Comparative Transcriptome Analysis of the Roots and Leaves of *Bupleurum Chinense* DC. Seedlings Under Drought Stress

Ya Fan
Shandong University of Traditional Chinese Medicine (TCM)
Yannan Geng
Affiliated Hospital of Shandong University of TCM
Hailu Cao
Hengde Bencao (Beijing) Agricultural Technology Co., LTD
Li Liu
Shandong University of Traditional Chinese Medicine (TCM)
Quanfang Zhang
Shandong Academy of Agricultural Sciences
Sun Yan
Shandong University of Traditional Chinese Medicine (TCM)
Du Kan
Shandong University of Traditional Chinese Medicine (TCM)
Ping Wang
Shandong University of TCM
Demin Gao (gdm607@126.com)
Shandong University of Traditional Chinese Medicine (TCM)

Research Article

Keywords: Bupleurum chinense, drought stress, transcriptome analysis, differentially expressed genes, biological pathway

Posted Date: December 7th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1088535/v1

Abstract

Background

drought stress is one of the important environmental factors affecting the quality and yield of medicinal materials, and is the main factor restricting the field production of *Bupleurum chinense*. *B. chinense* seedlings sensitive to low moisture, but there are few reports on the molecular mechanism of *B. chinense* seedlings under drought stress. Therefore, the transcriptome of the leaves and roots of *B. chinense* seedlings before and after drought were analyzed by Illumina sequencing technology and bioinformatics analysis.

Results

a total of 59.82 GB of clean data was obtained, and the unigenes were compared with Nr, Swissprot, String, GO, KEGG, and Pfam databases. Under drought stress, 3,737 and 6,816 differentially expressed genes (DEGs) were identified in leaves and roots of *B. chinense*, respectively. The obtained DEGs from leaves and roots were classified into 37, and 36 GO terms and were involved in 222 and 253 KEGG pathways, respectively. SSR analysis were obtained identified 33,728 loci, wherein dinucleotides accounted for the largest proportion. Genes involved in diterpenoid and unsaturated fatty acid biosynthesis were significantly over-expressed in roots under drought stress, suggesting these two cellular processes underpin the adaptation and resistance of *B. chinense* seedlings to drought stress.

Conclusions

the results provided a theoretical basis for further identification of the molecular mechanism of drought resistance and breeding of drought resistance varieties of *B. chinense.*

Background

During the growth and development of plants, they are often subjected to the stress of abiotic factors that lead to plant yield reduction or even death [1]. Drought stress is the most common abiotic stress affecting plants [2, 3], with the earliest effects on the roots as demonstrated with *Arabidopsis thaliana* [4]. Global warming means drought will occur more frequently, which will seriously affect the production of medicinal plants [5]. Transcriptional changes under drought stress can cause major molecular and physiological changes in plants, including rapid regulation of transcription and metabolism, regulation of osmotic potential, and decrease of leaf turgor pressure, ultimately slowing down or ceasing plant growth [6]. Drought stress also affects the synthesis of secondary metabolites of plants by changing gene transcription. Comparing differentially expressed genes (DEGs) of stressed and non-stressed plants can reveal important genes regulating drought resistance [7]. However, plant response to external drought is a very complex process involving the expression of many genes and the regulation of multiple signaling pathways [8].

Bupleurum chinense DC. is a perennial herb in the family Apiaceae. Its root, known as Radix Bupleuri, has been used in traditional Chinese medicine for more than 2000 years [9], where it has the effects of dispersing and reducing fever, soothing the liver and relieving depression, and lifting Yang Qi [10]. The main active components of Radix Bupleuri are saikosaponins a, c, and d (SSA, SSC, and SSD, respectively), which have immunomodulatory, anti-inflammatory, antioxidant, and hepatoprotective activities [11–13]. The synthesis and accumulation of saikosaponins are affected by environmental factors, such as soil fertility, light, and drought stress. It was found that application of nitrogen and phosphorus fertilizers increased the yields of SSA and SSD in arid and semi-arid regions [14], while low light activation of the antioxidant stress response pathway also increased the total amount of SSA and SSD yields [15]. Further studies showed that β-AS, P450, and UGT family genes are involved in the synthesis of saikosaponins, and regulation of these genes is expected to improve the yield of saikosaponins [16].

Sui carried out transcriptome analysis of *B. chinense* roots by 454GS-FLX pyrosequencing, providing a reference for exploring the important genes involved in biosynthesis and regulation of bioactive components in *B. chinense* [17]. Under drought stress, the activities of SOD, POD, and CAT in *B. chinense* roots significantly decreased, and the short-term expression of key enzymes encoding genes involved in the synthesis of secondary metabolites increased, suggesting that short-term pressure regulation may improve the quality of medicinal materials [18]. Subsequent studies found that squalene epoxidase and *β*-amyrin synthase genes of *B. chinense* may cooperate to affect the accumulation of saikosaponins in roots under drought stress [19]. These molecular and physiological studies were mostly related to the adult plant of *B. chinense*, while the study on the seedling has not been reported. Seedling is more sensitive to drought stress, because it is the critical period of *B. chinense* growth. However, transcriptome analysis of roots and leaves of *B. chinense* seedlings under drought stress has not been studied in detail.

In this study, we conducted transcriptional analysis of roots and leaves of *B. chinense* seedlings under drought stress, and identified DEGs and metabolic pathways related to drought stress response in the early growth stage. These data help us to better understand the biological pathways and mechanism of drought stress in *B. chinense*.

Results

Transcriptional assembly and sequencing results

The RNA extracted from 12 samples of *B. chinense* had $OD_{260/280}$ values all above 2.0, indicating that the quality of RNA was good and could be used for transcriptomic analysis (Table 1). A total of 59.82 Gb of clean data was obtained in this experiment (Table 2), and the average amount of clean data of each sample was 4.98 Gb, in which the percentage of Q30 bases was more than 95.48%, and the GC content was between 44.04% and 45.55%. The clean data of all samples were assembled *de novo* using Trinity, and the assembly results were optimized and evaluated. The number of unigenes assembled was 314,788, and the average length of N50 was 789 bp. The number unigenes with length from 1 to 600 bp was 229,595, accounting for 72.94% of the total, while 85,193 unigenes with length from 601 to 2000 bp, accounting for 27.06%. Transcriptome sequencing data is qualified and can be used for further analysis (Fig. 1).

Table 1
RNA quality test results of 12 *Bupleurum chinense* samples

Samples	Concentration (ng/μL)	$OD_{260/280}$	$OD_{260/230}$
BL-1a	771.1	2.18	2.07
BL-1b	734.1	2.17	2.32
BL-1c	650.6	2.14	2.24
BR-1a	194.4	2.15	1.82
BR-1b	319.2	2.12	2.24
BR-1c	412.3	2.09	2.03
BDL-1a	940.2	2.17	2.09
BDL-1b	1138.6	2.14	2.36
BDL-1c	801.5	2.17	2.27
BDR-1a	312.2	2.08	1.96
BDR-1b	275.5	2.07	1.92
BDR-1c	313.0	1.97	1. 94

Note: BL: The leaf of *B. chinense* BR: The root *B. chinense.* BDL: The drought leaf *B. chinense*. BDR: The drought root *B. chinense*.

