

# Expression Stabilities of Candidate Reference Genes for RT-qPCR under Different Stress Conditions in Soybean

Shuhua Ma<sup>1,9</sup>, Hongwei Niu<sup>1,9</sup>, Chunji Liu<sup>2</sup>, Jie Zhang<sup>1</sup>, Chunyan Hou<sup>1\*</sup>, Dongmei Wang<sup>1\*</sup>

**1** College of Life Science, Agricultural University of Hebei, Baoding, Hebei Province, China, **2** CSIRO Plant Industry, Queensland Bioscience Precinct, Australia

## Abstract

Due to its accuracy, sensitivity and high throughput, real time quantitative PCR (RT-qPCR) has been widely used in analysing gene expression. The quality of data from such analyses is affected by the quality of reference genes used. Expression stabilities for nine candidate reference genes widely used in soybean were evaluated under different stresses in this study. Our results showed that *EF1A* and *ACT11* were the best under salinity stress, *TUB4*, *TUA5* and *EF1A* were the best under drought stress, *ACT11* and *UKN2* were the best under dark treatment, and *EF1B* and *UKN2* were the best under virus infection. *EF1B* and *UKN2* were the top two genes which can be reliably used in all of the stress conditions assessed.

**Citation:** Ma S, Niu H, Liu C, Zhang J, Hou C, et al. (2013) Expression Stabilities of Candidate Reference Genes for RT-qPCR under Different Stress Conditions in Soybean. PLoS ONE 8(10): e75271. doi:10.1371/journal.pone.0075271

**Editor:** Ken Mills, Queen's University Belfast, United Kingdom

**Received:** May 12, 2013; **Accepted:** August 14, 2013; **Published:** October 4, 2013

**Copyright:** © 2013 Ma et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was supported by the National Natural Science Foundation of China (no. 30971706) and by the Natural Science Foundation of Hebei Province (no. C2008000321). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: houchunyan@126. Com (CH); dongmeiwang63@126. com (DW)

<sup>9</sup> These authors contributed equally to this work.

## Introduction

Due to its accurate quantification, high sensitivity and high throughput, real time quantitative PCR (RT-qPCR) has been widely used in analyzing gene expression, in determining number of gene copies and in detecting the presence of microbes in food products [1]. Reference genes are required in RT-qPCR analysis to minimize influences of RNA quality and quantity and efficiency of reverse transcription [2,3].

House-keeping genes are often selected as reference genes. The most commonly used reference genes include  $\beta$ -actin (*ACT*), glyceral-dehyde-3-phosphate dehydrogenase (*GAPDH*), 18S ribosomal RNA (*18S rRNA*), 25S ribosomal RNA (*25S rRNA*), polyubiquitin (*UBQ*), ubiquitin conjugating enzyme (*UBC*), translation elongation factor (*TEF*), cyclophilin (*CYP*), elongation factor 1-A (*EF1A*) and tubulin (*TUB*) etc. [3–5]. Ideal reference genes should provide stable expression in different plant tissues, at different stages of development or under different environments of experiments. However, it has been reported that many of the house-keeping genes provide stable expression under only certain environments and it is necessary to identify and select suitable reference genes for a given experiment [6]. There are also many reports showing that the use of 2 or 3 reference genes could be necessary as using a single reference gene could lead to significant error [7,8]. Efforts in identifying suitable reference genes have been reported in many crop species including *Arabidopsis thaliana* [9–11], *Brassica napus* [12], soybean [13], *Pisum sativum* [14], rice [5,15–16], *Platyedus orientalis* [17], coffee [18], *Gossypium hirsutum* [19], tomato [20,21] and water lily [22]. Several new reference genes have been identified [23–25]. It has been reported that *SKIP16* (SKIP/Ask-Interacting Protein 16), *MTP* (Metallo protease, Insulin degrading Enzyme), *UKN1* and *UKN2* (Hypothetical protein) all gave stable expressions at different stages of

development in soybean [13]. Investigating mechanisms of viral resistance in soybean has been one of our research foci. We have demonstrated that callose deposition at the plasmodesmata plays a critical role in host resistance to viral infection [26]. During the investigation of the physiological mechanism of callose deposition, we analysed the expression of callose synthase and  $\beta$ -1, 3-glucanase using RT-qPCR and observed that the expression levels of these genes were different with the use of different reference genes (data not shown). Based on this observation, we systematically investigated expression stabilities of eight widely used (including *ACT11*, *TUA5*, *CYP*, *EF1B*, *TUA4*, *TUB4*, *EF1A* and *ACT2/7*) and one recently reported (*UKN2*) reference genes under viral infection and stresses of darkening, salinity and drought, with the objective of identifying those which provide stable expressions in each of the environments assessed.

## Materials and Methods

### Plant Genotypes

Two soybean (*Glycine max* (L.) Merrill) cultivars Jidou 7 and Nannong 1138-2 were used in this study. The former was used for various treatments as described below and the latter was used for preparing inoculums of soybean mosaic virus (SMV). Seeds were germinated and grown in a greenhouse with a 14 h light/10 h dark cycle at a constant temperature of 25°C and 700  $\mu\text{mol photons m}^{-2} \text{s}^{-1}$ .

