tanshinone biosynthesis. *Plant J.* **2015**, *82*, 951–961. [[CrossRef](#)]
55. Zhang, K.; Qian, Q.; Huang, Z.; Wang, Y.; Li, M.; Hong, L.; Zeng, D.; Gu, M.; Chu, C.; Cheng, Z. GOLD HULL AND INTERNODE2 encodes a primarily multifunctional cinnamyl-alcohol dehydrogenase in rice. *Plant Physiol.* **2006**, *140*, 972–983. [[CrossRef](#)] [[PubMed](#)]
56. Mishra, N.K.; Chang, J.; Zhao, P.X. Prediction of membrane transport proteins and their substrate specificities using primary sequence information. *PLoS ONE* **2014**, *9*, e100278. [[CrossRef](#)] [[PubMed](#)]
57. He, M.; Yao, Y.; Li, Y.; Yang, M.; Wu, B.; Yu, D. Comprehensive transcriptome analysis reveals genes potentially involved in isoflavone biosynthesis in *Pueraria thomsonii* Benth. *PLoS ONE* **2019**, *14*, e0217593. [[CrossRef](#)]
58. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)]
59. Deng, Y.; Jianqi, L.I.; Songfeng, W.U.; Zhu, Y.; Chen, Y.; Fuchu, H.E. Integrated nr Database in Protein Annotation System and Its Localization. *Comput. Eng.* **2006**, *32*, 71–72.
60. Rolf, A.; Amos, B.; Wu, C.H.; Barker, W.C.; Brigitte, B.; Serenella, F.; Elisabeth, G.; Hongzhan, H.; Rodrigo, L.; Michele, M. UniProt: The Universal Protein knowledgebase. *Nucleic Acids Res.* **2004**, *32*, D115–D119.
61. Ashburner, M.; Ball, C.A.; Blake, J.A.; Botstein, D.; Butler, H.; Cherry, J.M.; Davis, A.P.; Dolinski, K.; Dwight, S.S.; Eppig, J.T.; et al. Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **2000**, *25*, 25–29. [[CrossRef](#)]
62. Tatusov, R.L.; Galperin, M.Y.; Natale, D.A.; Koonin, E.V. The COG database: A tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* **2000**, *28*, 33–36. [[CrossRef](#)] [[PubMed](#)]
63. Koonin, E.V.; Fedorova, N.D.; Jackson, J.D.; Jacobs, A.R.; Krylov, D.M.; Makarova, K.S.; Mazumder, R.; Mekhedov, S.L.; Nikolskaya, A.N.; Rao, B.S.; et al. A comprehensive evolutionary classification of proteins encoded in complete eukaryotic genomes. *Genome Biol.* **2004**, *5*, R7. [[CrossRef](#)] [[PubMed](#)]
64. Finn, R.D.; Bateman, A.; Clements, J.; Coggill, P.; Eberhardt, R.Y.; Eddy, S.R.; Heger, A.; Hetherington, K.; Holm, L.; Mistry, J.; et al. Pfam: The protein families database. *Nucleic Acids Res.* **2014**, *42*, D222–D230. [[CrossRef](#)] [[PubMed](#)]
65. Kanehisa, M.; Goto, S.; Kawashima, S.; Okuno, Y.; Hattori, M. The KEGG resource for deciphering the genome. *Nucleic Acids Res.* **2004**, *32*, D277–D280. [[CrossRef](#)]

66. Altschul, S.F.; Madden, T.L.; Schaffer, A.A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D.J. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **1997**, *25*, 3389–3402. [[CrossRef](#)]
67. Eddy, S.R. Profile hidden Markov models. *Bioinformatics* **1998**, *14*, 755–763. [[CrossRef](#)]
68. Saier, M.H., Jr.; Reddy, V.S.; Tsu, B.V.; Ahmed, M.S.; Li, C.; Moreno-Hagelsieb, G. The Transporter Classification Database (TCDB): Recent advances. *Nucleic Acids Res.* **2016**, *44*, D372–D379. [[CrossRef](#)]
69. Hafner, M.; Landgraf, P.; Ludwig, J.; Rice, A.; Ojo, T.; Lin, C.; Holoch, D.; Lim, C.; Tuschl, T. Identification of microRNAs and other small regulatory RNAs using cDNA library sequencing. *Methods* **2008**, *44*, 3–12. [[CrossRef](#)]
70. Wu, H.J.; Ma, Y.K.; Chen, T.; Wang, M.; Wang, X.J. PsRobot: A web-based plant small RNA meta-analysis toolbox. *Nucleic Acids Res.* **2012**, *40*, W22–W28. [[CrossRef](#)]
71. Meyers, B.C.; Axtell, M.J.; Bartel, B.; Bartel, D.P.; Baulcombe, D.; Bowman, J.L.; Cao, X.; Carrington, J.C.; Chen, X.; Green, P.J.; et al. Criteria for annotation of plant MicroRNAs. *Plant Cell* **2008**, *20*, 3186–3190. [[CrossRef](#)]
72. Zhang, B.; Pan, X.; Anderson, T.A. Identification of 188 conserved maize microRNAs and their targets. *FEBS Lett.* **2006**, *580*, 3753–3762. [[CrossRef](#)] [[PubMed](#)]
73. Zhang, B.H.; Pan, X.P.; Cox, S.B.; Cobb, G.P.; Anderson, T.A. Evidence that miRNAs are different from other RNAs. *Cell. Mol. Life Sci. CMLS* **2006**, *63*, 246–254. [[CrossRef](#)] [[PubMed](#)]
74. Anders, S.; Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **2010**, *11*, R106. [[CrossRef](#)] [[PubMed](#)]
75. Dai, X.; Zhuang, Z.; Zhao, P.X. psRNATarget: A plant small RNA target analysis server (2017 release). *Nucleic Acids Res.* **2018**, *46*, W49–W54. [[CrossRef](#)]
76. Chen, C.; Ridzon, D.A.; Broomer, A.J.; Zhou, Z.; Lee, D.H.; Nguyen, J.T.; Barbisin, M.; Xu, N.L.; Mahuvakar, V.R.; Andersen, M.R.; et al. Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res.* **2005**, *33*, e179. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).