Plant Gene 1 (2015) 35–42

Contents lists available at ScienceDirect

Plant Gene

journal homepage: www.elsevier.com/locate/plant-gene

Cloning and validation of reference genes for normalization of gene expression studies in pearl millet [*Pennisetum glaucum* (L.) R. Br.] by quantitative real-time PCR

Palakolanu Sudhakar Reddy *, Dumbala Srinivas Reddy, Kiran K. Sharma, Pooja Bhatnagar-Mathur, Vincent Vadez

International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, Hyderabad 502 324, Telangana, India

ARTICLE INFO

Article history:
Received 20 November 2014
Received in revised form 20 January 2015
Accepted 5 February 2015
Available online 12 February 2015

Keywords:
Normalization
qRT-PCR
Pearl millet
Gene expression stability
Reference gene

ABSTRACT

To facilitate gene expression studies in pearl millet (*Pennisetum glaucum* (L.) R. Br.), the key reference genes including *ACP*, *ACT*, *TUB*, *CYP*, *EF-1α*, *EIF4A*, *GAPDH*, *MDH*, *PP2C*, *UBC* and *S24* were selected based on the available literature, and their expression stabilities were studied to determine their suitability for normalizing gene expression in pearl millet. Sequence information of the reference genes were obtained from the closely related species and cloned from pearl millet using homology based cloning strategy. Further, expression stabilities were validated for their accurate expression in different tissues, genotypes and abiotic stress treatments using three statistical algorithms including geNorm, NormFinder and RefFinder. Analysis showed that while the expression of *EF-1α* and *EIF4A* was most stable in different plant tissues, *MDH* and *EIF4A* were stable under different abiotic stress conditions. Amongst the different genotypes of pearl millet tested, while *UBC* and *MDH* genes exhibited most stable expression, *MDH* and *ACP* showed greater stability in all samples set. Interestingly, the widely used reference genes *S24* and *TUB* were found to be least stable across all the tested samples. Pair-wise analysis showed that two reference genes were sufficient for proper normalization, except when analyzing the gene expression studies in all samples set. Results of this study can help in the selection of reference genes for quantitative real time PCR (qRT-PCR) normalization in pearl millet that will contribute towards more accurate and reliable quantification of transcripts in this important crop of the drylands.

1. Introduction

Quantification of gene expression levels has been determined by techniques including Northern blotting, semi-quantitative RT-PCR, and quantitative real time PCR (qRT-PCR). However, qRT-PCR has gained more attention over the rest owing to its ability for real-time detection of PCR products, high sensitivity, specificity, rapidity, and accuracy (Valasek and Repa, 2005; Park et al., 2008). Other advantages of qRT-PCR analysis include its ability to detect low-abundance transcripts (Guenin et al., 2009), and their evaluation without physical verification through agarose gel electrophoresis, thereby resulting in reduced experiment time and increased throughput. Consequently, it is widely applied in a number of applications including biotechnology, microbiology, diagnosis (Bankowski and Anderson, 2004), genetic detection (van Doorn et al., 2007), allelic discrimination (Suda et al., 2003), and analysis of genetically modified organisms (Brodmann et al., 2002; Holst-Jensen et al., 2003). The qRT-PCR technology depends on two kinds of strategies i.e., absolute quantification and relative quantification. Absolute quantification can be achieved by comparing the quantitative cycle (Cq) values of the test samples to a standard curve, whereas relative quantification describes a real-time PCR experiment in which the gene of interest in one sample (i.e., treated) is compared to the same gene in another sample (i.e., untreated/control). Relative quantities obtained from both unknown sample and control must be normalized with reference genes in such a way that the data become biologically meaningful. Data normalization can be done through the use of an endogenous unregulated reference gene transcript (Czechowski et al., 2005). The ideal reference gene should express stably across the developmental stages and variable experimental conditions (Bustin, 2002; Crismani et al., 2006) and follow the guidelines of Minimum Information for Publication of qRT-PCR Experiments (MIQE) (Bustin, 2010).

Abbreviations: PP2C, Protein Phosphatase 2C; UBC, Ubiquitin-Conjugating Enzyme; MDH, Malate Dehydrogenase; GAPDH, Glyceraldehyde-3-Phosphate Dehydrogenase; EF-1α, Elongation Factor 1A; EIF4A, Eukaryotic Initiation Factor 4A; ACP, Acyl Carrier Protein; qRT-PCR, Quantitative Real Time-Polymerase Chain Reaction.

* Corresponding author. Tel.: +91 40 30713594.
E-mail address: p.sudhakarreddy@cgiar.org (P.S. Reddy).

http://dx.doi.org/10.1016/j.plgene.2015.02.001

Reference genes are typically constitutive genes that have housekeeping function required for the maintenance of basic cellular function, and are expressed stably in all cells of an organism under normal and adverse conditions (Villasenor et al., 2011). Most frequently used reference genes in plant gene expression studies include *ACT*, *GAPDH*, *TUB*, *UBQ* and *18s*, *25s rRNA* (Czechowski et al., 2005; Jian et al., 2008; Cordoba et al., 2011). If expression of the reference gene is altered by the experimental conditions, the results obtained may be incorrect and lead to misinterpretation (Dheda et al., 2005; Gutierrez et al., 2008). Hence, the reference genes that have stable expression levels across tissues, developmental stages and conditions would represent the best system for normalization of the qRT-PCR data. Selection of such type of reference genes across the genotypes in the specific crop would be an extra advantage. Unfortunately, there are no universal reference gene(s) that are expressed at a constant level across species, experimental conditions and tissues (Czechowski et al., 2005; Die et al., 2010). Therefore, the use of suitable reference genes for normalization is essential for performing the qRT-PCR experiments.

Over the years, several statistical approaches such as geNorm (Vandesompele et al., 2002), NormFinder (Andersen et al., 2004) and Best-Keeper (Pfaffl et al., 2004) have been developed to assess the expression stability for an accurate selection of reference genes to validate gene expression. Extensive studies have been carried out to investigate the stability of reference genes in different abiotic stress conditions, tissues, and development stages of plant species such as *Arabidopsis* (Czechowski et al., 2005), *Oryza sativa* (Jain et al., 2006), *Solanum lycopersicum* (Exposito-Rodriguez et al., 2008), *Triticum aestivum* (Long et al., 2010), *Zea mays* (Manoli et al., 2012), *Brassica juncea* (Chandna et al., 2012), *Arachis hypogaea* (Reddy et al., 2013), *Glycine max* (Nakayama et al., 2014), *Populus euphratica* (Wang et al., 2014), *Brassica napus* (Yang et al., 2014), *Pennisetum glaucum* (Saha and Blumwald, 2014) and *Panicum virgatum* (Gimeno et al., 2014). The present study evaluated the expression stability of the eleven candidate reference genes in a wide range of variability covered in the samples which includes three sets of samples. Hence, the identification of reference genes, expressing stably across the environmental conditions, would provide the baseline for further gene expression studies in a dryland crop like pearl millet where a major emphasis is on its crop improvement for various abiotic constraints and nutritional enhancement.

