



De novo Transcriptome Assembly of a Chinese Locoweed (*Oxytropis ochrocephala*) Species Provides Insights into Genes Associated with Drought, Salinity, and Cold Tolerance

Wei He^{1†}, Huihui Zhuang^{1†}, Yanping Fu¹, Linwei Guo¹, Bin Guo¹, Lizhu Guo¹, Xihong Zhang² and Yahui Wei^{1*}

¹ Department of Biology, Northwest University, Xian, China, ² Grassland Station, Agriculture and Animal Husbandry Bureau, Zhongwei, China

OPEN ACCESS

Edited by:

Georgina Hernandez,
Universidad Nacional Autonoma de
Mexico, Mexico

Reviewed by:

Jose Luis Reyes,
Universidad Nacional Autonoma de
Mexico, Mexico
Paul Thomas Scott,
Terragen Biotech Pty Ltd, Australia

*Correspondence:

Yahui Wei
weiyahui@nwu.edu.cn

[†]These authors have contributed
equally to this work.

Specialty section:

This article was submitted to
Plant Genetics and Genomics,
a section of the journal
Frontiers in Plant Science

Received: 08 July 2015

Accepted: 19 November 2015

Published: 02 December 2015

Citation:

He W, Zhuang H, Fu Y, Guo L, Guo B,
Guo L, Zhang X and Wei Y (2015) De
novo Transcriptome Assembly of a
Chinese Locoweed (*Oxytropis
ochrocephala*) Species Provides
Insights into Genes Associated with
Drought, Salinity, and Cold Tolerance.
Front. Plant Sci. 6:1086.
doi: 10.3389/fpls.2015.01086

Background: Locoweeds (toxic *Oxytropis* and *Astragalus* species), containing the toxic agent swainsonine, pose serious threats to animal husbandry on grasslands in both China and the US. Some locoweeds have evolved adaptations in order to resist various stress conditions such as drought, salt and cold. As a result they replace other plants in their communities and become an ecological problem. Currently very limited genetic information of locoweeds is available and this hinders our understanding in the molecular basis of their environmental plasticity, and the interaction between locoweeds and their symbiotic swainsonine producing endophytes. Next-generation sequencing provides a means of obtaining transcriptomic sequences in a timely manner, which is particularly useful for non-model plants. In this study, we performed transcriptome sequencing of *Oxytropis ochrocephala* plants followed by a *de novo* assembly. Our primary aim was to provide an enriched pool of genetic sequences of an *Oxytropis* sp. for further locoweed research.

Results: Transcriptomes of four different *O. ochrocephala* samples, from control (CK) plants, and those that had experienced either drought (20% PEG), salt (150 mM NaCl) or cold (4°C) stress were sequenced using an Illumina HiSeq 2000 platform. From 232,209,506 clean reads 23,220,950,600 (~23 G nucleotides), 182,430 transcripts and 88,942 unigenes were retrieved, with an N50 value of 1237. Differential expression analysis revealed putative genes encoding heat shock proteins (HSPs) and late embryogenesis abundant (LEA) proteins, enzymes in secondary metabolite and plant hormone biosyntheses, and transcription factors which are involved in stress tolerance in *O. ochrocephala*. In order to validate our sequencing results, we further analyzed the expression profiles of nine genes by quantitative real-time PCR. Finally, we discuss the possible mechanism of *O. ochrocephala*'s adaptations to stress environment.

Conclusion: Our transcriptome sequencing data present useful genetic information of a locoweed species. This genetic information will underpin further research in elucidating the environmental acclimation mechanism in locoweeds and the endophyte-plant association.

Keywords: locoweeds, *Oxytropis ochrocephala*, transcriptome, stress, Illumina sequencing

INTRODUCTION

Toxic *Oxytropis* and *Astragalus* species (Fabaceae), collectively named locoweeds, pose a threat to livestock on grasslands, primarily in the U.S and China (Zhao et al., 2013). Locoweeds contain the indolizidine alkaloid swainsonine, an α -mannosidase inhibitor which causes over-accumulation of mannose rich oligosaccharide in the lysosomes which impairs the neural system of livestock (Tulsiani et al., 1988). The symptoms of affected animals include staggering, reduced browsing and water consumption, and even death (Ralphs et al., 2002). Locoweed consumption results in a deterioration of animal health and, as a consequence of weight loss, a reduction in meat productive rate. In addition to damaging livestock, the spread of locoweeds destabilizes grassland plant communities and undermines sustainable grassland management. In both the U.S and China, significant annual economic loss has been reported, in some cases up to 20 million USD per year (Creamer and Baucom, 2013; Zhao et al., 2013).

In China, locoweeds are widely distributed and often grow beyond temperate areas, where they encounter reduced rainfall, low temperature and high soil salinity. For example, *O. ochrocephala*, one of the most widely distributed locoweeds in China, is abundant on the grasslands of Qinghai, Gansu, Ningxia and Sichuan Provinces (Zhao et al., 2013), where it occurs in mountains, alpine meadows, open grasslands, and valleys. *O. ochrocephala*'s tolerance to environmental stresses contributes to it rapidly replacing local forage grass species in natural grassland plant communities [Wei He, personal observation]. Plants are often exposed to stressful environment, and have evolved to either avoid or tolerate these abiotic stress, by altering their gene expression, signal transduction, and cellular metabolism. In particular, the underlying gene expression changes trigger accumulation of protective proteins such as heat shock proteins (HSPs) (Wang et al., 2004) and late embryogenesis abundant (LEA) protein (Hundertmark and Hincha, 2008), and regulatory proteins, such as key enzymes in phytohormones synthesis (e.g., abscisic acid and ethylene; Peleg and Blumwald, 2011) and transcription factors that in turn regulate downstream stress inducible genes (Yamaguchi-Shinozaki and Shinozaki, 2006). Currently, the adaptation mechanisms of locoweeds that enable these plants to cope with environmental stress are still ambiguous.

On the other hand, although it has been proved that the production of swainsonine in locoweeds can be attributed primarily to their endophytes (Gardner et al., 2001; Ralphs et al., 2008; Pryor et al., 2009), it is unclear how the intrinsic properties of the plants contribute to the production of locoweed swainsonine. Oldrup et al. (2010) demonstrated in *O. sericea* plants subjected to certain stress conditions, such as drought and low pH, dry mass of its endophyte *U. oxytropis* declines, contrasting to elevated plant biomass and swainsonine concentration. Cook et al. (2013) observed that interactions between endophytes and locoweeds does offer the endophytes nutrients, hormones and some other signals in influencing their capacity to produce swainsonine. These studies imply that locoweeds, as the host of the endophytes, have a direct impact

on the production of swainsonine, but the interaction between locoweeds and their symbiotic endophytes remains unclear and needs further investigation.

To resolve these problems, in-depth research at the genetic level is necessary. However, to date limited genetic information of locoweeds is available from public databases (Chung et al., 2004; Archambault and Strömvik, 2011). Although, an EST dataset of a suppressive subtraction cDNA library enriched in genes from two temperate *Oxytropis* species has recently been made available from the NCBI, this only has 1245 ESTs and thus there is a need for further genetic information work in *Oxytropis*. Recent advances in sequencing technology, a reduction in costs, and the development of more data handling programs have resulted in more efforts in genome or transcriptome sequencing as a means of obtaining useful genetic profiles (Dugas et al., 2011; Wei et al., 2011; Liu et al., 2012; Zhou et al., 2012; Gross et al., 2013; Chen et al., 2014). In this study, we analyzed the transcriptomes of *O. ochrocephala* under drought, salt and cold stress. Our primary aim was to provide an enriched pool of genetic sequences of an undescribed *Oxytropis* species for further locoweed research. We use the data obtained to further discuss the possible mechanism of *O. ochrocephala*'s adaptations to stress environmental conditions. The sequences obtained in this study will underpin further research in elucidating the environmental acclimation mechanisms in locoweeds, and in serving as a reference transcriptome to advance the study of endophyte-plant association.

MATERIALS AND METHODS

Plant Material and RNA Extraction

Mature *O. ochrocephala* seeds were collected from Haiyuan, Ningxia Province (N 36°26'49.80," E 105°37'26.46," altitude: 2656) in July 2013. Seeds were scarified with sandpaper and then imbibed in deionized water for 12 h. All seeds were then placed onto wet filter papers in Petri dishes for germination. Germinated seeds were allowed to grow until the cotyledons emerged. Each seedling was then transferred into an individual pot (5 × 5 × 6 cm) containing a mixture of sand and peat (1:1), and placed into a growth chamber under controlled conditions (photo flux density of 300 $\mu\text{mol m}^{-2}\text{s}^{-1}$, 14/10 h day/night period; relative humidity of 55–60%; temperature of 25 \pm 2°C). All plants were watered weekly until 1 week before stress treatments.

A total of 120 robust 6-week old seedlings were selected and randomized evenly into four groups: (1) control, (2) drought, (3) cold, (4) salt. For the control and cold treatment, plants were watered as normal except that for the cold treatment 10 plants were transferred into a different growth chamber pre-cooled to 4°C. For the drought and salt treatments, 20% PEG-6000 (w/v, polyethylene glycol, Sangon, China) or 150 mM NaCl (Sangon, China) solutions were applied to irrigate the plants in every 4 h, respectively. The rest of growing conditions were as previously described above. A preliminary experiment was set up to test the effect of the stress treatments to the seedlings. Relative electrolyte leakage, malonaldehyde proline content (MDA), proline content,

and superoxide dismutase (SOD) and peroxidase (POD) activity were determined at 0, 6, 12, and 24 h post-treatment. Relative water content (RWC) and survival rate were determined 4 d after the treatments. Results showed that all the three stress treatments altered the physiological performance of the plants, and therefore demonstrated that they are effective in the laboratory in mimicking drought, salinity and cold stress in the field (Supplementary Figure 1 in **Data Sheet 1**).

