Selection of appropriate reference genes for quantitative real-time PCR in *Oxytropis ochrocephala* Bunge using transcriptome datasets under abiotic stress treatments

Huihui Zhuang, Yanping Fu, Wei He, Lin Wang and Yahui Wei *

Key Laboratory of Resource Biology and Biotechnology in Western China, Department of Life Science, Northwest University, Xi’an, China

**Background:** *Oxytropis ochrocephala* Bunge, an indigenous locoweed species in China, poses great threats to livestock on grasslands. There is a need for further genetic study in the plants per se, for understanding the basis of its acclimation mechanism in various unfavorable environmental conditions and to implement effective control measures. Quantitative real-time reverse transcription-polymerase chain reaction (qRT-PCR) is the most commonly used method for gene expression analysis. To facilitate gene expression studies and obtain more accurate qRT-PCR data, normalization relative to stable reference genes is required. The aim of this study was to select the most stable reference genes for transcriptional analysis in *O. ochrocephala*.

**Results:** We selected 12 candidate reference genes, 18S ribosomal RNA (18S RNA), actin2/7 (ACT7), β-actin (ACTB), actin101 (ACT101), actin11 (ACT11), β-tubulin (TUB), α-tubulin (TUA), glyceraldehyde-3-phosphate dehydrogenase-1 (GAPDH1), GAPDH2, metallothionein-like protein (MET), fructose-bisphosphate aldolase (FBA) and histone H3 (HIS), from the transcriptome datasets of *O. ochrocephala* and determined the suitability by analyzing their expression levels when exposed to a range of abiotic stress conditions. By employing software packages including geNorm, NormFinder and BestKeeper, HIS, ACT7, and ACT101 were assessed as the most suitable set for normalization in all samples. When normalized with the most stable reference genes, the expression patterns of the three target genes were in accordance with those in the transcriptome data, indicating that the reference genes selected in this study are suitable.

**Conclusions:** The study provided appropriate reference genes for accurate normalization in qRT-PCR analysis in *O. ochrocephala* and emphasized the importance of validating reference genes for gene expression analysis under specific experimental condition. The usage of inappropriate reference gene would cause misinterpretation.

**Keywords:** *Oxytropis ochrocephala* Bunge, reference genes, quantitative real-time PCR, transcriptome, abiotic stress
Introduction

*Oxytropis ochrocephala* Bunge (Fabaceae), a perennial grass species, grows aggressively on the grassland of northwestern China. It is the most widespread locoweed (the collective name for toxic *Oxytropis* and *Astragalus* species) in China. Locoweeds contain an indolizidine alkaloid swainsonine, an α-mannosidase inhibitor which causes over-accumulation of mannose rich oligosaccharide in the lysosomes and impairs the neural system of livestock, which has long been recognized as the principle for intoxication in animals (Ralphs et al., 2008). Although it has been proved that swainsonine is produced by its symbiotic fungal endophytes, the plants *per se* also have a great impact on the production of swainsonine (Pryor et al., 2009). Cook et al. (2013) observed that plants offer the endophytes nutrients, hormones and some other signals in influencing its capacity to produce the toxin. In addition, *O. ochrocephala* grows in high-altitude areas with deficient water, high soil saline content, and extreme low environmental temperature, indicating that the plant has the ability to cope with unfavorable environmental stresses. In order to better understand the role of the plant in production of swainsonine, the research on its stress response mechanism to harsh environment conditions is needed. Currently, genetic information of locoweeds for molecular biology research is limited in the public databases, which makes it more difficult for further in-depth research. A recently EST dataset of a suppressive subtraction cDNA library enriched in genes from two temperate *Oxytropis* species was made available in NCBI, but it only has 1245 ESTs and thus there is a need for further exploring the genetic information in *Oxytropis* (Chung et al., 2004; Archambault and Strömvik, 2011).

RNA sequencing (RNA-Seq), with the rapid development of the next-generation sequencing technology, has been applied prevalently on analyzing the transcriptomes of various species for a range of purposes (Wang et al., 2009; Li et al., 2015a; Stone and Storchova, 2015). The main outcome of RNA-Seq data is the identification of differentially expressed genes, while it was also used to search for reference genes. Any gene with a minimal expression level variation in every analyzed sample is considered as candidate reference gene. Meanwhile, qRT-PCR is commonly used to determine expression levels of genes and to validate transcriptomic data (Andersen et al., 2004; Caldana et al., 2007). For accurate qRT-PCR evaluation, it is necessary to select suitable reference genes as internal control under different experimental conditions because the starting material, RNA extraction, RT-PCR efficiency, and qRT-PCR efficiency can vary among experiments. In addition, gene expression can be highly tissue-specific and often varies based on the physiological status of the organism or experimental treatments. Thus, it is essential for an optimal reference gene or combination of genes for the specific experiment to normalize gene expression data.

The growing acquisition of plant genomes and transcriptomes provide a high-throughput approach to identify a set of reference genes. In the study of Czechowski et al. (2005), based on the Arabidopsis genome, hundreds of genes with stable expression levels were identified. Yang et al. (2014b) used a similar approach to mine 40 genes as candidate reference genes from the transcriptome datasets of *Brassica napus* and 14 reference genes were selected for further analysis with qRT-PCR in different tissues and under different experimental treatments. Reference genes are usually cellular maintenance genes, which regulate basic cellular functions and components, such as *Actin* (Nakayama et al., 2014), *Tubulin* (Zhang et al., 2013), GAPDH (Ling et al., 2014), 18S RNA (Yang et al., 2014a), *Ubiquitin* (UBQ) (Galeano et al., 2014), *Histone* (Wang et al., 2014; Yan et al., 2014). Recent studies have indicated that there is no universal reference gene for all experimental conditions and systematic expression stability analysis of reference genes under specific experimental conditions is necessary (Jian et al., 2008).