Pearl millet [*P. glaucum* (L.) R. Br.], an annual diploid crop belonging to the *Poaceae*, is the fifth most important cereal crop of semi-arid agricultural regions. It is a highly drought tolerant crop that can sustain at marginal land agro-ecosystems with higher temperature and low rainfall (200–600 mm) that is not suitable for other cereals. As pearl millet grows well in the semi-arid climates, there is high probability that it contains a large number of stress alleviating genes that provide stress adaptation to this crop. The limited amount of genome sequence information in pearl millet has restricted the progress in gene discovery, characterization, transcript profiling, and their application in crop improvement programs. Gene expression analysis by qRT-PCR in pearl millet is limited due to the lack of sequence information about genes, which can serve as suitable reference genes. In this study, we have cloned and evaluated eleven candidate reference genes from pearl millet including *PP2C* (*Protein Phosphatase 2C*), *UBC* (*Ubiquitin-Conjugating Enzyme*), *MDH* (*Malate Dehydrogenase*), *TUB* (*β-Tubulin*), *GAPDH* (*Glyceraldeyde-3-Phosphate Dehydrogenase*), *EF-1α* (*Elongation Factor 1-alpha*), *EIF4A* (*Eukaryotic Initiation Factor 4A*), *ACT* (*β-Actin*), *ACP* (*Acyl Carrier Protein*), *CYP* (*Cyclophilin*) and *S24* (*40S Ribosomal Protein*) that were studied in a wide range of tissues. Expression stability was assayed in diverse samples of pearl millet to study their suitability as best stable reference genes using geNorm, NormFinder and RefFinder tools. This study provides useful guidelines and benefit for the suitable reference gene selection for future gene expression studies in pearl millet and other related species of millet.

2. Materials and methods

2.1. Plant material and treatments

Pearl millet cultivar HHB67 was obtained from the Pearl millet Breeding Unit of ICRISAT, India. Pearl millet plants were grown in 8 inch pots with 4.0 kg of alfisol mixed with sand and compost mixture (3:2:1; 20% water holding capacity) under greenhouse conditions with 28 °C/20 °C day/night temperature. Different abiotic stress treatments (drought, salt, ABA, cold and heat) were subjected to the plants at 28 DAS. Drought stress was imposed by withholding the water for 10 days followed by leaf sample collection, while the salinity stress was imposed by bringing the pots containing plants to field capacity with a 250 mM NaCl solution, followed by leaf sampling 24 h after treatment. For ABA treatment, the seedlings were sprayed with 100 μM ABA solution, while for cold and heat stress treatments, the seedlings were kept at 4 °C and 45 °C, respectively for 4 h prior to tissue sampling. The plants were grown under normal conditions at 28 ± 1 °C, served as control. Different tissue samples including seedlings, leaf, panicle, seed and roots were collected under normal growth conditions at different growth stages. Five pearl millet cultivars including HHB67, H77, PRLT2/89-33, HRC1078 and HRC1086 were used for tissue sample collection under normal growth conditions. All experimental samples were divided into three subsets based on their nature. The organ and tissue subset comprised root, leaf, seedling, panicle and seed under normal growth conditions (five samples), the genotypes subset comprised five pearl millet cultivars including HHB67, H77, PRLT2/89-33, HRC1078 and HRC1086 (Five samples), and the abiotic stress subset comprised drought, salt, heat, low temperature and ABA stresses (Five samples). All samples set consist of all three subsets including different tissue, genotypes and abiotic stress subsets. Samples were collected from three different plants to provide as biological replicates. All the collected tissue samples were snap frozen in liquid nitrogen and stored at −80 °C until further use.

2.2. Total RNA isolation

Total RNA was isolated using NucleoSpin RNA plant kit (Macherey-Nagel, Duren, Germany) according to the manufacturer's instructions. RNA concentrations and purity were determined using a NanoVue plus spectrophotometer (GE health care, USA). RNA integrity was verified by electrophoresis in 1.4% agarose gel. Total RNA isolated from different tissues were diluted to 100 ng/μl concentrations and used for the qRT-PCR experiments.

2.3. Sequence retrieving and cloning of the candidate reference genes

While the pearl millet genome sequence has not been yet released, the sequence information of the proposed reference genes was not readily available in databases like NCBI and EBI. Hence, the available EST database was used for retrieving of the sequences of *PP2C*, *UBC*, *MDH*, *TUB*, *GAPDH*, and *EF-1α*. Other genes including *EIF4A*, *ACT*, *ACP*, *CYP* and *S24* were cloned using the homology based cloning method. Degenerate primers were designed based on the homologous genes of other closely related plant species (Supplementary Table 1). The cDNA was synthesized from 2 μg of total RNA using a SuperScript III first strand synthesis kit (Invitrogen, USA) and used as a template for the cloning of candidate reference genes using respective gene specific degenerate primers (Supplementary Table 1). PCR was carried out using 400 nM of each primer along with 200 μM of each dNTPs, 2U of *Taq* DNA polymerase and 2 μl of cDNA as template in a 50-μl reaction. PCR reaction conditions were as follows: 3 min denaturation at 94 °C, followed by 30 cycles of a denaturation step at 94 °C, annealing step at 59–63 °C and an extension step at 72 °C. Each step was 1 min long and the final extension step was at 72 °C for 10 min. The PCR amplified

fragments were cloned into the pCR4.0-TOPO vector (Invitrogen) according to the manufacturer's protocol and sequenced completely.

2.4. qRT-PCR primer designing

The qRT-PCR primers for eleven candidate genes were designed based on the available expressed sequence tags (EST) database and sequenced fragments (Table 1). The primers were designed using the Primer 3 Plus software (Untergasser et al., 2007) with the following parameters: product size of 90–170 base pairs, melting temperature (Tm) of 60–63 °C, length of 19–24 nucleotides and GC content of 45–55%. In order to verify the specificity of the reference genes, all the amplicons of eleven candidate genes after qRT-PCR were sequenced. Amplification efficiencies for all the primer pairs were calculated using a two-fold dilution series of pooled cDNA over at least five dilution points and measured in triplicate. The amplification efficiency (E) and correlation coefficients (R^2) of the primers were calculated from the slope of the line ($E = 10^{-1/\text{slope}} - 1$).