Table 2
Summary of RNA-Seq data of *Bupleurum chinense*

Samples	Raw reads number	Clean reads number	Clean bases (bp)	Q20 percentage (%)	Q30 percentage (%)
BL-1a	35,234,272	34,545,236 (98.04%)	5,037,919,736	98.85	95.51
BL-1b	39,567,542	38,911,926 (98.34%)	5,650,054,266	98.95	95.80
BL-1c	44,472,456	43,680,856 (98.22%)	6,364,577,275	98.92	95.72
BR-1a	37,035,734	36,399,496 (98.28%)	5,313,535,634	98.89	95.61
BR-1b	37,419,172	36,733,800 (98.17%)	5,321,935,283	98.90	95.67
BR-1c	36,962,800	36,285,118 (98.17%)	5,244,354,463	98.84	95.48
BDL-1a	26,185,802	25,734,308 (98.28%)	3,751,987,716	98.87	95.57
BDL-1b	36,947,406	36,338,764 (98.35%)	5,302,486,665	98.96	95.85
BDL-1c	37,896,440	37,165,906 (98.07%)	5,388,140,064	98.88	95.59
BDR-1a	38,098,270	37,457,442 (98.32%)	5,447,888,065	98.91	95.69
BDR-1b	36,410,500	35,717,738 (98.10%)	5,168,071,209	98.84	95.49
BDR-1c	43,882,736	43,128,824 (98.28%)	6,241,238,261	98.90	95.64

Note: BL: The leaf of *B. chinense* BR: The root *B. chinense*. BDL: The drought leaf *B. chinense*. BDR: The drought root *B. chinense*. Raw reads number: count the number of raw sequence data. Clean reads number: the filtered sequencing data. Clean bases: the total number of bases in clean data. Q20, Q30: respectively calculate the Phred value greater than 20, 30 bases accounted for the percentage of total bases (clean data).

Functional annotation of unigenes

The number of unigenes in each databases annotation is shown in Table 3. The Nr database' top hits for these unigenes were primarily from *Quercus suber* (18,435; 20.92%) of Fagaceae, *Daucus carota* subsp. sativus (16,810, 19.07%) of Apiaceae, and *Carpinus fangiana* (6,804, 7.72%) of Betulaceae (Fig. 2a).

Table 3
Functional annotation of *Bupleurum chinense* unigenes

Annotated in public database	Number of unigenes	Percentage (%)
GO	41025	13.03
KEGG	36452	11.58
Nr	88133	28.00
Pfam	46058	14.63
String	1001	0.32
Swissprot	41482	13.18
Common in all	320	0.10

In KOG Annotation (Fig. 2b), 926 sequences were divided into 25 categories, of which the three most common categories were translation, ribosomal structure and biogenesis; general function prediction; and posttranslational modification, protein turnover, and chaperones.

Analysis of DEGs

The transcriptome data of leaves and roots of *B. chinense* treated in the control group and the drought-stressed group were compared. A total of 3,737 DEGs were identified in the drought-treated leaves, of which 1,775 (47.50%) were up-regulated and

1,962 (52.50%) were down-regulated (Fig. 3a). A total of 6,816 DEGs were identified in the drought-treated roots, of which 4,058 (59.54%) were up-regulated and 2,758 (40.46%) were down-regulated (Fig. 3b).

GO annotation and enrichment analysis of DEGs

Go function enrichment analysis of DEGs obtained in the experiment showed that 1,225 and 1,815 DEGs were distributed in 4,265 and 4,470 GO terms in leaves and roots after drought stress, respectively.

According to the secondary graph of GO analysis (Fig. 4), DEGs in the leaves were enriched in 37 GO terms, and DEGs in the roots in 36 GO terms. Biological process (BP) had the most DEGs in roots and leaves, followed by molecular function (MF) and cellular component (CC). The main BP subterms for the DEGs in both roots and leaves after drought treatment were biological regulation, cellular process, metabolic process, and response to stimulus. Binding signals and catalytic activity were the most enriched MF subterms, and cellular anatomical entities were the main CC subterms. General transcription initiation factor activity, nutrient reserve activity, and toxin activity were the GO terms uniquely enriched in leaves. Pigmentation and small molecular sensor activity were the unique GO terms in roots. Twice as many genes were up-regulated in roots as in leaves.

KEGG enrichment analysis of DEGs

KEGG enrichment analysis was used to identify the DEGs to their biochemical metabolic pathways and signal transduction pathways. After drought stress, 335 and 470 DEGs were distributed in 222 and 253 metabolic pathways in leaves and roots, respectively.

KEGG enrichment analysis showed the significant DEGs in *B. chinense* seedlings after drought treatment. (Fig. 5). The common pathway in leaves and roots was carbon metabolism and plant hormone signal transduction.

Carbon fixation in photosynthetic organizations, biosynthesis of secondary metabolites, photosynthesis, and phenylpropanoid biosynthesis were the most enriched pathways in leaves after drought stress. The pathways significantly enriched in roots after drought stress included dieterpenoid biosynthesis, citrate cycle (TCA cycle), neurotrophin signaling pathway, microbial metabolism in different environments, PPAR signaling pathway, biosynthesis of unsaturated fatty acids, microbial metabolism in different environments, insulin signaling pathway, and pyruvate metabolism. Drought stress affected the biosynthesis of phenylpropanoid only in leaves, and diterpenoid and unsaturated fatty acids only in roots.

SSR analysis

SSR analysis of the transcriptome unigenes of *B. chinense* seedlings was conducted using MISA software (Table 4), and a total of 33,728 SSR loci were found. Among them, dinucleotide was the most abundant SSR, with a total of 16,235 (48.14%), followed by mononucleotide and trinucleotide, with 9,112 (27.02%) and 7,924 (23.49%) respectively. The number of tetranucleotide, pentanucleotide, and hexanucleotide was very small, with a total of 457 (1.35%). The proportion of dinucleotide repeats in *B. chinense* was high, which was consistent with previous studies of Apiaceae [20].

Table 4
SSR analysis of transcriptome unigenes of *Bupleurum chinense* seedlings

Type	Number of repeating units												
	5	6	7	8	9	10	11	12	13	14	15	other	Total
Mononucleotide	0	0	0	0	0	3995	1863	1084	659	502	334	675	9112
Dinucleotide	0	5248	3282	2536	1829	1255	791	513	110	108	110	453	16235
Trinucleotide	4464	1812	787	525	78	95	56	33	17	21	10	26	7924
Tetranucleotide	201	97	8	14	8	3	6	3	6	2	0	3	351
Pentanucleotide	32	3	4	1	1	2	0	0	0	0	0	0	43
Hexanucleotide	25	14	14	8	2	0	0	0	0	0	0	0	63

qPCR verification of DEGs

Seven DEGs with significant differences were selected for qPCR verification of the expression pattern in the roots and leaves of *B. chinense* (Table 5). The results showed that the transcriptome (Fig. 6-a) and qPCR (Fig. 6-b) experiments showed similar expression patterns, indicating that the transcriptome results are reliable for the identification of DEGs in this study.

