Reference gene selection for qRT-PCR assays in *Stellera chamaejasme* under abiotic stress and with hormone treatment based on transcriptome datasets

Xin Liu, Huirui Guan, Min Song, Yanping Fu, Xiaomin Han, Meng Lei, Jingyu Ren, Bin Guo, Wei He, Yahui Wei*

College of Life Science, Northwest University, Xi'an, Shaanxi, China

Key Laboratory of Resource Biology and Biotechnology in Western China, Northwest University, Xi'an, shaanxi, China

Corresponding Author:

Yahui Wei

Northwest University, No.229, TaiBai NorthRoad, Xi'an, shaanxi, 710069, China

Email address: weiyahui@nwu.edu.cn

Commented [AR1]: Replace by: subjected to abiotic stresses and hormone treatments

ABSTRACT

Background. *Stellera chamaejasme* Linn, an important poisonous plant of the China grassland, is toxic to humans and livestock. The rapid expansion of *S. chamaejasme* has greatly damaged the grassland ecology and, consequently, seriously endangered the development of animal husbandry. To draft efficient prevention and control measures, it has become more urgent to carry out research on its adaptive and expansion mechanisms in different unfavorable habitats at the genetic level. qRT-PCR is a widely used technique for studying gene expression at the transcript level; however, qRT-PCR requires reference genes (RGs) as endogenous controls for data normalization and only through ~~previous~~ RG selection ~~s~~ can we guarantee the reliability and robustness of expression studies and RNA-seq data analysis. Unfortunately, little research on the selection of RGs for gene expression data normalization in *S. chamaejasme* has been reported.

Method. In this study, 10 candidate RGs~~;~~ namely, *18S*, *60S*, *CYP*, *GAPCP1*, *GAPDH2*, *EF1B*, *MDH*, *SAND*, *TUA1*, and *TUA6*, were singled out from the transcriptome database of *S. chamaejasme*, and their expression stability under three abiotic stresses (drought, cold, and salt) and three hormone treatments (ABA, GA, and ETH) were estimated with geNorm, NormFinder, and BestKeeper.

Result. Our results showed that *GAPCP1* and *EF1B* were the best combination for the three abiotic stresses, whereas *TUA6* and *SAND*, *TUA1* and *CYP*, *GAPDH2* and *60S* were the best choices for ABA, GA and ETH treatment, respectively. Moreover, *GAPCP1* and *60S* were assessed to be the best combination for all samples, and *18S* was the least stable RG for use as an internal control ~~for~~ all of the experimental subsets. The expression patterns of two target genes (*P5CS2* and *GI*) further verified that the RGs that we selected were suitable for gene expression normalization.

Discussion. This work is the first attempt to comprehensively estimate the stability of RGs in *S. chamaejasme*. Our results provide suitable RGs for high-precision normalization in qRT-PCR analysis, thereby making it more convenient to analyze gene expression under these experimental conditions.

INTRODUCTION

Stellera chamaejasme Linn (Thymelaeaceae), a perennial herb and dominant plant of grassland desertification, is native to the northern and southwestern regions in China (Tseng, 1999; Sepulveda-Jimenez et al., 2005). The whole plant is toxic, and its main toxic component is isochamaejasmin, which can poison and kill cattle, sheep and other livestock (Shi & Wei, 2016). The rapid spread of *S. chamaejasme* speeds up the process of grassland desertification and also poisons a large number of livestock in pasturing areas, causing great damage and loss to the local grassland ecology and livestock husbandry (Shi & Wei, 2016). Thus, it is of fundamental importance to elucidate the mechanisms of the rapid spread and stress adaptation of *S. chamaejasme*. However, limited genome sequence information is available, which greatly hinders the study of stress functional genes, ultimately resulting in a slow advancement of prevention and control measures. For the above reasons, our group established local transcriptome data for *S. chamaejasme* seedlings at five different time points (300 mM NaCl treatment for 0 h, 3 h, 12 h, 24 h, and 72 h; three biological replicates) using the Illumina HiSeq 4000 sequencing platform.

Commented [AR2]: Insert: appropriate

Commented [AR3]: Insert … the programs (before geNorm)

Commented [AR4]: in

After transcriptome sequencing and data analysis, fragments per kilobase of exons per million fragments mapped (FPKM) converted from RSEM (RNA-Seq by Expectation Maximization) were used to estimate unigene expression, which in some cases led to a few false-positive results.

qRT-PCR is one of the most widely applied technologies to detect the expression levels of selected genes in many different samples (Huggett et al., 2005) because of its relatively accurate quantification, simplicity, specificity, high sensitivity and high throughput capacity (Qi et al., 2016; Wang et al., 2016a). In the relative quantitative method of qRT-PCR data processing, the choice of internal genes is particularly important, and small changes in reference gene (RG) stability will significantly influence the accuracy of the relative expression of target genes (Dheda et al., 2005). Generally speaking, an ideal RG should be an endogenous gene that does change in any of the tested tissues or under any of the experimental conditions (Derveaux, Vandesompele & Hellemans, 2010; Li et al., 2016a; Wang et al., 2016b). In cells, some endogenous housekeeper genes with consistent relative expression are often used as reference genes (Taylor et al., 2016).