2.5. Quantitative real-time PCR

All the qRT-PCR reactions were carried out on a Realplex (Eppendorf, Germany), in 96-well optical reaction plates. Reactions were performed in a total volume of 10 μl, containing 1 μl of RNA (100 ng), 400 nM of each primer, 5 μl of 2× one step SYBR RT-PCR buffer 4 (Takara, Japan) and 0.4 μl of prime script one step Enzyme Mix 2 (Takara, Japan) and made to 10 μl with RNase-free H_2O. The qRT-PCR cycling conditions were as follows: 42 °C for 5 min and 95 °C for 10 s (reverse transcription) followed by 40 cycles of 15 s at 95 °C, 15 s at 62 °C with fluorescent signal recording and 15 s at 72 °C. The melting curve analysis was included after 40 cycles to verify the primer specificity by heating from 58 °C to 95 °C with fluorescence measured within 20 min. All the samples were collected from the three independent plants and repeated the qRT-PCR experiment three times.

2.6. Data analysis

Expression levels of the eleven candidate reference genes were determined by threshold cycle (Cq) values which reflect the number of PCR cycles required for the amplification related fluorescent signal to cross a specific threshold line above the background level. The stability and suitability of eleven candidate reference genes were evaluated using three independent algorithms, i.e., geNorm (Vandesompele et al., 2002) NormFinder (Andersen et al., 2004) and RefFinder. Cq values were imported into the geNorm software program and the expression stability value (M) for each candidate gene was calculated and then the pair-wise variation (V) of this gene was compared with all other candidate reference genes. Larger stability value is considered to have less expression stability and smaller suggesting more stability. geNorm was used to calculate the pairwise variation ($V_n / V_n + 1$) for finding the optimal number of the genes required for the gene expression studies. The cut-off value for M is 0.15 (Vandesompele et al., 2002). For the NormFinder, Cq values were converted into relative quantities after correcting the PCR efficiencies. NormFinder ranks the stability of the candidate reference genes independently from each other. According to NormFinder analysis, the lowest stability value is the most stable gene expression within the gene set studied. The RefFinder is a web-based (http://www.leonxie.com/referencegene.php) tool, which integrates the currently available four major computational programs [geNorm (Vandesompele et al., 2002), NormFinder (Andersen et al., 2004), BestKeeper (Pfaffl et al., 2004) and comparative ΔCt method (Silver et al., 2006)] and calculates the geometric mean for the comprehensive ranking.

3. Results

3.1. Quality control

Total RNA samples isolated from the 45 tissues (including biological replicates) exhibited a high quality of RNA. The mean A260/280 ratio of the RNA samples was 2.01 ± 0.045 (range from 1.95 to 2.12) and reflected pure and protein-free RNA. To avoid erroneous results, only RNA samples with high quality were included in this study. The amplification efficiencies (E) and correlation coefficients (R^2) of the eleven candidate genes were generated using the slopes of the standard curves obtained by serial dilutions. The amplification efficiencies (E) of the reference genes ranged from 0.94 (*ACP* and *EF-1α*) to 1.01 (*CYP*). The correlation coefficient (R^2) values varied from 0.972 (*CYP*) to 0.998

Table 1
Details of candidate reference genes, their primer sequences, product size and amplicon characteristics.

S. No	Gene symbol	Gene name	Accession no.	Cellular function	Primers (F/R) (5′-3′)	Amplicon length (bp)	Tm (°C)	PCR efficiency	Regression coefficient (R^2)
1	PP2C	Protein phosphatase 2C	CD725527	Signal transduction	AGGCTCAGGAGAAGGTGCT GAGAACCTGG GAAGCTGGAC	127	88.6	0.97	0.998
2	UBC	Ubiquitin-conjugating enzyme	CD724586	Protein degradation	TTCAAACCTCCGAAGGTGTCTT GGCTCCACTGCTCTTTAAGAATG	100	80.1	0.98	0.992
3	MDH	Malate dehydrogenase	CD724779	Citric acid cycle and gluconeogenesis	AGAAGGCGCTTGCTTACTCAT CAGTTCTGGGTGAGGGAATCT	118	82.8	0.95	0.995
4	TUB	β-Tubulin	KM105955	Cytoskeleton structure protein	TGGTGTCACATGCTGCCTGA GGGTCAGCTC AGGGACAGTG	167	88.3	0.97	0.996
5	GAPDH	Glyceraldeyde-3-phosphate dehydrogenase	GQ398107	Glycolysis and gluconeogenesis	TGCCTTGCTCCCCTTGCTAA CAGCCCTTCCAC CTCTCCAG	139	85.2	1.00	0.997
6	EF-1α	Elongation factor 1-alpha	EF694165	Translation eukaryotic factor	AATGATCCGCTGCTGTAACAAG AGGCAATCTTGTCTGGGTTGTA	128	83.3	0.94	0.993
7	EIF4A	Eukaryotic initiation factor 4A	EU856535	Initiation phase of eukaryotic translation	ATCGTGAGCTTTACATCCATCG TATCCCTCAGGATACGGATGTC	105	85.3	0.97	0.997
8	ACT	β-Actin	KM105957	Cytoskeleton structure protein	TGTGAGCCATACCGTGCCAA GGCAGTGGTG GTGAAGGAGT	139	85.1	0.98	0.996
9	ACP	Acyl carrier protein	KM105958	Fatty acid and polyketide biosynthesis	ACAGTGTCCAACGAGTCAGCA AGGAGACGGTGGATCAGGTTTG	118	85.1	0.94	0.994
10	CYP	Cyclophilin	KM105959	Immunosuppressant, protein folding	TACAAGGGGTCGAGCTTCCAC TTCTCGCCGTAGATGGACTCC	104	88.5	1.01	0.972
11	S24	40S ribosomal protein	KM105960	Ribosomal protein	CCCCAGGAAGTGCTCTGCTA CATCAGCGTC ACCCTGAGCA	158	84.5	0.96	0.998