Table 5
Primer sequences used for qPCR

Gene name	Primer sequences (5′-3′)	Primer sequences (3′-5′)
DN2387_c0_g2	CTCAGCAAGACAGAATGCGG	ATCCTCCAAAGCCGAATCCC
DN10217_c0_g1	TCAGGCTCGACGGATTGATG	TCCGCCGACACAAATCTCAA
DN233_c1_g1	CTCCCTCCACCACCCATCTA	GCTGGGGTGAATAGAGAGGC
DN70694_c0_g1	CAACACACACACTCCCCTCA	CTCTGGCTGCATGCTTGTTG
DN1634_c0_g2	TCCACCAGACCAGCAACGAT	CACGGAGACGAAGCACAAGG
DN1227_c0_g2	GAGGAGCACAAGCAAAAGGC	GCCCTCCAGCTCCTTCAAAT
DN10971_c0_g1	CACCCTGTTGGAGTTGTGGA	ACACCACCACCACACACATA
Actin	CCCGATGGTCAAGTTATCACC	TTCCTGCAGCTTCCATTCCA

Discussion

In this study, a large number of useful gene sequences were obtained from leaves and roots transcriptome of *B. chinense* under drought stress using a high-throughput technique. These results provide abundant genetic resources for further analysis of functional annotation and metabolic pathway under drought stress. In addition, a total of 33,728 SSR loci were identified from the transcriptomic data of *B. chinense*, which provided a rich theoretical basis for further research on genetic diversity, molecular marker-assisted breeding, and genetic map construction for this species. The data also provides abundant genetic resources for further functional annotation and analysis of metabolic pathways associated with drought stress.

The genes plants over-express under drought stress fall into three categories. One is to express genes related to detoxification and antioxidant stress, such as superoxide dismutase synthesis genes. Secondly, genes are related to the synthesis of osmotic regulatory substances, such as proline synthesis genes. The third is to protect the expression of genes protecting the expression of biological macromolecules and membrane structure, such as chaperones proteins and post-embryogenic rich proteins [21–23]. In this study, the active genes involved in scavenging reactive oxygen and superoxide in both leaves and roots of *B. chinense*, such as superoxide dismutase activity, peroxidase activity, glutathione peroxidase activity and glutathione transferase were over-expressed. Among them, the genes involved in coding peroxidase activity have bidirectional regulation, indicating that peroxidase activity scavenging reactive oxygen species is one of the more important mechanisms of drought resistance [24].

Previous studies have shown that drought stress can regulate hormone synthesis and signal transduction *in vivo* [25–27]. Under drought stress, the genes encoding plant hormone signal transduction were significantly up-regulated in the leaves and roots of *B. chinense.* In addition, drought stress had a bidirectional regulation effect on DEGs encoding biosynthesis and metabolism of auxin, gibberellin, and cytokinin. For examples, drought stress significantly down-regulated expression of gibberellin20-oxidase (GA20ox) (TRINITY_DN2977_c0_g2), a key rate-limiting enzyme encoding gibberellin in *B. chinense*, especially in roots. GA20ox is a multifunctional enzyme that not only participates in GA biosynthesis, but also controls the synthesis of GA1 and GA4, and maintains the dynamic balance of gibberellin in cells [28, 29]. Spielmeyer found that the deletion of Os GA20ox2 gene in rice SD1 mutants resulted in reduced GA bioactivity in the stem and dwarfing of the plants [30]. However, the effects of drought stress on plant growth need in *B. chinense* to be further explored.

KEGG analysis showed that drought stress mainly affected photosynthesis and phylpropanoid synthesis in leaves and dieterpenoid synthesis and unsaturated fatty acids synthesis in roots of *B. chinense*. The genes involved in photosystem I (PS I) and photosystem II (PS II) were down-regulated under drought stress. This may be due to the inhibition of photosynthesis, the reduction of transcription and translation rate, the reduction of assimilate levels, and the degradation of related proteins and mRNA. This is consistent with the stress results of short-term drought on PS II studied in *Arabidopsis thaliana* [31]. Meanwhile, terpenoids were synthesized in root system to protect plants from drought stress [32]. It is noteworthy that after drought stress, the DEGs in roots of *B. chinense* were twice as much as in leaves, which may be due to the direct contact between roots and soil and the direct influence of soil moisture [33].

B. chinense leaves contain a lot of flavonoids, but the content of saikosaponins is low, while their roots are the opposite [19]. Previous studies have shown that genes encoding the phenylpropane pathway are involved in flavonoid synthesis [34], which also explains the significant expression of phenylpropanoid biosynthesis in leaves of *B. chinense* after drought stress. As pyruvate is a raw material for the synthesis of saikosaponins, the expression level of pyruvate directly affects the content of saikosaponins. After drought stress, pyruvate metabolism in the root system was significantly over-expressed, which is consistent with the previous study [35]. Drought stress promotes the transformation of primary metabolites into secondary metabolites in plants to resist damage caused by adversity stress. Appropriate drought stress is generally considered conducive to the accumulation of active components in plants and the improvement of the quality of medicinal materials [36], and this data for *B. chinense* reached the same conclusion.

Conclusion

In this study, high-throughput transcriptome sequencing technology was used to screen out the gene resources of *B. chinense* seedlings under drought stress. A total of 3,737 and 6,816 DEGs were identified in the control vs. drought stress of *B. chinense* leaves and roots, respectively. The DEGs from leaves and roots were classified into 37, and 36 GO terms and were involved in 222 and 253 KEGG pathways, respectively. SSR analysis were obtained identified 33,728 loci, wherein dinucleotides accounted for the largest proportion. The results provided a theoretical basis for further identification of the molecular mechanism of drought resistance and breeding of drought resistance varieties of *B. chinense.*

Materials And Methods

Material preparation

*B. chinen*se seeds (cultivated) were collected from Changzhi, Shanxi Province (E 112°56'8", N 35°56'47") and planted in botanical garden of Shandong University of traditional Chinese medicine (E116°35'07", N 36°33'20.98") in August 2020. The *B. chinen*se were identified by Dr. Lingchuan Xu (the professor specialized in pharmaceutical botany at Shandong University of TCM, China). The voucher specimens and their related information were deposited at the Herbarium of Shandong University of TCM, China (STCM2020100809). The Shandong University of Traditional Chinese Medicine provided all plant materials used in this study, and no specific permissions were required for the collection of those samples for research purposes following institutional, national and international guidelines. The soil moisture was about 20%, and the soil pH was 7.22. From August 2020 to April 2021, the daytime average temperature was 16.56 ℃, the night average temperature was 7.88 ℃, the maximum temperature was 35.00 ℃, and the minimum temperature was -19.00 ℃. When the seedlings grew to about 6-7 cm in the second year, they were transplanted them into a basin with a diameter of 20 cm and a height of 20 cm, and watered every day to ensure that the water content in the basin is 16.13 ± 3.38% (equivalent to 80% of the field capacity). After seven consecutive days of cultivation, the experiment group was continuously dried left unwatered and the control group was watered normally. When the leaves began to curl under drought stress, the growing seedlings were gently removed from the soil to maintain root integrity, washed with ultra-pure water, and put on autoclaved filter paper. All the leaves and roots from each plant were cut separately and placed in a low temperature test tube. After quick freezing in liquid nitrogen, samples were stored in refrigerator at -80 ℃ for RNA extraction. Each treatment was repeated three times.