Housekeeping genes (HKGs) generally refer to a class of highly conserved genes that have basic functionality in biochemistry metabolism in organisms (Fiume & Fletcher, 2012) and are normally expressed at relatively constant rates across different tissues (Warrington et al., 2000; Paolacci et al., 2009). However, several studies have found that the expression levels of HKGs vary to different degrees based on tissues, developmental stages, or experimental conditions (Thellin et al., 1999; Nicot et al., 2005; Wu et al., 2016). Therefore, it is necessary to select stably expressed HKGs as RGs before they are used to normalize target gene expression by qRT-PCR (Guenin et al., 2009; Gong et al., 2016). Up to date, many HKGs, such as 18S ribosomal RNA (*18S rRNA*), 28S ribosomal RNA (*28S rRNA*), β-actin (*ACT*), elongation factor 1-alpha (*EF1A*), glyceraldehyde-3-phosphate dehydrogenase (GAPDH), α or β tubulin (*TUB*), polyubiquitin (*UBQ*), cyclophilin (*CYP*), SAND protein family (*SAND*), malate dehydrogenase (*MDH*), glyceraledehyde-3-phosphate dehydrogenase of plastid 1 (*GAPCP1*) and so on, have been used to conducted studies for evaluating their stability under different experimental conditions (Demidenko, Logacheva & Penin, 2011; Chen et al., 2015; Cao, Wang & Lan, 2016; Ferraz Dos Santos et al., 2016; Wang et al., 2017).

As of now, there is no available internal control gene for qRT-PCR data normalization in *S. chamaejasme*, so we were unable to verify transcriptome sequencing results, analyze the expression patterns of salt or stress-related genes, or further clarify its spread mechanism. To solve this problem, in our study, we selected 10 candidate RGs based on the local salt *S. chamaejasme* transcriptome database and then determined their expression profiles in five different stages under various abiotic stresses (drought, salt, and cold) and with three hormone treatments (abscisic acid, ABA; gibberellin, GA; ethephon, ETH) by qRT-PCR and further evaluated their expression stabilities using three popular software packages**:** geNorm (Vandesompele et al., 2002), NormFinder (Andersen, Jensen & Orntoft, 2004) and BestKeeper (Pfaffl et al., 2004). The 10 candidate genes were *18S*, *60S*, *CYP*, *EF1B*, *GAPCP1*, *GAPDH2*, *MDH*, *SAND*, *TUA1*, and *TUA6*. Two target genes, Delta 1-pyrroline-5-carboxylate synthetase 2 (*P5CS*2), which encodes a crucial enzyme in the proline synthesis pathway under stress conditions by activating glutamate 5-kinase and glutamate-5-semialdehyde dehydrogenase (Strizhov et al., 1997), and GIGANTEA (*GI*), a circadian regulated gene whose protein product has not only been shown to regulate photoperiodic flowering and various developmental processes but has also been implicated in mediating cold stress and salinity stress responses (Cao,

Ye & Jiang, 2005; Penfield S, Hall A. 2009; Park, Kim & Yun, 2013; Li et al., 2017), were used to verify the selected RGs.

MATERIALS AND METHODS

Plant Materials and Stress Treatments

S. chamaejasme seeds were collected from Qilian, Qinghai province. After peeling, the seeds were treated with 98% H_2SO_4 for 9 – 11 min and were then rinsed for 30 min with running water and planted in individual pots (14.5 × 14.5 × 6.5 cm) filled with nutrition soil, vermiculite and perlite (6 : 1 : 1). Germinated seeds were grown 7 weeks and were then transferred to nurseries potted with double-layered filter paper for 3 days of adaptation cultivation. All of the nursery pots were placed in an artificial climate chamber at a temperature of 25 ± 2℃ during the day and 15 ± 2℃ at night, with a relative humidity of 50 – 55% and illumination intensity of 300 μmol $m^{-2}s^{-1}$ (14/ 10 h, day/ night). Three pots of 7-week-old seedlings (three biological replicates) with a consistent growth status for each group were chosen and treated with abiotic stresses and hormone treatments.

For drought and salt treatments, 20% PEG-6000 (w/v, Sangon, China) (Zhuang et al., 2015) and 300 mM NaCl (Sangon, China) (Wang et al., 2015) were applied to irrigate the seedlings, respectively. For cold stress, the seedlings in the nursery pots were shifted to another artificial climate chamber at 4℃. For hormone treatments, the leaves were sprayed with 0.1 mM ABA (Reddy et al., 2016; Wan et al., 2017), 0.1 mM GA (Li et al., 2016b), or 1.5 mM ETH (Wu et al., 2016). Seedlings were irrigated or sprayed every 12 hours during the course of the experiment. Complete seedlings were carefully collected at 0 h, 3 h, 12 h, 24 h, and 48 h after treatments; immediately frozen in liquid nitrogen; and stored at –80℃ refrigerator until total RNA isolation.

Total RNA Isolation and First Strand cDNA Synthesis

Five random individual plants, approximately 100 mg of seedlings in each sample, were used for total RNA isolation with a TRNzol reagent kit (TIANGEN, China). The concentration and 260/280 and 260/230 ratios of the RNA samples were detected with a Nano Drop ND-1000 Spectrophotometer (Nano Drop Technologies, USA), and the integrity of all of the RNA samples was verified by 1.0% (w/v) agarose gel electrophoresis (AGE). Subsequently, for reverse transcription PCR (RT-PCR) and qRT-PCR, a total of 3.0 μg of RNA was DNase I (Ambion, USA) treated and purified and then used to synthesize first strand cDNA by reverse transcription following (Roche, USA) in a 20 μl reaction system. Finally, cDNA, diluted 50-fold with ddH_2O, was used as the template for PCR amplification.