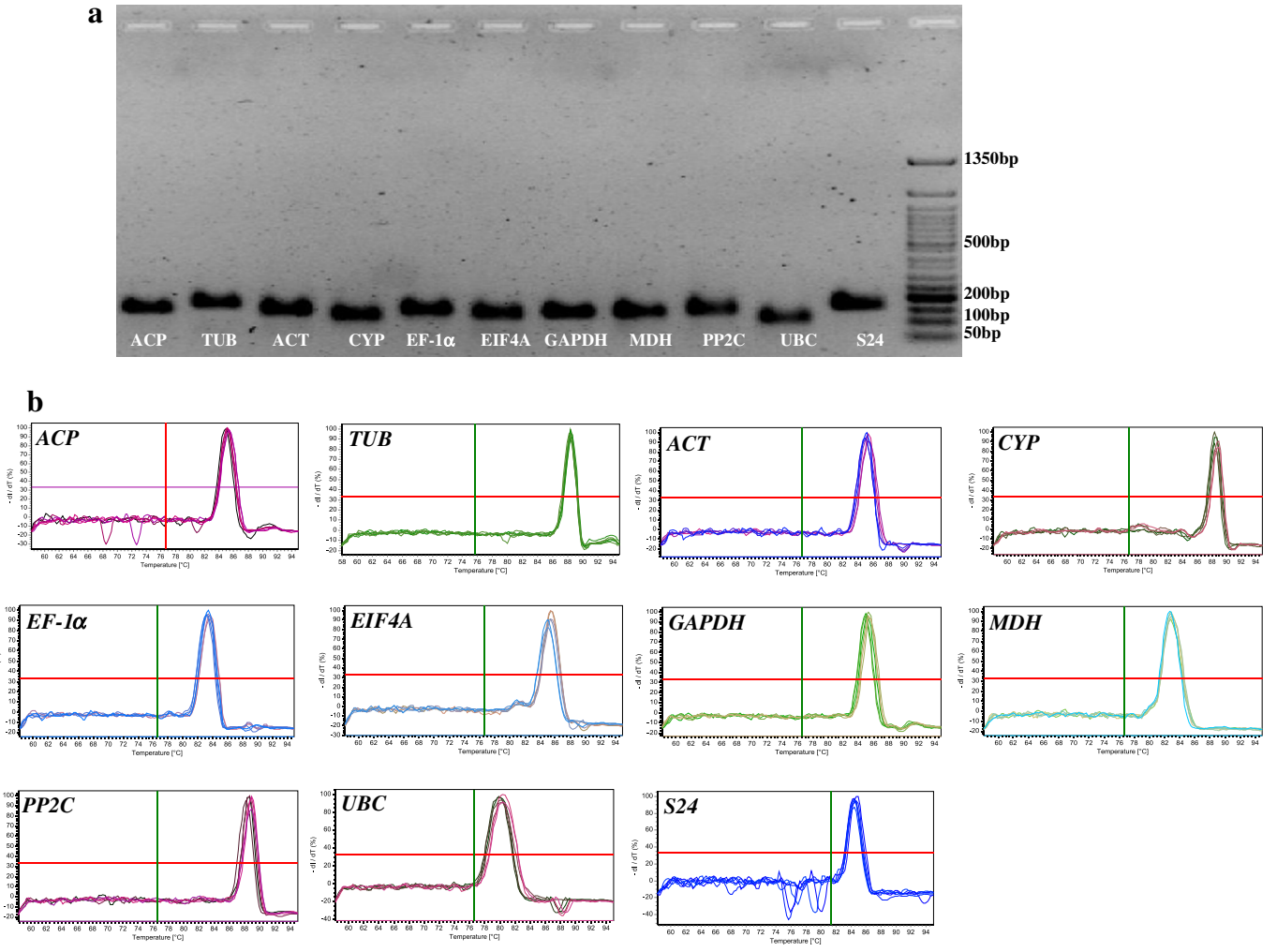

Fig. 1. Specificity of qRT-PCR amplicons. a 1.8% agarose gel electrophoresis showing amplification of a single product of the expected size for each reference gene. M represents 50 bp DNA Ladder. b Dissociation curves with single peaks generated from all amplicons.

(*S24* and *PP2C*). These results show that all primer pairs were suitable for qRT-PCR analysis (Table 1). Specificity of the primers was confirmed by agarose gel electrophoresis and melt curve analysis (Fig. 1a and b). The melting temperatures of all amplification products are listed in Table 1. These provided the good basis for the further analysis of validating reference genes.

3.2. Cloning of candidate reference genes and Cq variation

Eleven reference genes were selected for the present study, i.e., *ACP*, *ACT*, *TUB*, *CYP*, *EF-1α*, *EIF4A*, *GAPDH*, *MDH*, *PP2C*, *UBC* and *S24*. However, since, the sequences of *ACT*, *ACP*, *CYP*, *EIF4A* and *S24* reference genes were not available, homology based gene-cloning approach was used for cloning of these genes. Therefore, we cloned and sequenced partial sequences of these five reference genes, which have been used as reference genes in other plant species. Detailed information about the candidate reference genes, including their GenBank accession number, description, and function were listed in the Table 1. PCR and qRT-PCR amplifications were carried out with specific primers based on cDNA sequences of the eleven reference genes. The Cq value of each candidate reference gene in 45 different pearl millet samples was used to compare the range of expression levels. Cq value for the eleven candidates showed a wide range of variation, i.e., 14.74 (*TUB*) to 27.24 (*S24*) (Fig. 2). Majority of the Cq values were ranged between 18.12 and 22.03. However, the average Cq value for the each candidate reference gene did not show much variance (18.26 for *GAPDH* to 24.56 for *PP2C*) (Supplementary Table 2). The coefficient of variation (CV) values calculated for each reference gene indicated; lower gene expression variation in MDH and EIF4A (CV: 3.06 and 3.75 respectively) when compared to the rest and TUB showed highest variation in expression (CV: 11.26) (Supplementary Table 2, Fig. 2). The remaining candidate reference genes occupied the intermediate positions. This result shows that none of the candidate reference genes had constant expression levels across the pearl millet samples. Therefore, there is a need to validate the expression stabilities of these candidate reference genes to normalize the gene expression studies under specific conditions for selecting the suitable stable reference gene.

3.3. Expression stability of the candidate reference genes by geNorm analysis

Expression stabilities of the eleven candidate reference genes were validated using geNorm software (Vandesompele et al., 2002). According to this software, the gene with the lowest M value is considered as the most stable, while the M values > 1.5 are considered as unacceptable levels of expression variability. To determine the suitable stable reference genes under specific experimental conditions, the 14 samples were divided into four different subsets as described in the Materials