### Experiments Conducted

Following experiments were conducted when the unifoliate leaves of the seedlings were fully unfolded:

**Infection of the soybean seedlings with SMV:** The SMV inoculums were prepared by grinding leaves of infected soybean cv. Nannong 1138-2 to slurry with a pestle in a mortar with 0.1 M

phosphate buffer, pH 7.4. Carborundum was used as an abrasive. The unifoliate leaves of cv. Jidou 7 were inoculated by rubbing with a brush. The leaves were rinsed with distilled water immediately after inoculation. Leaf samples were taken at 0, 8, 24, and 48 hours post inoculation.

**Dark treatment:** the seedlings were transferred in to a box for this experiment. Leaf samples were taken at 0, 2, 4, and 6 hours respectively post the treatment.

**Salt treatment:** This experiment was conducted hydroponically. The seedlings were transferred to a complete nutrient solution containing 200 mM NaCl. Leaf samples were taken at 0, 2, 16, 20 hours post treatment.

**Drought treatment:** This experiment was also conducted hydroponically. The seedlings were transferred to a complete nutrient solution containing 15% PEG<sub>6000</sub>. Leaf samples were taken at 0, 2, 4, and 6 hours post treatment.

### RNA Isolation and cDNA Synthesis

RNA was extracted using RNaiso Plus (TaKaRa, Japan) according to the manufacturer's instructions, and genomic DNA was then eliminated using RNase-free DNase I (TaKaRa, Japan). The quantity and quality of the total RNA extracted was determined using a Nanophotometer p-class K5600. Only RNA samples meeting the following requirements were used in this study: Firstly, they should have an absorbance ratio at OD<sub>260</sub>/280 between 1.8 and 2.2; secondly, their absorbance ratio at OD<sub>260</sub>/230 should be about 2.0; thirdly, the ratio of 28S/18S ribosomal RNA should be between 1.5 and 2.0; and fourthly, they should have little smears on agarose gels. First strand cDNA was synthesized by reverse transcribing 500 ng of total RNA with PrimeScript<sup>®</sup>RT reagent Kit (Perfect Real Time) (TaKaRa, Japan). All cDNA were stored at -20°C until use.

### Primer Design and RT-qPCR

Primers for the 9 reference genes were designed using Primer 3 software (Table 1). These primer sets have a single melting temperature of 55°C and the amplicons amplified by them vary between 100–200 bp in length. The primers were synthesized by Shanghai Biological Engineering Technology Services Company.

RT-qPCR reactions were carried out in 96-well blocks with an Applied CHROMO4 Real-Time PCR system using SYBR premix Ex Taq II kit (TaKaRa, Japan). After 40 cycles, a melting curve analysis was carried out (60°C to 95°C) to verify the specificity of amplicons. Each amplification was repeated 3 times. The specificity of the amplicons was confirmed by the presence of a single peak.

### Statistical Analysis

The relative level of expression ( $Q$ ) for a given gene was calculated based on the formula  $Q = 2^{-\Delta Ct}$ . Each Ct value represents the average of three replicates.  $\Delta Ct$  equals to  $Ct_{\text{samples}} - Ct_{\text{min}}$  ( $Ct_{\text{min}}$  being the one with the lowest Ct value in all samples and  $Ct_{\text{samples}}$  represents each sample by Ct value) [15].

For selecting the appropriate reference genes, we used GeNorm [27] and NormFinder [28] in analysing the expression stability of the genes assessed. The  $Q$  value was imported into GeNormv3.5 and NormFinder for reference gene selection. Two parameters were obtained when  $Q$  values were loaded into GeNorm, an average expression stability value ( $M$  value) and an average pairwise variation of template normalization factor ( $V_{n/n+1}$  value). The  $M$  value of a gene was inversely correlated with its expression stability. The default setting for the cut-off value  $V$  (0.15) was used. Thus, if  $V_{n/n+1} \leq 0.15$ , it is not necessary to introduce  $n+1$  reference genes as the internal control [6]. Therefore, through

these two parameters we can analyze the most stable reference genes and obtain the required reference gene number.

Similar to GeNorm, NormFinder also generates  $Q$  from the original data via either standard curve (Figure S3) or  $\Delta Ct$  method and then use the software to obtain gene expression stable value  $M$  through analysis of variance and direct assessment of genetic stability.

With purified PCR product as a starting template, followed by five times of 10-fold serial dilutions ( $10^0$  (PCR product 500 times dilution),  $10^{-1}$ ,  $10^{-2}$ ,  $10^{-3}$ ,  $10^{-4}$ , respectively) creating a gradient of concentrations. Standard curve was drawn according to the results of RT-qPCR (Figure S3). The formula  $E = (10^{-1/\text{slope}} - 1) \times 100$  was used to calculate the amplification efficiencies of the genes and the calculation results were shown in Table 2. Primer amplification efficiency between 95%–105% can be used.

## Results

### Efficiency and Specificity of Designed Primer Pairs Used for RT-qPCR Amplification

The specificity of each primer pair can be assessed by RT-qPCR analysis. Figure S1 showed that each melting curve of the 9 reference genes assessed was characterized as a single peak (Figure S1), indicating that the primer pairs used are highly specific. The specificity of these primer pairs was also confirmed by running a 2% agarose gel (Figure S2).

Accurate analysis of RT-qPCR data requires that all primer pairs used should have the same amplification efficiency. The rate of amplification efficiency at 100% or close to it indicates that the reaction conditions are optimal and the results obtained should be highly repeatable. Experimentally, the amplification efficiency of each primer pair should be between 95–105%, thus it can be used for the next step of experiments. The amplification efficiencies for the nine primer pairs used in this study were shown in Table 2, and they were all between 97.65–110.77%. Correlation coefficients ( $R^2$ ) were all larger than 0.98, indicating that all of the primer pairs are highly specific and efficient in RT-qPCR amplification.

