



# Transcriptome sequencing and analysis reveals the molecular response to selenium stimuli in *Pueraria lobata* (willd.) Ohwi

Kunyu Guo<sup>1</sup>, Yiwei Yao<sup>2</sup>, Meng Yang<sup>2</sup>, Yanni Li<sup>2</sup>, Bin Wu<sup>2</sup> and Xianming Lin<sup>1</sup>

<sup>1</sup>Institute of Chinese Medicinal Materials, Hubei Academy of Agricultural Sciences, Enshi, China

<sup>2</sup>Chinese Academy of Medical Sciences and Peking Union Medical College, Institute of Medicinal Plant Development, Beijing, China

## ABSTRACT

*Pueraria lobata* (willd.) Ohwi is a consumable selenium-enriched plant used for medicinal purposes. The molecular response to selenium (Se) stimuli in *P. lobata* is currently unknown. We used RNA-Seq to identify potential genes involved in selenite metabolism and analyzed their expression profiles. We obtained a total of 150,567 unigenes, of which 90,961 were annotated, including 16 structural genes, 14 sulfate transporters, and 13 phosphate transporters that may be involved in Se metabolism, and 33 candidate structural genes involved in isoflavone biosynthesis. The genes with a |foldchange| >2 and q value <0.05 after sodium selenite treatment were identified as differentially expressed genes (DEGs). We obtained a total of 4,246 DEGs, which were enriched in GO terms that included “response to stimulus”, “response to stress”, “signal transduction”, “response to abiotic stimulus”, and “response to chemical”. Of the 4,246 DEGs, one sulfate transporter and five phosphate transporter genes involved Se metabolism, and nine structural genes involved in isoflavone biosynthesis were up-regulated. The expression patterns of 10 DEGs were selected randomly and validated using qRT-PCR. The Pearson Correlation Coefficient ( $r$ ) was 0.86, indicating the reliability of RNA-Seq results. 22 Reactive Oxygen Species (ROS) scavenging DEGs were found, 11 of which were up-regulated. 436, 624 transcription factors (TFs) correlated with structural genes were identified that may be involved in Se and isoflavone biosynthesis, respectively, using  $r$  ( $r > 0.7$  or  $r < -0.7$ ). 556 TFs were related to at least one sulfate and phosphate transporter. Our results provided a comprehensive description of gene expression and regulation in response to Se stimuli in *P. lobata*.

**Subjects** Bioinformatics, Molecular Biology, Plant Science

**Keywords** Transcriptome sequencing, Selenium stimuli, *Pueraria lobata* (willd.) ohwi, Molecule response

## INTRODUCTION

*Pueraria lobata* (Wild.) Ohwi, known as kudzu, is mostly found throughout East Asian countries (*He et al., 2019*). The dried roots of *P. lobata* have been used medicinally for centuries to treat influenza, muscle stiffness, and other ailments (*The Chinese Pharmacopoeia, 2005*) with its usage first recorded in the *Shengnong Bencao Jing*, the

Submitted 25 September 2019

Accepted 18 February 2020

Published 24 March 2020

Corresponding authors

Bin Wu, [bwu@implad.ac.cn](mailto:bwu@implad.ac.cn)

Xianming Lin, [lxm31@sina.com](mailto:lxm31@sina.com)

Academic editor

Ruslan Kalendar

Additional Information and  
Declarations can be found on  
page 13

DOI 10.7717/peerj.8768

© Copyright  
2020 Guo et al.

Distributed under  
Creative Commons CC-BY 4.0

OPEN ACCESS

first comprehensive work on Chinese Medicine. Isoflavonoids, specifically puerarin and daidzein, are the main bioactive compounds found in the roots of *P. lobata* (Du et al., 2010). Modern pharmacological studies have revealed important functions of the *P. lobata* root. Puerarin is known to protect the cardiovascular system and may prevent osteoporosis, liver injury, and inflammation (Wei, Chen & Xu, 2014) and daidzein decreases blood alcohol levels (Lowe et al., 2008). A series of medicine, and health and food products have been developed using *P. lobata* as the functional ingredient because of its efficacy (<http://db.pharmcube.com/database/cfda>).

Selenium (Se) is a trace element essential for humans and animals and is an active ingredient of glutathione peroxidase, which is an antioxidant and has anti-aging and immunity benefits in humans (Fairweather et al., 2011). A severe selenium deficiency can lead to Keshan disease and Kashin-Beck disease while an overdose can cause hair loss, headache, fatigue and discoloration of the nails (Fairweather et al., 2011). Plants absorb selenite- and selenate-Se from the soil and transform them into organic Se (Winkel et al., 2015), which is the form predominantly consumed by humans. Current research has only speculated on the transport and transformational processes of selenate and selenite in higher plant species (Winkel et al., 2015). This process can be roughly classified into two stages: selenocysteine (SeCys) formation and Se methylation. In the SeCys formation stage, selenate may be taken up via sulfate transporters (Terry, De Souza & Am, 2000) and selenite may be taken up via phosphate transporters (Barrow & Whelan, 2010; Terry, De Souza & Am, 2000) or Si-transporters (Zhao et al., 2010). Selenite may be further reduced to selenide by sulfite reductase or by glutathione disulfide reductase and glutathione oxidoreductase. Selenide may then be transformed to selenocysteine (SeCys) by cysteine synthase, or by selenide water dikinase, Se-cysteine-tRNA synthase, and cysteine-tRNA ligase. SeCys may be incorporated into Se-proteins (Aldwin Suryo & Davies, 2012; Hurst et al., 2013), functioning as an antioxidant. SeCys may be transformed into dimethyl diselenide (DMDSe) or selenomethionine (SeMet) during the Se methylation stage and eventually into dimethyl selenide (DMSe) via the methionine cycle by a series of enzymes. DMDSe and DMSe can be emitted to the atmosphere to alleviate the Se toxicification of the plant (Winkel et al., 2015).

Few studies have been performed on the molecular mechanism responsible for selenite stimuli in plants. One case study identified 14 selenite-responsive genes in *Astragalus racemosus* (Hung et al., 2012). Other research has shown that the antioxidative system was activated and photosynthesis and primary metabolism were enhanced under low selenite stress. High selenium stress has been shown to inhibit photosynthesis, primary metabolism, protein ubiquitination, and phosphorylation (Wang, Wang & Wong, 2012).

*P. lobata* is a selenium-enriched plant primarily composed of selenoproteins and selenopolysaccharides in which organic Se accounted for 82.42% of the total Se (Du et al., 2010). However, the molecular mechanism responsible for the Se stimuli response in *P. lobata* is unknown. RNA-seq was used to identify structural genes that may be involved in the biosynthesis of Se and isoflavonoids. We found that selenite may be taken up via phosphate transporters in *P. lobata*. Transcription factors (TFs) were found to

potentially regulate the structural genes in Se and isoflavonoid biosynthesis. Our results are foundational for engineering breeding cultivars of *P. lobata* with high Se compounds.