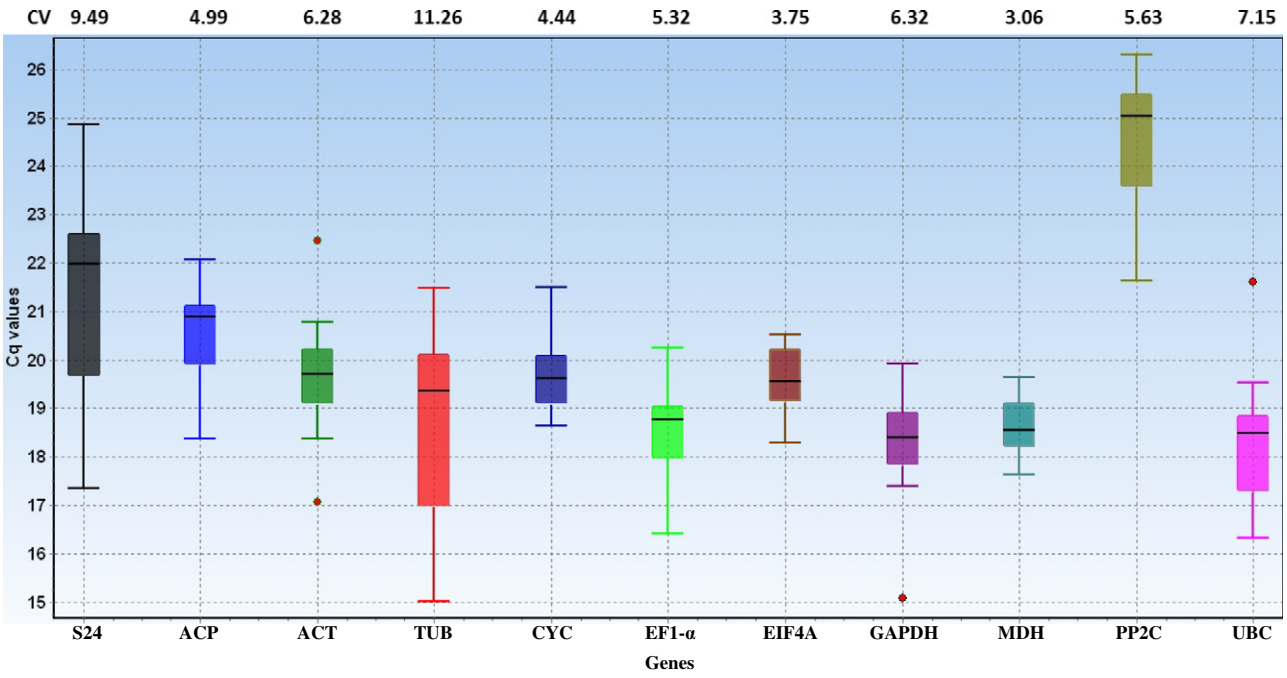

Fig. 2. Expression levels of candidate reference genes across all samples: Lines across the boxes depict the medians. Boxes indicate the interquartile range. Whiskers represent 95% confidence intervals; red dot indicate the presence of outliers. Coefficient of variance (CV) of each gene among all samples is given in percentage.

and methods section. In each sample set, the eleven reference genes were ranked from the most stable to the least stable (Fig. 3, Table 2). Under different abiotic stress conditions all the eleven reference genes showed the high expression stabilities with the threshold value of below 1.5. *PP2C* (M = 0.71) was the least stable gene and *MDH* and *EF-1α* (M = 0.24) were identified as the best pair of reference genes.

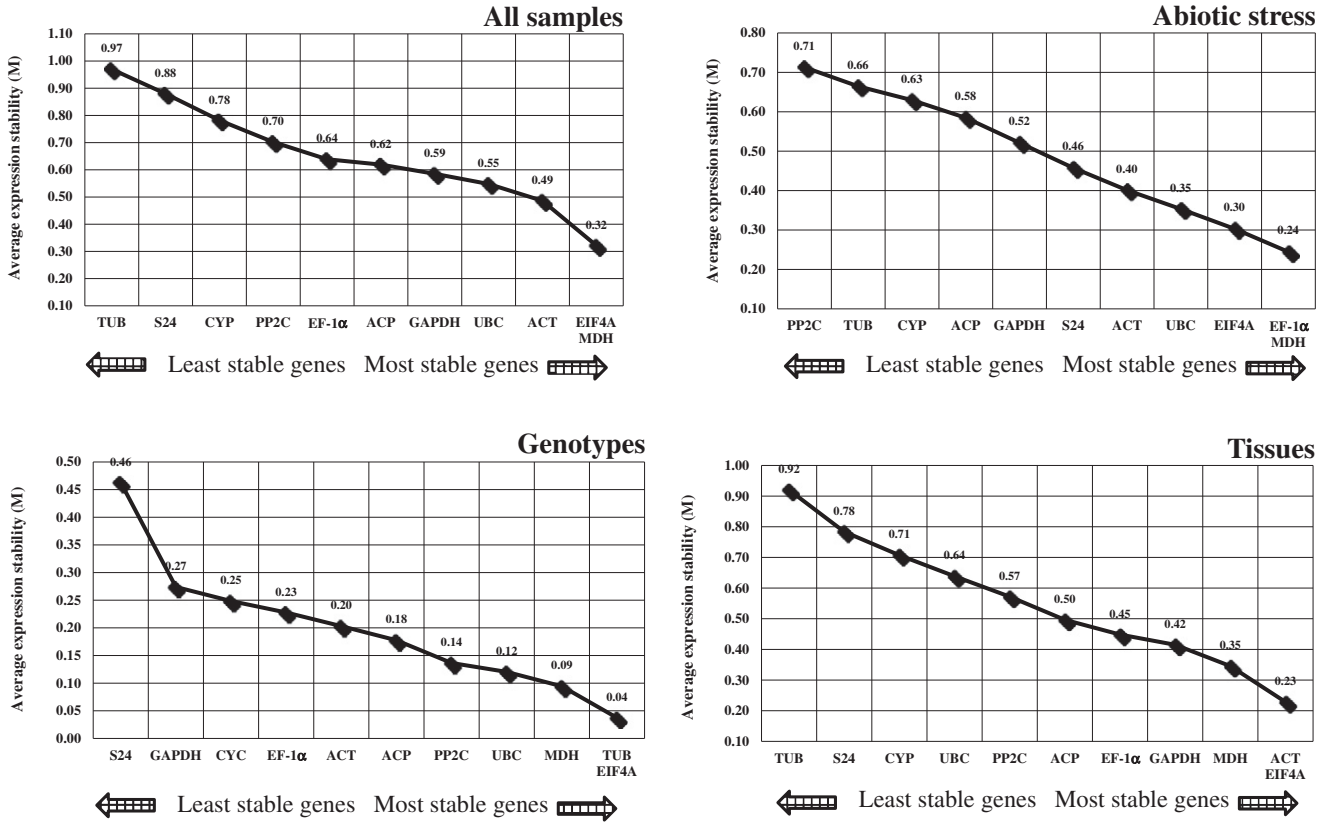

Fig. 3. Gene expression stability and rankings of eleven candidate reference genes using geNorm software. The average expression stability (M) was calculated following stepwise exclusion of the least stable gene across all the samples within an experimental set, the lowest M value indicates the most stable gene, while the highest value represents the most variable gene.

Table 2
Expression stability ranks of 11 candidate reference genes in different sets of pearl millet samples calculated using geNorm (GN) and NormFinder (NF) methods.