Commented [AR5]: delete

Commented [AR6]: delete

Candidate RG Selection and Primer Design

Ten candidate RGs from the local *S. chamaejasme* transcriptome database were selected by using local NCBI-blast (version 2.4.0+). The sequences of these genes were used to design the qRT-PCR primers using Primer 5.0, Oligo 7.60 and Beacon Designer 8.20 software with the following criteria: melting temperature ™ of 50 – 65℃, primer lengths of 17 – 25 bp, GC contents of 45 – 55% and product lengths of 90 – 300 bp. The specificity of all of the selected primer pairs was

observed via RT-PCR using the cDNA of control groups at 0 h as the template, and each gene fragmentation was underpinned by 2.0% (w/v) AGE and sequenced to ensure its reliability.

RT-PCR and qRT-PCR Analysis

To confirm the specificity of each primer that we designed, we performed RT-PCR in a 25 µl system using the Bio-Rad C1000 PCR system (Bio-Rad, USA). The reaction system was as follows: 2.5 µl of Ex Taq buffer, 2 µl of dNTPs, 0.125 µl of TaKaRa Ex Taq (TaKaRa, China), 2.0 µl of cDNA template, 0.5 µl of the forward primer (10 µM), 0.5 µl of the reverse primer (10 µM) and 17.375 µl of sterilized water. The RT-PCR reaction parameters were: 95℃ for 3 min, 40 cycles at 95℃ for 30 s, 58℃ for 30 s, 72℃ for 20 s, and 72℃ for 5 min. The amplification products were evaluated by 2.0% (w/v) AGE. To further confirm that the amplicon corresponded with what we desired, target products contained in the agarose gel were extracted using a TIANgel Midi Purification Kit (TIANGEN, China) and then sequenced.

qRT-PCR reactions were carried out with the Fast Start Universal SYBR GreenMaster (Roche, Germany) on a Bio-Rad CFX96 Real-Time PCR system (Bio-Rad, USA) in accordance with the manufacturer's instructions. Reactions were conducted at 95℃ for 3 min as an initial activation, followed by 40 cycles at 95℃ for 10 s, 58℃ for 10 s, and 72℃ for 20 s. ~~After 40 cycles,~~ The melting curves, ranging from 58℃ to 95℃, were determined to check the specificity of the amplicons. In the negative control group, qRT-PCR was performed using water instead of cDNA as the template. Three technical replicates were analyzed for each biological sample, and the final Ct values for each set of samples were the average of three biological replicates. A total of 45 cDNA samples from five time points in the control groups were used to determine the mean amplification efficiency (E) of each primer pair ~~by calculating~~ with the LinRegPCR program (Ruijter et al., 2009; Zhuang et al., 2015; Vavrinova, Behuliak & Zicha, 2016; Wu et al., 2016).

Data Analysis of Gene Expression Stability

Three different types of statistical tools: geNorm (version 3.5), NormFinder (version 0.953) and BestKeeper (version 1.0), were applied to rank the expression stability of the RGs across all of the experimental sets. For geNorm and NormFinder, the raw Ct values calculated by the CFX equipment^{TM} software were converted into the relative quantities using the formula $2^{-\Delta Ct}$ (Δ Ct = each corresponding Ct value – lowest Ct value) for gene expression profiling. For BestKeeper, the raw Ct values and amplification efficiencies estimated by the LinRegPCR program were used to calculate the coefficient of variation (CV) and standard deviation (SD). The RG with the lowest CV ± SD value was identified as the most stable gene. geNorm software was also used to determine the proper RG numbers with pair wise variation (Vn/Vn+1, n refers to the RGs number) between two sequential normalization factors.

Validation of Reference Genes

To test the accuracy of the ~~analysis~~ results ~~from this study~~, the geometric mean from the sort results of geNorm, NormFinder, and BestKeeper in each subset were used to calculate the comprehensive ranking of the candidate genes. The smaller the comprehensive ranking results, the better the gene expression stability. Then, the combination of the top two best RGs, best ranked RG and worst ranked RG were used to standardize the expression of two target genes, i.e.,

Commented [AR7]: refer amounts or concentrations, not µl; water volume is not needed as the recation volume is referred above

Commented [AR8]: replace by: ….to the target sequence, PCR ……

Commented [AR9]: where and how?

Commented [AR10]: denaturation

P5CS2 and *GI*, under different experimental conditions. Furthermore, the expression levels of *P5CS2* and *GI* under salt stress calculated by the combination of the top two best RGs were also compared with the FPKM values in the *S. chamaejasme* transcriptome database.

RESULTS

Selection of Candidate RGs and Target Genes

After comparing the reported RGs in other species with the local transcriptome database of *S. chamaejasme* using the local BLast program, ten RGs and two target genes were chosen to perform the gene normalization studies. The results showed that the E value of each blast gene indicated high homology. The untranslated region (UTR) of these full-length unigene sequences were used to design the specific primers for RT-PCR and qRT-PCR. The unigene ID, NCBI accession number, gene symbol, gene name, homolog locus of 10 candidate RGs and two target genes, and E value compared with those of the homologous genes are listed in **Table 1**.

Commented [AR11]: lower case

Verification of the Primer Specificity and qRT-PCR Amplification Efficiency

The specificity of each primer was tested by 2.0% AGE, sequencing and melting curves analysis, which provided the expected amplicon length (**Supplemental figure 1)** and single peak melting curves (**Supplemental figure 2)**. The primer sequences, amplicon size, product Tm, amplification efficiencies, and other relevant information are given in **Table 2**. The amplification product length of PCR varied from 94 bp to 267 bp. The Tm for all PCR products spanned from 76.0℃ for *MDH* to 83.5℃ for *GAPCP1*. The E-values of these genes were between 1.824 (*MDH*) and 1.930 (*GAPDH2*), and the linear correlation coefficients (R^2) varied from 0.994 (*SAND*) to 0.998 (*CYP*). In conclusion, we had every reason to believe that all of these specificity and efficiency estimates of the amplification were reliable for further analysis.