Transcriptome analysis

RNA extraction and Library Construction

Total RNA was extracted from leaves and roots of *B. chinense* using Trizol® Reagent (Invitrogen, San Diego, USA). The concentration and purity of the extracted RNA was measured with a Nanodrop 2000 (Thermo Scientific, Massachusetts, USA), and its integrity confirmed by agarose gel electrophoresis. The cDNA libraries were constructed with Illumina's Truseq ^{TM} RNA sample prep kit (Illumina, San Diego, USA). The naming of roots and leaves before and after drought is shown in Table 1.

Transcriptome assembly and sequence analysis

The original sequences obtained by Illumina-Hiseq 2500 were processed by removing low-quality reads containing more than 5% unknown bases [37]. After quality control, the high-quality sequences were assembled by Trinity (http://trinityrnaseq.sourceforge.net/, version number: trinityrnaseq-r2013-02-25) to obtain transcripts, and the longest of each gene cluster was identified as a unigene for subsequent analysis.

Bioinformatics analysis

The unigenes were compared with six databases: NCBI non-redundant protein sequences (Nr), a manually annotated and reviewed protein sequence database (Swissprot), search tool for the retrieval of interacting genes database (String), gene ontology (GO), kyoto encyclopedia of genes and genomes (KEGG), and protein family (Pfam). The result with the highest similarity to the sequence was selected as the annotation [38]. Blast2GO (http://www.balst2go.com) was used to perform GO functional classification annotation of the DEGs, and the Kyoto Encyclopedia of Genes and Genomes (KEGG) (http://www.genome.jp/kegg/) was used to analyze their specific metabolic pathways [39, 40], using $q < 0.05$ as the enrichment standard of the KEGG pathways. The COG method was used to remove the alignment sequences with e_values greater than $1e^{-5}$, and select the sequences with the best alignment results for annotation [41].

The gene expression levels were calculated using the Reads Per kb Per Million Reads (RPKM) method [42]. RPKM value is not only used to analyze sequencing saturation, but also as a measure of gene expression [43]. In view of the biological duplication of *B. chinense* samples in this experiment, therefore, genes with fold change ≤ 2 and false discovery rate (FDR) Plant Mol Biol Rep ≤ 0.05 were used as screening conditions for DEGs [44].

Verification of DEGs

In order to verify the transcriptome data, seven DEGs with $2 \leq |\log_2 ratio| \leq 10$ and high expression levels were selected from the DEGs of leaves and roots of *B. chinense* under drought stress for qPCR verification [45]. The actin gene of *B. chinense* was used as internal control. Specific primers were designed with Primer Premier 5.0. The reaction mixtures contained 10 μL of 2 × T5 Fast qPCR Mix (SYBR Green I), 0.8 μL each of 10μM forward and reverse primers, and 1 1μL of cDNA, were supplemented with ddH_2O to reach a total volume of 20 μL. The procedure of qPCR was as follows: pre-deformation at 95 ℃for 1 min; followed by 40 cycles at 95 ℃for 15 s, 60 ℃for 15 s, and 72 ℃for 30 s; and analysis and verification of the dissolution curve at 60 ℃to 95 ℃. Each sample was repeated three times. he Actin gene of *B. chinense* was used as internal control. $2^{-\Delta\Delta CT}$ was used to assess the relative changes of gene expression ([46]).

Availability of data and materials

The data sets supporting the results of this article are available in the NCBI GenBank repository (Bioproject: PRJNA780763) under the link https://www.ncbi.nlm.nih.gov/Traces/study/?acc=PRJNA780763.

Abbreviations

SS: saikosaponins

SOD: superoxide dismutase

POD: peroxidase

CAT: catalase

DEG: differentially expressed genes

Nr: NCBI non-redundant protein sequences

Swissprot: a manually annotated and reviewed protein sequence database

String:search tool for the retrieval of interacting genes database

GO: gene ontology

KEGG: kyoto encyclopedia of genes and genomes

Pfam: protein family

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

Not applicable.

Competing interests

All the authors declare that we have no competing interests.

Funding

This work was supported by grants from the National Science and Technology to Boost Economy 2020 (No. SQ2020YFF0426286) and (No. SQ2020YFF0426541).

Authors' contributions

GDM conceived the study. GYN, CHL and ZQF provided *Bupleurum chinense* seeds, DK, LL and FY cultivated seedlings of Bupleurum.SY and FY completed the experimental part, and FY wrote the manuscript. All authors have read and approved the manuscript.

Acknowledgements

We are particularly grateful to professor Bu Xun of Shandong Academy of Agricultural Sciences for her assistance in molecular biology.

References

1. Nakashima K, Ito Y, Yamaguchi-Shinozaki K. Transcriptional regulatory networks in response to abiotic stresses in Arabidopsis and grasses. Plant Physiol. 2009;149(1):88–95. doi:10.1104/pp.108.129791
2. Sallam A, Alqudah AM, Dawood MFA, Baenziger PS, Börner A. Drought Stress Tolerance in Wheat and Barley: Advances in Physiology, Breeding and Genetics Research. *Int J Mol Sci*. 2019;20(13):3137. Published 2019 Jun 27.