### Expression Levels of the Candidate Reference Genes Assessed

Ct value represents the number of cycles that the PCR product appear to effectively increase [29,30]. By calculating Ct value, we can compare the mRNA abundance of each of the tested genes. The higher the Ct value, the lower the mRNA level of the gene; Conversely, the lower the Ct value, the higher the mRNA level of the gene. Average Ct values for most of the nine reference genes under different treatments varied between 18 and 20, whereas the full range varied between 18 and 25 (Figure 1). The expression levels of *TU44* and *UKN2* were lower than those of the other 7 genes, and their Ct values were 22 and 24, respectively. The Ct values of the reference genes were variable under different treatment conditions. The differences of Ct values (coefficient of variation) indicate the stability of the reference gene expression, the greater the coefficient of variation of the gene, the more unstable of the gene expression. The coefficient of variation of *ACT2/7* was high above six cycles and those of *EF1B* and *UKN2* were low (approximately two cycles). Calculating the Ct value for determining the reference gene expression stability in different samples is important to select the internal reference standard.

### GeNorm Analysis

GeNorm software can analyze and determine the most stable reference gene by analysing the stability of the reference gene

**Table 1.** Primer sequences and related information for each candidate reference gene.

Genesymbol name	Gene	NCBI Accession No.	Arabidopsis homolog locus	Primer sequence (5'-3')		Size (bp)	Function
<i>ACT11</i>	Actin 11	LOC100781831	AT3G12110	ATTTTGACTGAGCGTGGTTATTCC	GCTGGTCCTGGCTGTCTCC	126	Cytoskeletal structural protein
<i>TUA5</i>	Alpha Tubulin	LOC732582	AT5G19780	TGCCACCATCAAGACTAAGAGG	ACCACCAGGAACAACAGAAGG	103	Structural constituent of cytoskeleton
<i>CYP</i>	Cyclophilin	LOC100500498	AT2G21130	ACGACGAAGACGGAGTGG	CGACGACGACAGGCTTGG	130	Protein folding
<i>EF1B</i>	Elongation factor 1-beta	LOC100500082	AT5G12110	CCACTGCTGAAGAAGATGATGATG	AAGGACAGAAGACTTGCCACTC	134	Translational elongation
<i>TUA4</i>	alpha Tubulin	LOC100781185	AT1G50010	CATACCCTAGAATCCATTTTC	TGTACTTTCCGTGACGAG	156	Structural constituent of cytoskeleton
<i>TUB4</i>	beta Tubulin	LOC100798849	AT5G12250	TGGCGTCCACATTCATTG	GAATCCATCTCGTCCAT	137	Structural constituent of cytoskeleton
<i>EF1A</i>	Eukaryotic elongation factor 1-alpha	LOC100785429	AT5G60390	GACCTTCTCGTTTCTCGCA	CGAACCTCTCAATCACACGC	162	Translational elongation
<i>ACT2/7</i>	Actin2/7	LOC100789000	AT5G09810	CTCCCTCAGCACCTTCCAA	GGTCCAGCTTTCACACTCCAT	119	Cytoskeletal structural protein
<i>UKN2</i>	Hypothetical protein	LOC100789577	AT4G33380	TGTGCTCTGTGAAGAGATTG	TCATAATCTGTGTGCAGTTC	156	Unkown

doi:10.1371/journal.pone.0075271.t001

expression (M value) in different samples. The default value suggested by GeNorm is  $M = 1.5$ , the reference gene, of which the M value is less than 1.5, is considered to be used as an internal reference gene. The larger the M value, the worse the stability. Figure 2A shows the results of the analysis of all 16 samples under the four stress conditions. As shown, *EF1B* and *UKN2* gave the lowest M values ( $M = 0.328$ ), indicating that the expression stability of the two reference genes were the best among the nine tested genes. *ACT2/7* showed the highest M value ( $M = 1.172$ ), indicating that it was the most unstable gene.

The stabilities of the nine tested reference genes were not consistent under different treatment conditions. Under salt

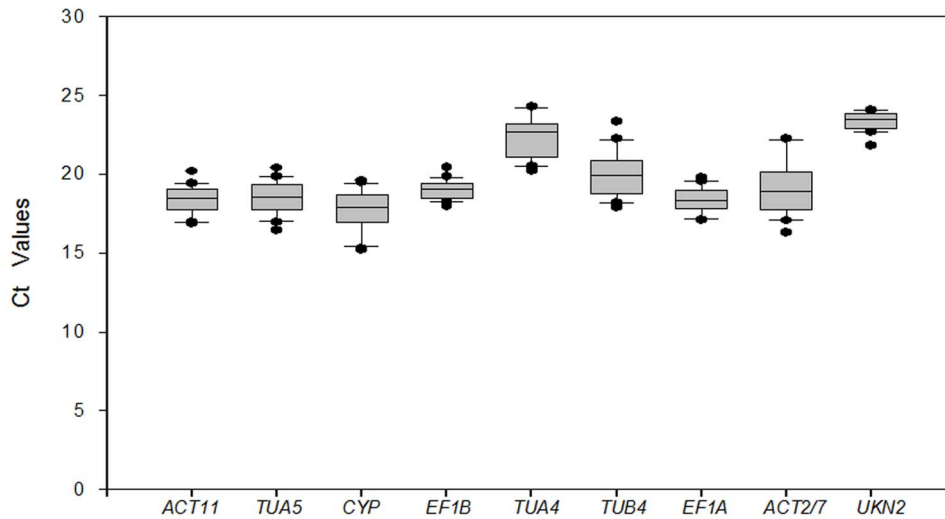
treatment, the most stable genes were *ACT11* and *EF1A* and the most unstable one was *ACT2/7* (Figure 2B); under drought treatment, the most stably expressed genes were *TUA5* and *TUB4* and the most unstable one was *CYP* (Figure 2C); under dark treatment, *ACT11* and *UKN2* were the most stably expressed genes while *TUB4* was the most unstable one (Figure 2D); under virus stress, *EF1B* and *UKN2* were the most stably expressed genes while *CYP* the most unstably expressed one (Figure 2E). The M values for all of these tested reference genes were all lower than the default value of  $M = 1.5$ .