Expression Profiles of Candidate RGs

Boxplot analysis of the Ct values of different reference genes in all of the experimental samples was performed using origin 2017 software (**Fig. 1**). The results demonstrated that the mean Ct values of the ten candidate RGs presented a relatively wide field, from 19.26 to 30.76. *60S* showed the least expression variation, while *18S* exhibited the highest variation, with the Ct values ranging from 15.58 to 22.59. Since the Ct values are negatively related to the gene expression levels, the smaller the Ct value, the higher the gene expression level. As **Fig. 1** shows, *18S* was the highest-expressed RG for its lowest mean Ct value (15.58), and *GAPCP1* had the lowest expression level on account of its maximum mean Ct value (32.58).

Analysis of Gene Expression Stability

geNorm analysis. geNorm calculates the gene expression stability measure M value as the average pairwise variation V for the RG and other tested RGs (Vandesompele et al., 2002). The smaller the M value, the more stable the gene, and vice versa. In our study, the M values of the 10 candidate RGs of *S. chamaejasme* calculated by geNorm software were below 1.5 in all of the experimental settings (**Fig. 2**), suggesting that these genes should be considered relatively stable. As described in **Fig. 2A, Fig. 2B** and **Fig. 2C**, *GAPCP1* and *EF1B* under drought stress, *GAPCP1* and *60S* under cold stress, and *EF1B*and *60S* under salt stress were the most stable RGs

with the lowest M values of 0.07, 0.05 and 0.19, respectively. At the same time, in the ABA (**Fig. 2D**), GA (**Fig. 2E**), and ETH (**Fig. 2F**) treatment groups, *SAND* and *TUA6*, *TUA1* and *CYP*, *GAPDH*2 and *60S* were considered to be the most stable genes with the lowest M values of 0.21, 0.22 and 0.19, respectively. In addition, for all of the sample sets (**Fig. 2G**), *GAPCP1* and *CYP* were suggested to be two most stable RGs. On the contrary, *18S* was the least stable gene in all of the sets except for ETH treatment, in which *TUA6* was the least stable gene.

NormFinder analysis. NormFinder provides a stability value for each gene by analyzing expression data obtained through qRT-PCR, which is a direct measurement for estimating expression variation when the gene is used for normalization (Dheda et al., 2005). The orders based on the stability values calculated by NormFinder (**Table 3**) were similar to those determined by geNorm. The stability ranking results under cold stress and in the GA treatment subsets were completely consistent with the results determined through geNorm; meanwhile, *TUA6* and *18S* were the two least stable genes for ETH treatment and the rest of the treatments. For the cold stress group, *GAPCP1* and *EF1B* were the most two stable RGs (also ranked first by geNorm). For the salt stress group, *GAPCP1* and *TUA1* were the two most stable RGs, which was different from the geNorm results. For all samples, ABA-treated and ETH-treated subsets, NormFinder suggested that *GAPCP1* and *60S*, *TUA1* and *SAND*, *GAPDH2* and *60S* were the most stable RGs, respectively, which were not exactly the same as the geNorm analysis results.

Commented [AR12]: replace by: the two most

BestKeeper analysis. BestKeeper evaluates the RG expression stability by calculating the CV and SD of the average Ct values. A lower CV value indicates more stable RG expression, and candidate RGs with SD values greater than 1.0 were judged to be unstable and should be avoided for gene expression normalization (Guenin et al., 2009). As shown in **Table 4**, under drought stress and for all of the sample subsets, *TUA1* had the lowest CV ± SD values of 0.52 ± 0.16 and 0.53 ± 0.16 and was considered to be the most stable RG. Under the cold stress condition and salt stress and ABA treatment subsets, *EF1B*, which had the lowest CV ± SD values of 1.16 ± 0.31, 1.35 ± 0.36 and 1.04 ± 0.27, respectively, was identified as the best RG. In the GA treatment subset, *TUA6* had the lowest CV ± SD value of 0.82 ± 0.22 and was the most stable. In the ETH treatment subset, BestKeeper suggested that *GAPDH2* was the most stable RG with the lowest CV ± SD value of 0.68 ± 0.18. Additionally, only a few genes had a SD value greater than 1.0, indicating that most of the candidate RGs were relatively stable. Except for the ETH treatment subset, the most unstable RG among all of the experimental settings was *18S*, which was the same as the results of geNorm and NormFinder.

Commented [AR13]: should be part of Material and Methods; not here

Determination of the Optimal Number of RGs

At the suggestion of the geNorm Service tool, the critical value Vn/Vn+1 to determine the optimal RG number for qRT-PCR normalization is 0.15, below which the inclusion of an additional RG is not required (Vandesompele et al., 2002). As **Fig. 3** shows, the V2/3 values of all of the experimental groups were less than 0.15, which indicated that a two RG combination would be sufficient to use for normalization.

Comprehensive Stability Analysis of RGs

Table 5 and **Fig. 4** summarize and rank the determination results obtained from the geNorm, NormFinder and Bestkeeper programs. Based on the analysis ~~results~~, *GAPCP1* and *EF1B* were the most stable RGs under three abiotic stresses; thus, *TUA6* and *SAND*, *TUA1* and *CYP*,

GAPDH2 and *60S* were the best RG combinations under the ABA, GA and ETH treatments, respectively. Still, *18S* was the most unstable RG under all of the experimental conditions.