Plants materials were collected from all individuals in the four groups at 3, 6, and 12 h after the treatments. The three time-points were selected considering that changes at the transcriptomic level are much more rapid than those at the physiological level, as suggested by previous study in *Arabidopsis* (Kreps et al., 2002). Each sample consisted of 10 seedlings to ensure adequate RNA for RNA-seq. Each whole seedling was removed from the soil, washed and dried, and then immersed in liquid nitrogen immediately. RNA was extracted according to the method of Gasic et al. (2004). To eliminate DNA contamination, total RNA was DNase treated and purified according to the protocol of the manufacturer (Ambion, USA). The integrity of all RNA samples was assessed using 1% denatured agarose gel electrophoresis. RNA purity and concentration were both determined on a NanoDrop ND-1000 Spectrophotometer (NanoDrop Technologies, USA).

Library Preparation and Sequencing

A total of 3 μ g RNA was prepared from pooling an equal amount of the three post-treatment RNA samples (3, 6, and 12 h for each treatment). Strand non-specific sequencing libraries were generated using NEBNext[®] Ultra[™] RNA Library Prep Kit for Illumina (NEB, USA) following the instructions of the manufacturer. Four index codes were added as indicators for sequences from each sample. Briefly, mRNA was purified from total RNA using oligo-dT attached magnetic beads. Fragmentation was carried out using divalent cations under elevated temperature. First strand cDNA was reverse transcribed using random hexamer primers, with second strand cDNA synthesized subsequently. Remaining overhangs were blunted via exonuclease/polymerase activities. After adenylation of 3' ends of DNA fragments, NEBNext Adaptor with hairpin loop structure were ligated to prepare for hybridization. In order to select cDNA fragments with a preference of 150~200 bp in length, the library fragments were purified with the AMPure XP system (Beckman Coulter, USA). Then 3 μ l USER Enzyme (NEB) was used with size-selected, adaptor-ligated cDNA at 37°C for 15 min followed by 5 min at 95°C before PCR. PCR was performed with Phusion High-Fidelity DNA polymerase, Universal PCR primers and Index Primer. Finally, PCR products were purified with the AMPure XP system (Beckman Coulter, USA). Library quality was assessed with the Agilent Bioanalyzer 2100 system (Agilent Technologies, USA).

The clustering of the index-coded samples was performed on a cBot Cluster Generation System using TruSeq PE Cluster Kit v3-cBot-HS (Illumina, USA) according to the instructions of the manufacturer. After cluster generation, the library preparations were sequenced on an Illumina HiSeq 2000 platform (Illumina, USA) and paired-end reads were generated.

Sequence Annotation

Raw data (raw reads, NCBI Accession: PRJNA292613) were initially processed to generate clean data (clean reads). At this step Q20, Q30, GC-content and sequence duplication levels of the clean data were calculated. *De novo* assembly of the transcriptome was accomplished using Trinity version r20140413p1 (Grabherr et al., 2011) using the Butterfly option, with all parameters set to default except for min_kmer_cov set as 2. Gene function and classification was annotated based on the following databases: Nr, Nt (NCBI non-redundant nucleotide sequences), GO (Gene Ontology), Swiss-Prot, Pfam (Protein family), KOG (euKaryotic Ortholog Groups) and KEGG (Kyoto Encyclopedia of Genes and Genomes).

NCBI blast version 2.2.31+ was employed for annotation against Nr, Nt, SwissProt and KOG databases. The cut-off value was set to $1e^{-5}$ for Nr, Nt and Swiss-Prot, and $1e^{-3}$ for KOG. In order to calculate the percent length coverage of unigenes, NCBI blast 2.2.31+ was also used to determine the number of unique top matching proteins that align across more than X% of the lengths of the unigenes (http://trinityrnaseq.sourceforge.net/analysis/full_length_transcript_analysis.html). Pfam annotation was retrieved using HMMER 3.0 package (Finn et al., 2011), with a cut-off value of 0.01. The Nr and Pfam hits were imported into Blast2GO v2.5 for GO classification (Götz et al., 2008), with a cut-off value of $1e^{-6}$. KEGG annotation was obtained from the KEGG Automatic Annotation Server (Kanehisa et al., 2008). For species similarity analysis, species hits in Nr annotation with the lowest *e*-values were summarized and the results integrated.

Differential Expression Analysis

We used the RSEM package version 1.2.0 (Li and Dewey, 2011) to analyze the read counts, which were then converted to FPKM (Fragments Per Kilobase of transcript per Millions fragments sequenced), a commonly accepted estimate for the expression level of unigenes (Trapnell et al., 2010). Prior to differential gene expression analysis, for each sequenced library, we adjusted the read counts with the edgeR program package version 3.0.8 (Robinson et al., 2010) through one scaling normalized factor. Differential expression analysis between the control and each of the three treatment samples was performed using the edgeR package with BCV (Biological coefficient of variation) set as 0.2 (dispersion = 0.04) for exactTest and other parameters using default settings. The threshold for significantly differential expression was set as $q \leq 0.005$ and $|\log_2(\text{fold change})| \geq 1$.

GO Enrichment and KEGG Pathway Enrichment Analysis

We performed GO enrichment analysis of the differentially expressed genes (DEGs) was implemented using the GOseq R packages version 1.10.0 based on the Wallenius non-central hyper-geometric distribution (Young et al., 2010), which can adjust for gene length bias in DEGs. KEGG is a database resource for understanding high-level functions and utilities of the biological system, such as the cell, the organism and the ecosystem, from molecular-level information, especially large-scale molecular datasets generated by genome sequencing and other high-throughput experimental technologies (<http://www.>

genome.jp/kegg/). We used KOBAS software version 2.0 (Mao et al., 2005) to test the statistical enrichment of differential expression genes in KEGG pathways (corrected $P < 0.05$).

Quantitative Real-Time PCR Determination of Gene Expression

We performed stress treatments and RNA extraction as described above. For cDNA synthesis, 2.5 μ g of total RNA was used for reverse transcription with random hexmers according to the instructions of the manufacturer (Thermo Scientific, USA). The cDNA was then diluted 30-fold with nuclease-free water for qRT-PCR. qRT-PCR was performed using a Bio-Rad CFX96 Real-Time PCR system (Bio-Rad, USA). The expression levels of nine genes involved in stress responses were determined (primer sequences, as well as for housekeeping genes were listed in Supplementary Table 1 in **Data Sheet 1**). The qRT-PCR reactions were performed using the FastStart Universal SYBR GreenMaster system according to the instructions of the manufacturer (Roche, Germany). The qRT-PCR condition was set as recommended by the manufacturer: 95°C for 10 min, 40 cycles at 95°C for 30 s, 58°C for 30 s, and 72°C for 30 s with a dissolving curve followed. Three biological replications and three technical replications, respectively, were used. Relative expression levels were normalized using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001).

RESULTS

Pair-End Sequences and Assembly

Transcriptomes of four different *O. ochrocephala* samples, including the control (CK), drought (20% PEG), cold (4°C), and salt (150 mM NaCl) stress were sequenced. A total number of 232,209,506 clean reads were generated from sequence reads of the four transcriptomes (NCBI Accession: PRJNA292613), corresponding to 23,220,950,600 (23 G) nucleotides (nt) (Supplementary Table 2 in **Data Sheet 1**). The average Q20 and Q30 values are 97.61 and 93.06%, respectively; the average GC content is 42.8% (Supplementary Table 2 in **Data Sheet 1**).

A *de novo* assembly was applied employing all the transcripts from the four treatments. After assembly, 217,270 transcripts (238,301,188 nt) and 118,596 unigenes (87,242,224 nt) were retrieved (**Table 1**, Assembled sequences of all unigenes were included in **Data Sheet 2**). Because a proportion of the total unigenes were identified to have non-plant origins after a blast search against the Nr database, these unigenes and the corresponding transcripts were removed from the transcript and

unigene pools for downstream analyses. After this adjustment, of all 88,942 unigenes the median length is 368 nt, and N50 is 1237. The length distribution of all the transcripts and unigenes showed that as the length of unigenes increase, the number of unigenes decrease (Supplementary Figure 2 in **Data Sheet 1**); approximately 17,111 (19%) unigenes are longer than 1 kb.

Sequence Annotation and Similarity Analysis

Sequence annotation against seven different public nucleotide/protein databases resulted in the successful annotation of 40,059 (45.03%) unigenes in at least one of the databases, while approximately 55% unigenes remain unmapped in any existing database (**Table 2**). According to Nr annotation (Nr annotation of unigenes were included in **Data Sheet 3**), 37,363 (42.01%) unigenes were mapped, whilst the remaining unmapped unigenes mainly consist of those ones with lengths ≤ 1000 bp (49,464 out of 51,579, 95.91%). We calculated the percent length coverage of unigenes matching translated sequences in Nr database (**Table 3**). This showed that there are 14,501 (38.8%) unigenes that are represented by nearly full-length transcripts, having $>80\%$ alignment coverage.

We further analyzed the annotated unigenes for their organism similarity to other plant species in the NCBI Nr database. Species hits from Nr were selected and the most abundant 12 species are shown (**Figure 1**). The top five matches

TABLE 2 | Summary of unigenes annotation against public available databases.

	Number of unigenes	Percentage (%)
Annotated in Nr	37,363	42.01
Annotated in Nt	24,328	27.35
Annotated in KO	10,465	11.76
Annotated in SwissProt	21,844	24.55
Annotated in Pfam	24,552	27.60
Annotated in GO	24,983	28.08
Annotated in KOG	11,596	13.03
Annotated in all Databases	5198	5.84
Annotated in at least one Database	40,059	45.03
Total Unigenes	88,942	100

Nr, NCBI non-redundant protein sequences; Nt, NCBI non-redundant nucleotide sequences; KO, KEGG Ortholog database; Swiss-Prot, A manually annotated and reviewed protein sequence database; Pfam, Protein family; GO, Gene Ontology; KOG/COG, Clusters of Orthologous Groups of proteins.