In order to get more reliable results from RT-qPCR, it is generally recommended to use two or more reference genes.

**Table 2.** Efficiency of designed primer pairs used for RT-qPCR amplification.

Gene symbol	Arabidopsis homolog locus	Tm (°C)	PCR efficiency (%)	Regression coefficient (R2)
<i>ACT11</i>	AT3G12110	82.82	103.89	0.999
<i>TUA5</i>	AT5G19780	80.97	103.89	0.998
<i>CYP</i>	AT2G21130	87.23	97.65	0.996
<i>EF1B</i>	AT5G12110	82.05	110.77	0.999
<i>TUA4</i>	AT1G50010	83.28	106.78	0.986
<i>TUB4</i>	AT5G12250	83.66	100.59	0.998
<i>EF1A</i>	AT5G60390	81.97	105.68	0.998
<i>ACT2/7</i>	AT5G09810	81.15	99.99	0.995
<i>UKN2</i>	AT4G33380	83.98	100.59	0.997

doi:10.1371/journal.pone.0075271.t002



**Figure 1. Candidate reference gene expression levels in different samples.** Expression data displayed as Ct values for each reference gene in all samples. A line across the box is depicted as the median. The box indicates the 25th and 75th percentiles. Whiskers represent the maximum and minimum values.

doi:10.1371/journal.pone.0075271.g001

GeNorm software could analyze the pairwise variation value of the normalization factor ( $V$ ) [19]. By using the  $V$  factor, the appropriate number of reference genes could be determined under different treatments. Therefore, It is proposed that 0.15 could be used as the cut-off value for  $V$ , below which the inclusion of an additional control gene is not required; that is, if  $V_{n/n+1} < 0.15$ , it is not necessary to use  $n+1$  reference genes as internal controls. The analysis results were shown in Figure 3, under the salt stress (Figure 3B) and viral infection (Figure 3E), the  $V_{2/3} < 0.15$ , the third reference gene failed to significantly expressed standardization factor differences, indicating that under these two treatments chosen the two most stable expressed genes, *EF1A* and the *ACT11* (Figure 2B), and *EF1B* and *UKN2* (Figure 2E), were enough to be used as accurate standardized controls. Under the drought treatment (Figure 3C),  $V_{2/3} > 0.15$ , it was necessary to introduce a third reference gene, which means that the introduction of the third gene could enable a smaller  $V$  value. This was proved to be true by the analysis result,  $V_{3/4} < 0.15$ , therefore the three genes, *TUB4*, *TUA5* and *EF1A* were chosen as reference genes under this condition (Fig. 2C), which suggested that group cooperation of internal references can get more accurate data. Under dark treatment (Figure 3D), although the value of  $V_{3/4}$  is smaller, but  $V_{2/3} < 0.15$ , therefore *ACT11* and *UKN2* (Figure 2D) would be sufficient used as cooperation internal references as accurately standardized controls. Overall analysis of all the samples,  $V_{2/3}$  and  $V_{3/4}$  values were less than 0.15 (Figure 3E), which indicated that *EF1B* and *UKN2* can be used as the standardization of gene expressions under various stress treatments in this study (Figure 2E).

### NormFinder Analysis

In this study, NormFinder software was also used to determine the best reference gene as a standardization of RT-qPCR. Like GeNorm software, lower  $M$  values indicate the higher stability of a given gene. Results obtained through NormFinder software are shown in Table 3.