Reference Gene Validation

As shown in **Fig. 5** and **Fig. 6**, when the best RG combinations were used for performing normalization, the expression levels of *P5CS2* and GI were affected by different treatments. A sustained increase in expression level of *P5CS2* was observed after drought stress, and a peak point was observed at 48 h (**Fig. 5A**). A tendency of first increase, after ~~the~~ downward, and then upward in the transcript level of *P5CS2* appeared after cold and GA treatments (**Fig. 5B and 5E**). Additional, upregulated expression of *P5CS2* was observed after salt and ABA treatments, and reached the maximum value at 12 h and 24h following a decrease (**Fig. 5C and 5D**). Whereas, *P5CS2* expression was first downregulated at 3 h after ETH treatment and then began to continuous increase, reaching the maximum at 48 h (**Fig. 5F**). The maximal expression levels of *GI* under drought **(Fig. 6A)**, cold **(Fig. 6B)**, salt **(Fig. 6C)**, ABA **(Fig. 6D)** and ETH **(Fig. 6E)** treatments also appeared prominent changes, which were 4.66-fold, 29.22-fold, 2.10-fold, 6.45-fold and 2.45-fold higher than those of the control group, while there was no significant difference under GA treatment **(Fig. 6E)**.

Compared with the best RG combinations for normalization of P5CS and G1, similar expression patterns were obtained when the most stable single genes, *GAPCP1* (drought and cold), *EF1B* (salt), *TUA6* (ABA), *TUA1* (GA) and *GAPDH2* (ETH), were used ~~the~~ for normalization under the above treatments. However, different expression patterns were generated and the expression levels of *P5CS2* and *GI* were overestimated when the least stable gene, *18S*, was selected as the RG for normalization.

In particular, as shown in **Fig. 5C**, under salt treatment, when the RG combination (*GAPCP1* and *EF1B*) was selected for normalization, gene expression of *P5CS2* gradually increased from 0 h, reached the maximum at 24 h, and then began to slightly decline at 48 h. In the same way, the expression levels of *GI* increased at first, then decreased and maintained a lower level until 48 h (**Fig. 6C**). The expression trends of *P5CS2* and *GI* over the first 24 hours were generally consistent with those of RNA-seq (**Supplemental figure 3**), which further validated the accuracy and reliability of our experimental results.

DISCUSSION

qRT-PCR is currently viewed as a powerful technique that can be used to quantify target gene expression. The accuracy of qRT-PCR directly depends on the stability of the internal genes used. The use of inappropriate RGs for normalizing qRT-PCR data will lead to deviations in the results (Shivhare & Lata, 2016). In this study, three programs, geNorm, NormFinder and Bestkeeper, were used to select optimum RGs for six different experimental conditions. ~~As described in results above, three types of data analysis software detected different RGs as stable internal controls for qRT-PCR normalization based on different mathematical methods or parameters.~~ The ten potential RGs exhibited differential stability in response to different stresses. Taking ABA treatment as an example, in the experimental subset, geNorm software ranked *SAND* as the best RG, and NormFinder regarded *TUA1* as the most stable RG. However, Bestkeeper identified *EF1B* as the best RG according to its lowest CV value. This means that the three types of software generated different results, and a solution was not found. Our study

Commented [AR14]: Additionally,

Commented [AR15]: Not necessary

carried out a comprehensive analysis and provided ultimate stability ordering results by ranking the geometric means of the three software analysis results, which is a common strategy for evaluating the expression stability of RGs reported in previous scientific papers.

EF1B catalyzes the exchange of guanosine diphosphate (GDP) bound to the G-protein, elongation factor 1-alpha (*EF1A*), for guanosine triphosphate (GTP), an important step in the elongation cycle of protein biosynthesis. It has been considered to be one of the most stable reference genes during drought and salt stresses (Wan et al. 2017) and other stress conditions (Ma et al., 2013). In our study, *EF1B* was ranked as one of the two best RGs under drought and salt stresses according to the comprehensive ranking results, which was the same as in *Stipa grandis* (Wan et al. 2017). In addition, *EF1B* performed a better expression stability under three abiotic stresses than those in the three hormone treatments. The results showed that there are no universal RGs that are stably expressed in all biological materials and/or under all trial conditions.

The expression stability of two homologous RGs, *TUA1* and *TUA6*, were estimated in our study. According to the ~~above analysis~~ results, the stability ranking of *TUA1* was always better than that of *TUA6* under all conditions except the ABA treatment, under which *TUA6* exhibited better expression stability. Nevertheless, it is notable that the homologous RGs showed different rank orders in each subset, and in most cases, *TUA1* showed better expression stability than *TUA6*. ~~For the homolog genes, they had similar coding sequences from different gene loci. This means that we need to ensure primer specificity when we amplify these homogenous sequences from their gene families. In our study, the primer sets were designed in the UTR to avoid the conserved domain, which to a large extent enabled gene specific amplification.~~

18S is a frequently used HKG and is widely used for normalization in qRT-PCR analysis. Wang et al. (2017) reported that *18S rRNA* was the most stably expressed gene under diverse heavy metals stresses in tea plants; Huang et al.(2017) also found 18S rRNA was the most stable gene under UV irradiation and hormonal stimuli in *Baphicacanthus cusia*. However, Our analysis ~~results~~ suggested that *18S* was the most unstable RG in all of the experiment groups because of its excessively high expression level. In comparison with the best RG combination and the most stable RG, when *18S* was selected as a RG to validate the expression of the two target genes *P5CS2* and *GI*, their expression patterns were significantly overestimated, which was consistent with the findings in *Oxytropis ochrocephala* (Zhuang et al., 2015) and *rice* (Bevitori et al., 2014).