## MATERIALS AND METHODS

### Plant growth conditions and sodium selenite treatment

Mature stems were cut from one year-old *P. lobata* cv. GuiYege NO. 1 plants from Huazhong Medicinal Botanical Garden of the Institute of Chinese Medicinal Materials, Hubei Academy of Agricultural Sciences (Enshi, China). The stems were transplanted into sterilized vermiculite in plastic pots (16 cm × 16 cm) and grown in a greenhouse under a light/dark period of 16 h/8 h at 25 °C to control soil moisture and temperature. *P. lobata* seedlings with five leaves were treated with 200 mL sodium selenite at 0, 1, 5, 15, 25, 35, 55 mg/L, respectively, using an even spray on the vermiculite in the pots. The control was treated with distilled water. Every treatment was replicated three times.

### Measurement of physiological and biochemical indexes

The root length was measured on the 9th day after the sodium selenite treatment using a vernier caliper. Biochemical indexes were measured on the day of treatment and then then the 1st, 3th, 5th, 7th and 9th days of treatment thereafter. The total selenium content was measured using hydride atomic fluorescence spectrometry ([Fernanda et al., 2016](#)). The activity of superoxide dismutase (SOD) was determined as described by [Zhang et al. \(2012b\)](#) and the content of malondialdehyde (MDA) was assayed using the protocol from a previous study ([Feng et al., 2009b](#)). The correlation analyses of the differences between sampling days and concentrations of SOD and MDA were performed by SPSS 19 ([Qiu et al., 2017](#)).

### RNA extraction, library construction, and sequencing

Samples were collected on the day of treatment and the 1st, 3th, 5th, 7th and 9th days following the treatment with 25 mg/L sodium selenite. The roots were washed with distilled water, blotted with dry filter paper, and immediately frozen in liquid nitrogen. Total RNA was extracted, treated, and measured following the protocol from our previous study ([Wu et al., 2015](#)). Total RNA was extracted using the TRizol reagent according to the manufacturer's instructions and was digested to eliminate the residual genomic DNA using RNase-free DNase. The Agilent Technologies 2100 Bioanalyzer was used to measure and quantify the total RNA.

RNA-Seq libraries were constructed and sequenced as follows: the mRNA of each sample was separated from the total RNA using oligo (dT) magnetic beads; the mRNA were then cleaved into short fragments; the short fragments were used as templates and the first-strand cDNA was synthesized by reverse transcriptase and random primers. The RNA templates were then removed and the second-strand cDNA was synthesized using dNTPs, DNA polymerase I, and RNase H. These short, double-cDNA fragments were purified with DNA clean beads. The short cDNA fragments were ligated with the Illumina paired-end adaptors and purified with DNA clean beads following end reparation and A-tailing. PCR was used to selectively enrich DNA fragments with adapter molecules on both ends and

to create the final cDNA library. The quality of the cDNA library was measured using the Agilent 2100 Bioanalyzer. The libraries were then sequenced from the 5' and 3' ends using the Illumina sequencing system.

### **De novo transcriptome assembly and annotation**

RNA-Seq reads were *de novo* assembled using Trinity to obtain high-quality transcription sequences with the default parameters (Grabherr *et al.*, 2011). The assembly unigenes were compared with sequences in the CDD database using BLAST (Marchler *et al.*, 2013), KOG (Koonin *et al.*, 2004), COG (Tatusov *et al.*, 2000), NR and NT (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>), PFAM (Finn *et al.*, 2016), Swissprot (Boeckmann *et al.*, 2005), TrEMBL (Boeckmann *et al.*, 2003), GO (Ashburner *et al.*, 2000) and KEGG (Kanehisa *et al.*, 2004) for annotations with *E*-value  $10^{-5}$  as the cutoff.

### **Illumina data processing and gene expression profiles quantification**

Raw reads were processed using in-house Perl scripts in FASTQ format. Clean reads were obtained after removing low-quality reads and those containing adapters and poly-N using Trimmomatic (Bolger, Lohse & Usadel, 2014). Gene expression levels were determined by TPM (Transcripts Per Kilobase of exon model per Million mapped read) to calculate the fold change. The read count was calculated using Salmon (Patro *et al.*, 2017) and was loaded into DESeq2 (Anders & Huber, 2010) to calculate the *P* value for differential expression analysis. Genes that had a  $|\text{foldchange}| \geq 2$  and *q* value  $< 0.05$  identified by DESeq were defined as differentially expressed genes (DEGs).

### **TFs identification**

Sequences of the unigenes were compared to the sequences in the iTAK database to analyze TFs in *P. lobata* using the default parameters (Zheng *et al.*, 2016).

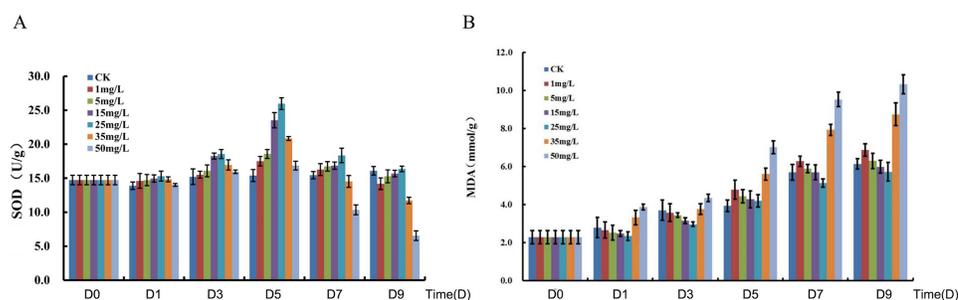
### **Validation of differential expression genes by qRT-PCR**

qRT-PCR were performed as described by our previous paper (Wu *et al.*, 2015). 10 randomly selected genes were validated. Gene-specific primers were designed by Primer5 and the primer sequences are listed in Table S1. The 40S ribosomal protein S8 was used as an internal control (He *et al.*, 2019). The Pearson Correlation Coefficient (*r*) was calculated between the mean results of qRT-PCR and RNA-Seq.

## **RESULTS**

### **Physiological and biochemical index of *P. lobata* after sodium selenite treatment**

The total Se content in *P. lobata* seedlings showed an increase and followed by a decrease with an increased sodium selenite concentration; plant leaves accumulated more Se than the stems. The total Se content in the seedling was lower when the concentration of sodium selenite was less than 5 mg/L, however the total Se content was maximized when the concentration was 25 mg/L. The accumulation of selenium in *P. lobata* seedlings began to decline when the concentration was more than 35 mg/L (Fig. S1). The root length tended to be the same under the Se accumulation and the sodium selenite treatments (Fig. S2).