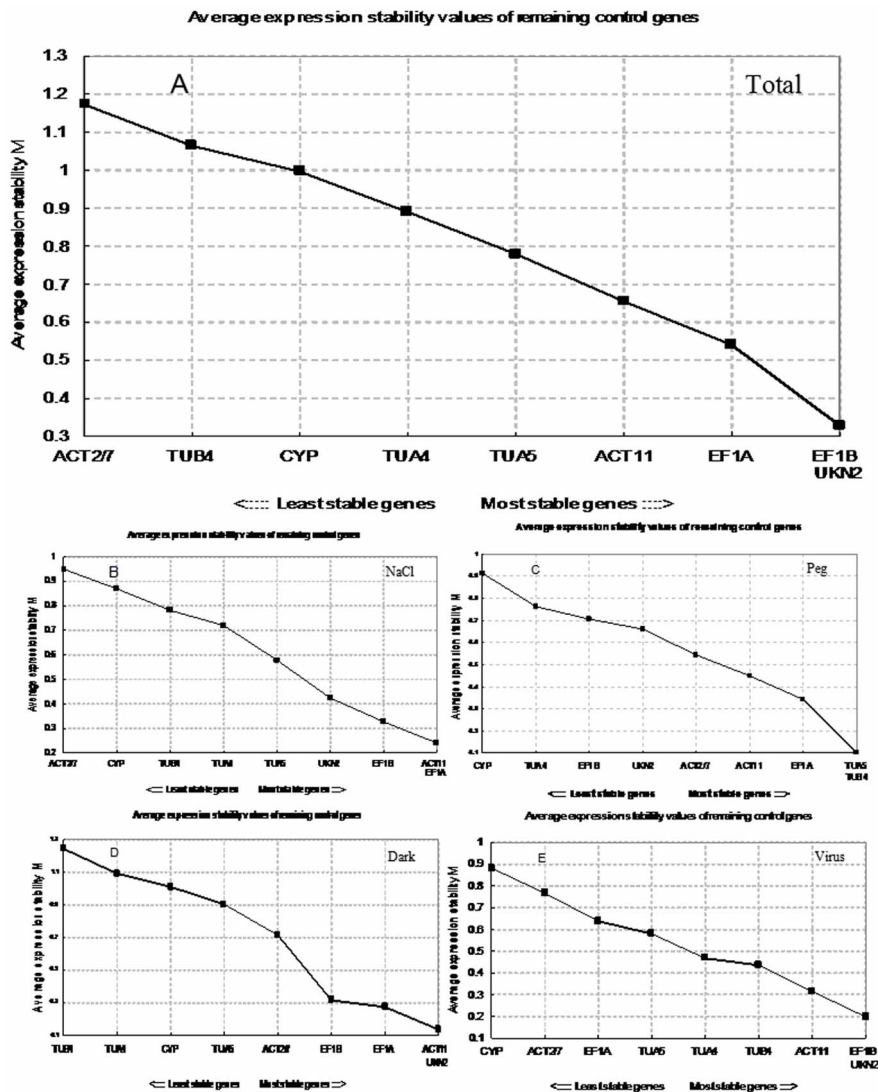
From Table 3, we found that the most stably expressed genes in all samples were *ACT11* and *EF1A*, the same as those obtained by GeNorm software, and *ACT2/7* was calculated to be the most

unstable expressed gene. Under salt treatment, *EF1A* and *EF1B* had the lowest value of  $M$  indicating that the two genes were the most stably expressed ones, and *ACT2/7* and *CYP* were the most unstable ones. Under drought treatment and virus infection, *TUB4* was the most stable expressed gene and *CYP* was the most unstable one. Under dark treatment, the most stably expressed genes were *ACT2/7* and *CYP*. Their  $M$  values were 0.327 and 0.518, respectively. *TUB4* was the most unstably expressed gene under this treatment. Overall, the results obtained through the two GeNorm and NormFinder softwares were not very consistent. It is possible that the inconsistent is due to the different calculation methods used between the two softwares.

### Discussion

The selection of suitable reference genes is critical in obtaining accurate results from RT-qPCR analysis. The expression stabilities of nine of the possible candidate genes, including *ACT11*, *TUA5*, *CYP*, *EF1B*, *TUA4*, *TUB4*, *EF1A*, *ACT2/7* and *UKN2*, were investigated under different stress conditions in a study reported in this paper. Our results showed that the most stably expressed genes were *EF1A* and *ACT11* under salinity treatment, were *TUB4*, *TUA5* and *EF1A* under drought treatment, were *ACT11* and *UKN2* under dark treatment, and were *EF1B* and *UKN2* under virus infection. Overall *EF1B* and *UKN2* gave the best expression stabilities under each of the four treatments assessed.

Studying its expression is a critical component in determining the function of a given gene in molecular biology and such studies help us to understand the growth and development of different species. RT-qPCR is the most commonly used method in such studies and the use of appropriate reference genes is essential in accurately determining the expression quantity of a given gene [31]. Ideal reference genes are those which give constant expression levels. However, such genes may not exist as plant growth is affected by environments. Different results can be obtained with the use of different reference genes and inaccurate assessment of gene expression could be obtained if suitable reference genes are not used. For example, Mafra et al. analysed the relative expression levels of *WRKY70* (transcription factor) in citrus challenged with fungal pathogens and found that, when



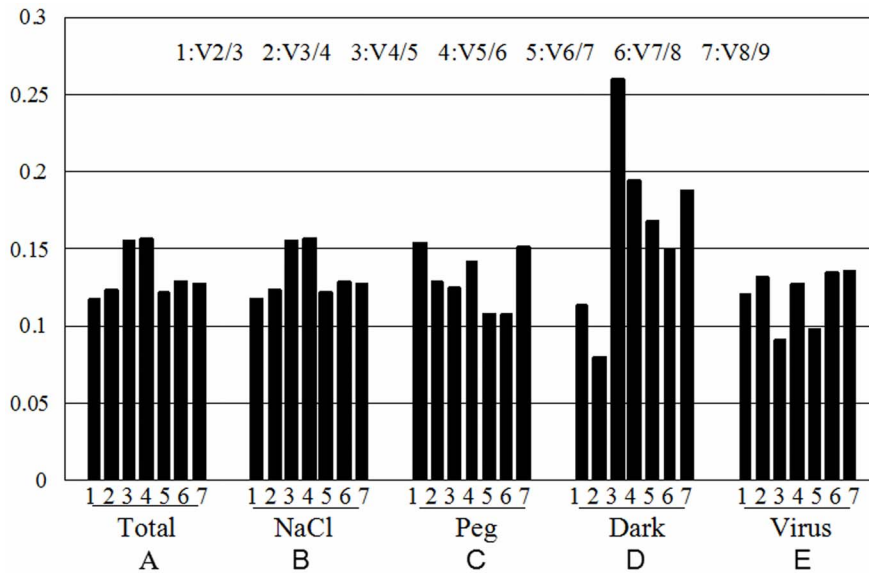
**Figure 2. Gene expression stability and ranking of 9 reference genes as calculated by GeNorm.** Expression stability and ranking of 9 reference genes calculated with GeNorm in all samples (A), NaCl-treated (B), PEG-treated (C), Dark-treated (D), Virus-treated (E); A lower average expression stability M value indicates more stable expression. doi:10.1371/journal.pone.0075271.g002