doi:10.3390/ijms20133137

3. Lamers J, van der Meer T, Testerink C. How Plants Sense and Respond to Stressful Environments. *Plant Physiol*. 2020;182(4):1624–1635. doi:10.1104/pp.19.01464
4. Bashir K, Rasheed S, Matsui A, Iida K, Tanaka M, Seki M. Monitoring Transcriptomic Changes in Soil-Grown Roots and Shoots of Arabidopsis thaliana Subjected to a Progressive Drought Stress. *Methods Mol Biol*. 2018;1761:223–230. doi:10.1007/978-1-4939-7747-5_17
5. Alhaithloul HA, Soliman MH, Ameta KL, El-Esawi MA, Elkelish A. Changes in Ecophysiology, Osmolytes, and Secondary Metabolites of the Medicinal Plants of *Mentha piperita* and *Catharanthus roseus*Subjected to Drought and Heat Stress. *Biomolecules*. 2019;10(1):43. Published 2019 Dec 27. doi:10.3390/biom10010043
6. Bashir K, Matsui A, Rasheed S, Seki M. Recent advances in the characterization of plant transcriptomes in response to drought, salinity, heat, and cold stress. *F1000Res*. 2019;8:F1000 Faculty Rev-658. Published 2019 May 14. doi:10.12688/f1000research.18424.1
7. Li T, Wang R, Zhao D, Tao J. Effects of drought stress on physiological responses and gene expression changes in herbaceous peony (*Paeonia lactiflora* Pall.). *Plant Signal Behav*. 2020;15(5):1746034. doi:10.1080/15592324.2020.1746034
8. Wilkinson S, Davies WJ. ABA-based chemical signalling: the co-ordination of responses to stress in plants. *Plant Cell Environ*. 2002;25(2):195–210. doi:10.1046/j.0016-8025.2001.00824.x
9. Sun P, Li Y, Wei S, et al. Pharmacological Effects and Chemical Constituents of Bupleurum. *Mini Rev Med Chem*. 2019;19(1):34–55.
10. Chinese Pharmacopoeia Commission Pharmacopoeia of the People's Republic of China, I, China Medical Sciences and Technology Press, Beijing (2020), p. 293 (in Chinese)
11. Li X, Li X, Huang N, Liu R, Sun R. A comprehensive review and perspectives on pharmacology and toxicology of saikosaponins. *Phytomedicine*. 2018;50:73–87. doi:10.1016/j.phymed.2018. 09.174
12. Du ZA, Sun MN, Hu ZS. Saikosaponin a Ameliorates LPS-Induced Acute Lung Injury in Mice. *Inflammation*. 2018;41(1):193–198. doi:10.1007/s10753-017-0677-3
13. Wang X, Yang G. Saikosaponin A attenuates neural injury caused by ischemia/reperfusion. *Transl Neurosci*. 2020;11(1):227–235. Published 2020 Jul 2. doi:10.1515/tnsci-2020-0129
14. Zhu ZB, Liang ZS, Han RL, Xin Wang. Impact of fertilization on drought response in the medicinal herb *Bupleurum chinense* DC.: Growth and saikosaponin production. *Ind. Crop Prod*. 2008;29:629–633. doi:10.1016/j.indcrop.2008.08.002.
15. Gong J, Liu M, Xu S, Jiang Y, Pan Y, Zhai Z, et al.. Effects of light deficiency on the accumulation of saikosaponins and the ecophysiological characteristics of wild *Bupleurum chinense* DC. in China. *Ind Crop Prod*.2017;99:179–188. doi: 10.1016/j.indcrop.2017.01.040
16. Yu M, Chen H, Liu SH, Li YC, Sui C, Hou DB, Wei JH. Differential expression of genes involved in saikosaponin biosynthesis between *Bupleurum chinense* DC. and *Bupleurum scorzonerifolium* Willd. Front Genet. 2020;11:583245. doi: 10.3389/fgene.2020.583245
17. Sui C, Chen M, Xu J, Wei J, Jin Y, Xu Y, Sun J, Gao K, Yang C, Zhang Z, Chen S, Luo H. Comparison of root transcriptomes and expressions of genes involved in main medicinal secondary metabolites from *Bupleurum chinense* and *Bupleurum scorzonerifolium*, the two Chinese official Radix bupleuri source species. *Physiol Plant.* 2015;153(2):230–42. doi: 10.1111/ppl.12254.
18. Yang L, Zhao Y, Zhang Q, et al. Effects of drought-re-watering-drought on the photosynthesis physiology and secondary metabolite production of Bupleurum chinense DC. *Plant Cell Rep*. 2019;38(9):1181–1197. doi:10.1007/s00299-019-02436-8
19. Yang LL, Yang L, Yang X, Zhang T, Lan YM, Zhao Y, Han M, Yang LM. Drought stress induces biosynthesis of flavonoids in leaves and saikosaponins in roots of *Bupleurum chinense* DC. Phytochemistry. 2020;177:112434. doi: 10.1016/j.phytochem.2020.112434.
20. Tulsani NJ, Hamid R, Jacob F, et al. Transcriptome landscaping for gene mining and SSR marker development in Coriander (*Coriandrum sativum* L.). *Genomics*. 2020;112(2):1545–1553. doi:10.1016/j.ygeno.2019.09.004

21. Shinozaki K, Yamaguchi-Shinozaki K. Gene networks involved in drought stress response and tolerance. *J Exp Bot*. 2007;58(2):221–227. doi:10.1093/jxb/erl164
22. Alexandersson E, Danielson JA, Råde J, et al. Transcriptional regulation of aquaporins in accessions of Arabidopsis in response to drought stress. *Plant J*. 2010;61(4):650–660. doi:10.1111/j.1365-313X.2009.04087.x
23. Li H, Yao W, Fu Y, Li S, Guo Q. De novo assembly and discovery of genes that are involved in drought tolerance in Tibetan Sophora moorcroftiana. *PLoS One*. 2015;10(1):e111054. Published 2015 Jan 5. doi:10.1371/journal.pone.0111054
24. Meng HL, Zhang W, Zhang GH, et al. Unigene-based RNA-seq provides insights on drought stress responses in Marsdenia tenacissima. *PLoS One*. 2018;13(11):e0202848. Published 2018 Nov 30. doi:10.1371/journal.pone.0202848
25. Shinozaki K, Yamaguchi-Shinozaki K, Seki M. Regulatory network of gene expression in the drought and cold stress responses. *Curr Opin Plant Biol*. 2003;6(5):410–417. doi:10.1016/s1369-5266(03)00092-x
26. Ullah A, Manghwar H, Shaban M, et al. Phytohormones enhanced drought tolerance in plants: a coping strategy. *Environ Sci Pollut Res Int*. 2018;25(33):33103–33118. doi:10.1007/s11356-018-3364-5
27. Gupta A, Rico-Medina A, Caño-Delgado AI. The physiology of plant responses to drought. *Science.* 2020;368(6488):266–269. doi:10.1126/science.aaz7614
28. Yamaguchi S. Gibberellin metabolism and its regulation. *Annu Rev Plant Biol*. 2008;59:225–251. doi:10.1146/annurev.arplant.59.032607.092804
29. Davière JM, Achard P. Gibberellin signaling in plants. *Development*. 2013;140(6):1147–1151. doi:10.1242/dev.087650
30. Spielmeyer W, Ellis MH, Chandler PM. Semidwarf (sd-1), "green revolution" rice, contains a defective gibberellin 20-oxidase gene. *Proc Natl Acad Sci U S A*. 2002;99(13):9043–9048. doi:10.1073/pnas.132266399
31. Chen YE, Liu WJ, Su YQ, et al. Different response of photosystem II to short and long-term drought stress in Arabidopsis thaliana. *Physiol Plant*. 2016;158(2):225–235. doi:10.1111/ppl.12438
32. Kleine S, Müller C. Drought stress and leaf herbivory affect root terpenoid concentrations and growth of Tanacetum vulgare. *J Chem Ecol*. 2014;40(10):1115–1125. doi:10.1007/s10886-014-0505-2
33. Li C, Dong J, Zhang X, et al. Gene expression profiling of Bothriochloa ischaemum leaves and roots under drought stress. *Gene*. 2019;691:77–86. doi:10.1016/j.gene.2018.12.038
34. Vom Endt D, Kijne JW, Memelink J. Transcription factors controlling plant secondary metabolism: what regulates the regulators?. *Phytochemistry*. 2002;61(2):107–114. doi:10.1016/s0031-9422(02)00185-1
35. Zhang Y, Zhou ZY, Xia PG, Liang ZS, Liu SL, Liu ZL. *Zhongguo Zhong Yao Za Zhi*. 2016;41(4):643–647. doi:10.4268/cjcmm20160416
36. Chaves MM, Oliveira MM. Mechanisms underlying plant resilience to water deficits: prospects for water-saving agriculture. *J Exp Bot*. 2004;55(407):2365–2384. doi:10.1093/jxb/erh269
37. Choudhary S, Naika MBN, Sharma R, Meena RD, Singh R, Lal G. Transcriptome profiling of coriander: a dual purpose crop unravels stem gall resistance genes. *J Genet*. 2019;98:19.
38. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9(4):357–9. doi: 10.1038/nmeth.1923.
39. Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*. 2005;21(18):3674–3676. doi:10.1093/bioinformatics/bti610
40. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000;28(1):27–30. doi:10.1093/nar/28.1.27
41. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*. 2011;12:323. doi: 10.1186/1471-2105-12-323.
42. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods*. 2008;5(7):621–8. doi: 10.1038/nmeth.1226.
43. Peng Q, Fang X, Zong X, et al. Comparative transcriptome analysis of Bambusa pervariabilis × Dendrocalamopsis grandis against Arthrinium phaeospermum under protein AP-toxin induction. *Gene*. 2020;725:144160.