A few statistical algorithms have been developed to identify which reference gene(s) is (are) suitable for normalization under a given experiment condition for a specific species. These programs have been successfully employed to determine the stability of reference gene expression and identify stable reference genes for various plant species, such as tall fescue (Yang et al., 2015), moss (Li et al., 2015b), switchgrass (Gimeno et al., 2014), and *Caragana korshinskii* (Yang et al., 2014c). Because of the different algorithms, the rankings provided by different programs are often not completely identical, and Jacob et al. (2013) recommend that more than two algorithms should be used for reference gene stability evaluation.

In this study, we selected 12 candidate reference genes based on the abiotic transcriptome datasets of *O. ochrocephala* by RNA-Seq (unpublished data) and further validated the expression stability of these genes by qRT-PCR and evaluated them using popular software packages including geNorm (Vandesompele et al., 2002), NormFinder (Andersen et al., 2004) and BestKeeper (Pfaffl et al., 2004). Three target genes, 1-aminocyclopropane-1-carboxylate oxidase (*ACO*), calcium binding protein (*CaBP*), and cinnamate-4-monoxygenase (*C4H*), were used to validate the effectiveness of the selected reference genes. The three genes *ACO*, *CaBP*, and *C4H* represent the characteristics of stress-responsive genes in plants. Under drought stress, *ACO* gene is up-regulated to increase the production of ethylene which is a key plant hormone to regulate plant growth and development (Chen et al., 2005). *CaBP* gene is a member of the calcium signaling system, which regulates an important pathway in plant cold acclimation (Yang et al., 2010). *C4H* is involved in the accumulation of soluble phenylpropanoids which act as the radical scavenger, helping to reduce cell oxidative damage caused by biotic and abiotic stress (Hemm et al., 2004). This is the first report which has determined the reference genes in *O. ochrocephala* and the other *Oxytropis* species. In particular, this work provides the basis for further research in exploring the mechanism of adaption *O. ochrocephala* to stress environments.

Materials and Methods

Plant Materials and Stress Treatments

Seeds of *O. ochrocephala* were collected from Haiyuan, Ningxia province. Seeds were pre-treated with concentrated sulphuric acid for 8 min, then washed three times with distilled water and placed onto wet filter papers in Petri dishes for germination. Germinated seeds were kept growing until cotyledons emerged...
and then transferred into individual pots (5 × 5 × 6 cm) filled with sand and peat (1:1), and grown in the growth chamber under controlled conditions at a photo flux density of 300 µmol m⁻²s⁻¹ (14/10 h, day/night period), at a relative humidity of 55–60%, and a temperature of 25 ± 2°C (Tang et al., 2007). Each group of 10 pots of 4-week-old seedlings with consensus growth status were selected and treated by drought, cold and salt stress, respectively. Three biological replicates were included for each treatment and control.

For drought treatment, 20% PEG-6000 solution (w/v, polyethylene glycol, Sangon, China) was applied to irrigate the plants in the soil every 4 h, simulating osmotic stress which is equivalent to −0.54 MPa, according to Jones and Gorham (1983); For cold treatment, seedlings were transferred into another growth chamber at 4°C with the rest growing conditions being the same. For salt treatment, 150 mM NaCl (Sangon, China) solution was applied to irrigate the plants. Entire seedlings were collected at different times (0, 3, 6, 12 h) from the onset of treatment and immediately frozen in liquid nitrogen for RNA extraction.

**Total RNA Extraction and cDNA Synthesis**

Total RNA was extracted according to the modified CTAB method (Gasic et al., 2004). To eliminate DNA contamination, total RNA was DNase I (Ambion, USA) treated and purified according to the manufacturer’s protocol. The integrity of total RNA samples were verified by 1.2% (w/v) agarose gel electrophoresis, and the quantity and quality of RNA samples were measured with the NanoDrop ND-1000 Spectrophotometer (NanoDrop Technologies, USA).

In order to perform qRT-PCR, cDNA was synthesized by reverse transcription using 2.5 µg total RNA in a 20 µl reaction volume according to the manufacturer’s instructions (Thermo Scientific, USA). The cDNA was diluted 10-fold with nuclease-free water for qRT-PCR.

**Transcriptome Data Mining for Candidate Reference Genes**

We performed transcriptome sequencing of *O. ochrocephala* using Illumina paired-end sequencing technology on an Illumina Hi-Seq™ 2000 platform for the four samples (control, drought, cold and salt). After assembly and annotation, a software package, RSEM (RNA-Seq by Expectation Maximization) (Li and Dewey, 2011), was employed to analyze the read counts, which were then converted to fragments Kilobase of exon model per million mapped reads (FPKM values), a commonly accepted estimate for the expression level of unigenes (Trapnell et al., 2010). To estimate expression stability of each gene, the following indices of FPKM values, including mean expression value (MV), standard deviation (SD) and coefficient of variation (CV) value (dividing SD by MV) were calculated according to the methodology described by de Jonge et al. (2007). Genes with lower CV values were considered to be stably expressed and the cut-off CV value was set as lower than 0.3, suggested by Czechowski et al. (2005) and Narsai et al. (2010).

Twelve candidate reference genes were selected based on gene annotation to NR database. Sequences of the candidate reference genes were used to design primers using Primer5 software. The criteria for primer design were set as follows: melting temperature (Tm) in a range of 55–60°C, primer lengths of 18–24 bp, GC contents of 45–55% and amplicon lengths of 95–282 bp.