Two target genes, *P5CS2* and *GI*, were used to verify the stability of the selected RGs for gene expression normalization. Strizhov et al. (1997) stated that expression of *Arabidopsis thaliana P5CS* (*AtP5CS*) is root and leaf specific and can be regulated by salinity, drought and ABA. The same experimental results were reproduced in our experiments. We also found that *P5CS* could be efficiently expressed during the later period of cold stress, which may be a supplement to previous findings. The induction mechanism remains to be further studied. Cao, Ye and Jiang (2005) reported that *GI* is induced by cold stress, but not by salt, mannitol, and ABA. By contrast, Park, Kim and Yun (2013) and Kim et al. (2013) claimed that GI, as a negative regulator, participated in the regulation of salt stress in *Arabidopsis* by interacting with salt overly sensitive 2 (SOS2). Moreover, Riboni et al. (2016) revealed that ABA affects flowering through two independent regulatory mechanisms: activation of *GI* and constant (*CO*) functions upstream of the florigen genes and down-regulation of the suppressor of overexpression of CO1

Commented [AR16]: Irrelevant; most importantly, bring forward to the discussion a comparison on the use of these 2 genes in previous studies

Commented [AR17]: Also found that

Commented [AR18]: Lower case

(*SOC1*) signaling. Our findings indicated that the gene expression of *GI* not only changed under salt stress and cold stress but also underwent a significant change under drought, ABA and ETH treatments. We have reason to believe that these mechanisms will be revealed with future in-depth experiments.

There is no doubt that it is necessary to select suitable RGs and/or RG combinations for gene normalization studies to obtain more accurate and reliable results. Combined with all of the validation results above, we observed that, in most cases, *P5CS2* and *GI* showed similar response patterns when normalized by the RG combinations ~~and were the most stable RGs~~, but some differences still emerged. Unfortunately, we could not tell which choice was better for normalization. However, to eliminate the small variations caused by technical protocols in qRT-PCR, two or more RGs are often required to correct for non-specific experimental variation (Thellin et al., 1999; Bustin et al., 2009). In this study, two RG combinations, whose V2/3 values were less than 0.15 across all of the experimental subsets~~, which was in line with the geNorm software results, completely meet the requirements of normalization~~.

Commented [AR19]: Insert: most stable RG combinations

Commented [AR20]: Finalize the sentence (and thus the section) with the selected combinations

CONCLUSIONS

This study represents the first attempt to comprehensively analyze the stability of RGs for use as internal controls in qRT-PCR analysis of target gene expression in *S. chamaejasme* under three abiotic stresses and three hormone treatments by combining results from ~~the~~ three different methods. The results indicated that the stability of an identical gene was not exactly the same under different treatments, and the stability ranking of the RGs calculated by three parameters was not identical under the same treatment. As a result, it makes sense to carry out a comprehensive analysis of the results of the three procedures. Moreover, it may be a better choice to select a combination of two or more RGs as an effective internal control to further improve the accuracy and reliability of gene expression normalization under different stresses. In conclusion, this study provides a guideline to select a valid RG combination that can ensure more accurate qRT-PCR-based gene expression quantification and basic data to facilitate future molecular studies on gene expression in *S. chamaejasme* and ~~the~~ other Thymelaeaceae species (Che et al., 2016).

ACKNOWLEDGEMENTS

The authors are particularly grateful to Jiakun Dai, Na Fan, Shilan Feng, and other members of our research group for their helpful comments/suggestions to improve the experimental design.

REFERENCES

Andersen CL, Jensen JL, and Orntoft TF. 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 Research* 64:5245-5250. DOI: 10.1158/0008-5472.CAN-04-0496

Bevitori R, Oliveira MB, Grossi-de-Sa MF, Lanna AC, da Silveira RD, and Petrofeza S. 2014. Selection of optimized candidate reference genes for qRT-PCR normalization in rice (*Oryza sativa* L.) during *Magnaporthe oryzae* infection and drought. *Genetics and Molecular Research* 13:9795-9805. DOI: 10.4238/2014.November.27.7

Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, and Wittwer CT. 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clinical Chemistry* 55:611-622. DOI: 10.1373/clinchem.2008.112797

Cao J, Wang L, and Lan H. 2016. Validation of reference genes for quantitative RT-PCR normalization in *Suaeda aralocaspica*, an annual halophyte with heteromorphism and C4 pathway without Kranz anatomy. *PeerJ* 4:e1697. DOI: 10.7717/peerj.1697

Cao S, Ye M, and Jiang S. 2005. Involvement of GIGANTEA gene in the regulation of the cold stress response in *Arabidopsis*. *Plant Cell Reports* 24:683-690. DOI: 10.1007/s00299-005-0061-x

Che JX, Shi JL, Lu Y, and Liu YL. 2016. Validation of reference genes for normalization of gene expression by qRT-PCR in a resveratrol-producing entophytic fungus (*Alternaria* sp. MG1). *AMB Express* 6:106. DOI: 10.1186/s13568-016-0283-z

Chen Y, Hu B, Tan Z, Liu J, Yang Z, Li Z, and Huang B. 2015. Selection of reference genes for quantitative real-time PCR normalization in creeping bentgrass involved in four abiotic stresses. *Plant Cell Rep* 34:1825-1834. DOI: 10.1007/s00299-015-1830-9