	All samples		Abiotic stress		Genotypes		Tissue	
Gene	GN	NF	GN	NF	GN	NF	GN	NF
MDH	1	3	1	1	3	1	3	7
EIF4A	2	4	3	4	1	6	1	2
ACT	3	7	5	8	7	8	2	4
UBC	4	6	4	10	4	2	8	8
GAPDH	5	5	7	3	10	10	4	5
ACP	6	1	8	5	6	7	6	1
EF-1α	7	2	2	2	8	3	5	3
PP2C	8	8	11	11	5	5	7	6
CYP	9	9	9	7	9	9	9	9
S24	10	10	6	6	11	11	10	10
TUB	11	11	10	9	2	4	11	11

Similarly, reference genes with different genotypes showed a lesser variation since all the M values were far below the acceptable limit, where *S24* (M = 0.46) and *GAPDH* (0.27) was the least stable, while *EIF4A* and *TUB* (0.082) were the stably expressed genes. In a sample set of different tissues, geNorm identified *EIF4A* and *ACT* (M = 0.23) as the best-ranked candidate reference genes, whereas *TUB* was the least stable gene (M = 0.92). In the category of all samples set, *MDH* (0.43) and *EIF4A* (0.32) were identified as the best pair of reference genes and *TUB* and *S24* (0.97 and 0.88) were considered as the least stable genes (Fig. 3 and Table 2).

3.4. Optimal number of internal candidate genes for normalization

To generate accurate and reliable results, two or more genes are required for qRT-PCR experiment. Therefore, we used geNorm algorithm to find the optimal number of suitable reference genes required for proper normalization by step wise calculation of the pairwise variation ($V_n / V_n + 1$) between two sequential normalization factors (NFs), which measures the effect of increasing reference genes required for normalization. geNorm software uses 0.15 as the cutoff value below which the inclusion of an additional reference gene is not required (Vandesompele et al., 2002). In our analysis, for the all the sample sets except the all sample set, V2 / 3 value was below 0.15 (the recommended cut-off value), which suggested that two reference genes would be sufficient for qRT-PCR analysis related experiments in pearl millet (Fig. 4). This suggests that the optimal number of reference genes for normalization in the different category of pearl millet tissues was two and addition of the third reference gene showed no significant effect on normalization of gene expression (Fig. 4). However, in all samples set, the pairwise variation V2 / 3 value was higher than 0.15 (0.184), which indicated that two reference genes were not sufficient for normalization, and hence it needs a third gene.

3.5. NormFinder analysis

NormFinder is another excel-based statistical algorithm for finding the suitable reference gene among a set of candidate references and experimental conditions. This approach ranks the set of reference genes according to their stability value under a given experimental design. The results of gene expression stability analysis are shown in Table 3. The results shown with NormFinder was similar to the one obtained from geNorm to some extent (Table 2). In the case of all the experimental results and different tissue category subsets, *ACP* showed the lowest stability value (0.31) and *TUB* showed the highest stability value (1.3) (Table 2). This meant that *ACP* was the best stably expressed gene under this category. In the genotype subset, *MDH* and *UBC* had the lowest stability values (0.067 and 0.082, respectively) while *S24* showed the highest stability value (1.6). *MDH* and *EF-1α* (0.31 and 0.32) showed the lowest stability values, which suggested that these two genes were considered as the most stable genes under abiotic stress conditions. These results agreed with the results of geNorm (Fig. 3 and Table 2). Overall, NormFinder predicts that the reference genes *ACP*, *MDH* and *GAPDH* were the most stable genes in all sample set, whereas *CYP*, *PP2C* and *S24* showed the least stable genes in the four groups, similar to the results obtained by geNorm analysis.

3.6. RefFinder analysis

The RefFinder analysis revealed that *UBC* and *MDH* were most stable genes in genotype samples subset; similarly *EF-1α* and *EIF4A* were ranked as top in different tissue sample subset. While, *MDH* and *EIF4A* genes were the most stable genes under abiotic stress conditions, whereas *MDH* and *ACP* were observed to be highly stable in all samples

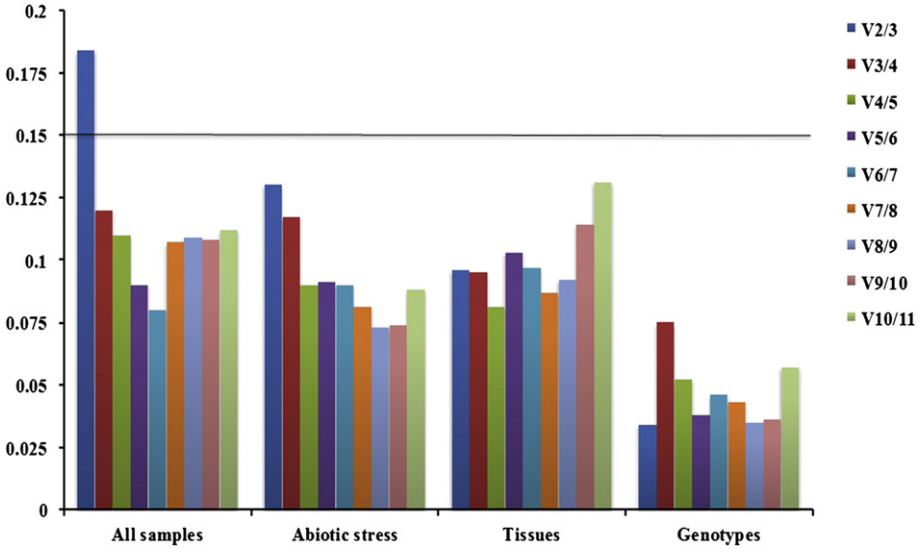

Fig. 4. Determination of the optimal number of reference genes for geNorm analysis. The pairwise variation ($V_n / V_n + 1$) was analyzed for the normalization factors NF*n* and NF*n* + 1 by geNorm program to determine (V < 0.15) the optimal number of reference genes.

Table 3
Expression stability values and ranking for pearl millet candidate reference genes calculated using NormFinder, in all samples, abiotic stress treated samples, tissues and genotypes sample sets. Lower average expression stability value indicates more stable expression.

Rank	All samples	Abiotic Stress	Genotypes	Tissues
1	ACP (0.311)	MDH (0.3153)	MDH (0.0667)	ACP (0.2529)
2	EF-1α (0.3402)	EF-1α (0.3213)	UBC (0.0818)	EIF4A (0.2532)
3	MDH (0.448)	GAPDH (0.4077)	EF-1α (0.1203)	EF-1α (0.262)
4	EIF4A (0.4872)	EIF4A (0.4104)	TUB (0.122)	ACT (0.4032)
5	GAPDH (0.4876)	ACP (0.48)	PP2C (0.131)	GAPDH (0.4558)
6	UBC (0.5149)	S24 (0.5137)	EIF4A (0.1462)	PP2C (0.5345)
7	ACT (0.5983)	CYP (0.5747)	ACP (0.1703)	MDH (0.5776)
8	PP2C (0.6888)	ACT (0.5917)	ACT (0.2762)	UBC (0.6896)
9	CYP (1.1223)	TUB (0.6165)	CYP (0.2777)	CYP (1.0169)
10	S24 (1.1358)	UBC (0.6173)	GAPDH (0.4643)	S24 (1.0257)
11	TUB (1.2706)	PP2C (0.8386)	S24 (1.592)	TUB (1.5177)

subset. The comprehensive ranking also revealed that the *S24* gene was the least stable gene in all the sample subsets. Other candidate genes such as *TUB* in different samples and in all samples subsets; *ACT* in abiotic stress and *CYC* in genotypes sample subset were also found as the least stable genes in this analysis (Table 4).