**RT-PCR and qRT-PCR Analysis**

To verify the specificity of all the primer sets, PCR was performed using pooled cDNA as templates, and the PCR products were examined by 2% (w/v) agarose gel electrophoresis. The amplicons should appear as a single band with the correct size. To confirm that the PCR product did not contain multiple amplicons of the same size, which were amplified from possibly existing several genes in a gene family, eight individual PCRs were performed for each gene and the products were sequenced. Then the genes were further analyzed by qRT-PCR.

qRT-PCR reactions were performed in 20 µl system on a Bio-Rad CFX96 Real-Time PCR system (Bio-Rad, USA) with a reaction contained 10 µl of FastStart Universal SYBR GreenMaster (Roche, Germany), 2 µl of diluted cDNA template, 1 µl of each primer (10 µM). For each gene, a no template control (NTC) was included using water instead of cDNA as template. The qRT-PCR reactions were conducted following the fast thermal cycles: 95°C for 10 min, 40 cycles at 95°C for 30 s, 58°C for 30 s, and 72°C for 30 s. After 40 cycles, the dissociation curve was performed to confirm the specificity of each primer again by heating up the product from 60°C to 95°C. The Rn (normalized reporter) threshold was automatically selected to obtain the cycle threshold (Ct) values. The final Ct value of each sample was the mean of three biological replicates and three technical replicates. The mean amplification efficiency of each primer pair was checked by the LinRegPCR program (Ruijter et al., 2009).

**Data Analysis of Gene Expression Stability**

Three different types of programs, geNorm (version 3.5), BestKeeper and NormFinder were used to rank the stability of the selected reference genes across all the experimental sets (Xia et al., 2014; Yang et al., 2014c; Galli et al., 2015). For geNorm and NormFinder, the raw Ct value of each gene was converted into the relative quantities by the formula 2⁻ΔCt (ΔCt = each corresponding Ct value - lowest Ct value). For BestKeeper, the Ct value was used to calculate the standard deviation (SD) and the coefficient of variation (CV). The most stable gene exhibits the lowest CV ± SD value. Genes with SD value greater than 1 were considered to be unacceptable as reference genes. The pairwise variation (Vn/Vn+1) between two sequential normalization factors was obtained by geNorm software for determine the optimal number of reference genes needed to normalize. The recommended cut-off threshold is 0.15, below which an additional control gene is not required for normalization (Vandesompele et al., 2002). The overall recommended rankings of the best reference genes were obtained using the results of three algorithms.

**Validation of Reference Genes**

The expression pattern of three target genes *ACO, CaBP*, and *CH4* were analyzed using the most stable and least stable...
reference genes after normalization across three experimental sets, drought stress (for ACO), cold stress (for CaBP) and salt stress (for C4H). To validate the results, the expression levels of the target genes in qRT-PCR were compared with the FPKM values in RNA-seq data in each sample (Figure S2). The sequences of these genes were obtained from the transcriptome data (Table S3). The amplification efficiencies of the target genes were also estimated by the LinRegPCR program. The average Ct value was calculated from three biological and technical replicates and used for relative expression analyses. The relative expression data were calculated according to the $2^{-\Delta\Delta CT}$ method and presented as fold change (Livak and Schmittgen, 2001).

**Results**

**Identification of Putative Reference Genes Based on Transcriptome Datasets**

Based on the FPKM value, we selected 12 genes with low CV values (below 0.3) (18S, ACT101, ACTB, ACT7, ACT11, GAPDH1, GAPDH2, TUB, TUA, MET, FBA, and HIS) and these genes have been reported as stable genes in other species. The sequences of the 12 reference genes were obtained from the transcriptome data. Gene symbol, gene length and function description were shown in Table 1. We used full length unigene sequences from transcriptome data to design the specific primers for qRT-PCR. Primer pairs were designed to avoid the conserved sequences from transcriptome data to design the specific primer sets. Gene symbol number length numbers FPKM FPKM FPKM

**Expression Profiles of Reference Genes**

The primer specificities were confirmed by agarose gel electrophoresis, sequencing and melting curves, which showed the single amplicon of the expected size and the single peak melting curves (Figure S1, sequencing result not shown). The qRT-PCR products ranged from 95 to 282 bp. The amplification efficiencies of these reference genes were spanning from 1.8 to 2, and the linear $R^2$ (correlation coefficients) ranged from 0.995 to 0.999. The primer sequences and relevant amplification information were presented in Table 2.

**geNorm Analysis**

Average pairwise expression ratios (M) of the 12 reference genes revealed a minimum of 15.93 ± 2.56 (18S) (mean ± SD) and a maximum of 30.34 ± 2.20 (MET) for highest and lowest expression levels, which showed the expression level of MET was almost 44-fold lower than 18S. The genes with higher SD of Ct values indicated more variable expression compared to these with lower SD. HIS showed the smallest variation in gene expression (23.37 ± 0.79), while 18S with the most variable levels of expression (Figure 1).

**TABLE 1 | Description of the candidate reference genes in transcriptome datasets of O. ochrocephala.**

| Gene symbol | Gene ID | Accession number | Gene length (bp) | NR description | NR accession numbers | CK_FPKM | PEG_FPKM | COLD_FPKM | SALT_FPKM | CV |
|-------------|---------|------------------|------------------|----------------|---------------------|---------|----------|-----------|-----------|-----|
| 18S         | comp82308_c0 | KR822224         | 7829             | ribosomal RNA gene [Medicago truncatula] | XP_003614382.1 | 702.31 | 619.06 | 650.75 | 642.99 | 0.054 |
| ACT101      | comp81236_c0 | KR822225         | 1849             | Actin 101 [Glycine max] | XP_003547582.1 | 343.95 | 343.4 | 395.6 | 321.39 | 0.089 |
| ACT7        | comp79791_c0 | KR822227         | 1936             | Actin 2/7 [Glycine max] | XP_003523242.1 | 409.7 | 433.17 | 541.08 | 414.85 | 0.137 |
| ACT11       | comp81236_c3 | KR822228         | 1167             | Actin 11 [Medicago truncatula] | XP_003622019.1 | 178.49 | 168.43 | 192.51 | 164.35 | 0.071 |
| TUA         | comp80981_c0 | KR822229         | 1970             | Alpha-tubulin [Medicago truncatula] | XP_003629736.1 | 204.66 | 175.09 | 156.08 | 156.78 | 0.132 |
| TUB         | comp49473_c0 | KR822230         | 1969             | Beta-Tubulin [Medicago truncatula] | XP_003603765.1 | 169 | 135.68 | 129.07 | 121.21 | 0.152 |
| GAPDH1      | comp83553_c0 | KR822231         | 2197             | GAPDH1 [Medicago truncatula] | XP_003603851.1 | 403.17 | 385.18 | 276.76 | 346.27 | 0.159 |
| GAPDH2      | comp90007_c0 | KR822232         | 1731             | GAPDH2 [Medicago truncatula] | XP_003620182.1 | 1094.27 | 1141.33 | 968.64 | 1140.51 | 0.075 |
| MET         | comp71233_c0 | KR822233         | 681              | metallothionein-like protein [Glycine max] | XP_003520973.1 | 1976.38 | 1789.19 | 2507.53 | 2023.19 | 0.148 |
| FBA         | comp67830_c0 | KR822234         | 1851             | Fructose-bisphosphate aldoklase [Glycine max] | XP_003537836.1 | 1479.71 | 1708.19 | 1020.81 | 1267.67 | 0.215 |
| HIS         | comp67746_c0 | KR733680         | 955              | Histone H3 [Zea mays] | DAA45649.1 | 654.76 | 620.72 | 509.92 | 616.37 | 0.104 |
TABLE 2 | List of primer sequences of the candidate reference genes and their amplification efficiencies.