Demidenko NV, Logacheva MD, and Penin AA. 2011. Selection and validation of reference genes for quantitative real-time PCR in buckwheat (*Fagopyrum esculentum*) based on transcriptome sequence data. *PLoS One* 6:e19434. DOI: 10.1371/journal.pone.0019434

Derveaux S, Vandesompele J, and Hellemans J. 2010. How to do successful gene expression analysis using real-time PCR. *Methods* 50:227-230. DOI: 10.1016/j.ymeth.2009.11.001

Dheda K, Huggett JF, Chang JS, Kim LU, Bustin SA, Johnson MA, Rook GAW, and Zumla A. 2005. The implications of using an inappropriate reference gene for real-time reverse transcription PCR data normalization. *Analytical Biochemistry* 344:141-143. DOI: 10.1016/j.ab.2005.05.022

Ferraz Dos Santos L, Santana Silva RJ, Oliveira Jordao do Amaral D, Barbosa de Paula MF, Ludke Falcao L, Legavre T, Alves RM, Marcellino LH, and Micheli F. 2016. Selection of reference genes for expression study in pulp and seeds of *Theobroma grandiflorum* (Willd. ex Spreng.) Schum. *PLoS One* 11:e0160646. DOI: 10.1371/journal.pone.0160646

Fiume E, and Fletcher JC. 2012. Regulation of *Arabidopsis* embryo and endosperm development by the polypeptide signaling molecule CLE8. *Plant Cell* 24:1000-1012. DOI: 10.1105/tpc.111.094839

Gong L, Yang Y, Chen Y, Shi J, Song Y, and Zhang H. 2016. LbCML38 and LbRH52, two reference genes derived from RNA-Seq data suitable for assessing gene expression in *Lycium barbarum* L. *Scientific Reports* 6:37031. DOI: 10.1038/srep37031

Guenin S, Mauriat M, Pelloux J, Van Wuytswinkel O, Bellini C, and Gutierrez L. 2009. Normalization of qRT-PCR data: the necessity of adopting a systematic, experimental conditions-specific, validation of references. *Journal of Experimental Botany* 60:487-493. DOI: 10.1093/jxb/ern305

Huang Y, Tan H, Yu J, Chen Y, Guo Z, Wang G, Zhang Q, Chen J, Zhang L, and Diao Y. 2017. Stable internal reference genes for normalizing real-time quantitative PCR in

Baphicacanthus cusia under hormonal stimuli and UV irradiation, and in different Plant organs. *Front Plant Sci* 8:668. DOI: 10.3389/fpls.2017.00668

Huggett J, Dheda K, Bustin S, and Zumla A. 2005. Real-time RT-PCR normalisation; strategies and considerations. *Genes and Immunity* 6:279-284. DOI: 10.1038/sj.gene.6364190

Kim WY, Ali Z, Park HJ, Park SJ, Cha JY, Perez-Hormaeche J, Quintero FJ, Shin G, Kim MR, Qiang Z, Ning L, Park HC, Lee SY, Bressan RA, Pardo JM, Bohnert HJ, and Yun DJ. 2013. Release of SOS2 kinase from sequestration with GIGANTEA determines salt tolerance in *Arabidopsis*. *Nat Commun* 4:1352. DOI: 10.1038/ncomms2357

Li J, Jia H, Han X, Zhang J, Sun P, Lu M, and Hu J. 2016a. Selection of reliable reference genes for gene expression analysis under abiotic stresses in the desert biomass willow, *Salix psammophila*. *Frontiers in Plant Science* 7:1505. DOI: 10.3389/fpls.2016.01505

Li MY, Song X, Wang F, and Xiong AS. 2016b. Suitable reference genes for accurate gene expression analysis in parsley (*Petroselinum crispum*) for abiotic stresses and hormone stimuli. *Frontiers in Plant Science* 7:1481. DOI: 10.3389/fpls.2016.01481

Li S, Ying Y, Secco D, Wang C, Narsai R, Whelan J, and Shou H. 2017. Molecular interaction between PHO2 and GIGANTEA reveals a new crosstalk between flowering time and phosphate homeostasis in *Oryza sativa*. *Plant, Cell & Environment* 40:1487-1499. DOI: 10.1111/pce.12945

Ma S, Niu H, Liu C, Zhang J, Hou C, and Wang D. 2013. Expression stabilities of candidate reference genes for RT-qPCR under different stress conditions in soybean. *PLoS One* 8:e75271. DOI: 10.1371/journal.pone.0075271

Nicot N, Hausman JF, Hoffmann L, and Evers D. 2005. Housekeeping gene selection for real-time RT-PCR normalization in potato during biotic and abiotic stress. *Journal of Experimental Botany* 56:2907-2914. DOI: 10.1093/jxb/eri285

Paolacci AR, Tanzarella OA, Porceddu E, and Ciaffi M. 2009. Identification and validation of reference genes for quantitative RT-PCR normalization in wheat. *Bmc Molecular Biology* 10:11. DOI: 10.1186/1471-2199-10-11

Park HJ, Kim WY, and Yun DJ. 2013. A role for GIGANTEA: keeping the balance between flowering and salinity stress tolerance. *Plant Signaling and Behavior* 8:e24820. DOI: 10.4161/psb.24820

Penfield S, Hall A. 2009. A role for multiple circadian clock genes in the response to signals that break seed dormancy in *Arabidopsis*. *Plant Cell* 21:1722-32. DOI: 10.1105/tpc.108.064022.