**Figure 1** Measurement of SOD (A) and MDA (B). D0, D1, D3, D5, D7 and D9 represent different day after sodium selenite treatment. Columns of different colors represent concentrations of sodium selenite.

Full-size DOI: [10.7717/peerj.8768/fig-1](https://doi.org/10.7717/peerj.8768/fig-1)

SOD content showed an initial increase and then decrease. SOD reached its maximum content at 25 mg/L, 5 days after treatment, and then decreased gradually (Fig. 1A). The content of MDA tended to increase with the sodium selenite concentration and treatment time (Fig. 1B). Correlation analyses were conducted on the differences between sampling days and concentrations of SOD. MDA was also conducted. Results for SOD activity revealed that the treatment concentration of sodium selenite was negatively correlated on the 7th day and significantly negatively correlated on the 9th day (Table S2). The treatment from the 1st day was positively correlated with those of the 3rd, 5th, and 7th days. The treatment from the 3rd day was significantly positively correlated with that of the 5th day. The treatment from the 5th day was positively correlated with that of the 7th day. The treatment from the 7th day was significantly positively correlated with that of the 9th day. Results showed that the treatment concentration of sodium selenite was positively correlated the MDA content from the 1st day of treatment and was significantly positively correlated with the treatments from the 5th, 7th and 9th days, respectively (Table S3). There was a significant positive correlation between the treatment from the 1st day and those of the 3rd, 5th, 7th and 9th days, respectively.

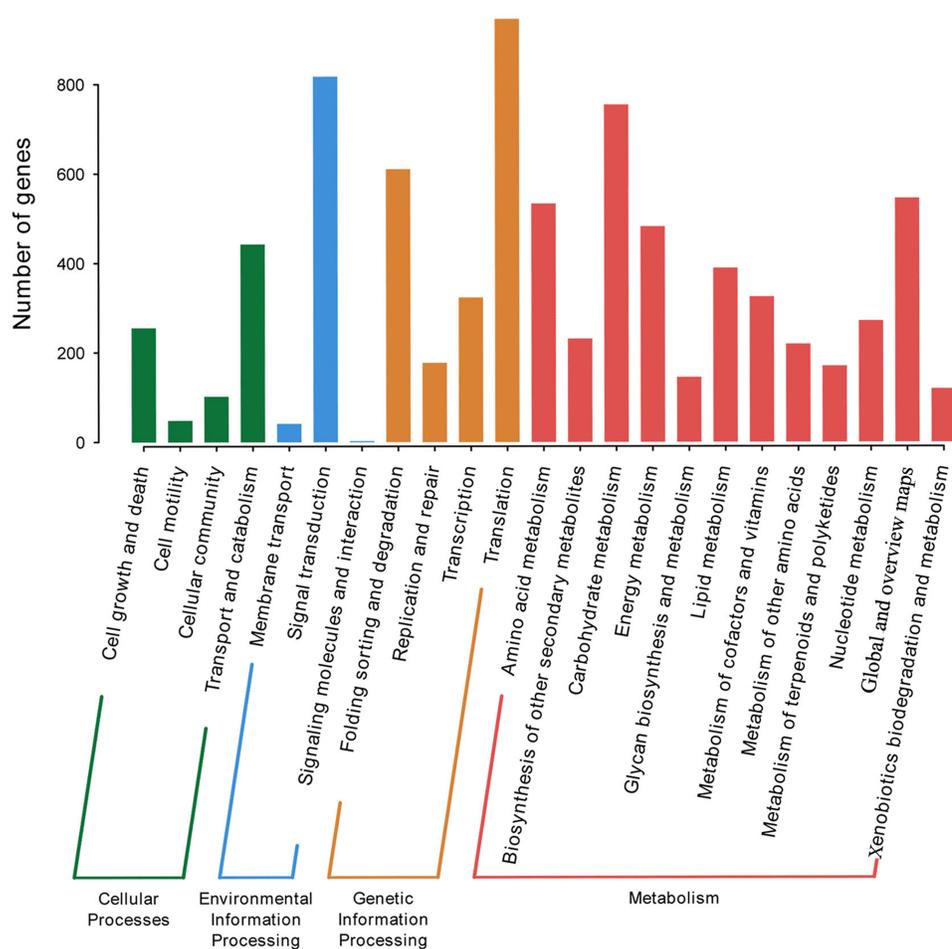
### RNA-Seq sequencing and assembly

The 18 mRNA samples from days 0, 1, 3, 5, 7, 9 after sodium selenite treatment were sequenced on the Illumina platform. We obtained more than 6 Gb of data from every sample for the downstream analysis. The transcriptome data was assembled and 150,567 unigenes were obtained using Trinity (Grabherr et al., 2011). The N50 length and average read length of the assembly sequence were 1,742 bp and 962.46 bp, respectively. The length distribution and GC content of the unigenes was calculated (Fig. S3).

### Gene function annotation and categorization

The assembled 150,567 unigenes were compared with sequences found in the CDD, KOG, Nr, Nt, Pfam, Swissport, TrEMBL, GO, and KEGG by BLAST (version 2.2.26) databases using default parameters. There were 46, 439, 33, 401, 59,266, 69,142, 31,796, 58,101, 58,917, 64,796, and 5,415 resulting annotations found (Table S4). 90,961 total unigenes were annotated.