| Name   | Forward primer sequence [5′–3′] | Reverse primer sequence [5′–3′] | Amplicon Size (bp) | Product TMa (◦C) | qRT-PCR Efficiencyb | R2   |
|--------|---------------------------------|---------------------------------|--------------------|------------------|---------------------|------|
| 18S    | CACCTTAGGGGCGTCTG           | GGTGATACCCGTCGTGAGAC          | 282                | 87.5–88          | 1.807               | 0.998|
| ACT101 | GGCTCGATATGGTGTGTA          | GTCTGGCTTTCAGGAACCAAA         | 232                | 79.5–80          | 1.901               | 0.997|
| ACTB   | GGGCTCCATGGAGAACCC           | TCTGGGCGTACCAATAGCTT          | 186                | 81–81.5          | 1.987               | 0.995|
| ACT7   | CTCCACAGAGAGAAGTCTAGT        | GCCGAAATATACACCTCGAGT         | 266                | 80–80.5          | 1.908               | 0.995|
| ACT11  | TCCCTTCTGACTCCCTCA           | GACCTGACATGGGGAAAC           | 168                | 81.5–82          | 1.907               | 0.995|
| TUA    | GAAAGTTATGGAGGAGTTGGA       | TAACAATGGTATGGGAAGTATATA       | 260                | 77.5–78.5        | 1.834               | 0.997|
| TUB    | TAGCTCTTCTCCCTCGCTAT       | CAATAAAGCGATATGGTAGG          | 188                | 80–80.5          | 1.961               | 0.996|
| GAPDH 1| CCAAATACAGGTGTCCTACTCT      | CACTGGTTCTCCCCCTACA           | 211                | 80–80.5          | 1.920               | 0.998|
| GAPDH 2| TCTATTCTGGGTGTGGACTTTTT     | CCTCCCCTGCCTCCCA             | 162                | 77.5–78          | 1.901               | 0.997|
| MET    | GGTGGAGATGATGAAGTGA          | TTCTATGCGCAGTAGATGAGTGG       | 251                | 75–76            | 1.910               | 0.997|
| FBA    | GAATCATTTTCTGGGCAAGT        | ATGATTTCTCGGCAACCTT           | 224                | 76–77            | 1.902               | 0.999|
| HIS    | ATGGGCTTTGGGCGTCTTGA        | CTTATGCGACTAGCACAGAA          | 95                 | 76.5–77          | 2.047               | 0.997|

aThe melting temperature was calculated by Bio-Rad CFX Manager software.
bThe mean qRT-PCR amplification efficiency and correlation coefficients (R2) for each primer pair were determined by LinReg PCR software.

FIGURE 1 | Cycle threshold (Ct) values of 12 candidate reference genes across all samples. The final Ct value of each sample was the mean of three biological and technical replicates. Box graph indicates the interquartile range. A line across the box showed as the median. Lower and upper dashes represent the maximum and minimum values.

geNorm, ACT101, and HIS were the most stable reference genes in total samples. In contrast, TUA and 18S were the least stable reference genes. Under each subset of the treatment, the two best reference genes in drought stress were ACT101 and HIS with the lowest M value. The most preferred genes for normalization in the cold stress were ACT7 and HIS. As for the salt stress, GAPDH2 and HIS were the most stable reference genes. In general, HIS was the most stable gene in total or single treatment (Figure 2).

The optimal number of the reference genes required for accurate normalization was determined by pairwise variation (Vn/Vn+1). In the subsets of drought and cold stress, the V2/3 value was below 0.15 (0.144 and 0.107, respectively), which suggested that two reference genes should be used for normalization. In the salt treatment subset, three reference genes were sufficient for accurate normalization, as the V3/4 value was lower than 0.15. When total samples were considered, the pairwise variation V7/8 value was the lowest (0.165) but still above 0.15 (Figure 3).

NormFinder Analysis
NormFinder ranks genes based on the stability value for each reference gene. More stable gene expression has the lower stability value. In the subset of drought stress, TUB and ACTB were the most stable. In cold stress, HIS and ACT101 were the most stable. In salt stress, the top three stably expressed genes were GAPDH2, ACT101, and HIS. When evaluated the total experimental samples, ACT7 and FBA were the top ranked genes (Table 3). In contrast, TUA was the least stable reference gene under drought stress, ACT11 was the least stable under cold stress and 18S was the least stably expressed gene under salt stress. The rank in NormFinder was slightly different from that in geNorm. Genes considered as the most stable by geNorm (ACT101 and HIS) ranked third and forth by NormFinder, respectively.