Pfaffl MW, Tichopad A, Prgomet C, and Neuvians TP. 2004. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper--Excel-based tool using pair-wise correlations. *Biotechnology Letters* 26:509-515.

Qi S, Yang L, Wen X, Hong Y, Song X, Zhang M, and Dai S. 2016. Reference gene selection for RT-qPCR analysis of flower development in *Chrysanthemum morifolium* and *Chrysanthemum lavandulifolium*. *Frontiers in Plant Science* 7:287. DOI: 10.3389/fpls.2016.00287

Reddy DS, Bhatnagar-Mathur P, Reddy PS, Sri Cindhuri K, Sivaji Ganesh A, and Sharma KK. 2016. Identification and validation of reference genes and their impact on normalized gene expression studies across cultivated and wild *Cicer* species. *PLoS One* 11:e0148451. DOI: 10.1371/journal.pone.0148451

Riboni M, Robustelli Test A, Galbiati M, Tonelli C, and Conti L. 2016. ABA-dependent control of GIGANTEA signalling enables drought escape via up-regulation of FLOWERING LOCUS T in *Arabidopsis thaliana*. *Journal of Experimental Botany* 67:6309-6322. DOI: 10.1093/jxb/erw384

Ruijter JM, Ramakers C, Hoogaars WM, Karlen Y, Bakker O, van den Hoff MJ, and Moorman AF. 2009. Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Research* 37:e45. DOI: 10.1093/nar/gkp045

Shi Z, and Wei Y. 2016. *Important poisonous plants of china grassland*. Beijing: China Agriculture Press.

Shivhare R, and Lata C. 2016. Selection of suitable reference genes for assessing gene expression in pearl millet under different abiotic stresses and their combinations. *Scientific Reports* 6:23036. DOI: 10.1038/srep23036

Strizhov N, Abraham E, Okresz L, Blickling S, Zilberstein A, Schell J, Koncz C, and Szabados L. 1997. Differential expression of two P5CS genes controlling proline accumulation during salt-stress requires ABA and is regulated by ABA1, ABI1 and AXR2 in *Arabidopsis*. *The Plant Journal* 12:557-569.

Taylor CM, Jost R, Erskine W, and Nelson MN. 2016. Identifying stable reference genes for qRT-PCR normalisation in gene expression studies of narrow-leafed lupin (Lupinus angustifolius L.). *PLoS One* 11:e0148300. DOI: 10.1371/journal.pone.0148300

Thellin O, Zorzi W, Lakaye B, De Borman B, Coumans B, Hennen G, Grisar T, Igout A, and Heinen E. 1999. Housekeeping genes as internal standards: use and limits. *Journal of Experimental Botany* 75:291-295.

Tseng YC. 1999. Flora Reipublicae Popularis Sinicae. Beijing: Science Press. p 397.

Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, and Speleman F. 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology* 3:RESEARCH0034.

Vavrinova A, Behuliak M, and Zicha J. 2016. The importance of the selection of appropriate reference genes for gene expression profiling in adrenal medulla or sympathetic ganglia of spontaneously hypertensive rat. *Physiological Research* 65:401-411.

Wan D, Wan Y, Yang Q, Zou B, Ren W, Ding Y, Wang Z, Wang R, Wang K, and Hou X. 2017. Selection of reference genes for qRT-PCR analysis of gene expression in *Stipa grandis* during environmental stresses. *PLoS One* 12:e0169465. DOI: 10.1371/journal.pone.0169465

Wang ML, Li QH, Xin HH, Chen X, Zhu XJ, and Li XH. 2017. Reliable reference genes for normalization of gene expression data in tea plants (*Camellia sinensis*) exposed to metal stresses. *PLoS One* 12:e0175863. DOI: 10.1371/journal.pone.0175863

Wang X, He J, Wang W, Ren M, Gao S, Zhao G, Wang J, and Yang Q. 2016a. Validation of internal reference genes for relative quantitation studies of gene expression in human laryngeal cancer. *PeerJ* 4:e2763. DOI: 10.7717/peerj.2763

Wang X, Ma X, Huang L, and Zhang X. 2015. Identification of the valid reference genes for quantitative RT-PCR in annual ryegrass (*Lolium multiflorum*) under salt stress. *Molecules* 20:4833-4847. DOI: 10.3390/molecules20034833

Wang Y, Chen Y, Ding L, Zhang J, Wei J, and Wang H. 2016b. Validation of reference genes for gene expression by quantitative real-time RT-PCR in stem segments spanning primary to secondary growth in *Populus tomentosa*. *PLoS One* 11:e0157370. DOI: 10.1371/journal.pone.0157370

Warrington JA, Nair A, Mahadevappa M, and Tsyganskaya M. 2000. Comparison of human adult and fetal expression and identification of 535 housekeeping/maintenance genes. *Physiological Genomics* 2:143-147.

Wu J, Zhang H, Liu L, Li W, Wei Y, and Shi S. 2016. Validation of reference genes for RT-qPCR studies of gene expression in preharvest and postharvest Longan fruits under different experimental conditions. *Frontier in Plant Science* 7:780. DOI: 10.3389/fpls.2016.00780

Zhuang H, Fu Y, He W, Wang L, and Wei Y. 2015. Selection of appropriate reference genes for quantitative real-time PCR in *Oxytropis ochrocephala* Bunge using transcriptome datasets under abiotic stress treatments. *Frontiers in Plant Science* 6:475. DOI: 10.3389/fpls.2015.00475