4. Discussion

Gene expression patterns are important determinant for understanding the biological processes during developmental and environmental stresses. Several methods have been employed to determine gene expression levels. The most widely used technique is qRT-PCR, which is a potent, accurate, simple and sensitive tool for detecting gene expression levels. For the qRT-PCR analysis, selection of suitable internal controls is very important to obtain reliable, proper and accurate data. The ideal reference genes should have relatively stable expression, irrespective of the nature of the sample (Crismani et al., 2006). Widely used reference genes such as *ACT*, *TUB* and *EF-1α* are used as internal controls regardless of their variance and instability in the expression levels under different experimental conditions (Cordoba et al., 2011). Therefore, for accurate normalization, qRT-PCR data must be normalized with one or more suitable and stable reference genes. Hence, validation is extremely important for gene expression studies in pearl millet to avoid the unnecessary errors in qRT-PCR analysis. To evaluate the best set of candidate reference genes for different individual samples, tissues, genotypes and abiotic stress conditions in pearl millet, three statistical algorithms, geNorm, NormFinder and RefFinder were used for finding the expression stability of eleven candidate reference genes. Due to distinct statistical calculations, some inconsistency is expected between the two methods (Table 3). The top four stable genes were almost similar in the two approaches of each sample subset, but some differences were found in their ranking order.

In the present study, we cloned and tested the expression stabilities of eleven commonly used reference genes, i.e., *ACP*, *ACT*, *TUB*, *CYP*, *EF-1α*, *EIF4A*, *GAPDH*, *MDH*, *PP2C*, *UBC*, and *S24* that have been previously reported as the most stable candidate reference genes in different plant species. After determination of primer amplification efficiencies, eleven candidate transcripts were selected for evaluation of their normalization potential in diverse samples of pearl millet, which was broadly divided into four experimental subsets. In all sample subset, *MDH*, *EIF4A* and *ACT* were identified as the top three reference genes using geNorm, while *MDH* was ranked third, *EIF4A* was fourth and *ACT* was seventh using NormFinder. According to RefFinder, *MDH* and *ACP* were suggested as the most stable genes among the eleven reference genes under the category of the all sample subset. The *MDH* and *EIF4A* would be the best optimal reference gene pair for the gene expression study under different abiotic stresses. For the genotypes subset, the top three optimal reference genes were *EIF4A*, *TUB* and *MDH* as calculated by geNorm, whereas NormFinder identified *MDH*, *UBC* and *EF-1α* as the best set of reference genes. RefFinder suggested that *UBC* and *MDH* would be the best pair for the gene expression studies in the different genotypes of pearl millet (Table 4). In the case of different tissues subset, based on the RefFinder results, we can conclude that the *EF-1α* and *EIF4A* pair would be the best for gene expression study in the different tissues of pearl millet (Table 4). Among all the tested reference genes, *S24* holds a specific attention that was ranked as the least stable gene in both algorithms geNorm and NormFinder. Therefore, the use of *S24* as reference gene should be avoided in the qRT-PCR related experiments in pearl millet in the future. Previous study in pearl millet suggested that *PP2A*, *TIP41* and *UBC2* were most stable under diverse experimental conditions (Saha and Blumwald, 2014), but in the present study we found the *EIF4A* and two new candidate reference genes, which were not validated in the previous study such as MDH and ACP, were found to be the most stable genes. Our results were well fitted with the studies in the *Theobroma cacao*, where *MDH*, *GAPDH*, and *ACP* were the most stable genes in various tissues (Pinheiro et al., 2011).

Many studies have shown that the application of more than one reference gene would provide greater accuracy in the qRT-PCR experiments in the plants (Vandesompele et al., 2002; Le et al., 2012). geNorm software determines the optimal number of candidate reference genes for normalization of gene expression data by calculating the pairwise variation between two sequential normalization factors containing an increased number of reference genes. In this study, three out of four experimental sets, except the all samples subset, showed pairwise variation V2/3 values below 0.15, which indicates that combination of two-reference gene was enough for the optimal normalization. This suggests that addition of the third candidate reference gene was necessary to normalize gene expression in all samples set of pearl millet.

In conclusion, our results suggested that different set of reference genes should be applied according to the experimental condition. *MDH* and *ACP* were the most stable in all the samples, *MDH* and *EIF4A* for abiotic stress subsets, *UBC* and *MDH* genes for the genotypes, and *EF-1α* and *EIF4A* for the different tissues subsets. Hence, these genes

Table 4
Expression stability ranks of 11 candidate reference genes in different sets of pearl millet samples calculated using RefFinder tool.

	All samples		Abiotic stress		Genotypes		Tissues	
Rank	Genes	Geomean of ranking values	Genes	Geomean of ranking values	Genes	Geomean of ranking values	Genes	Geomean of ranking values
1	MDH	1.57	MDH	1.63	UBC	1.93	EF-1α	2.21
2	ACP	2.45	EIF4A	2.21	MDH	2.06	EIF4A	2.21
3	EIF4A	2.78	ACP	3.6	EIF4A	3.31	ACP	3.22
4	EF-1α	4.6	EF-1α	4.12	TUB	4.36	MDH	3.66
5	GAPDH	4.73	GAPDH	4.61	ACP	5.03	ACT	4.23
6	CYP	5.2	PP2C	6.04	GAPDH	5.2	GAPDH	5.14
7	ACT	5.45	CYP	6.24	EF-1α	5.38	CYP	5.2
8	UBC	5.63	TUB	6.4	ACT	5.86	UBC	5.86
9	PP2C	8.24	UBC	7.14	PP2C	7.67	PP2C	7.97
10	TUB	10.24	ACT	7.93	CYP	8.41	S24	10
11	S24	10.74	S24	10	S24	11	TUB	11

could be recommended as optimal reference genes in qRT-PCR analysis-related experiments of pearl millet. Since *S24*, *PP2C* and *TUB* showed poor stabilities in all the experimental sample groups, they couldn't be used as reference genes for gene expression normalization. The present study provides useful guidelines for the selection of candidate reference genes for future gene expression studies in the pearl millet and other close relatives of the millets.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.plgene.2015.02.001.

Acknowledgments

This work was supported in part through a fellowship and research grant under the INSPIRE Faculty Fellowship Award Grant No. IFA-LSPA-06 and Young Scientist Scheme SB/YS/LS-12/2013 by the Department of Science and Technology, Government of India. This work has been undertaken as part of the CGIAR Research Program on Dryland Cereals. ICRISAT is a member of CGIAR Consortium.