BestKeeper Analysis
The BestKeeper program determines the stability ranking of the reference genes based on the CV ± SD values. In the drought stress set, HIS (2.17 ± 0.51) and ACT101 (2.55 ± 0.67)
Zhuang et al. Selection of reference genes in O. ochrocephala

FIGURE 2 | Expression stability values (M) of 12 candidate reference genes calculated by geNorm. Lower M values indicate more stable expression. Ranking of the gene expression stability was performed in all the samples and each abiotic stress samples. The least stable genes were on the left and the most stable genes on the right. (A) All tested samples (B) Drought stress (C) Cold stress (D) Salt stress.

with lowest CV ± SD values were identified as stable genes, which were consistent with the results of geNorm. In the cold stress set, only ACT7 and HIS showed SD < 1, which were considered as the most stable genes. In the salt stress treatment, HIS (2.24 ± 0.51), GAPDH2 (3.13 ± 0.85), and ACTB (3.26 ± 0.90) were identified as the best reference genes for normalization (Table 4).

Reference Gene Validation
Data generated by the three algorithms across different experimental sets were further compared (Table 5). The best and worst ranked candidate reference genes were selected for normalizing three target genes ACO, CaBP, and C4H under different experimental conditions. Under drought treatment, the expression level of ACO was up-regulated 2.8-fold when normalized using the two stable genes (HIS and ACT101), while the expression level was overestimated (232-fold) when normalized using the least stable combination (TUA and 18S). When HIS and ACT101 were used as the reference genes, the expression level of ACO was generally identified with the expression profile in RNA-seq (Figure S2). In the same way, under cold stress, the expression level of CaBP was 3.4-fold higher than control when using the most stable reference genes (ACT7 and HIS) as the internal control. By contrast, the expression pattern of CaBP was down-regulated when the least stable genes were used (ACT11 and TUA). Based on the result of geNorm, three reference genes were needed for normalization under salt stress condition. The stable gene combination (GAPDH2, HIS, and ACT101) was used to analyze the transcript abundance of C4H under salt stress. C4H was 3.4-fold up-regulated evaluated by the stable genes and there was a bias when the worst genes (18S+GAPDH1+ACT11) were used for normalization (Figure 4).

Discussion
An accurate and careful expression analysis of stress related genes can bring valid information to reveal more details about the transcriptional networks in stress response process. Therefore, it is required that any transcriptional study be based on accurate selection of suitable reference genes for normalization to ensure the accuracy of the analysis. To date, few studies have undertaken a comparison and selection of reference genes in O. ochrocephala or even in any Oxytropis species. As a result, this hinders relevant work in characterization of genes involved in stress responses in O. ochrocephala. We had performed the first large-scale transcriptome data for O. ochrocephala (unpublished), consisting of 118,596 unigenes, which served as the source of the reference gene selection. The study demonstrated that the RNA-Seq data are useful source for candidate reference genes screening and represented an important strategy for large-scale reference gene selection for non-model plants. By using three different computer
Zhuang et al. Selection of reference genes in O. ochrocephala

FIGURE 3 | Pairwise variation (V) of 12 candidate reference genes calculated by geNorm to determine the optimal number of reference genes for accurate normalization. Asterisk indicates the optimal number of reference genes required for normalization.

TABLE 3 | Expression stability of the candidate reference genes calculated by NormFinder.

| Rank | Gene   | Stability | Gene   | Stability | Gene   | Stability | Gene   | Stability | Gene   | Stability |
|------|--------|-----------|--------|-----------|--------|-----------|--------|-----------|--------|-----------|
| 1    | ACT7   | 0.602     | TUB    | 0.523     | HIS    | 0.804     | GAPDH2 | 0.520     |
| 2    | FBA    | 0.619     | ACTB   | 0.978     | ACT101 | 0.881     | ACT101 | 0.582     |
| 3    | ACT101 | 0.809     | ACT7   | 1.111     | ACTB   | 0.915     | HIS    | 0.591     |
| 4    | HIS    | 0.857     | MET    | 1.270     | ACT7   | 1.287     | TUB    | 0.936     |
| 5    | ACTB   | 0.925     | ACT101 | 1.290     | GAPDH2 | 1.422     | ACTB   | 1.039     |
| 6    | TUB    | 1.006     | FBA    | 1.299     | FBA    | 1.477     | ACT7   | 1.232     |
| 7    | GAPDH2 | 1.133     | GAPDH2 | 1.460     | TUB    | 1.607     | TUA    | 1.539     |
| 8    | ACT11  | 1.347     | HIS    | 1.582     | GAPDH1 | 1.719     | FBA    | 1.778     |
| 9    | MET    | 1.529     | GAPDH1 | 1.887     | MET    | 1.752     | ACT11  | 1.858     |
| 10   | GAPDH1 | 1.947     | ACT11  | 1.919     | 18S    | 1.790     | MET    | 1.880     |
| 11   | TUA    | 2.079     | 18S    | 1.933     | TUA    | 2.054     | GAPDH1 | 2.313     |
| 12   | 18S    | 2.360     | TUA    | 2.239     | ACT11  | 2.262     | 18S    | 2.551     |

Lower expression stability value indicates more stable expression.

Amongst protein-coding genes, HIS was indicated as a good potential internal control under drought treatment. This result was consistent with a previous study in strawberry, which confirmed that HIS gene was the most stable reference genes under osmotic stresses and salt stress (Galli et al., 2015). However, in liverwort, as evaluated by geNorm and NormFinder, HIS gene showed poor expression stability except for the group involving hormone treatment (Saint-Marcoux et al., 2015). The result demonstrates that in specific species reference genes are regulated differently, each species may have its own stably expressed genes which should be determined for each subset of experimental conditions. MET and FBA were relatively uncommon reference genes compared with the frequently used ones (commonly used). Several reports have indicated that the expression levels of traditional reference genes vary considerably and the newly discovered reference genes perform better than the traditional reference genes in Arabidopsis (Dekkers et al., 2012), Brassica napus L. (Yang et al., 2014b), and tea (Hao et al., 2014). While in this study, neither FBA nor MET ranked the top three most stable genes under tested experimental conditions. The commonly used reference genes outperformed the novel ones in O. ochrocephala. Moreover, based on transcriptome datasets, more novel and stable reference gene could be identified from other O. ochrocephala samples in the further study.