GeNorm was used, the use of the two most variable reference genes (*CYP* and *TUB*) resulted in an 42 folds increase in the relative transcript abundance of this gene [32]. However, the use of the three most stable reference genes (*DIMI*, *GAPC2* and *PTB1*) resulted in only a four folds increase in the relative transcript abundance of this gene. Similar results were also obtained by Le et al. in soybean shoots [33]. When *SUB12*, *60 s* and *Fbox* were used as reference genes, the relative transcript abundance of *GmNAC19* was 1.12, 18.66 and 13.93, respectively, when samples were taken at 10 h post dehydration stress. In an analysis of  $\beta$ -1,3-glucanase *BG1* in soybean leaves following virus infection, we found that the relative transcript abundance of this gene was different with the use of different reference genes and that its abundance also varied at different times of post inoculation (unpublished). When *TUA4* was used as the reference gene, the relative transcript abundance of *BG1* peaked at 48 h post inoculation (76 folds). However, when the most stable reference gene *UKN2* identified in the study was used, the relative abundance of *BG1* did not peak until 144 h post inoculation

when its abundance was 36.7 folds of that of the control (data unpublished). These data showed that the selection of suitable reference genes is critical in RT-qPCR analysis.

In the study, nine candidate reference genes (eight of them have been widely used) were analysed under four biotic and abiotic stresses. Results obtained from analysing the RT-qPCR data by using the GeNorm software showed that *UKN2* and *EF1B* were the most stably expressed reference genes, which was followed by *EF1A*. However, the most stably expressed reference gene was *ACT11*, which was followed by *EF1A* and *TUA4*, when the data were analyzed by using the NormFinder software. Both methods found that *ACT27* was the most unstable gene, thus we should avoid to use it as the reference gene in determining the expression quantity of a given gene through RT-qPCR analysis in soybean.

Samples from different stress treatments were also analysed in this study. The results showed that *EF1A* and *TUB4* were the most stably expressed genes under stresses of salinity and water, and *ACT27* and *CYP* were the most unstable ones under these conditions (Fig. 2; Table 3). Different results were obtained when



**Figure 3. Determination of the optimal number of reference genes for normalization by pairwise variation (V) using GeNorm.** The pairwise variation (V) to determine the optimal number of reference gene for accurate normalization in all samples (A), NaCl-treated (B), PEG-treated (C), Dark-treated (D), Virus-treated (E). It is the representative of the V2/3, V3/4, V4/5, V5/6, V6/7, V7/8, V8/9 form one to seven. doi:10.1371/journal.pone.0075271.g003

the dark treatment data were analyzed using the two different softwares. The most stable gene was *ACT2/7* when the data were analysed by NormFinder. However it was *UKN2* when GeNorm was used, although *TUB4* was found to be the most unstable gene using either of the two softwares.

Similarly, different results were also obtained in analyzing the data from the experiments of viral infection using the two different softwares, GeNorm and NormFinder. The most stable gene was

*TUB4* when the data were analyzed by NormFinder, but it was changed to be *UKN2* when GeNorm was used, although *CYP* was found to be the most unstable gene using either of the two softwares (Fig. 2; Table 3). These results showed that the nine candidate genes have different expression stabilities under different conditions of stresses, and using the two different softwares may lead to different selection of the best reference genes. Thus it is very important to selectively exploit different reference genes for different experiments by combining two or more software analyses.

There have numerous reports on the identification of suitable reference genes for RT-qPCR analysis. Jian [34] reported a study in which ten of the most widely used reference genes were analyzed using GeNorm. It was found that *EF1B* and *CYP2* were the most stable ones among all the samples examined; that *EF1B* and *CYP2* were the most stable ones across the different tissues at a given stage of plant development, that *ACT2/7* and *TUA* were the best ones across the different stages of plant development, and that *ACT11* and *EF1B* were the best ones across the different lighting periods examined. These results were slightly different from what we obtained in this paper, which is very likely due to the different materials and/or stress conditions used.

Hu et al. (2009) studied the expression stability of 14 reference genes under different conditions. Using GeNorm and NormFinder softwares, they found that *SKIP16*, *UKN1* and *UKN2* were the most stable ones among all of the samples examined [13]. Considering the fact that *UKN2* was also the best reference gene in our study, it is highly recommended that *UKN2* could be safely used for RT-qPCR analysis as a reference gene. There are many reports claiming that the expression stability of *ACT2/7* was poor. However, its expression stability was good in different tissues of *Arabidopsis thaliana* and *Platycladus orientalis* [17] and under low-temperature (cold) stress.