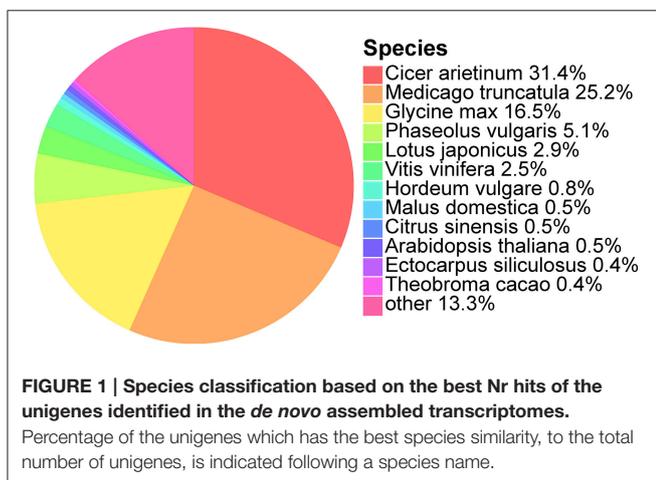
TABLE 1 | Summary of *de novo* assembly of transcriptome sequencing reads without reference genome.

	Median length (nt)	Max length (nt)	N50	N90	Total number	Total nucleotides (nt)
Transcripts	658 (706)	15,782 (15,782)	1893 (1920)	440 (466)	217,270 (182,430)	238,301,188 (206,141,699)
Unigenes	379 (368)	15,782 (15,782)	1316 (1237)	280 (273)	118,596 (88,942)	87,242,224 (62,980,053)

nt, nucleotide; N50 (N90), defined as the length for which the collection of all sequences of that length or longer contains at least half (90%) of the sum of the lengths of all sequences. Non-plant transcripts and unigenes were removed from each pool and adjusted numbers were indicated in brackets.

TABLE 3 | Distribution of percent length coverage of unigenes matching translated sequences in the Nr database.

Coverage	Unigene count	Percentage (%)	Accumulated unigene count	Accumulated percentage (%)
100% ≥ X > 90%	12,233	32.7	12,233	32.7
90% ≥ X > 80%	2268	6.1	14,501	38.8
80% ≥ X > 70%	1917	5.1	16,418	43.9
70% ≥ X > 60%	2136	5.7	18,554	49.6
60% ≥ X > 50%	2195	5.9	20,749	55.4
50% ≥ X > 40%	2359	6.3	23,108	61.7
40% ≥ X > 30%	2844	7.6	25,952	69.3
30% ≥ X > 20%	3747	10.0	29,699	79.3
20% ≥ X > 10%	5026	13.5	34,725	92.8
10% ≥ X > 0%	2638	7.1	37,363	100



are *Cicer arietinum* (31.4%), *Medicago truncatula* (25.2%), *Glycine max* (16.5%), *Phaseolus vulgaris* (5.1%), and *Lotus japonicas* (2.9%). All of these five species belong to the Fabaceae, the same family as *O. ochrocephala*, and explains up to 82.1% of all the annotated unigenes. The remaining unigenes (17.9%) are similar to other 401 plant species.

Functional Classification of the Unigenes

In order to study the putative function of the unigenes we identified, we classified all unigenes and their deduced protein into different functional groups, by employing GO and KOG annotation, and KEGG pathway classification. GO annotation is the most widely used classification to group genes by their putative function, and similarly, KOG and KEGG pathway classifications also provided useful information in clustering putative genes or proteins by their functions. We present GO annotation here as an example of gene functional classification (Figure 2). In GO annotation, 24,983 out of 88,942 unigenes (28.08%) are successfully assigned into 46 functional groups, all of which belong to one of the basic three ontologies: biological process, cellular component, and molecular function (Figure 2). The highest proportion of these unigenes is characterized by “cellular process,” “metabolic process,” “binding,” “catalytic

process,” “cell and cell part.” KOG and KEGG pathways are shown in Supplementary Figures 3, 4 in Data Sheet 1.

Differentially Expressed Genes Under Stress Conditions

In order to determine the patters of differentially expressed genes (DEGs) in *O. ochrocephala* under drought, salinity and cold stress, normalized read counts of all unigenes and gene expression levels, estimated as FPKM, were determined. This showed distinct expression patterns among the treatments and control (Figures 3, 4). The most distinct gene expression pattern was exhibited by cold stressed plants, followed by those plants that had experienced saline and drought stress. Through, GO and KEGG pathway enrichment, DEGs were further categorized into different functional groups (Table 4, Data Sheets 4, 5). Most of the genes relate to enzymes in plant metabolite biosynthesis pathways or transcription factors, and are up-regulated in the three stress treatments. However, we observed that more genes are involved in cold stress than those in drought and saline stress, especially drought stress.

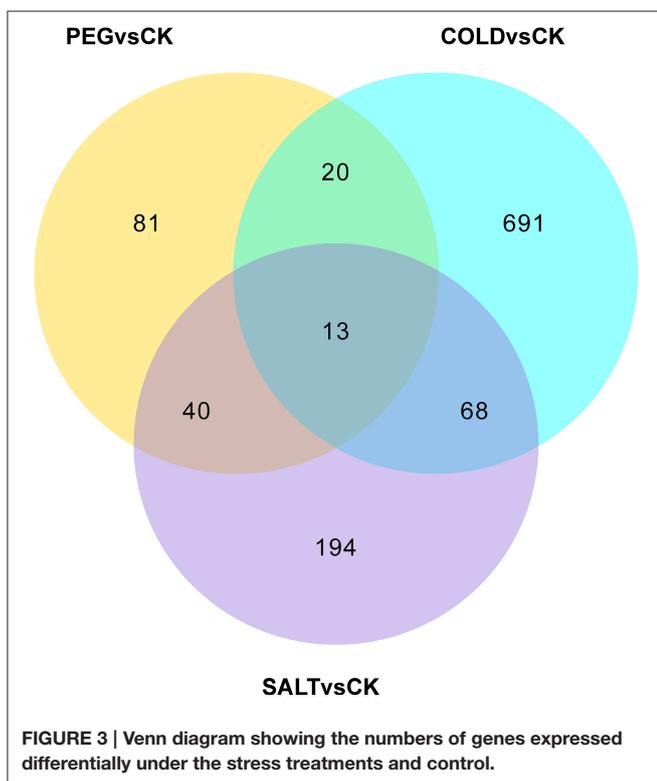
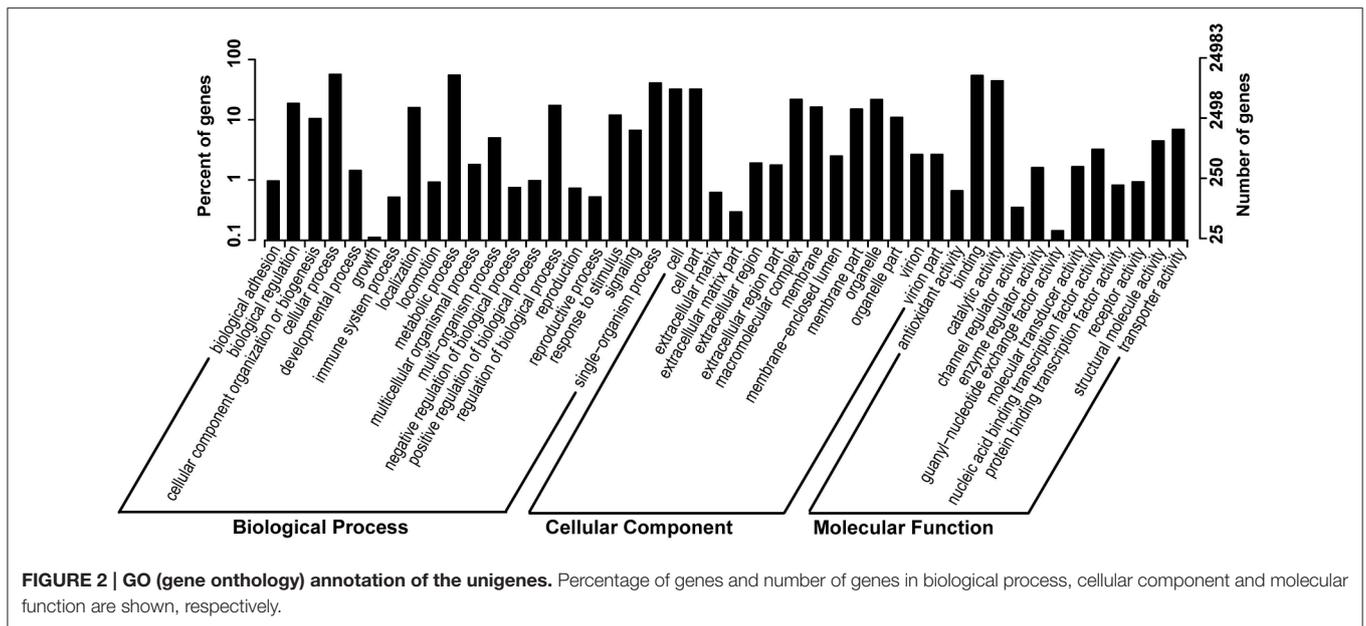
In order to validate the transcriptomic sequencing results from DEGs, we further performed qRT-PCR (quantitative Real-Time PCR) analyses on nine genes involved in the three stress responses (Figure 5). These genes were selected based on the fact that either their expression patterns change remarkably according to the DEG data or their functions had been well established in previous studies in plant stress responses to drought, salinity and cold. All the genes we tested are up-regulated in response to the three stress treatments, respectively. In general, the differential expression profiles from the qRT-PCR results showed similar patterns to FPKM fold change. Nevertheless, the expression fold changes in four genes are more drastic in the qRT-PCR results than those in the DEGs.

DISCUSSION

Sequencing and Annotation

In this study we performed RNA sequencing and report a *de novo* assembly of a common locoweed species in China, *O. ochrocephala*. In total, 88,942 unigenes with an N50 of 1237 were obtained, of which 19% are longer than 1 kb. Furthermore, of the 37,363 unigenes annotated to Nr, 38.8% are near full-length sequences matching >80% of the matched protein sequences. Our data are comparable to many other transcriptome studies (Garg et al., 2011; Xie et al., 2012; Fu et al., 2013; Farrell et al., 2014; Ge et al., 2014), and indicate satisfactory sequencing and assembly quality. Nevertheless, those un-annotated unigenes demonstrate that in *Oxytropis* spp, sequencing depth has not yet been fully reached, and some of these sequences may have functions specific to the genus.