The actin gene(s), the most frequently used reference gene (Zhong et al., 2011; Galli et al., 2013), also showed poor consistency in Brassica napus L, soybean and peach (Tong et al., 2009; Ma et al., 2013; Yang et al., 2014b). It is notable that among the four actin genes tested in our study, ACT101 and ACT2/7 ranked head of ACTB and ACT11. Similarly, gene expression levels varied between the homologous genes TUA and TUB, and between GAPDH1 and GAPDH2. In the present study, the
commonly used reference gene TUA and GAPDH1 displayed variable expression pattern but TUB and GAPDH2 were more stable in each subset. This result was consistent with an earlier study in soybean, in that the expression of Actin2/7 was also more stable than Actin11 in different tissues (Jian et al., 2008) and similarly in rice, UBQ5 was the most stably expressed, whereas UBQ10 exhibited the least stable expression in a given set of tissue samples (Jain et al., 2006). These results suggest that the expression pattern and stability of members from the same gene family may be variable in the same experiment. For these homolog genes, the fluctuated Ct values suggest different levels of expression and they have the same CDS (coding sequence) from different gene loci. This highlights the need to ensure primer specificity at the outset of any gene expression experiment because of the similarity of coding sequence in the genes within the same gene family. In our experiment, primer sets were designed in the UTR to avoid the conserved domain, which enabled the gene specific amplification (Table S2).

Our results indicated that 18S was the least stable reference gene in almost all the experimental condition. The difference of Ct between 18S and HIS (stable gene) was 7, suggesting over 20-fold of gene expression between the two genes. When 18S was validated as a reference gene for normalization the target genes ACO, the expression pattern was obviously overestimated.

### TABLE 4 | Expression stability of 12 candidate reference gene calculated by BestKeeper.

| Rank | Gene | ALL SD | ALL CV | Drought SD | Drought CV | Cold SD | Cold CV | Salt SD | Salt CV |
|------|------|--------|--------|------------|------------|---------|---------|---------|---------|
| 1    | HIS  | 0.68   | 2.91   | 0.51       | 2.17       | 0.53    | 2.27    | 0.51    | 2.24    |
| 2    | ACT101 | 0.71  | 2.72   | 0.67       | 2.55       | 0.97    | 4.10    | 0.85    | 3.13    |
| 3    | FBA  | 0.82   | 2.95   | 1.00       | 4.01       | 1.10    | 3.92    | 0.90    | 3.26    |
| 4    | ACT7 | 0.91   | 3.89   | 1.11       | 4.05       | 1.11    | 3.90    | 0.94    | 3.73    |
| 5    | TUB  | 1.36   | 4.94   | 1.17       | 4.09       | 1.12    | 3.73    | 1.05    | 3.52    |
| 6    | GAPDH2 | 1.42  | 4.99   | 1.31       | 4.60       | 1.19    | 4.78    | 1.08    | 3.76    |
| 7    | ACTB | 1.42   | 4.94   | 1.53       | 5.36       | 1.22    | 4.50    | 1.33    | 5.66    |
| 8    | MET  | 1.52   | 5.02   | 1.80       | 5.97       | 1.64    | 5.99    | 1.54    | 5.95    |
| 9    | ACT11 | 1.58  | 5.55   | 2.00       | 7.30       | 1.91    | 13.85   | 1.67    | 5.99    |
| 10   | 18S  | 1.86   | 11.67  | 2.24       | 7.91       | 2.03    | 7.40    | 2.21    | 7.88    |
| 11   | TUA  | 1.96   | 7.20   | 2.31       | 13.78      | 2.22    | 8.29    | 2.59    | 10.33   |
| 12   | GAPDH1 | 2.55  | 9.54   | 2.40       | 7.76       | 2.61    | 9.20    | 3.17    | 23.74   |