Ideal reference genes are those which give constant expression levels. However, such genes may not exist as plant growth is affected by environments. Different results can be obtained with the use of different reference genes and inaccurate assessment of gene expression could be resulted if suitable reference genes are not used. In this study, we investigated the expression stabilities of

**Table 3. Ranking of candidate reference genes in order of their expression stability as calculated by NormFinder.**

	Total(A)	NaCl(B)	Peg(C)	Dark(D)	Virus(E)
1	<i>ACT11</i>	<i>EF1A</i>	<i>TUB4</i>	<i>ACT2/7</i>	<i>TUB4</i>
M value	0.436	0.210	0.092	0.327	0.096
2	<i>EF1A</i>	<i>EF1B</i>	<i>ACT11</i>	<i>CYP</i>	<i>TUA4</i>
M value	0.436	0.269	0.139	0.518	0.212
3	<i>TUA4</i>	<i>ACT11</i>	<i>EF1A</i>	<i>EF1B</i>	<i>UKN2</i>
M value	0.475	0.325	0.200	0.528	0.301
4	<i>TUA5</i>	<i>UKN2</i>	<i>TUA5</i>	<i>EF1A</i>	<i>EF1B</i>
M value	0.495	0.427	0.210	0.534	0.322
5	<i>EF1B</i>	<i>TUA5</i>	<i>ACT2/7</i>	<i>TUA5</i>	<i>TUA5</i>
M value	0.505	0.435	0.452	0.594	0.349
6	<i>UKN2</i>	<i>TUA4</i>	<i>UKN2</i>	<i>TUA4</i>	<i>ACT11</i>
M value	0.506	0.494	0.551	0.693	0.357
7	<i>TUB4</i>	<i>TUB4</i>	<i>EF1B</i>	<i>UKN2</i>	<i>EF1A</i>
M value	0.653	0.525	0.583	0.698	0.454
8	<i>CYP</i>	<i>CYP</i>	<i>TUA4</i>	<i>ACT11</i>	<i>ACT2/7</i>
M value	0.742	0.749	0.594	0.728	0.795
9	<i>ACT2/7</i>	<i>ACT2/7</i>	<i>CYP</i>	<i>TUB4</i>	<i>CYP</i>
M value	0.950	0.777	0.935	1.155	0.835

doi:10.1371/journal.pone.0075271.t003

nine candidate genes under different stress conditions. Our results showed that the expression of each of these genes varied to some degree among the different treatments and none of them fits the definition of house-keeping genes. In order to get reliable results in gene expression in soybean, two or more reference genes need to be used in RT-qPCR analysis. In addition, we also showed that similar results were obtained from using both GeNorm and NormFinder softwares, and the top four reference genes detected among all of the samples by these two softwares were similar. These results demonstrate that the reference genes selected from this study were reliable and they form a solid base for conducting functional gene expression analysis in soybean.

## Supporting Information

**Figure S1 Dissociation curve data for the 9 reference genes.**  
(TIF)

**Figure S2 RT-qPCR amplification specificity of the 9 reference genes.** Amplification fragments were separated by 2% agarose gel electrophoresis.  
(TIF)

**Figure S3 RT-qPCR standard curve of the 9 reference genes.**  
(TIF)

## Acknowledgments

We thank Dr. Haijian Zhi (University of Nan Jing Agricultural University, China) for providing appropriate virus strains and advice on their propagation.

## Author Contributions

Conceived and designed the experiments: DW CH CL. Performed the experiments: SM HN. Analyzed the data: SM HN JZ DW CH. Wrote the paper: SM CL.