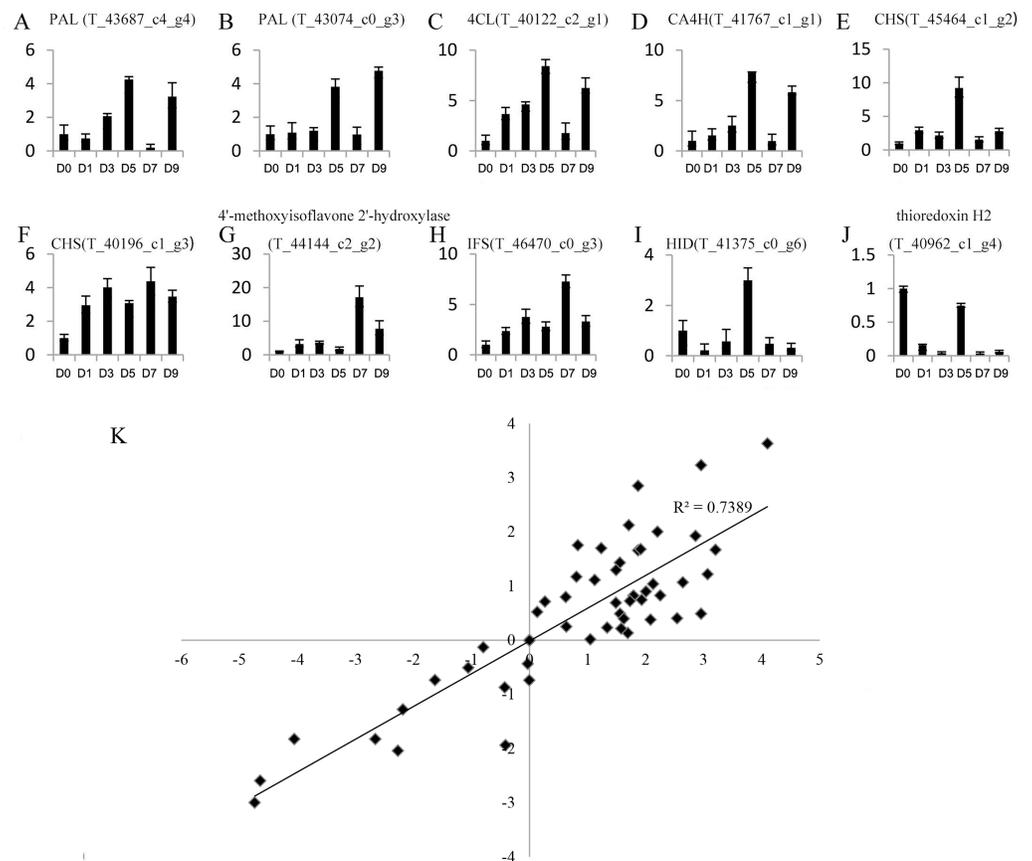
**Figure 3** KEGG categories of *P. lobata* unigenes.

[Full-size](#) [DOI: 10.7717/peerj.8768/fig-3](https://doi.org/10.7717/peerj.8768/fig-3)

### Comprehensive analysis of DEGs

Genes with a  $|\text{foldchange}| > 2$  and  $q$  value  $< 0.05$  after sodium selenite treatments for different lengths of time were identified as differentially expressed genes (DEGs). A total of 4,246 DEGs were obtained (Table S7). 10 DEGs were selected randomly for analysis by qRT-PCR (Figs. 4A–4J) to validate the reliability of the RNA-Seq results (Table S8). The results between RNA-Seq and qRT-PCR were analyzed. The Pearson Correlation Coefficient ( $r$ ) is approximately 0.86 (Fig. 4K), which indicates that the results of RNA-Seq and qRT-PCR are positively linearly related. The expression levels in RNA-Seq reflect the differences in gene expression after selenium treatment.

1,538 DEGs were identified after the first day of selenite treatment, including 465 up-regulated genes and 1,073 down-regulated genes. 1,386 DEGs were identified on the third day of treatment, including 559 up-regulated genes and 827 down-regulated genes. 1,348 DEGs were identified on the fifth day of treatment, including 407 up-regulated genes and 941 down-regulated genes. 2,510 DEGs were identified on the seventh day of treatment, including 883 up-regulated genes and 1,627 down-regulated genes. 1,556 DEGs



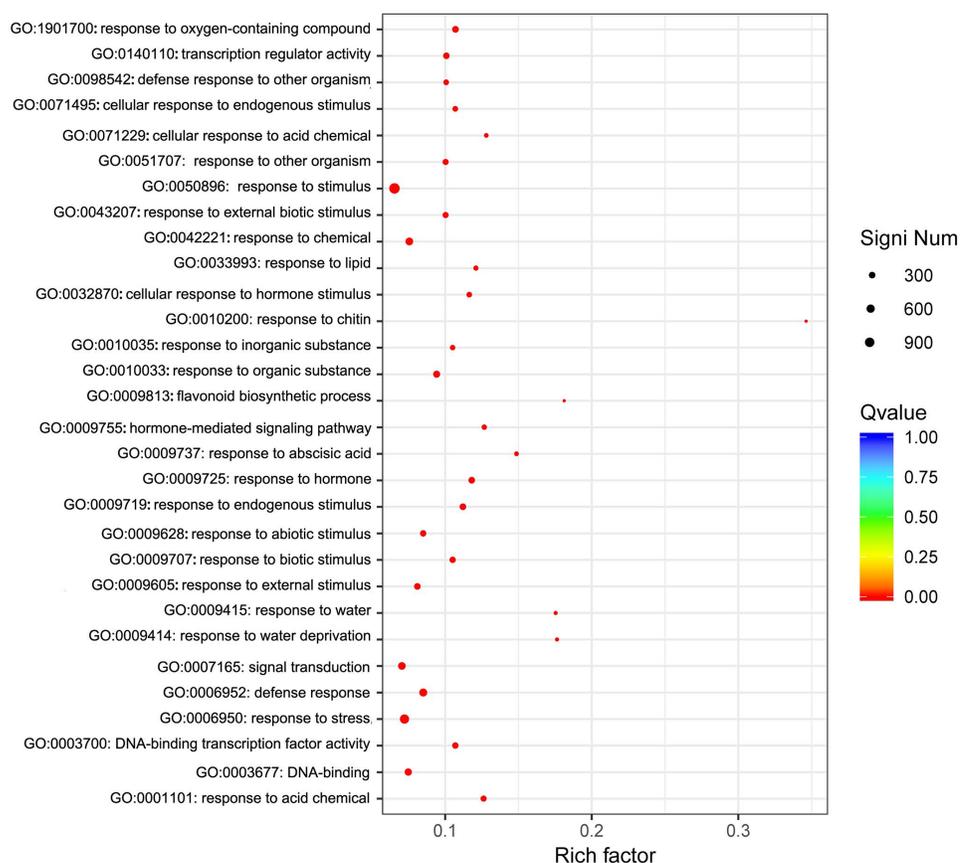
**Figure 4** Expression profiles of 10 unigenes by qRT-PCR (A–J) and the correlation of the results between qRT-PCR and RNA-Seq (K). (A–J) Fold changes of the unigenes are shown. The average expression levels in the controls were set to 1. Error bars represent standard error. (K) The correlation between the qRT-PCR and RNA-Seq expression levels. The X-axis represents log<sub>2</sub> fold change in the expression levels found by qRT-PCR. The Y-axis indicates the log<sub>2</sub> value of the expression level fold change from RNA-Seq. Full-size [DOI: 10.7717/peerj.8768/fig-4](https://doi.org/10.7717/peerj.8768/fig-4)

**Table 1** The number of up- and down-regulated DEGs based comparison with 0 day. As control (|fold-change| >2; q value <0.05).

Comparison	Up-regulated	Down-regulated	Total
D1 vs D0	465	1073	1538
D3 vs D0	559	827	1,386
D5 vs D0	407	941	1,348
D7 vs D0	883	1,627	2,510
D9 vs D0	1,032	524	1,556

were identified on the ninth day of treatment, including 1,032 up-regulated genes and 524 down-regulated genes (Table 1).