### TABLE 5 | Expression stability ranking of the 12 candidate reference genes.

| Method                               | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   | 10  | 11  | 12  |
|--------------------------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| A. RANKING ORDER UNDER ALL STRESS (BETTER–GOOD–AVERAGE) | HIS | ACT101 | FBA | ACT7 | TUB | GAPDH2 | ACTB | MET | ACT11 | 18S | TUA | GAPDH1 |
| BestKeeper                           | HIS | ACT101 | FBA | ACT7 | TUB | GAPDH2 | ACTB | MET | ACT11 | 18S | TUA | GAPDH1 |
| NormFinder                           | ACT7 | FBA | ACT101 | HIS | ACTB | TUB | GAPDH2 | ACTB | MET | ACT11 | 18S | TUA | GAPDH1 |
| geNorm                               | ACT101/HIS | ACT7 | FBA | ACTB | TUB | GAPDH2 | MET | ACT11 | 18S | TUA | GAPDH1 |
| Comprehensive ranking                | ACT7 | HIS | ACT101 | FBA | ACTB | TUB | GAPDH2 | MET | ACT11 | 18S | TUA | GAPDH1 |
| B. RANKING ORDER UNDER DROUGHT STRESS (BETTER–GOOD–AVERAGE) | HIS | ACT101 | FBA | ACT7 | TUB | TUA | ACTB | MET | GAPDH2 | 18S | ACT11 | |
| BestKeeper                           | HIS | ACT101 | FBA | ACT7 | TUB | TUA | ACTB | MET | GAPDH2 | 18S | ACT11 | |
| NormFinder                           | ACT101 | ACT7 | TUB | ACTB | MET | ACT101 | FBA | GAPDH2 | HIS | ACT11 | 18S | TUA |
| geNorm                               | ACT101/HIS | ACT7 | FBA | ACTB | TUB | GAPDH2 | ACTB | MET | GAPDH2 | 18S | TUA | ACT11 |
| Comprehensive ranking                | HIS | ACT101 | ACT7 | TUB | ACTB | MET | GAPDH2 | MET | ACT11 | 18S | TUA | ACT11 |
| C. RANKING ORDER UNDER COLD STRESS (BETTER–GOOD–AVERAGE) | HIS | ACT101 | GAPDH2 | FBA | MET | ACT101 | TUB | ACTB | 18S | TUA | GAPDH1 | ACT11 |
| BestKeeper                           | HIS | ACT101 | GAPDH2 | FBA | MET | ACT101 | TUB | ACTB | 18S | TUA | GAPDH1 | ACT11 |
| NormFinder                           | HIS | ACT101 | ACT7 | GAPDH2 | TUB | GAPDH2 | MET | ACT101 | TUB | GAPDH1 | 18S | TUA | ACT11 |
| geNorm                               | ACT7/HIS | FBA | MET | ACTB | GAPDH2 | ACT101 | TUB | GAPDH1 | 18S | TUA | ACT11 |
| Comprehensive ranking                | HIS | ACT101 | GAPDH2 | FBA | MET | ACTB | TUB | GAPDH1 | 18S | TUA | ACT11 |
| D. RANKING ORDER UNDER SALT STRESS (BETTER–GOOD–AVERAGE) | HIS | GAPDH2 | ACTB | ACT101 | MET | FBA | ACT7 | TUB | TUA | ACT11 | GAPDH1 | 18S |
| BestKeeper                           | HIS | GAPDH2 | ACTB | ACT101 | MET | FBA | ACT7 | TUB | TUA | ACT11 | GAPDH1 | 18S |
| NormFinder                           | GAPDH2 | ACT101 | HIS | TUB | ACTB | ACT7 | TUB | FBA | ACT11 | MET | GAPDH1 | 18S |
| geNorm                               | GAPDH2/HIS | ACT101 | ACTB | ACT7 | FBA | MET | TUA | TUB | ACT11 | GAPDH1 | 18S |
| Comprehensive ranking                | GAPDH2 | HIS | ACT101 | ACTB | ACT7 | FBA | MET | TUA | ACT11 | GAPDH1 | 18S |
consistent with findings in broomrape (González-Verdejo et al., 2008) and rice (Bevitori et al., 2014). According to Jain et al. (2006), the reason why 18S is not suitable for normalization in real-time PCR analysis is the excessive high expression level compared with the target genes, which make it difficult to analyze the qRT-PCR data by subtracting the baseline. Nonetheless, in the absence of better choices, 18S is still frequently used as the reference genes for normalizing gene expression because it is independent of developmental stages and external stimuli (Gantasala et al., 2013; Huang et al., 2014).

In order to avoid the erroneous data that may be triggered by using single reference gene, normalization with multiple reference genes is becoming the common way. To evaluate the optimal number of genes required for normalization, geNorm showed V7/8 was the lowest value in total treatment samples. However, using seven reference genes to normalize is not feasible in practice. The V value (pairwise variation) threshold 0.15 should not be considered as an absolute value but rather a suggested one. Several reports have even reported higher V values in some species (De Ketelaere et al., 2006; Wan et al., 2010) and it is dependent on the consideration of the research purpose. Vandesompele et al. (2002) proposed that at least three genes are required for more reliable normalization. Considering these results, three top ranked reference genes HIS, ACT7, and ACT101 were appropriate for gene expression normalization under drought, cold and salt conditions. Besides, using different reference genes may lead to opposite or misunderstanding results in normalization of target genes.

To conclude, the current work presented here shows that the transcriptome data provide an effective approach to identify a set of reference genes and the reference genes, HIS, ACT7, and ACT101, were stably expressed under abiotic stress treatments, which should facilitate future studies on gene expression in *O. ochrocephala*.

**Author Contributions**

Designed and performed the experiments: HZ, LW; Analyzed data and wrote the paper: HZ, YF; Conducted the experiments: YF, WH; Contributed reagents/materials/fund support: WH, YW.

**Acknowledgments**

This work was supported by Special Fund for Agro-scientific Research in The Public Interest (Grant No. 201203062), Natural Science Foundation of China (Grant No. 31402133), Special Fund for Scientific Research from Shaanxi Provincial Department of Education (Grant No. 14JK1725), and Key Laboratory Fund for Scientific Research from Shaanxi Provincial Department of Education (Grant No. 14JS099).

**Supplementary Material**

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2015.00475
References

Andersen, C. L., Jensen, J. L., and Ørntoft, T. F. (2004). Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. Cancer Res. 64, 5245–5250. doi: 10.1158/0008-5472.CAN-04-0496

Arachchambatta, A., and Sivakum, M. V. (2011). PR-10, defensin and cold dehydrin genes are among those over expressed in Oxytropsis (Fabaceae) species adapted to the arctic. Funct. Integr. Genomics 11, 497–505. doi: 10.1007/s10142-011-0223-6

Bevitori, R., Oliveira, M. B., Grossi-de-Sa, M. F., Lanna, A. C., da Silveira, R. D., and Petrofesa, S. (2010). Selection and validation of reference genes for quantitative real-time RT-PCR analysis in switchgrass (Panicum virgatum L.) during Magnaporthe oryzae infection and drought. Plant Physiol. 153, 2215–2217. doi: 10.1104/pp.110.152215

Hemmi, M. R., Rider, S. D., Ogas, J., Murrey, D. J., and Chapple, C. (2004). Light induces phenylpropanoid metabolism in Arabidopsis roots. Plant J. 38, 765–778. doi: 10.1111/j.1365-313X.2004.02089.x

Huang, I. K., Yan, H., Jiang, X., Zhang, X., Zhang, Y., Huang, X., et al. (2014). Evaluation of candidate reference genes for normalization of quantitative RT-PCR in switchgrass under various abiotic stress conditions. Bioenergy Res. 7, 1201–1211. doi: 10.1007/s12155-014-9457-1

Jacob, F., Guertler, R., Naim, S., Nixdorf, S., Fedier, A., Hacker, N. F., et al. (2013). Careful selection of reference genes is required for reliable performance of RT-qPCR in human normal and cancer cell lines. PLoS ONE 8:e59180. doi: 10.1371/journal.pone.0059180

Jones, R. W., and Gorham, J. (1983). “Osmoregulation,” in Physiological Plant Ecology III, eds O. L. Lange, P. S. Nobel, C. B. Osmond, and H. Ziegler (Berlin: Springer-Verlag), 35–58.