## References

- Guo Y, Chen SS, Guo WZ, Li J (2009) Advance in fluorescent quantitative PCR and its applications. *Prog Vet Med* 30 (2): 78–82 (in Chinese with English abstract).
- Suzuki T, Higgins PJ, Crawford DR (2000) Control selection for RNA quantitation. *Biotechniques* 29: 332–337.
- Bustin SA (2002) Quantification of mRNA using real time reverse transcription PCR (RT-PCR): trends and problems. *Journal of Molecular Endocrinology* 29: 23–39.
- Dhedea K, Huggett JF, Bustin SA, Johnson MA, Rook G, et al. (2004) Validation of housekeeping genes for normalizing RNA expression in real-time PCR. *Biotechniques* 37: 112–119.
- Kim BR, Nam HY, Kim SU, Kim SI, Chang YJ (2003) Normalization of reverse transcription quantitative-PCR with housekeeping genes in rice. *Biotechnology Lett* 25: 1869–1872.
- Andersen CL, Jensen JL, 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* 2004, 64: 5245–5250.
- Tricarico C, Pinzani P, Bianchi S, Paglierani M, Distante V, et al. (2002) Quantitative real-time reverse transcription polymerase chain reaction: normalization to rRNA or single housekeeping genes is inappropriate for human tissue biopsies. *Anal Biochemistry* 309 (2): 293–300.
- Radonic A, Thulke S, Mackay IM, Landt O, Siegert W, et al. (2004) Guideline to reference gene selection for quantitative real-time PCR. *Biochemical and Biophysical Research Communications* 313 (4): 856–862.
- Hong SM, Bahn SC, Lyu A, Jung HS, Ahn JH (2010) Identification and Testing of Superior Reference Genes for a Starting Pool of Transcript Normalization in Arabidopsis. *Plant and Cell Physiology* 51 (10): 1694–1706.
- Remans T, Smeets K, Opdenakker K, Mathijsen D, Vangronsveld J, et al. (2008) Normalisation of real-time RT-PCR gene expression measurements in Arabidopsis thaliana exposed to increased metal concentrations. *Planta* 227 (6): 1343–1349.
- Czechowski T, Stitt M, Altmann T, Udvardi MK, Scheible WR (2005) Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. *Plant Physiol* 139: 5–17.
- Chen X, Truksa M, Shah S, Weselake RJ (2010) A survey of quantitative real-time polymerase chain reaction internal reference genes for expression studies in Brassica napus. *Analytical Biochemistry* 405 (1): 138–140.
- Hu R, Fan C, Li H, Zhang Q, Fu Y-F (2009) Evaluation of putative reference genes for gene expression normalization in soybean by quantitative real-time RT-PCR. *BMC Molecular Biology* 10: 93.
- Die JV, Román B, Nadal S, González-Verdejo CI (2010) Evaluation of candidate reference genes for expression studies in Pisum sativum under different experimental conditions. *Planta* 232: 145–153.
- Li QF, Jiang MY, Yu HX, Xin SW, Gu MH, et al (2008) Selection of internal reference genes for quantitative RT-PCR analysis of total RNA from endosperm of rice (*Oryza sativa* L.). *Journal of Yangzhou University (agriculture and Life Science Edition)* 29(2): 61–66.
- Jain M, Nijhawan A, Tyagi AK, Khurana JP (2006) Validation of housekeeping genes as internal control for studying gene expression in rice by quantitative real-time PCR. *Biochemical and Biophysical Research Communications* 345 (2): 646–651.
- Chang E, Shi S, Liu J, Cheng T, Xue L, et al. (2012) Selection of Reference Genes for Quantitative Gene Expression Studies in *Platycladus orientalis* (Cupressaceae) Using Real-Time PCR. *PLOS ONE* 7 (3): e33278.
- Barsalobres-Cavallari CF, Severino FE, Maluf MP, Maia IG (2009) Identification of suitable internal control genes for expression studies in *Coffea arabica* under different experimental conditions. *BMC Mol Biol* 10: 1.
- Artico S, Nardeli SM, Brillhante O, Grossi-de-Sa MF, Alves-Ferreira M (2010) Identification and evaluation of new reference genes in *Gossypium hirsutum* for accurate normalization of real-time quantitative RT-PCR data. *BMC Plant Biol* 10: 49.
- Lovdal T, Lillo C (2009) Reference gene selection for quantitative real-time PCR normalization in tomato subjected to nitrogen, cold, and light stress. *Analytical Biochemistry* 387 (2): 238–242.
- Yang L, Pan A, Jia J, Ding J, Chen J, et al (2005) Validation of a tomato-specific gene, LAT52, used as an endogenous reference gene in qualitative and real-time quantitative PCR detection of transgenic tomatoes. *Agricultural and Food Chemistry* 5: 183–190.
- Luo HL, Chen SM, Wan HJ, Chen FD, Gu CS (2010) Candidate reference genes for gene expression studies in water lily. *Analytical and Bioanalytical Chemistry* 404 (1): 100–102.
- Brunner AM, Yakovlev IA, Strauss SH (2004) Validating internal controls for quantitative plant gene expression studies. *BMC Plant Biology* 4: 14.
- Burton RA, Shirley NJ, King BJ, Harvey AJ, Fincher GB (2004) The Cesa Gene Family of Barley. Quantitative Analysis of Transcripts Reveals Two Groups of Co-Expressed Genes. *Plant Physiology* 134 (1): 224–236.
- Volkov RA, Panchuk II, Schoffl F (2003) Heat-stress dependency and developmental modulation of gene expression: the potential of house-keeping genes as internal standards in mRNA expression profiling using real-time RT-PCR. *Journal of Experimental Botany* 54 (391): 2343–2349.
- Li WL, Zhao YS, Liu CJ, Yao GB, Wu SS, et al. (2012) Callose deposition at plasmodesmata is a critical factor in restricting the cell-to-cell movement of Soybean mosaic virus. *plant cell report* 31 (5): 905–916.
- GeNorm Software. Available: <http://medgen.ugent.be/~jvdesomp/geNorm>.
- NormFinder Software. Available: <http://www.mdll.dk/publication/snornfinder.html>.
- Bustin SA (2000) Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *J Mol Endocrinology* 25 (2): 169–193.
- Scharlaken B, de Graaf DC, Goossens K, Brunain M, Peelman LJ, et al. (2008) Reference gene selection for insect expression studies using quantitative real-time PCR: The head of the honeybee, *Apis mellifera*, after a bacterial challenge. *Insect Sci* 8 (33): 1–10.
- Hong SY, Seo PJ, Yang MS, Xiang FN, Park CM (2008) Exploring valid reference genes for gene expression studies in *Brachypodium distachyon* by real-time PCR. *BMC Plant Biol* 8: 112.
- Mafra V, Kubo KS, Alves-Ferreira M, Ribeiro-Alves M, Stuart RM, et al. (2012) Reference Genes for Accurate Transcript Normalization in Citrus Genotypes under Different Experimental Conditions. *PLOS ONE* 7 (2): e31263.
- Le DT, Aldrich DL, Valliyodan B, Watanabe Y, Ha CV, et al. (2012) Evaluation of Candidate Reference Genes for Normalization of Quantitative RT-PCR in Soybean Tissues under Various Abiotic Stress Conditions. *PLOS ONE* 7 (9): e46487.
- Jian B, Liu B, Bi YR, Hou WS, Wu CX, et al. (2008) Validation of internal control for gene expression study in soybean by quantitative real-time PCR. *BMC Mol Biol* 9: 59.