It is interesting that in GO annotation, very few unigenes were observed within the term “growth.” Indeed, seedlings of *O. ochrocephala* in this study had reduced growth rates when established from seeds and subsequently grew slowly in a greenhouse. Different attempts were made to change the growth conditions by altering the combination of light, temperature, moisture, and seasonal time of planting, and by

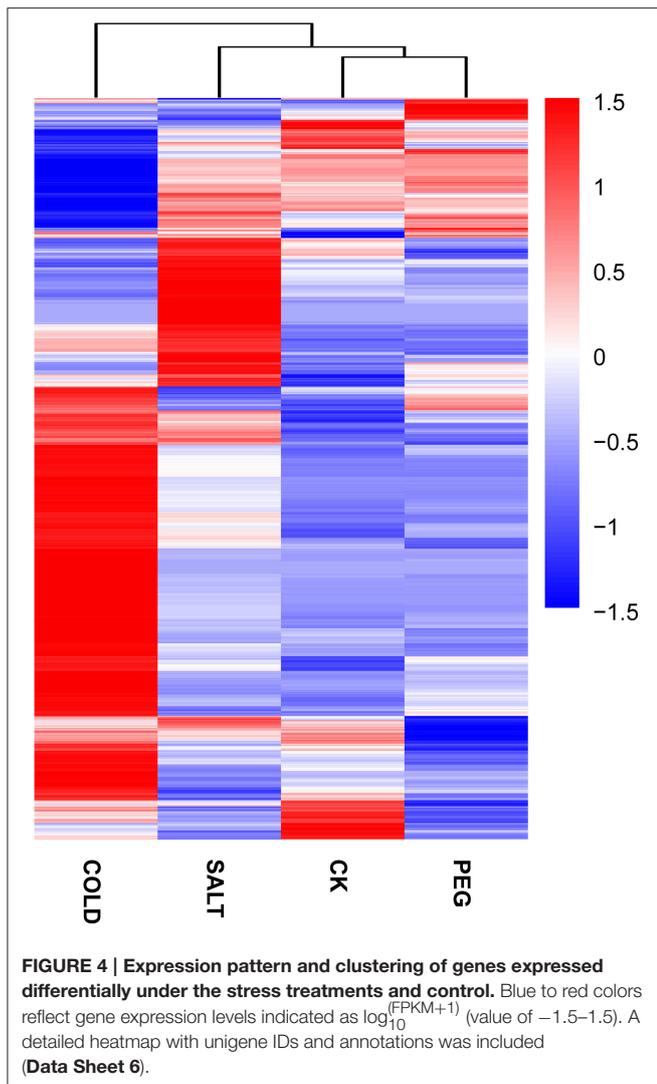


using seed accessions from various sites in China. None of these trials produced optimal seedling growth. In addition, we have observed that *O. ochrocephala* plants grown on grasslands with intensive grazing have reduced plant height than the plants in a reserve park in Ningxia Province with lower grazing pressure. Moreover, when field collected seeds develop into seedlings in the laboratory, Ningxia accession always outperformed Qinghai

accession (data not shown). We suggest that the slow growth rate of *O. ochrocephala* is likely to be attributed to its life cycle strategy, in which the syntheses of toxins and antinutrients are trade-off to growth. As defensive compounds, secondary metabolites are costly in nature and their production may occur at the expense of plants' growth (Vincent et al., 2005; Treutter, 2006). Ecological theory predicts that plants may invest less energy and resources into growth than into accumulating defensive metabolites as a strategy to deter herbivores, to subsequently outcompete faster growing intra- and/or inter-specific individuals (Blumenthal, 2006; Fine et al., 2006).

Differential Expression Patterns of Stress Response Genes by Sequencing

DEG analysis revealed genes expressed differentially under drought, cold, and saline stress conditions in *O. ochrocephala*. There are more genes involved in cold stress response than those in drought and saline stress, suggesting that *O. ochrocephala* plants are not as sensitive to drought and saline stress as they are to cold stress. These genes in different functional groups are highly likely to be involved in stress response in *O. ochrocephala*. Firstly, the expression levels of protective proteins are altered, such as HSP and LEA. HSPs are a group of stress-induced proteins, which have been shown to act as a chaperone of proteins that regulate growth and development, thus preventing proteins from aggregating and denaturing (Kotak et al., 2004; Nishizawa et al., 2006). We identified 11 HSP unigenes in *O. ochrocephala*. One unigene is closely related to HSP83, one to HSP81, one to HSP70, and the others are annotated as either "small heat shock protein" or "heat shock factors." These unigenes have top BLAST hits against sequences from legumes such as *Cicer arietinum*, *Medicago truncatula*, and *Glycine max*, to non-legumes such as *Ziziphus jujube*. In addition, three unigenes, with the highest similarity to sequences in *Medicago truncatula*, *Medicago sativa*,



and *Glycine max*, respectively, were all annotated to an 18.2 kDa class I heat shock protein. These results suggest that there may be a diversity of HSP genes in *O. ochrocephala*. In this study, DEGs of all HSPs were only found in the cold treatment and they are down-regulated by on average 6.1-fold. This is in contrast to many studies, in which HSPs usually show changes with both increase and decrease (Hu et al., 2009). However, no transcripts of HSPs were found with significantly differential expression profiles in the drought and saline treatments, probably indicating that the expression of HSPs in these two treatments are not drastically altered.

LEA proteins are ubiquitous proteins that are highly stress inducible. LEA proteins are rapidly synthesized in those vegetative tissues experiencing water deficiency and seeds in desiccation in both desiccation insensitive and resurrection plants (Hundertmark and Hinch, 2008). LEA proteins may prevent molecular denaturation, by facilitating hydration in target proteins and other macromolecules (Wang et al., 2004). Several functional studies have confirmed a protective role of LEA proteins and dehydrins against drought and osmotic stress

tolerance (NDong et al., 2002). In *O. ochrocephala*, many LEA genes express abundant transcripts in the three abiotic stresses, and except one unigene, all the others are up-regulated.

Secondly, as regulatory proteins, key enzymes in phytohormones synthesis and transcription factors are up-regulated by the three stress treatments. Secondary metabolites possess diverse protective or signaling functions in plant defense. In addition, they help certain plants gain complete edge over the others in a community through a mechanism known as allelopathy (Hierro and Callaway, 2003). In our transcriptome dataset, DEGs were identified in the biosynthesis pathway of flavonoids, alkaloid, phenolics, lignin, coumarins, anthocyanins, cyanogenic glycosides, terpenoids, and steroidal compounds (Table 4). Many of the DEGs are in the flavonoid pathway, one of the best-characterized metabolic pathways in the secondary metabolism in plants (Treutter, 2006). Previous studies suggested that certain flavonoids play a key role in mediating the ROS (reactive oxygen species) homeostasis in plants as scavengers, and therefore, enhance plants' tolerance abiotic stress. In our study, several unigenes encoding key enzymes in the flavonoid biosynthesis pathway showed differential expression patterns in stress conditions, including phenylalanine ammonia lyase (comp91892_c0, comp82145_c0, and comp87406_c0), cinnamate 4-hydroxylase (comp76428_c0), 4-coumarate CoA ligase (comp60088_c0), chalcone synthase (comp91058_c2) and chalcone reductase (comp76128_c0). The differential expression patterns of the genes involved in flavonoid synthesis suggest that flavonoids may be active in the *O. ochrocephala* response to environmental stress. This result agrees with previous studies that have found these enzymes are up-regulated in stressed plants (Sallaud et al., 1995; Treutter, 2006; Gao et al., 2008).

Plant hormones are known to be involved in plant responses to various stresses (Shinozaki and Yamaguchi-Shinozaki, 2006). For instance, abscisic acid (ABA) is essential for various stress responses and the endogenous ABA level changes drastically in ABA dependent stress responses. Genes upstream of and in the ABA synthesis pathway can be up-regulated under water deficiency (Nambara and Marion-Poll, 2005). We found that some key enzymes in ABA biosynthesis such as 9-cisepoxycarotenoid dioxygenase (NCED) (comp86683_c0) and alcohol dehydrogenase (comp86382_c0) were both up-regulated and 8'-hydroxylase (comp78994_c0), the enzyme for the oxidative catabolism of ABA, was down-regulated under high salinity conditions. This confirms the role of ABA in regulating the transcriptional profile in *O. ochrocephala*. Bioactive gibberellins (GAs) control diverse aspects of growth and development, including seed germination, stem elongation, leaf expansion, and flower and seed development (Richards et al., 2001). We found that one unigene (comp80804_c0) encoding ent-kaurenoic acid oxidase, the key enzyme in GA biosynthesis, was down-regulated by exposure to cold, whilst a unigene encoding GA 2-oxidase (comp62108_c0) in the GA deactivation pathway was up-regulated. Both processes may result in a decreased endogenous level of bioactive GA, similar to other plant species (Achard et al., 2008). Cytokinin activity promotes cell expansion through cell wall loosening (Thomas et al., 1981). The control of cytokinin levels may be an important

TABLE 4 | Putative genes involved in stress tolerance in *Oxytropis ochrocephala*.