Li, B., and Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data without a reference genome. BMC Bioinformatics 12:323. doi: 10.1186/1471-2105-12-323

Li, H., Yao, W., Fu, Y., Li, S., and Guo, Q. (2015a). De novo assembly and discovery of genes that are involved in drought tolerance in Tibetan Sophora moorcroftiana. PLoS ONE 10:e011054. doi: 10.1371/journal.pone.0110154

Li, X. S., Zhang, D. Y., Li, H. Y., Gao, B., Yang, H. L., Zhang, Y. M., et al. (2015b). Characterization of reference genes for RT-qPCR in the desert moss Syntrichia caninervis in response to abiotic stress and desiccation/rehydration. Front. Plant Sci. 6:38. doi: 10.3389/fpls.2015.00038

Ling, H., Wu, Q., Guo, J., Xu, L., and Que, Y. (2014). Comprehensive selection of reference genes for gene expression normalization in sugarcane by real time quantitative RT-PCR. PLoS ONE 9:e97469. doi: 10.1371/journal.pone.0097469

Narsai, R., Ivanova, A., Ng, S., and Whelan, J. (2014). Identification of reference genes for Oryza sativa in Arabidopsis chartacea var. chartacea and phylogeography of endangered Oryza species. Mol. Ecol. 13, 3657–3673. doi: 10.1111/j.1365-294X.2004.02360.x

Pfaffl, M. W., Tichopad, A., Prgomet, C., and Neuvians, T. P. (2004). Determination of stable housekeeping genes, differentially regulated target genes, and quantification of mRNA input datasets. BMC Plant Biol. 5, 84305.47

Pryor, B., Creamer, R., Shoemaker, R., McLain-Romero, J., and Hambleton, S. (2009). Undiflum, a new genus for endophytic Embothrium oxytropis and parasitic Helminthosporium bormussleri on legumes. Botany 87, 178–194. doi: 10.1139/B08-130
Ralphs, M. H., Creamer, R., Baucom, D., Gardner, D. R., Welsh, S. L., Graham, J. D., 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

Ruijter, J., Ramakers, C., Hoogaars, W., Karlen, Y., Bakker, O., Moorman, A. F. M., et al. (2009). Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Res.* 37, e45–e45. doi: 10.1093/nar/gkp045

Saint-Marcoux, D., Proust, H., Dolan, L., and Langdale, J. A. (2015). Analysis of multiple transcriptomes of the African oil palm (*Elaeis guineensis*) for gene expression studies in the biofuel plant. *PLoS ONE* 10:e110454. doi: 10.1371/journal.pone.0110454

Yan, X., Dong, X., Zhang, W., Yin, H., Xiao, H., Chen, P., et al. (2014). Reference gene selection for quantitative real-time PCR normalization in *Reaumuria soongorica*. *PLoS ONE* 9:e104124. doi: 10.1371/journal.pone.0104124

Yan, C., Pan, H., Liu, Y., and Zhou, X. (2014a). Selection of reference genes for expression analysis using quantitative real-time PCR in the pea aphid, *Acyrthosiphon pisum* (Harris) (Hemiptera, Aphididae). *PLoS ONE* 9:e110454. doi: 10.1371/journal.pone.0110454

Yang, H., Liu, J., Huang, S., Guo, T., Deng, L., and Hua, W. (2014b). Selection and evaluation of novel reference genes for quantitative reverse transcription PCR (qRT-PCR) based on genome and transcriptome data in *Brassica napus*. *Gene* 538, 113–122. doi: 10.1016/j.gene.2013.12.057

Yang, Q., Yin, J., Li, G., Qi, L., Yang, F., and Wang, R. (2014c). Reference gene selection for qRT-PCR in *Caragana korshinskii* Kom. under different stress conditions. *Mol. Biol. Rep.* 41, 2325–2334. doi: 10.1007/s11033-014-3068-9

Yan, T., Chaudhuri, S., Yang, L., Du, L., and Poovaiah, B. (2010). A calcium/calmodulin-regulated member of the receptor-like kinase family confers cold tolerance in plants. *J. Biol. Chem.* 285, 7119–7126. doi: 10.1074/jbc.M109.035659

Zhang, L., He, L. L., Fu, Q. T., and Zhu, Z. F. (2011). Selection of reliable reference genes for expression studies in *Paspalum euphratica* using real-time quantitative PCR. *Mol. Genet. Genomics* 379, 113–122. doi: 10.1007/s00438-010-0600-7

Zhang, L., He, L. L., Fu, Q. T., and Xu, Z. F. (2013). Selection of reliable reference genes for gene expression studies in the biofuel plant *Jatropha curcas* using real-time quantitative PCR. *Int. J. Mol. Sci.* 14, 24338–24354. doi: 10.3390/ijms14122438

Zhong, H. Y., Chen, J. W., Li, C. Q., Chen, L., Wu, J. Y., Chen, J. Y., et al. (2011). Selection of reliable reference genes for expression studies by reverse transcription quantitative real-time PCR in *Itch** under different experimental conditions. *Plant Cell Rep.* 30, 641–653. doi: 10.1007/s00299-010-0992-8

**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 Zhuang, Fu, He, Wang 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.*