doi:10.1016/j.gene.2019.144160

44. Liu B, Ou C, Chen S, et al. Differentially Expressed Genes between Carrot Petaloid Cytoplasmic Male Sterile and Maintainer during Floral Development. *Sci Rep*. 2019;9(1):17384. Published 2019 Nov 22. doi:10.1038/s41598-019-53717-x
45. Song X, Gao J, Peng H. Transcriptomic dynamics changes related to anthocyanin accumulation in the fleshy roots of carmine radish (*Raphanus sativus* L.) characterized using RNA-Seq. *PeerJ*. 2021;9:e10978. Published 2021 Apr 7. doi:10.7717/peerj.10978
46. Song X, Sun P, Yuan J, et al. The celery genome sequence reveals sequential paleo-polyploidizations, karyotype evolution and resistance gene reduction in apiales. *Plant Biotechnol J*. 2021;19(4):731–744. doi:10.1111/pbi.13499

Figures

Figure 1

Size distribution of the unigenes in Bupleurum chinense.

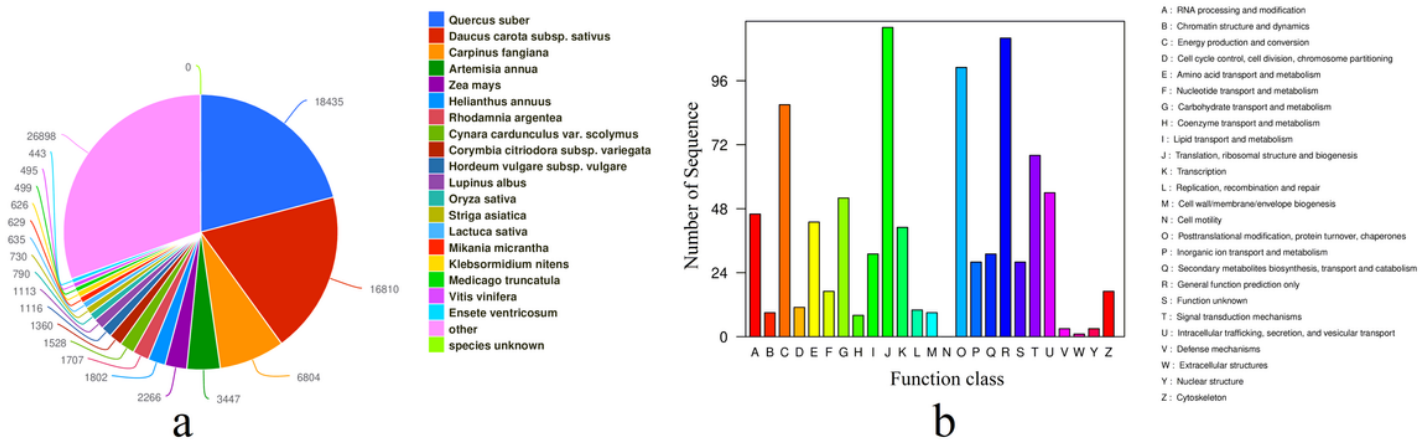

Figure 2

Transcriptome annotation of unigenes in Bupleurum chinense in Nr and COG database. a. Nr database. Each section in the figure represents a species. b. COG database. The abscissa represents the functional classification of unigenes, and the ordinate represents the number of unigenes with this function.