Gene ID	Annotation/Species	Sample	log ₂ (fold change)	q-value
HEAT SHOCK PROTEINS				
comp25961_c0	Heat shock protein 83 [<i>Medicago truncatula</i>]	Cold	-1.09	1.58E-12
comp68481_c0	Heat shock protein 70 [<i>Ziziphus jujuba</i>]		-1.80	1.99E-03
comp70341_c0	18.2 kDa class I heat shock protein [<i>Glycine max</i>]		-2.14	4.81E-04
comp74708_c0	Small heat shock protein C4 [<i>Glycine max</i>]		-4.09	1.61E-03
comp80742_c0	Heat shock protein 81-1 [<i>Theobroma cacao</i>]		-2.53	4.63E-03
comp81398_c0	22.7 kDa class IV heat shock protein [<i>Medicago truncatula</i>]		-2.69	2.08E-04
comp83116_c0	26.5 kDa heat shock protein [<i>Medicago truncatula</i>]		-3.88	3.82E-08
comp83893_c0	18.2 kDa class I heat shock protein [<i>Cicer arietinum</i>]		-2.22	1.01E-04
comp85605_c0	Heat shock protein [<i>Glycine max</i>]		-3.09	2.75E-08
comp87890_c0	Heat shock protein (hsp17.9) [<i>Pisum sativum</i>]		-2.47	2.99E-06
comp90848_c0	18.2 kDa class I heat shock protein [<i>Medicago truncatula</i>]		-2.67	1.39E-06
LEA PROTEINS				
comp83922_c0	Late embryogenesis abundant domain-containing protein [<i>Arabidopsis thaliana</i>]	Drought	-2.99	1.01E-03
comp76465_c0	Late embryogenesis abundant protein [<i>Medicago truncatula</i>]		2.97	4.61E-03
comp77463_c0	Y2K4 dehydrin variant G3, partial [<i>Medicago sativa</i>]		1.89	4.43E-03
comp77835_c0	Late embryogenesis abundant protein group 9 protein [<i>Arachis hypogaea</i>]	Salt	4.74	3.24E-10
comp80885_c0	Late embryogenesis abundant protein-1 [<i>Cicer arietinum</i>]		5.92	1.22E-25
comp80885_c1	Late embryogenesis abundant protein-2 [<i>Caragana jubata</i>]		3.66	1.76E-07
comp81820_c0	Soybean seed maturation polypeptides [<i>Glycine max</i>]		2.98	9.26E-08
comp84138_c0	Late embryogenesis abundant protein [<i>Medicago truncatula</i>]		6.34	1.10E-16
comp77463_c0	Y2K4 dehydrin variant G3 [<i>Eriobotrya japonica</i>]		4.26	9.51E-43
comp72368_c0	Putative late-embryogenesis protein-like [<i>Ammopiptanthus mongolicus</i>]	Cold	2.39	5.17E-06
comp80481_c0	Late embryogenesis abundant protein 1 [<i>Medicago truncatula</i>]	Cold	2.53	7.67E-18
comp70033_c0	Late embryogenesis protein-like protein [<i>Glycine max</i>]		2.21	1.27E-04
METABOLITES				
comp76128_c0	Chalcone reductase 1 [<i>Astragalus membranaceus</i>]	Drought	1.31	2.47E-13
comp76729_c0	Isoflavone reductase-like protein [<i>Medicago truncatula</i>]	Salt	1.05	9.63E-16
comp83244_c0	6a-hydroxymaackiain methyltransferase [<i>Pisum sativum</i>]		1.53	6.67E-04
comp76428_c0	Cinnamate 4-hydroxylase [<i>Astragalus membranaceus</i>]		1.12	8.33E-09
comp60088_c0	4-coumarate CoA ligase mRNA, complete cds [<i>Medicago truncatula</i>]		1.04	1.30E-05
comp67410_c0	Caffeoyl-CoA O-methyltransferase [<i>Caragana korshinskii</i>]		1.07	3.13E-08
comp76128_c0	Chalcone reductase 1 [<i>Astragalus membranaceus</i>]		1.42	1.61E-15
comp84899_c0	Isoliquiritigenin 2'-O-methyltransferase [<i>Medicago truncatula</i>]		1.40	6.34E-34
comp91892_c0	Phenylalanine ammonia-lyase [<i>Astragalus mongholicus</i>]		1.43	3.85E-13
comp91892_c0	Phenylalanine ammonia-lyase [<i>Astragalus mongholicus</i>]	Cold	1.82	4.12E-55
comp82145_c0	Phenylalanine ammonia-lyase [<i>Astragalus mongholicus</i>]		2.26	3.05E-05
comp87406_c0	Phenylalanine ammonia lyase [<i>Astragalus membranaceus</i>]		2.69	8.18E-32
comp81759_c0	Isoflavone 7-O-methyltransferase [<i>Glycyrrhiza echinata</i>]		1.71	3.84E-03
comp83244_c0	6a-hydroxymaackiain methyltransferase 1 [<i>Pisum sativum</i>]		1.58	3.53E-06
comp91058_c2	Chalcone synthase 1 [<i>Astragalus membranaceus</i>]		1.99	3.67E-16
comp86596_c0	Malonyl-CoA:isoflavone 7-O-glucoside-6"-O-malonyltransferase [<i>Glycine max</i>]		1.77	3.84E-03
comp82821_c0	Laccase-7-like [<i>Cicer arietinum</i>]		1.31	4.35E-04
comp91112_c1	Anthocyanin 3'-O-beta-glucosyltransferase [<i>Medicago truncatula</i>]		-1.75	1.84E-30
comp77150_c0	Beta-glucosidase 12-like isoform X3 [<i>Glycine max</i>]		2.69	3.68E-04
comp62108_c0	Gibberellin 2-beta-dioxygenase 1 [<i>Medicago truncatula</i>]		1.95	4.55E-04
comp80680_c0	Zeaxanthin epoxidase [<i>Cicer arietinum</i>]		-1.90	5.72E-04
comp76729_c0	Isoflavone reductase-like protein [<i>Medicago truncatula</i>]		1.05	9.63E-16

(Continued)

TABLE 4 | Continued

Gene ID	Annotation/Species	Sample	log ₂ (fold change)	q-value	
HORMONE RELATED					
comp81166_c0	1-aminocyclopropane-1-carboxylate oxidase [<i>Cicer arietinum</i>]	Ethylene	Drought	2.85	4.72E-06
comp73104_c0	Allene oxide cyclase [<i>Medicago truncatula</i>]	JA		3.83	1.12E-06
comp86382_c0	Alcohol dehydrogenase [<i>Medicago truncatula</i>]	ABA		1.86	4.54E-03
comp84731_c4	PR1-b gene for putative basic PR1 protein [<i>Cicer arietinum</i>]	SA		1.58	3.47E-12
comp86683_c0	9-cis-epoxycarotenoid dioxygenase [<i>Caragana korshinskii</i>]	ABA	Salt	3.21	3.08E-09
comp80804_c0	Ent-kaurenoic acid hydroxylase [<i>Medicago truncatula</i>]	GA	Cold	-1.12	6.90E-14
comp62108_c0	Gibberellin 2-oxidase [<i>Medicago truncatula</i>]			1.95	4.55E-04
comp80680_c0	Zeaxanthin epoxidase [<i>Cicer arietinum</i>]	ZR		-1.90	5.72E-04
comp87210_c0	Cytokinin-O-glucosyltransferase [<i>Medicago truncatula</i>]			1.80	1.46E-03
comp84200_c0	Auxin-induced protein [<i>Ricinus communis</i>]	IAA		2.35	5.24E-30
comp72368_c0	Indole-3-acetic acid-induced protein [<i>Glycine max</i>]			2.39	5.17E-06
comp78994_c0	Abscisic acid 8'-hydroxylase [<i>Medicago truncatula</i>]	ABA		-2.18	1.15E-04
comp79621_c0	ABA-inducible protein [<i>Caragana jubata</i>]			1.25	9.63E-06
comp82821_c0	Laccase-7-like [<i>Cicer arietinum</i>]	Cytokinin		1.31	4.35E-04
TRANSCRIPTION FACTORS					
comp87209_c3	Homeodomain leucine zipper protein [<i>Cicer arietinum</i>]	bZI	Salt	7.15	2.04E-05
comp67029_c1	bZIP transcription factor [<i>Caragana korshinskii</i>]			1.39	1.33E-05
comp72293_c0	Homeobox-leucine zipper protein [<i>Medicago sativa</i>]			2.42	9.97E-05
comp81505_c0	Transcription factor bHLH25-like [<i>Glycine max</i>]	bHLH		2.56	2.06E-04
comp85725_c0	Transcription factor bHLH25-like [<i>Glycine max</i>]			2.08	1.15E-03
comp71824_c0	NAC transcription factor-like protein [<i>Medicago truncatula</i>]	NAC		2.29	9.09E-13
comp84372_c0	R2R3-MYB transcription factor [<i>Medicago truncatula</i>]	MYB		2.32	5.58E-05
comp76034_c0	WRKY transcription factor 33 [<i>Cicer arietinum</i>]	WRK	cold	2.07	1.39E-04
comp82399_c1	WRKY transcription factor 40-like isoform 1 [<i>Glycine max</i>]	WRKY		1.83	1.02E-03
comp87093_c1	WRKY transcription factor 41 [<i>Cicer arietinum</i>]	WRK	Cold	1.78	1.96E-03
comp69006_c0	ER transcription factor 1B-like [<i>Glycine max</i>]	ER-TF		2.97	5.50E-09
comp70325_c0	ER transcription factor ERF061 [<i>Medicago truncatula</i>]	ER-TF		2.38	3.23E-03
comp77817_c0	ER transcription factor 11 [<i>Medicago sativa</i>]	ER-TF		1.13	1.58E-06
comp83387_c0	ER transcription factor [<i>Cicer arietinum</i>]	ER-TF		1.36	4.69E-05
comp86762_c0	ER transcription factor 7-like [<i>Cicer arietinum</i>]	ER-TF		1.40	9.05E-07
comp90492_c1	ER transcription factor ERF023-like [<i>Cicer arietinum</i>]	ER-TF		-2.01	3.78E-03
comp70325_c0	ER transcription factor ERF061 [<i>Medicago truncatula</i>]	ER-TF		2.37	1.47E-05
comp89496_c2	ER transcription factor ERF098-like [<i>Glycine max</i>]	ER-TF		2.72	2.96E-05
comp90492_c1	ER transcription factor ERF023-like [<i>Cicer arietinum</i>]	ER-TF		-2.01	3.57E-04
comp71742_c1	AP2/ERF and B3 domain-containing transcription repressor [<i>Cicer arietinum</i>]	ER-TF		2.18	5.51E-04
comp56812_c1	ER transcription factor 1B-like [<i>Cicer arietinum</i>]	ER-TF		3.75	9.79E-04
comp23350_c0	bZIP transcription factor [<i>Glycine m</i>]	bZIP		1.20	3.99E-09
comp85721_c0	bZIP transcription factor bZIP105 [<i>Glycine max</i>]			1.39	1.51E-07
comp87209_c3	Homeobox-leucine zipper protein ATHB-40 [<i>Cicer arietinum</i>]			7.22	5.28E-06
comp69091_c0	Glycine max bZIP transcription factor bZIP35 [<i>Glycine max</i>]			2.10	2.10E-04
comp85114_c0	Transcription factor bHLH25 [<i>Medicago truncatula</i>]	bHLH		-1.62	4.87E-03
comp83156_c0	Transcription factors MYC4 [<i>Medicago truncatula</i>]			3.00	1.67E-06
comp91054_c0	Transcription factor bHLH96 [<i>Cicer arietinum</i>]			-1.87	2.92E-03
comp90383_c1	NAC domain protein NAC2 [<i>Cicer arietinum</i>]	NAC		1.71	2.92E-03
comp83816_c0	NAC domain protein NAC1 [<i>Glycine max</i>]			2.46	5.91E-05
comp71824_c0	NAC domain protein [<i>Medicago truncatula</i>]			2.21	3.36E-05
comp82153_c0	MYB transcription factor [<i>Medicago truncatula</i>]	MYB		4.52	4.07E-18
comp80312_c0	MYB family transcription factor-like protein [<i>Medicago truncatula</i>]			2.19	4.83E-03