We focused on the functional enrichment of the 4246 DEGs. GO enrichment showed that “response to stimulus”, “response to stress”, “signal transduction”, “response to abiotic stimulus”, and “response to chemical” were the top five classifications (Fig. 5), which

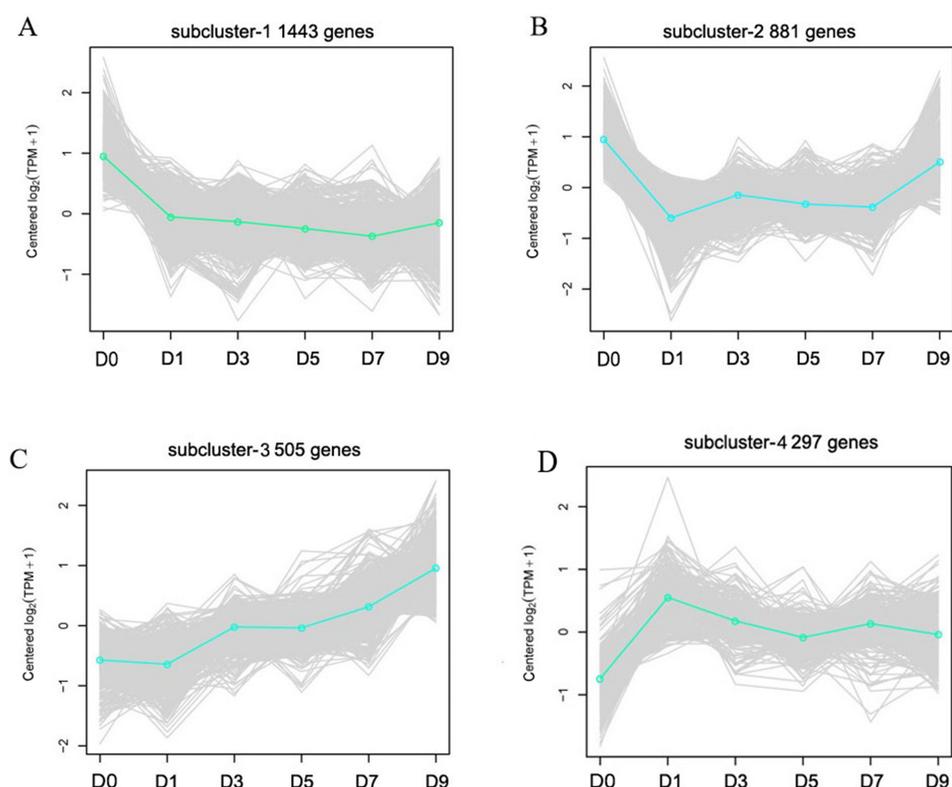


**Figure 5** GO enrichment of DEGs. The X-axis represents the enrichment factor and  $\log_{10}$  of the Q-value. The Y-axis indicates the different GO terms.

Full-size DOI: [10.7717/peerj.8768/fig-5](https://doi.org/10.7717/peerj.8768/fig-5)

indicated that sodium selenite was a stress molecule. KEGG enrichment indicated that sodium selenite affected plant hormone signal transduction, MAPK signaling pathway-plant, starch and sucrose metabolism, ribosome biogenesis in eukaryotes, flavonoid biosynthesis, and isoflavonoid biosynthesis (Fig. S5).

Cluster analysis was performed on gene expression profiles of RNA-seq libraries from 18 samples according to Ernst's STEM algorithms arithmetic to assess the dynamic changes of DEGs after sodium selenite treatment (Ernst & Bar-Joseph, 2006). Four representative expression patterns were obtained (Fig. 6) and classified as subclusters. 1,443 genes were down-regulated as a whole after selenium treatment in subcluster 1; these were mainly involved in DNA metabolism, material transport, synthesis of primary metabolites, and stress-related processes. 881 genes in subcluster 2 were down-regulated on the first day and then up-regulated at another time, and were mainly involved in disease resistance, stress tolerance, material transport and cytoskeleton construction. 505 genes in subcluster 3 were up-regulated as a whole after selenium treatment; these were mainly involved in the secondary metabolism synthesis, response to toxin metabolism, and response to stress. 297



**Figure 6** Representative time-course profile clusters. (A–D) Four subclusters 1–4 are shown. D0, D1, D3, D5, D7 and D9 represent different day after sodium selenite treatment. For each time point and each gene, the centered log<sub>2</sub> (TPM + 1) is indicated the mean expression profile for each cluster.

Full-size [DOI: 10.7717/peerj.8768/fig-6](https://doi.org/10.7717/peerj.8768/fig-6)

genes were up-regulated and then down-regulated on the first day in subcluster 4, and were mainly involved in regulating material transport, DNA repair, and resistance reactions.

### DEGs overlap at five different time points

DEGs were analyzed continuously after sodium selenite treatment. When compared with the control (Fig. 7), 160 DEGs were observed, including 101 up-regulated DEGs and 59 down-regulated DEGs (Table S9). Among the up-regulated DEGs, 9 genes (T\_39191\_c1\_g1, T\_39355\_c1\_g4, T\_41331\_c0\_g6, T\_42572\_c0\_g2, T\_44141\_c2\_g4, T\_44939\_c1\_g2, T\_44939\_c1\_g3, T\_44939\_c1\_g5, T\_45070\_c2\_g1) are potentially involved in the biosynthesis of isoflavones; 1 gene (T\_40246\_c1\_g1) encoded the selenium binding protein, one gene (T\_42307\_c2\_g3) encoded glutathione s-type transferase and two genes (T\_44116\_c2\_g1, T\_44116\_c2\_g3) encoded phosphate transporters. Among those down-regulated genes, two MYB transcription factors, MYB39 (T\_36389\_c0\_g1) and R2R3-MYB (T\_41871\_c2\_g5), one thioredox protein (T\_44132\_c2\_g1), and other genes that encoded TMV disease-resistant proteins and ion transport proteins were found.



increase followed by a decrease in selenium content. A maximum selenium content was achieved at a concentration of 25 mg/L of sodium selenite. The selenium content in leaves was higher than in the stems, which was consistent with results from other studies ([Han et al., 2013](#); [Han et al., 2015a](#); [Han et al., 2015b](#)). We measured the activity of SOD in *P. lobata* treated with different concentrations of sodium selenite. The results showed that an increase of treatment time did not significantly change the activity of SOD in the control group. However, an increased concentration of sodium selenite caused an increase in the activity of SOD in *P. lobata* followed by a decrease. The trend was most obvious on the fifth day of treatment when the concentration was 25mg/L. MDA content was determined and the results showed that MDA in *P. lobata* decreased and then increased with longer treatment times and greater concentrations. Our results indicate that the activity of SOD and the content of MDA in *P. lobata* could be increased using a treatment of low concentration sodium selenite. These results were consistent with results from previous studies ([Han et al., 2013](#); ([Han et al., 2015a](#); [Han et al., 2015b](#)). The root length, MDA content, and SOD activity under different concentrations of sodium selenite further supported 25 mg/L as a suitable treatment concentration.