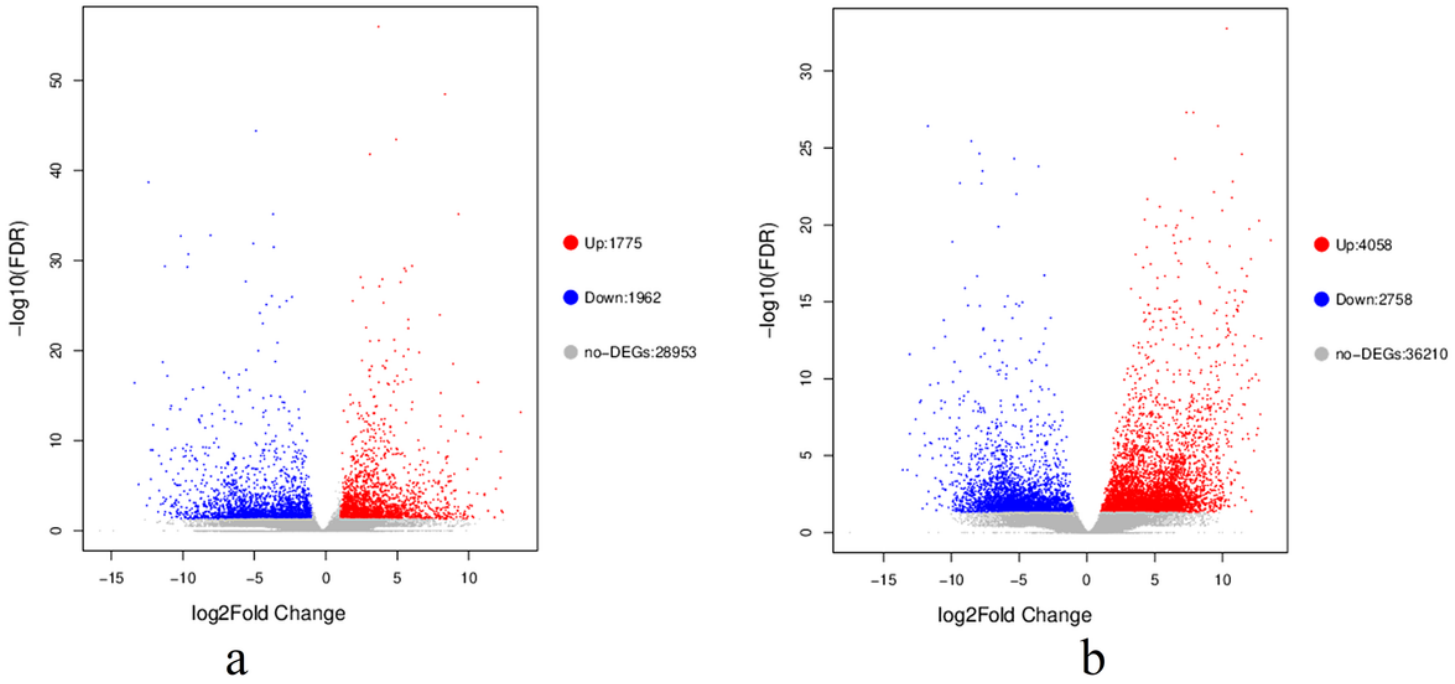

Figure 3

Volcano plot of differentially expressed genes (DEGs) of Bupleurum chinense. a. Leaves. b. Roots. FDR means false discovery rate. The abscissa is the fold change value of unigenes expression difference between samples, and the ordinate is the statistical test value of gene or transcript expression difference, i.e. P value. The higher the P value is, the more significant the expression difference is. The X and Y values are logarithmized.

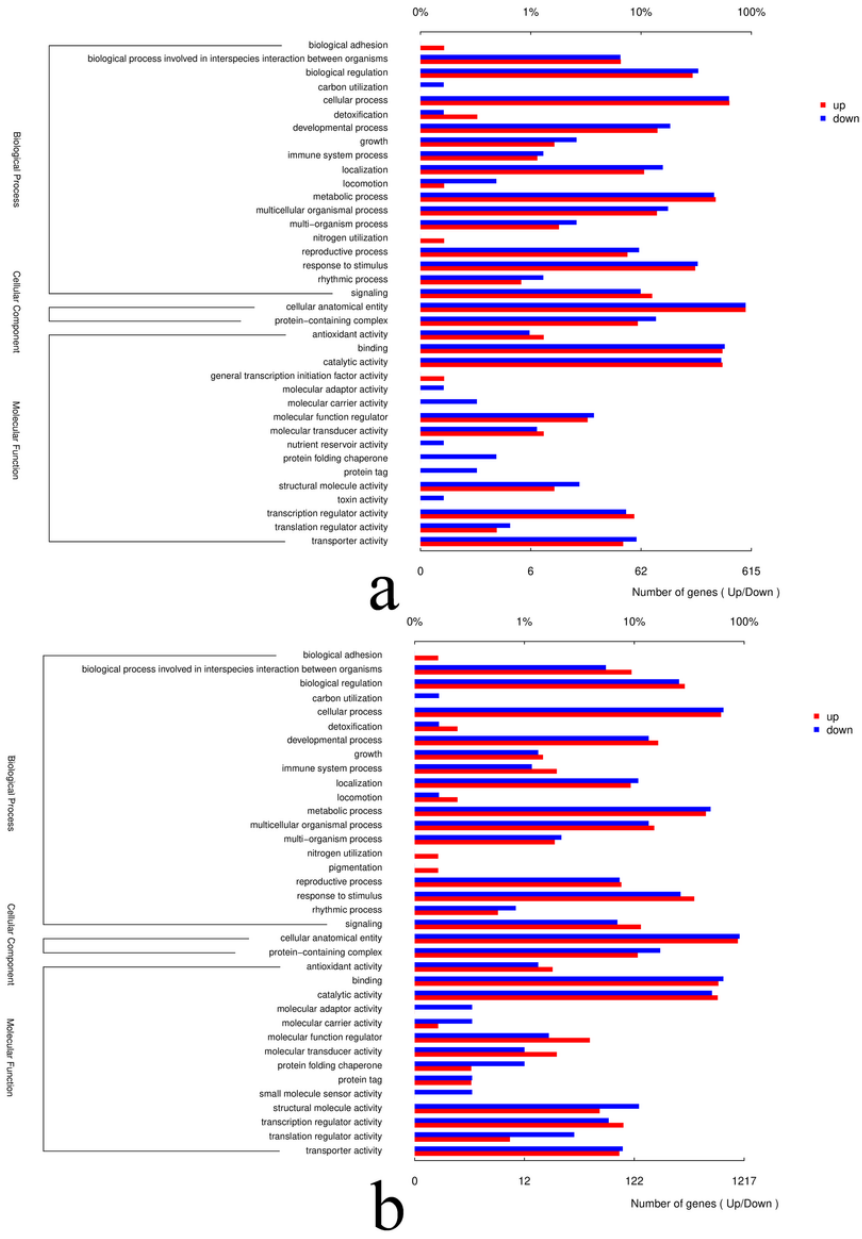

Figure 4

GO functional annotation histogram of differentially expressed genes (DEGs) of Bupleurum chinense after drought treatment. a. leaves. b. roots. The abscissa at the bottom indicates the number of genes annotated to a certain GO term, and the top indicates the ratio of the number of genes annotated to a certain GO term to the total number of genes annotated by GO database. The ordinate represents the primary and secondary classifications of GO.