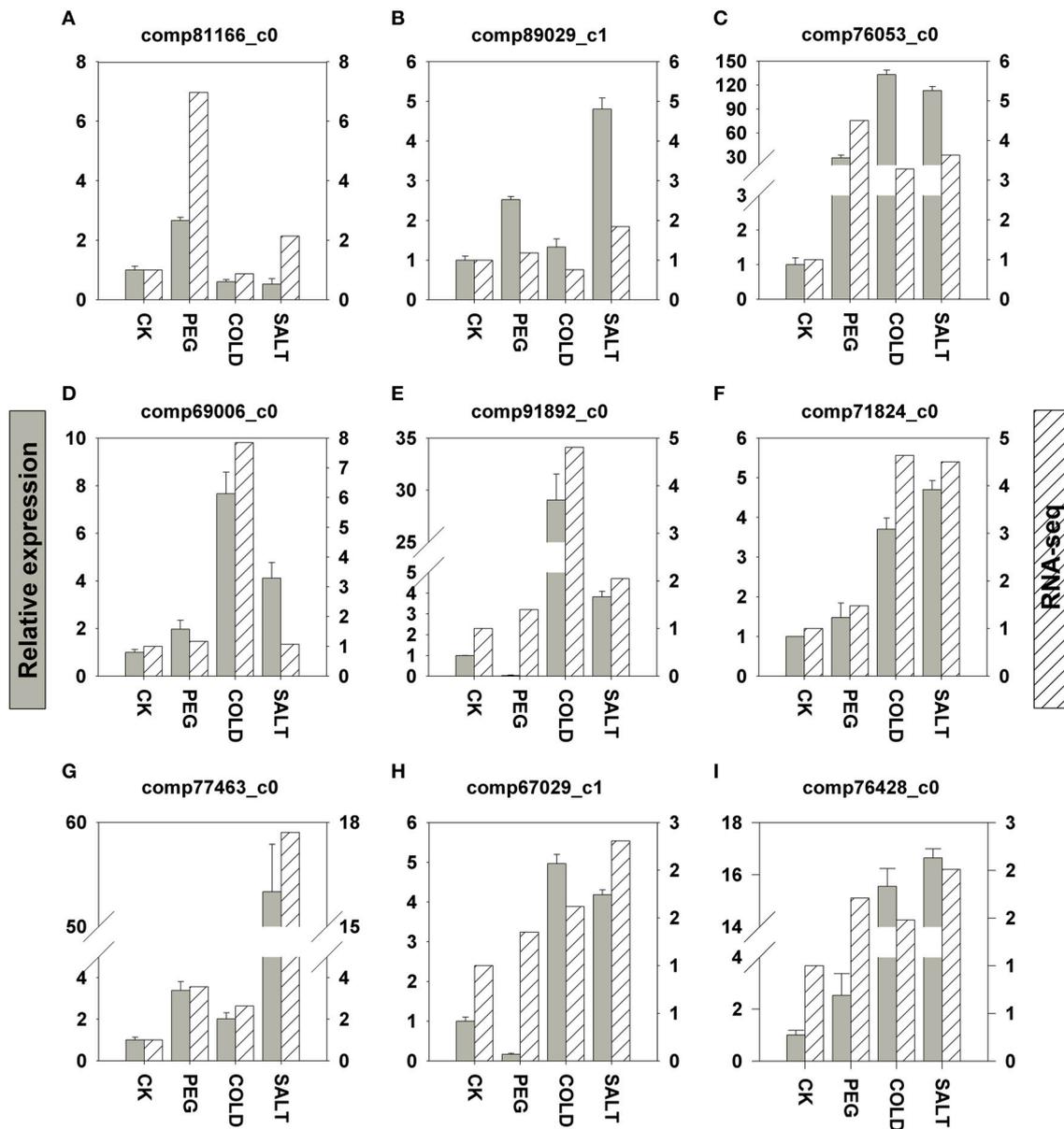


FIGURE 5 | qRT-PCR validation of nine DEGs related to drought (A–C), cold (D–F), and saline (G–I). Y axes in the graphs show normalized fold changes in qRT-PCR (left) and FPKM in RNA-seq (right). qRT-PCR expression levels were normalized against *ACTIN* and *HISTONE*. Error bars represent one standard error of the mean ($n = 3$). CK: control sample. **(A)** comp81166_c0: 1-aminocyclopropane-1-carboxylate oxidase. **(B)** comp89029_c1: galactinol synthase. **(C)** comp76053_c0: peroxidase. **(D)** comp69006_c0: ethylene-responsive transcription factor. **(E)** comp91892_c0: phenylalanine ammonia-lyase. **(F)** comp71824_c0: NAC transcription factor. **(G)** comp77463_c0: Y2K4 dehydrin variant G3. **(H)** comp67029_c1: bZIP transcription factor. **(I)** comp76428_c0: cinnamate 4-hydroxylase.

mechanism through which cell expansion is controlled in *O. ochrocephala*. Our data suggested that cytokinin degradation in the apoplast may be affected by exposure to cold. A cytokinin-O-glucosyltransferase gene (comp87210_c0) was specifically up-regulated in the cold and saline treatments, along with a laccase gene (comp82821_c0). The former enzyme reversibly converts cytokinins to ortho-glucosyl derivatives, and temporarily reduces cytokinin activity (Martin et al., 1999). The latter enzyme has been proposed to function together in the apoplast to degrade cytokinins (Galuszka et al., 2005). Taken together, changes in

plant hormone levels inevitably influence the production of secondary metabolites, and it is conceivable that cross-talk among these hormones as well as other signal molecules function together to endow *O. ochrocephala* plants with resistance to stress.

Transcription factors also respond to stress signals at the early stage during stress. First, the WRKY transcription factors, which may be involved in the initial steps of the defense-response signaling pathway, often act as repressors as well as activators for transcription, and members of this family function in abiotic

stresses such as drought and cold (Rushton et al., 2010; Tripathi et al., 2013). In *O. ochrocephala*, unigenes (comp76034_c0 and comp87093_c1) annotated as WRKY transcription factors are up-regulated during cold exposure. Second, we found a total of 12 ethylene-responsive (ER) transcription factor genes belonging to the AP2/EREBP family to be a response to cold in *O. ochrocephala*. Ethylene-responsive (ER) transcription factors, which are related to the C-repeat binding transcription factors (CBFs), have been observed to be expressed in plants in response to abiotic stresses (Lee et al., 2007). For example, CBFs were identified as a key responder to low temperature stress (Zhou et al., 2011). In the CBF dependent pathway, the CRT/DRE cis-element in *COR* (cold response genes) is recognized by the CBFs, and as a result, the downstream genes are activated. This results in chilling and freezing tolerance in plants (Gilmour et al., 1998). In *O. ochrocephala*, we identified two genes (comp22730_c0 and comp85560_c1) homologous to *COR*. In Arabidopsis, *CBF1* encodes an AP2 domain containing a transcriptional activator that binds to the C-repeat/DRE in response to low temperature and water deficiency (Gilmour et al., 2000; Haake et al., 2002). However, the CBF genes and their targets *COR* genes in *O. ochrocephala* are induced by cold stress only, not by drought. This suggests that there may be different mechanisms of dehydration responses in *O. ochrocephala* and Arabidopsis. Third, in this study, we found that three basic leucine zipper (bZIP) family genes encoding AREB (ABA-responsive element binding) and ABA-inducible protein (comp79621_c0) are up-regulated by cold and saline stress. These genes are known to be induced by ABA and are involved in drought stress response in an ABA-dependent manner in other plant species (Kim, 2005). Other important transcript factors such as bHLH, NAC, MYB exhibit similar changes in expression in response to saline and cold stress.

qPCR Validation of Nine Genes Associated with Drought, Cold and Salinity Stress Response

In gene expression analysis, qRT-PCR is widely used to verify differential expression patterns measured by microarray and transcriptomic sequencing. In this study we tested the expression of five genes discussed above, including genes encoding phenylalanine ammonia-lyase and cinnamate 4-hydroxylase (Figures 5E,I) in flavonoids biosynthesis pathway, and transcription factors (TF) such as ethylene-responsive TF, NAC TF, and bZIP TF (Figures 5D,F,H). In addition, we further determined four genes that are well characterized in plant stress responses. 1-aminocyclopropane-1-carboxylate (ACC) oxidase (Figure 5A) is the last enzyme in ethylene production in plant related to ethylene regulation (Morgan and Drew, 1997; Kim et al., 1998); galactinol synthase (Figure 5B) determines the synthesis RFOs (raffinose family oligosaccharides), a group of compatible solutes in plant osmotic adjustment in drought and salinity stress (Sengupta et al., 2015); Y2K4 dehydrin variant G3 (Figure 5G) helps plants gain freezing tolerance as shown in alfalfa (Rémus-Borel et al., 2010); peroxidase (POD, Figure 5C and Supplementary Figure 1E in Data Sheet 1). Functions as the first tier of defense against reactive oxygen species (ROS)

generated under drought and other stress conditions (Veljovic-Jovanovic et al., 2006). The nine genes are all up-regulated in drought, saline and cold stress responses in *O. ochrocephala*. The consistency between our qRT-PCR and the DEGs results indicate that candidate genes in stress response in *O. ochrocephala* revealed by RPKM can be a useful resource for further validation.