The transcriptome of *P. lobata* seedlings was sequenced after treatment with 25 mg/L sodium selenite and the DEGs were identified. 150,567 unigenes were obtained, of which 4,246 were DEGs. The enriched GO terms of DEGs respond to stimulus, stress, signal transduction, abiotic stimulus, and chemicals, which are similar to the responses of sodium selenate in *Astragalus chrysochlorus* ([Özgür et al., 2015](#)). This result is consistent with those of previous studies that the molecular mechanisms of plants that absorb selenite and selenate share some common pathways ([Li, McGrath & Zhao, 2008](#); [Yu & Gu, 2008](#)). Sodium selenite may be taken up by phosphate transporters, because the expression profiles of 5 phosphate transporters genes were up-regulated after applying sodium selenite, and 2 genes (T\_44116\_c2\_g1, T\_44116\_c2\_g3) were continuously up-regulated. Phosphorus is known to be one of the macronutrients required for plant growth and development and a phosphate transporter is needed to absorb phosphate fertilizer. Therefore, we suggest that the absorption of sodium selenite and phosphate fertilizer may compete with one another. The use of these two fertilizers at the same time is contraindicated for the cultivation of Se-enriched plants.

ROS can damage cell membranes and other components and harm plants ([Nandini & Samir, 2016](#)). Low selenium stress has been reported to activate the antioxidant system to scavenge ROS and enhance a plant's ability to withstand a variety of oxidative stresses ([Wang, Wang & Wong, 2012](#)). We identified DEGs encoding major ROS scavenging enzymes, including SOD, CAT, APX, GPX, GST, and MDHAR ([Apel & Hirt, 2004](#)). We found that 1 SOD, 1 APX, and 9 GST s were up-regulated when compared with the control, which might enhance the anti-oxidation mechanisms after sodium selenite treatment in *P. lobata* ([Table S7](#)). GST (T\_42307\_c2\_g3) was up-regulated consistently after selenite treatment.

To address the question of whether sodium selenite treatment in *P. lobate* would promote the biosynthesis of isoflavone, we analyzed the expression profiles of isoflavone-related structural genes and found that 9 genes were up-regulated ([Table S6](#)), indicating

that isoflavone biosynthesis was promoted after 25 mg/L sodium selenite treatment. These results are consistent with previous studies reporting that the total flavonoid and chlorogenic acid contents were enhanced when the application dose did not exceed 2.0 mg/kg (Li, Sun & Liu, 2010). These results suggest that the content of Se and other active compounds could be enhanced in medicinal plants by applying a suitable concentration of sodium selenite.

We identified the role of TFs in regulating selenium metabolism for the first time. 436 TFs were related with structural genes in Se metabolism (Table S12) and 556 TFs were co-expressed with sulfate transporters or phosphate transporters (Table S13). Our results provided a comprehensive overview of the gene expression and regulation in response to selenium stimuli in *P. lobata*.

## CONCLUSION

RNA-Seq was used to identify the molecular mechanism involved in the response to Se stimuli in *P. lobata*. We found 150,567 unigenes, of which 90,961 were annotated. Among these annotated genes, 16 structural genes, 14 sulfate transporters, and 13 phosphate transporters were potentially involved in Se metabolism and 32 structural genes were related to isoflavone biosynthesis. We obtained 4,246 DEGs, of which one sulfate transporter and five phosphate transporter genes involved in Se metabolism and nine structural genes involved in isoflavone biosynthesis were up-regulated. Furthermore, twenty-two ROS scavenging DEGs were identified, of which 11 were up-regulated. Finally, we identified some TFs that are potentially involved in Se metabolism and isoflavone biosynthesis. In order to cultivate Se-enriched plants, sodium selenite and phosphate fertilizer should not be applied at the same time. The content of Se and active compounds may be enhanced in plants used for medicinal purposes by applying a suitable concentration of sodium selenite.

## ADDITIONAL INFORMATION AND DECLARATIONS

### Funding

This study is supported by Special Project of Technological Innovation in Hubei province (2017AKB075) and the Science and Technology Program Research and Development Project of Enshi (D20180016). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

### Grant Disclosures

The following grant information was disclosed by the authors:

Special Project of Technological Innovation in Hubei Province: 2017AKB075.

Science and Technology Program Research and Development Project of Enshi: D20180016.

### Competing Interests

The authors declare there are no competing interests.

## Author Contributions

- Kunyuan Guo, Yiwei Yao, Meng Yang, Yanni Li, Bin Wu and Xianming Lin conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.

## Data Availability

The following information was supplied regarding data availability:

The raw sequence data are available at the NCBI SRA database: [PRJNA516771](https://www.ncbi.nlm.nih.gov/sra/PRJNA516771).

## Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.8768#supplemental-information>.

## REFERENCES

- Aldwin Suryo R, Davies MJ. 2012. Selenium-containing amino acids as direct and indirect antioxidants. *Lubmb Life* **64**:863–871 DOI [10.1002/iub.1084](https://doi.org/10.1002/iub.1084).
- Anders S, Huber W. 2010. Differential expression analysis for sequence count data. *Genome Biology* **11**(10):Article R106 DOI [10.1186/gb-2010-11-10-r106](https://doi.org/10.1186/gb-2010-11-10-r106).
- Apel K, Hirt H. 2004. Reactive oxygen species: metabolism, oxidative stress, and signal transduction. *Annual Review of Plant Biology* **55**:373–399 DOI [10.1146/annurev.arplant.55.031903.141701](https://doi.org/10.1146/annurev.arplant.55.031903.141701).
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. 2000. Gene ontology: tool for the unification of biology. *Nature Genetics* **25**:25–29 DOI [10.1038/75556](https://doi.org/10.1038/75556).
- Barrow NJ, Whelan BR. 2010. Testing a mechanistic model. VIII. The effects of time and temperature of incubation on the sorption and subsequent desorption of selenite and selenate by a soil. *European Journal of Soil Science* **40**(1):29–37 DOI [10.1111/j.1365-2389.1989.tb01251.x](https://doi.org/10.1111/j.1365-2389.1989.tb01251.x).
- Boeckmann B, Bairoch A, Apweiler R, Blatter MC, Estreicher A, Gasteiger E, Martin MJ, Michoud K, O'Donovan C, Phan I, Pilbout S, Schneider M. 2003. The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Research* **31**(1):365–370 DOI [10.1093/nar/gkg095](https://doi.org/10.1093/nar/gkg095).
- Boeckmann B, Blatter MC, Famiglietti L, Hinz U, Lane L, Roechert B, Bairoch A. 2005. Protein variety and functional diversity: swiss-Prot annotation in its biological context. *Comptes Rendus Biologies* **328**(10):882–899 DOI [10.1016/j.crv.2005.06.001](https://doi.org/10.1016/j.crv.2005.06.001).
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**(15):2114–2120 DOI [10.1093/bioinformatics/btu170](https://doi.org/10.1093/bioinformatics/btu170).
- Du G, Zhao HY, Zhang QW, Li GH, Yang FQ, Wang Y, Li YC, Wang YT. 2010. A rapid method for simultaneous determination of 14 phenolic compounds in