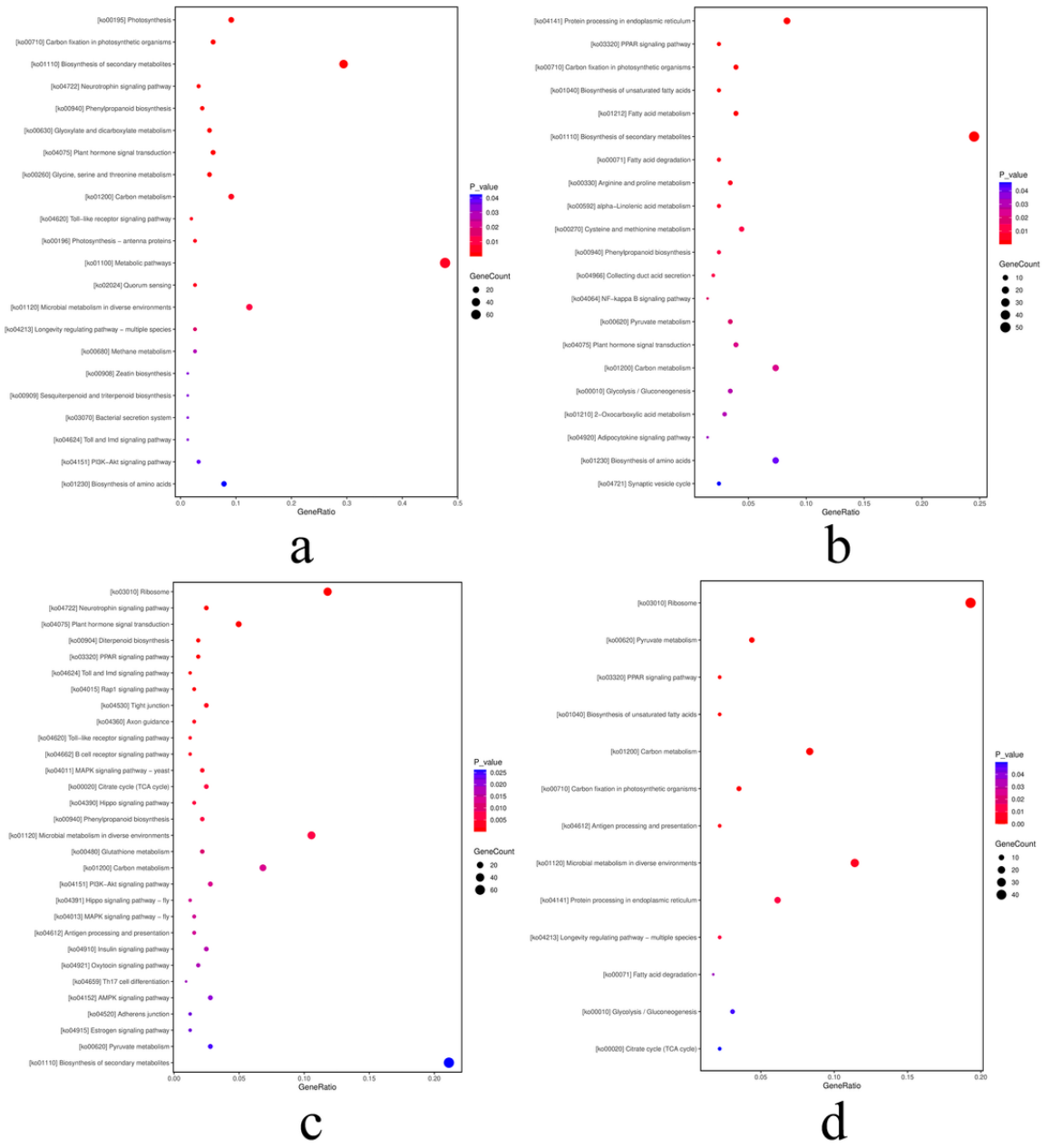

Figure 5

Bubble diagram of KEGG enrichment pathway of significant DEGs in Bupleurum chinense seedlings after drought treatment. a. Up-regulated DEGs in leaves. b. Down-regulated DEGs in leaves. c. Up-regulated DEGs in roots. d. Down-regulated DEGs in roots. The abscissa represents the enrichment rate, and the formula is as follows, Enrich factor = GeneRatio/BgRatio. The ordinate represents the pathway type of the KEGG. The color indicates that the significance of enrichment, that is, the P-value. The darker the color is, the more significant the enrichment is. The right color gradient represents the size of P value.

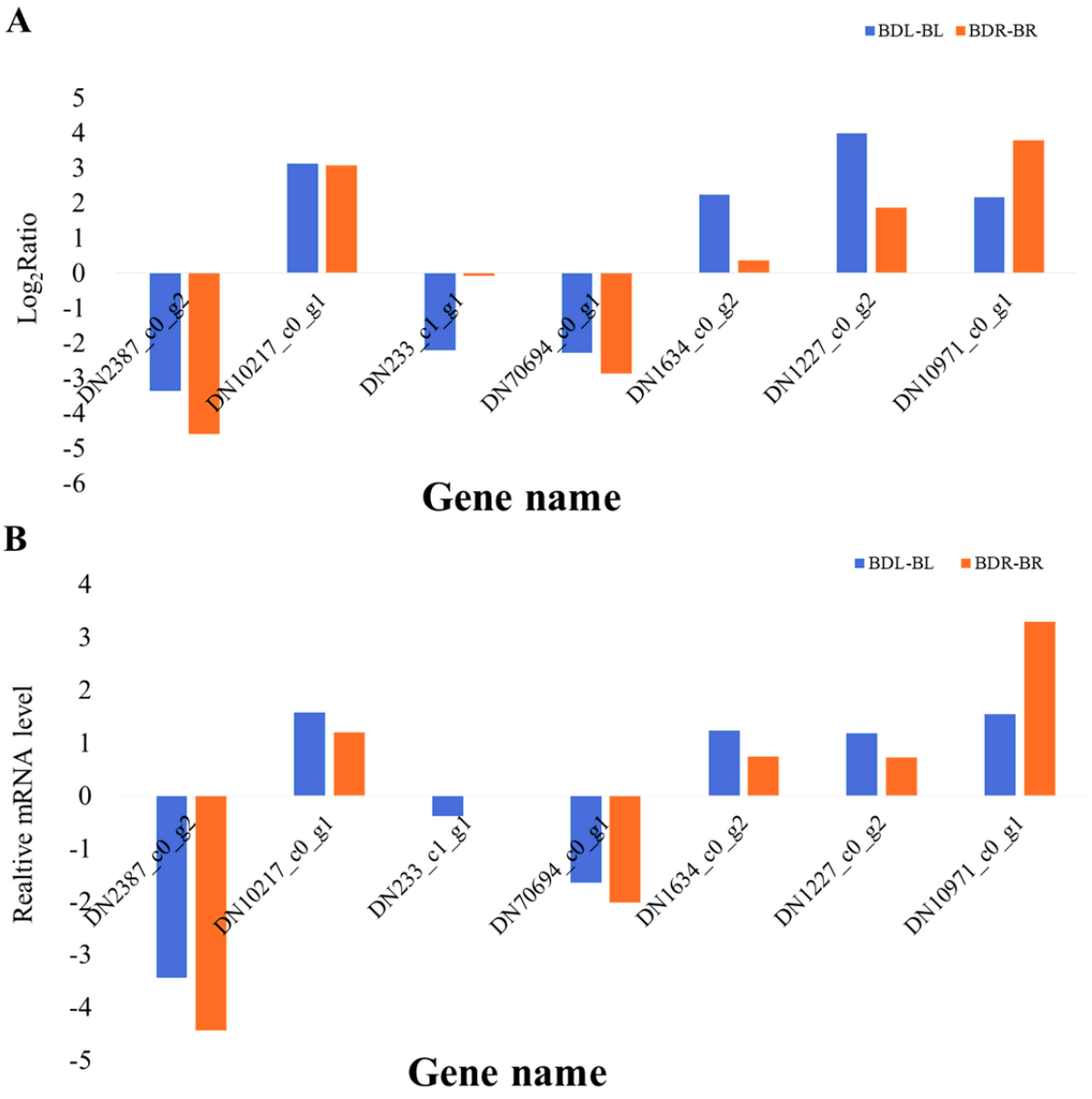

Figure 6

Transcriptome and qPCR analysis of seven DEGs in Bupleurum chinense seedlings. a. 7 DEGs transcriptomes. The ordinate represents the logarithmic value of the expression multiples of the experimental group and the control group with a base of 2. b. 7 DEGs based on qPCR. BDL-BL: comparison of leaves after and before drought stress. BDR-BR: comparison of roots after and before drought stress.