CONCLUSION

In this study, we performed the first transcriptome sequencing analysis for a typical locoweed species, *O. ochrocephala*. This will significantly enrich available genetic information for locoweed research. In addition, by analyzing differential expression patterns of the genes under the three abiotic stress conditions, we identified multiple genes which are likely to be involved in drought, saline and cold acclimation in *O. ochrocephala*. Further, research is needed for functional analyses of these candidate genes to understand the molecular mechanism of adaptation in locoweeds to variant environments.

AUTHOR CONTRIBUTIONS

WH and HZ designed the experiment, prepared samples for RNA-seq, analyzed the data and prepared the manuscript. YF helped in data interpretation and manuscript preparation. BG and LG analyzed the data and performed qRT-PCR analysis. XZ and LG collected plant materials. YW conceived the project, designed the experiments, supervised the analysis, and critically revised the manuscript. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

We thank the reviewers and Dr. Derek Dunn for their constructive suggestions. This work was supported by the National Natural Science Foundation of China [31402133] to WH, Special Fund for Agro-scientific Research in the Public Interest of China [201203062] to YW, Science Fund of the Department of Education in Shaanxi Province [14JK1725] to WH, and research fund from Northwest University [PR12082 and NH13030] to WH, and Key Laboratory Program of the Department of Education in Shaanxi Province [14JS099] to BG.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpls.2015.01086>

Data Sheet 1 | Supplementary tables and figures.

Data Sheet 2 | Assembled sequences of all unigenes.

Data Sheet 3 | Nr (non-redundant) annotation of unigenes.

Data Sheet 4 | DEG (differential expressed genes) with annotation against Nr.

Data Sheet 5 | GO (gene ontology) enrichment of DEGs.

Data Sheet 6 | Hcluster heatmap with annotation and unigene IDs.

REFERENCES

- Achard, P., Gong, F., Cheminant, S., Alioua, M., Hedden, P., and Genschik, P. (2008). The cold-inducible CBF1 factor-dependent signaling pathway modulates the accumulation of the growth-repressing DELLA proteins via its effect on gibberellin metabolism. *Plant Cell* 20, 2117–2129. doi: 10.1105/tpc.108.058941
- Archambault, A., and Strömviik, M. V. (2011). PR-10, defensin and cold dehydrin genes are among those over expressed in *Oxytropis* (Fabaceae) species adapted to the arctic. *Funct. Integr. Genomics* 11, 497–505. doi: 10.1007/s10142-011-0223-6
- Blumenthal, D. M. (2006). Interactions between resource availability and enemy release in plant invasion. *Ecol. Lett.* 9, 887–895. doi: 10.1111/j.1461-0248.2006.00934.x
- NDong, C., Danyluk, J., Wilson, K. E., Pocock, T., Huner, N. P., and Sarhan, F. (2002). Cold-regulated cereal chloroplast late embryogenesis abundant-like proteins. Molecular characterization and functional analyses. *Plant Physiol.* 129, 1368–1381. doi: 10.1104/pp.001925
- Chen, J., Tian, Q., Pang, T., Jiang, L., Wu, R., Xia, X., et al. (2014). Deep-sequencing transcriptome analysis of low temperature perception in a desert tree, *Populus euphratica*. *BMC Genomics* 15:326. doi: 10.1186/1471-2164-15-326
- Chung, M., Gelembiuk, G., and Givnish, T. J. (2004). Population genetics and phylogeography of endangered *Oxytropis campestris* var. *chartacea* and relatives: arctic-alpine disjuncts in eastern North America. *Mol. Ecol.* 13, 3657–3673. doi: 10.1111/j.1365-294X.2004.02360.x
- Cook, D., Grum, D. S., Gardner, D. R., Welch, K. D., and Pfister, J. A. (2013). Influence of endophyte genotype on swainsonine concentrations in *Oxytropis sericea*. *Toxicon* 61, 105–111. doi: 10.1016/j.toxicon.2012.10.018
- Creamer, R., and Baucom, D. (2013). Fungal endophytes of locoweeds: a commensal relationship? *J. Plant Physiol. Pathol.* 2, 1–4. doi: 10.4172/jppp.1000104
- Dugas, D. V., Monaco, M. K., Olson, A., Klein, R. R., Kumari, S., Ware, D., et al. (2011). Functional annotation of the transcriptome of *Sorghum bicolor* in response to osmotic stress and abscisic acid. *BMC Genomics* 12:514. doi: 10.1186/1471-2164-12-514
- Farrell, J. D., Byrne, S., Paina, C., and Asp, T. (2014). De novo assembly of the perennial ryegrass transcriptome using an RNA-Seq strategy. *PLoS ONE* 9:e103567. doi: 10.1371/journal.pone.0103567
- Fine, P. V., Miller, Z. J., Mesones, L., Irazuzta, S., Appel, H. M., Stevens, M. H. H., et al. (2006). The growth-defense trade-off and habitat specialization by plants in Amazonian forests. *Ecology* 87, S150–S162. doi: 10.1890/0012-9658(2006)87[150:TGTAHS]2.0.CO;2
- Finn, R. D., Clements, J., and Eddy, S. R. (2011). HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* 39, w29–w37. doi: 10.1093/nar/gkr367
- Fu, N., Wang, Q., and Shen, H. L. (2013). De novo assembly, gene annotation and marker development using Illumina paired-end transcriptome sequences in celery (*Apium graveolens* L.). *PLoS ONE* 8:e57686. doi: 10.1371/journal.pone.0057686
- Galuszka, P., Frébortová, J., Luhová, L., Bilyeu, K. D., English, J. T., and Frébort, I. (2005). Tissue localization of cytokinin dehydrogenase in maize: possible involvement of quinone species generated from plant phenolics by other enzymatic systems in the catalytic reaction. *Plant Cell Physiol.* 46, 716–728. doi: 10.1093/pcp/pci074
- Gao, S., Ouyang, C., Wang, S., Xu, Y., Tang, L., and Chen, F. (2008). Effects of salt stress on growth, antioxidant enzyme and phenylalanine ammonia-lyase activities in *Jatropha curcas* L. seedlings. *Plant Soil Environ.* 54, 374–381.
- Gardner, D. R., Molyneux, R. J., and Ralphs, M. H. (2001). Analysis of swainsonine: extraction methods, detection, and measurement in populations of locoweeds (*Oxytropis* spp.). *J. Agric. Food Chem.* 49, 4573–4580. doi: 10.1021/jf010596p
- Garg, R., Patel, R. K., Tyagi, A. K., and Jain, M. (2011). De novo assembly of chickpea transcriptome using short reads for gene discovery and marker identification. *DNA Res.* 18, 53–63. doi: 10.1093/dnares/dsq028
- Gasic, K., Hernandez, A., and Korban, S. S. (2004). RNA extraction from different apple tissues rich in polyphenols and polysaccharides for cDNA library construction. *Plant Mol. Biol. Report.* 22, 437–438. doi: 10.1007/BF02772687
- Ge, X., Chen, H., Wang, H., Shi, A., and Liu, K. (2014). De Novo assembly and annotation of *Salvia splendens* transcriptome using the Illumina platform. *PLoS ONE* 9:e87693. doi: 10.1371/journal.pone.0087693
- Gilmour, S. J., Sebolt, A. M., Salazar, M. P., Everard, J. D., and Thomashow, M. F. (2000). Overexpression of the Arabidopsis CBF3 transcriptional activator mimics multiple biochemical changes associated with cold acclimation. *Plant Physiol.* 124, 1854–1865. doi: 10.1104/pp.124.4.1854
- Gilmour, S. J., Zarka, D. G., Stockinger, E. J., Salazar, M. P., Houghton, J. M., and Thomashow, M. F. (1998). Low temperature regulation of the Arabidopsis CBF family of AP2 transcriptional activators as an early step in cold-induced COR gene expression. *Plant J.* 16, 433–442. doi: 10.1046/j.1365-313x.1998.00310.x
- Götz, S., García-Gómez, J. M., Terol, J., Williams, T. D., Nagaraj, S. H., Nueda, M. J., et al. (2008). High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res.* 36, 3420–3435. doi: 10.1093/nar/gkn176
- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29, 644–652. doi: 10.1038/nbt.1883
- Gross, S. M., Martin, J. A., Simpson, J., Abraham-Juarez, M. J., Wang, Z., and Visel, A. (2013). De novo transcriptome assembly of drought tolerant CAM plants, *Agave deserti* and *Agave tequilana*. *BMC Genomics* 14:563. doi: 10.1186/1471-2164-14-563
- Haake, V., Cook, D., Riechmann, J., Pineda, O., Thomashow, M. F., and Zhang, J. Z. (2002). Transcription factor CBF4 is a regulator of drought adaptation in Arabidopsis. *Plant Physiol.* 130, 639–648. doi: 10.1104/pp.006478
- Hierro, J. L., and Callaway, R. M. (2003). Allelopathy and exotic plant invasion. *Plant Soil* 256, 29–39. doi: 10.1023/A:1026208327014
- Hu, W., Hu, G., and Han, B. (2009). Genome-wide survey and expression profiling of heat shock proteins and heat shock factors revealed overlapped and stress specific response under abiotic stresses in rice. *Plant Sci.* 176, 583–590. doi: 10.1016/j.plantsci.2009.01.016
- Hundertmark, M., and Hinch, D. K. (2008). LEA (late embryogenesis abundant) proteins and their encoding genes in *Arabidopsis thaliana*. *BMC Genomics* 9:118. doi: 10.1186/1471-2164-9-118
- Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., et al. (2008). KEGG for linking genomes to life and the environment. *Nucleic Acids Res.* 36, D480–D484. doi: 10.1093/nar/gkm882
- Kim, S. Y. (2005). The role of ABF family bZIP class transcription factors in stress response. *Physiol. Plant.* 126, 519–527. doi: 10.1111/j.1399-3054.2005.00601.x
- Kim, Y. S., Choi, D., Lee, M. M., Lee, S. H., and Kim, W. T. (1998). Biotic and abiotic stress-related expression of 1-aminocyclopropane-1-carboxylate oxidase gene family in *Nicotiana glutinosa* L. *Plant Cell Physiol.* 39, 565–573.
- Kotak, S., Port, M., Ganguli, A., Bicker, F., and Koskull-Döring, V. (2004). Characterization of C-terminal domains of Arabidopsis heat stress transcription factors (Hsfs) and identification of a new signature combination of plant class A Hsfs with AHA and NES motifs essential for activator function and intracellular localization. *Plant J.* 39, 98–112. doi: 10.1111/j.1365-313X.2004.02111.x
- Kreps, J. A., Wu, Y., Chang, H. S., Zhu, T., Wang, X., and Harper, J. F. (2002). Transcriptome changes for Arabidopsis in response to salt, osmotic, and cold stress. *Plant Physiol.* 130, 2129–2141. doi: 10.1104/pp.008532
- Lee, H. E., Shin, D., Park, S. R., Han, S. E., Jeong, M. J., Kwon, T. R., et al. (2007). Ethylene responsive element binding protein 1 (StEREBP1) from *Solanum tuberosum* increases tolerance to abiotic stress in transgenic potato plants. *Biochem. Biophys. Res. Commun.* 353, 863–868. doi: 10.1016/j.bbrc.2006.12.095
- Li, B., and Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12:323. doi: 10.1186/1471-2105-12-323
- Liu, G., Li, W., Zheng, P., Xu, T., Chen, L., Liu, D., et al. (2012). Transcriptomic analysis of 'Suli' pear (*Pyrus pyrifolia* white pear group) buds during the dormancy by RNA-Seq. *BMC Genomics* 13:700. doi: 10.1186/1471-2164-13-700
- Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262
- Mao, X., Cai, T., Olyarchuk, J. G., and Wei, L. (2005). Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. *Bioinformatics* 21, 3787–3793. doi: 10.1093/bioinformatics/bti430