- Radix Puerariae using microwave-assisted extraction and ultra high performance liquid chromatography coupled with diode array detection and time-of-flight mass spectrometry. *Journal of Chromatography A* **1217**(5):705–714 DOI [10.1016/j.chroma.2009.12.017](https://doi.org/10.1016/j.chroma.2009.12.017).
- Ernst J, Bar-Joseph Z. 2006.** STEM: a tool for the analysis of short time series gene expression data. *BMC Bioinformatics* **7**:191 DOI [10.1186/1471-2105-7-191](https://doi.org/10.1186/1471-2105-7-191).
- Fairweather TSJ, Bao Y, Broadley MR, Collings R, Ford D, Hesketh JE, Hurst R. 2011.** Selenium in human health and disease. *Antioxidants Redox Signaling* **14**(7):1337–1383 DOI [10.1089/ars.2010.3275](https://doi.org/10.1089/ars.2010.3275).
- Feng R, Wei C, Tu S, Wu F. 2009b.** Effects of Se on the uptake of essential elements in *Pteris vittata* L. *Plant Soil* **325**(1):123–132 DOI [10.1007/s11104-009-9961-9](https://doi.org/10.1007/s11104-009-9961-9).
- Fernanda AS, Lindomar AP, Antonio MS, Laura F, Vctor C, Sergio LCF. 2016.** Development of a MSFIA system for sequential determination of antimony, arsenic and selenium using hydride generation atomic fluorescence spectrometry. *Talanta* **156-157**:29–33 DOI [10.1016/j.talanta.2016.04.063](https://doi.org/10.1016/j.talanta.2016.04.063).
- Finn RD, Coghill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangradorvegas A. 2016.** The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Research* **44**(D1):D279–D285 DOI [10.1093/nar/gkv1344](https://doi.org/10.1093/nar/gkv1344).
- Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng QD. 2011.** Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* **29**(7):644–652 DOI [10.1038/nbt.1883](https://doi.org/10.1038/nbt.1883).
- Han D, Li XH, Xiong SL, Tu SX, Chen ZG, Li JP, Xie ZJ. 2013.** Selenium uptake, speciation and stressed response of *Nicotianatabacum* L. *Environmental and Experimental Botany* **95**(8):6–14 DOI [10.1016/j.envexpbot.2013.07.001](https://doi.org/10.1016/j.envexpbot.2013.07.001).
- Han D, Xiong SL, Tu SX, Liu JC, Chen C. 2015a.** Interactive effects of selenium and arsenic on growth, antioxidant system, arsenic and selenium species of *Nicotianatabacum* L. *Environmental and Experimental Botany* **117**:12–19 DOI [10.1016/j.envexpbot.2015.04.008](https://doi.org/10.1016/j.envexpbot.2015.04.008).
- Han R, Takahashi H, Nakamura M, Yoshimoto N, Suzuki H, Shibata D, Yamazaki M, Saito K. 2015b.** Transcriptomic landscape of *Pueraria lobata* demonstrates potential for phytochemical study. *Frontiers in Plant Science* **6**:Article 426 DOI [10.3389/fpls.2015.00426](https://doi.org/10.3389/fpls.2015.00426).
- He M, Yao Y, Li Y, Yang M, Wu B, Yu D. 2019.** Comprehensive transcriptome analysis reveals genes potentially involved in isoflavone biosynthesis in *Pueraria thomsonii* Benth. *PLOS ONE* **14**(6):e0217593 DOI [10.1371/journal.pone.0217593](https://doi.org/10.1371/journal.pone.0217593).
- He XZ, Blount JW, Ge S, Tang Y, Dixon RA. 2011.** A genomic approach to isoflavone biosynthesis in kudzu (*Pueraria lobata*). *Planta* **233**(4):843–855 DOI [10.1007/s00425-010-1344-1](https://doi.org/10.1007/s00425-010-1344-1).

- Hung CY, Holliday BM, Kaur H, Yadav R, Kittur FS, Xie J. 2012. Identification and characterization of selenate- and selenite-responsive genes in a Se-hyperaccumulator *Astragalus racemosus*. *Molecular Biology Reports* **39**(7):7635–7646 DOI [10.1007/s11033-012-1598-8](https://doi.org/10.1007/s11033-012-1598-8).
- Hurst R, Collings R, Harvey LJ, King M, Hooper L, Bouwman J, Gurinovic M, Fairweather-Tait SJ. 2013. EURRECA—estimating selenium requirements for deriving dietary reference values. *C R C Critical Reviews in Food Technology* **53**(10):1077–1096 DOI [10.1080/10408398.2012.742861](https://doi.org/10.1080/10408398.2012.742861).
- Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M. 2004. The KEGG resource for deciphering the genome. *Nucleic Acids Research* **32**:D277–D280 DOI [10.1093/nar/gkh063](https://doi.org/10.1093/nar/gkh063).
- Koonin EV, Fedorova ND, Jackson JD, Jacobs AR, Krylov DM, Makarova KS, Mazumder R, Mekhedov SL, Nikolskaya AN, Rao BS. 2004. A comprehensive evolutionary classification of proteins encoded in complete eukaryotic genomes. *Genome Biology* **5**:R7–R7 DOI [10.1186/gb-2004-5-2-r7](https://doi.org/10.1186/gb-2004-5-2-r7).
- Li HF, McGrath SP, Zhao FJ. 2008. Selenium uptake, translocation and speciation in wheat supplied with selenate or selenite. *New Phytologist* **178**:92–102 DOI [10.1111/j.1469-8137.2007.02343.x](https://doi.org/10.1111/j.1469-8137.2007.02343.x).
- Li J, Li CF, Gou JB, Wang X, Fan RY, Zhang YS. 2016a. An alternative pathway for formononetin biosynthesis in *Pueraria lobata*. *Frontiers in Plant Science* **7**:Article 861 DOI [10.3389/fpls.2016.00861](https://doi.org/10.3389/fpls.2016.00861).
- Li J, Li CF, Gou JB, Zhang YS. 2016b. Molecular cloning and functional characterization of a novel isoflavone 3'-O-methyltransferase from *Pueraria lobata*. *Frontiers in Plant Science* **7**:Article 793 DOI [10.3389/fpls.2016.00793](https://doi.org/10.3389/fpls.2016.00793).
- Li YM, Sun YX, Liu DH. 2010. Effects of Applying Selenium on Contents of Total Flavonoid, Chlorogenic Acid and Selenium in Flower of *Chrysanthemum morifolium*. *Soil* **42**(4):618–623 DOI [10.13758/j.cnki.tr.2010.04.025](https://doi.org/10.13758/j.cnki.tr.2010.04.025).
- Lowe ED, Gao GY, Johnson LN, Keung WM. 2008. Structure of daidzin, a naturally occurring anti-alcohol-addiction agent, in complex with human mitochondrial aldehyde dehydrogenase. *Journal of Medicinal Chemistry* **51**(15):4482–4487 DOI [10.1021/jm800488j](https://doi.org/10.1021/jm800488j).
- Marchler BA, Zheng C, Chitsaz F, Derbyshire MK, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Lanczycki CJ. 2013. CDD: conserved domains and protein three-dimensional structure. *Nucleic Acids Research* **41**:348–352 DOI [10.1093/nar/gks1243](https://doi.org/10.1093/nar/gks1243).
- Nandini Y, Samir S. 2016. Reactive oxygen species, oxidative stress and ROS scavenging system in plants. *Journal of Chemical and Pharmaceutical Research* **8**:595–604 DOI [10.1371/journal.pone.0217593](https://doi.org/10.1371/journal.pone.0217593).
- Özgür Ç, Neslihan TK, Ş Arı, Baohong Z. 2015. De Novo Transcriptome Assembly and Comparative Analysis Elucidate Complicated Mechanism Regulating *Astragalus chrysochlorus* Response to Selenium Stimuli. *PLOS ONE* **10**(10):e0135677 DOI [10.1371/journal.pone.0135677](https://doi.org/10.1371/journal.pone.0135677).

- Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. 2017.** Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods* **14**(4):417–419 DOI [10.1038/nmeth.4197](https://doi.org/10.1038/nmeth.4197).
- Qiu ZX, Sha ZS, Che XM, Wang MY. 2017.** Correlation analysis of ADAMTS-4, VCAM-1, and TAK1 expression in cartilage tissue from spine tuberculosis. *Genetics and Molecular Research* **16**:Article 3 DOI [10.4238/gmr16038961](https://doi.org/10.4238/gmr16038961).
- Tatusov RL, Galperin MY, Natale DA, Koonin EV. 2000.** The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Research* **28**(1):33–36 DOI [10.1093/nar/28.1.33](https://doi.org/10.1093/nar/28.1.33).
- Terry N, De Souza MP, Am TAZ. 2000.** Selenium in higher plants [Review]. *Annual Review of Plant Physiology & Plant Molecular Biology* **51**:401–432 DOI [10.1146/annurev.arplant.51.1.401](https://doi.org/10.1146/annurev.arplant.51.1.401).
- The Chinese Pharmacopoeia. 2005.** *The chemical industry publishing house*. Beijing: Chinese Pharmacopoeia Commission, 203.
- Wang X, Fan R, Li J, Li C, Zhang Y. 2016.** Molecular Cloning and Functional Characterization of a Novel (Iso)flavone 4', 7-O-diglucoside Glucosyltransferase from *Pueraria lobata*. *Frontiers in Plant Science* **7**:387 DOI [10.3389/fpls.2016.00387](https://doi.org/10.3389/fpls.2016.00387).
- Wang X, Li C, Zhou C, Li J, Zhang Y. 2017.** Molecular characterization of the C-glucosylation for pueraria biosynthesis in *Pueraria lobata*. *Plant Journal* **90**(3):535–546 DOI [10.3389/fpls.2016.00387](https://doi.org/10.3389/fpls.2016.00387).
- Wang YD, Wang X, Wong YS. 2012.** Proteomics analysis reveals multiple regulatory mechanisms in response to selenium in rice. *Journal of Proteomics* **75**(6):1849–1866 DOI [10.1016/j.jprot.2011.12.030](https://doi.org/10.1016/j.jprot.2011.12.030).
- Wei SY, Chen Y, Xu XY. 2014.** Progress on the pharmacological research of puerarin: a review. *Chinese Journal of Natural Medicines* **12**(6):407–414 DOI [10.1016/S1875-5364\(14\)60064-9](https://doi.org/10.1016/S1875-5364(14)60064-9).
- Winkel LHE, Vriens B, Jones GD, Schneider LS, Bañuelos GS. 2015.** Selenium cycling across soil-plant-atmosphere interfaces: a critical review. *Nutrients* **7**(6):4199–4239 DOI [10.3390/nu7064199](https://doi.org/10.3390/nu7064199).
- Wu B, Long Q, Gao Y, Wang Z, Shao T, Liu Y, Li Y, Ding W. 2015.** Comprehensive characterization of a time-461 course transcriptional response induced by autotoxins in *Panax ginseng* using RNA-Seq. *BMC Genomics* **462**(16):1010 DOI [10.1186/s12864-015-2151-7](https://doi.org/10.1186/s12864-015-2151-7).
- Yu XZ, Gu JD. 2008.** Differences in uptake and translocation of selenate and selenite by the weeping willow and hybrid willow. *Environmental Science & Pollution Research* **15**(6):499–508 DOI [10.1007/s11356-008-0036-x](https://doi.org/10.1007/s11356-008-0036-x).
- Zhang M, Hu C, Zhao X, Tan Q, Sun X, Cao A, Cui M, Zhang Y. 2012b.** Molybdenum improves antioxidant and osmotic-adjustment ability against salt stress in Chinese cabbage (*Brassica campestris* L. ssp. *Pekinensis*). *Plant Soil* **355**(1):375–383 DOI [10.1007/s11104-011-1109-z](https://doi.org/10.1007/s11104-011-1109-z).
- Zhao XQ, Mitani N, Yamaji N, Shen RF, Ma JF. 2010.** Involvement of silicon influx transporter OsNIP2;1 in selenite uptake in rice. *Plant Physiology* **153**(4):1871–1877 DOI [10.1104/pp.110.157867](https://doi.org/10.1104/pp.110.157867).

**Zheng Y, Jiao C, Sun H, Rosli HG, Pombo MA, Zhang P, Banf M. 2016.** iTAK: A Program for Genome-wide Prediction and Classification of Plant Transcription Factors, Transcriptional Regulators, and Protein Kinases. *Molecular Plant* **9(12)**:1667–1670  
[DOI 10.1016/j.molp.2016.09.014](https://doi.org/10.1016/j.molp.2016.09.014).