- Martin, R. C., Mok, M. C., and Mok, D. W. (1999). Isolation of a cytokinin gene, *ZOG1*, encoding zeatin O-glucosyltransferase from *Phaseolus lunatus*. *Proc. Natl. Acad. Sci. U.S.A.* 96, 284–289. doi: 10.1073/pnas.96.1.284
- Morgan, P. W., and Drew, M. C. (1997). Ethylene and plant responses to stress. *Physiol. Plant.* 100, 620–630. doi: 10.1111/j.1399-3054.1997.tb03068.x
- Nambara, E., and Marion-Poll, A. (2005). Abscisic acid biosynthesis and catabolism. *Annu. Rev. Plant Biol.* 56, 165–185. doi: 10.1146/annurev.arplant.56.032604.144046
- Nishizawa, A., Yabuta, Y., Yoshida, E., Maruta, T., Yoshimura, K., and Shigeoka, S. (2006). Arabidopsis heat shock transcription factor A2 as a key regulator in response to several types of environmental stress. *Plant J.* 48, 535–547. doi: 10.1111/j.1365-3113X.2006.02889.x
- Oldrup, E., McLain-Romero, J., Padilla, A., Moya, A., Gardner, D., and Creamer, R. (2010). Localization of endophytic *Undifilum* fungi in locoweed seed and influence of environmental parameters on a locoweed *in vitro* culture system. *Botany* 88, 512–521. doi: 10.1139/B10-026
- Peleg, Z., and Blumwald, E. (2011). Hormone balance and abiotic stress tolerance in crop plants. *Curr. Opin. Plant Biol.* 14, 290–295. doi: 10.1016/j.pbi.2011.02.001
- Pryor, B., Creamer, R., Shoemaker, R., McLain-Romero, J., and Hambleton, S. (2009). *Undifilum*, a new genus for endophytic *Embellisia oxytropis* and parasitic *Helminthosporium bornmuelleri* on legumes. *Botany* 87, 178–194. doi: 10.1139/B08-130
- Ralphs, M., Creamer, R., Baucom, D., Gardner, D., Welsh, S., Graham, J., et al. (2008). Relationship between the endophyte *Embellisia* spp. and the toxic alkaloid swainsonine in major locoweed species (*Astragalus* and *Oxytropis*). *J. Chem. Ecol.* 34, 32–38. doi: 10.1007/s10886-007-9399-6
- Ralphs, M. H., Graham, J. D., and James, L. F. (2002). A close look at locoweed poisoning on shortgrass prairies. *Rangelands* 24, 30–34.
- Rémus-Borel, W., Castonguay, Y., Cloutier, J., Michaud, R., Bertrand, A., Desgagnés, R., et al. (2010). Dehydrin variants associated with superior freezing tolerance in alfalfa (*Medicago sativa* L.). *Theor. Appl. Genet.* 120, 1163–1174. doi: 10.1007/s00122-009-1243-7
- Richards, D. E., King, K. E., Ait-ali, T., and Harberd, N. P. (2001). How gibberellin regulates plant growth and development: a molecular genetic analysis of gibberellin signaling. *Annu. Rev. Plant Biol.* 52, 67–88. doi: 10.1146/annurev.arplant.52.1.67
- Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi: 10.1093/bioinformatics/btp616
- Rushton, P. J., Somssich, I. E., Ringler, P., and Shen, Q. J. (2010). WRKY transcription factors. *Trends Plant Sci.* 15, 247–258. doi: 10.1016/j.tplants.2010.02.006
- Sallaud, C., El-Turk, J., Breda, C., Buffard, D., De Kozak, I., Esnault, R., et al. (1995). Differential expression of cDNA coding for chalcone reductase, a key enzyme of the 5-deoxyflavonoid pathway, under various stress conditions in *Medicago sativa*. *Plant Sci.* 109, 179–190. doi: 10.1016/0168-9452(95)04179-X
- Sengupta, S., Mukherjee, S., Basak, P., and Majumder, A. L. (2015). Significance of galactinol and raffinose family oligosaccharide synthesis in plants. *Front. Plant Sci.* 6:656. doi: 10.3389/fpls.2015.00656
- Shinozaki, K., and Yamaguchi-Shinozaki, K. (2006). Gene networks involved in drought stress response and tolerance. *J. Exp. Bot.* 58, 221–227. doi: 10.1093/jxb/erl164
- Thomas, J., Ross, C. W., Chastain, C. J., Koomanoff, N., Hendrix, J. E., and Van Volkenburgh, E. (1981). Cytokinin-induced wall extensibility in excised cotyledons of radish and cucumber. *Plant Physiol.* 68, 107–110. doi: 10.1104/pp.68.1.107
- Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., et al. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28, 511–515. doi: 10.1038/nbt.1621
- Treutter, D. (2006). Significance of flavonoids in plant resistance: a review. *Environ. Chem. Lett.* 4, 147–157. doi: 10.1007/s10311-006-0068-8
- Tripathi, P., Rabara, R. C., and Rushton, P. J. (2013). A systems biology perspective on the role of WRKY transcription factors in drought responses in plants. *Planta* 239, 255–266. doi: 10.1007/s00425-013-1985-y
- Tulsiani, D., Broquist, H., James, L., and Touster, O. (1988). Production of hybrid glycoproteins and accumulation of oligosaccharides in the brain of sheep and pigs administered swainsonine or locoweed. *Arch. Biochem. Biophys.* 264, 607–617. doi: 10.1016/0003-9861(88)90327-X
- Veljovic-Jovanovic, S., Kukavica, B., Stevanovic, B., and Navari-Izzo, F. (2006). Senescence and drought-related changes in peroxidase and superoxide dismutase isoforms in leaves of *Ramonda serbica*. *J. Exp. Bot.* 57, 1759–1768. doi: 10.1093/jxb/erl007
- Vincent, D., Lapierre, C., Pollet, B., Cornic, G., Negroni, L., and Zivy, M. (2005). Water deficits affect caffeate O-methyltransferase, lignification, and related enzymes in maize leaves. A proteomic investigation. *Plant Physiol.* 137, 949–960. doi: 10.1104/pp.104.050815
- Wang, W., Vinocur, B., Shoseyov, O., and Altman, A. (2004). Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. *Trends Plant Sci.* 9, 244–252. doi: 10.1016/j.tplants.2004.03.006
- Wei, W., Qi, X., Wang, L., Zhang, Y., Hua, W., Li, D., et al. (2011). Characterization of the sesame (*Sesamum indicum* L.) global transcriptome using Illumina paired-end sequencing and development of EST-SSR markers. *BMC Genomics* 12:451. doi: 10.1186/1471-2164-12-451
- Xie, F., Burklew, C. E., Yang, Y., Liu, M., Xiao, P., Zhang, B., et al. (2012). *De novo* sequencing and a comprehensive analysis of purple sweet potato (*Ipomoea batatas* L.) transcriptome. *Planta* 236, 101–113. doi: 10.1007/s00425-012-1591-4
- Yamaguchi-Shinozaki, K., and Shinozaki, K. (2006). Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses. *Annu. Rev. Plant Biol.* 57, 781–803. doi: 10.1146/annurev.arplant.57.032905.105444
- Young, M. D., Wakefield, M. J., Smyth, G. K., and Oshlack, A. (2010). Method Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* 11:R14. doi: 10.1186/gb-2010-11-2-r14
- Zhao, M., Gao, X., Wang, J., He, X., and Han, B. (2013). A review of the most economically important poisonous plants to the livestock industry on temperate grasslands of China. *J. Appl. Toxicol.* 33, 9–17. doi: 10.1002/jat.2789
- Zhou, M., Shen, C., Wu, L., Tang, K., and Lin, J. (2011). CBF-dependent signaling pathway: a key responder to low temperature stress in plants. *Crit. Rev. Biotechnol.* 31, 186–192. doi: 10.3109/07388551.2010.505910
- Zhou, Y., Gao, F., Liu, R., Feng, J., and Li, H. (2012). *De novo* sequencing and analysis of root transcriptome using 454 pyrosequencing to discover putative genes associated with drought tolerance in *Ammopiptanthus mongolicus*. *BMC Genomics* 13:266. doi: 10.1186/1471-2164-13-266

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 He, Zhuang, Fu, Guo, Guo, Guo, Zhang and Wei. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.