References

Andersen, C.L., Jensen, J.L., Orntoft, 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.

Bankowski, M.J., Anderson, S.M., 2004. Real-time nucleic acid amplification in clinical microbiology. Clin. Microbiol. Newsl. 26, 9–15.

Brodmann, P.D., Ilg, E.C., Berthoud, H., Herrmann, A., 2002. Real-time quantitative polymerase chain reaction methods for four genetically modified maize varieties and maize DNA content in food. J. AOAC Int. 85, 646–653.

Bustin, S.A., 2002. Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. J. Mol. Endocrinol. 29, 23–39.

Bustin, S.A., 2010. Why the need for qPCR publication guidelines?—the case for MIQE. Methods 50, 217–226.

Chandna, R., Augustine, R., Bisht, N.C., 2012. Evaluation of candidate reference genes for gene expression normalization in *Brassica juncea* using real time quantitative RT-PCR. PLoS One 7, e36918.

Cordoba, E.M., Die, J.V., Gonzalez-Verdejo, C.I., Nadal, S., Roman, B., 2011. Selection of reference genes in *Hedysarum coronarium* under various stresses and stages of development. Anal. Biochem. 409, 236–243.

Crismani, W., Baumann, U., Sutton, T., Shirley, N., Webster, T., Spangenberg, G., Langridge, P., Able, J.A., 2006. Microarray expression analysis of meiosis and microsporogenesis in hexaploid bread wheat. BMC Genomics 7, 267.

Czechowski, T., Stitt, M., Altmann, T., Udvardi, M.K., Scheible, W.R., 2005. Genome-wide identification and testing of superior reference genes for transcript normalization in *Arabidopsis*. Plant Physiol. 139, 5–17.

Dheda, K., Huggett, J.F., Chang, J.S., Kim, L.U., Bustin, S.A., Johnson, M.A., Rook, G.A., Zumla, A., 2005. The implications of using an inappropriate reference gene for real-time reverse transcription PCR data normalization. Anal. Biochem. 344, 141–143.

Die, J.V., Roman, B., Nadal, S., Gonzalez-Verdejo, C.I., 2010. Evaluation of candidate reference genes for expression studies in *Pisum sativum* under different experimental conditions. Planta 232, 145–153.

Exposito-Rodriguez, M., Borges, A.A., Borges-Perez, A., Perez, J.A., 2008. Selection of internal control genes for quantitative real-time RT-PCR studies during tomato development process. BMC Plant Biol. 8, 131.

Gimeno, J., Eattock, N., Van Deynze, A., Blumwald, E., 2014. Selection and validation of reference genes for gene expression analysis in switchgrass (*Panicum virgatum*) using quantitative real-time RT-PCR. PLoS One 9, e91474.

Guenin, S., Mauriat, M., Pelloux, J., Van Wuytswinkel, O., Bellini, C., Gutierrez, L., 2009. Normalization of qRT-PCR data: the necessity of adopting a systematic, experimental conditions-specific, validation of references. J. Exp. Bot. 60, 487–493.

Gutierrez, L., Mauriat, M., Guenin, S., Pelloux, J., Lefebvre, J.F., Louvet, R., Rusterucci, C., Moritz, T., Guerineau, F., Bellini, C., Van Wuytswinkel, O., 2008. The lack of a systematic validation of reference genes: a serious pitfall undervalued in reverse transcription-polymerase chain reaction (RT-PCR) analysis in plants. Plant Biotechnol. J. 6, 609–618.

Holst-Jensen, A., Ronning, S.B., Lovseth, A., Berdal, K.G., 2003. PCR technology for screening and quantification of genetically modified organisms (GMOs). Anal. Bioanal. Chem. 375, 985–993.

Jain, M., Nijhawan, A., Tyagi, A.K., Khurana, J.P., 2006. Validation of housekeeping genes as internal control for studying gene expression in rice by quantitative real-time PCR. Biochem. Biophys. Res. Commun. 345, 646–651.

Jian, B., Liu, B., Bi, Y., Hou, W., Wu, C., Han, T., 2008. Validation of internal control for gene expression study in soybean by quantitative real-time PCR. BMC Mol. Biol. 9, 59.

Le, D.T., Aldrich, D.L., Valliyodan, B., Watanabe, Y., Ha, C.V., Nishiyama, R., Guttikonda, S.K., Quach, T.N., Gutierrez-Gonzalez, J.J., Tran, L.S., Nguyen, H.T., 2012. Evaluation of candidate reference genes for normalization of quantitative RT-PCR in soybean tissues under various abiotic stress conditions. PLoS One 7, e46487.

Long, X.Y., Wang, J.R., Ouellet, T., Rocheleau, H., Wei, Y.M., Pu, Z.E., Jiang, Q.T., Lan, X.J., Zheng, Y.L., 2010. Genome-wide identification and evaluation of novel internal control genes for Q-PCR based transcript normalization in wheat. Plant Mol. Biol. 74, 307–311.

Manoli, A., Sturaro, A., Trevisan, S., Quaggiotti, S., Nonis, A., 2012. Evaluation of candidate reference genes for qPCR in maize. J. Plant Physiol. 169, 807–815.

Nakayama, T.J., Rodrigues, F.A., Neumaier, N., Marcelino-Guimaraes, F.C., Farias, J.R., de Oliveira, M.C., Borem, A., de Oliveira, A.C., Emygdio, B.M., Nepomuceno, A.L., 2014. Reference genes for quantitative real-time polymerase chain reaction studies in soybean plants under hypoxic conditions. Genet. Mol. Res. 13, 860–871.

Park, Y., Kim, J., Choi, J.R., Song, J., Chung, J.S., Lee, K.A., 2008. Evaluation of multiplex PCR assay using dual priming oligonucleotide system for detection mutation in the Duchenne muscular dystrophy gene. Korean J. Lab. Med. 28, 386–391.

Pfaffl, M.W., Tichopad, A., Prgomet, C., Neuvians, T.P., 2004. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper—Excel-based tool using pair-wise correlations. Biotechnol. Lett. 26, 509–515.

Pinheiro, T.T., Litholdo Jr., C.G., Sereno, M.L., Leal Jr., G.A., Albuquerque, P.S., Figueira, A., 2011. Establishing references for gene expression analyses by RT-qPCR in *Theobroma cacao* tissues. Genet. Mol. Res. 10, 3291–3305.

Reddy, D.S., Bhatnagar-Mathur, P., Cindhuri, K.S., Sharma, K.K., 2013. Evaluation and validation of reference genes for normalization of quantitative real-time PCR based gene expression studies in peanut. PLoS One 8, e78555.

Saha, P., Blumwald, E., 2014. Assessing reference genes for accurate transcript normalization using quantitative real-time PCR in pearl millet [*Pennisetum glaucum* (L.) R. Br]. PLoS One 9, e106308.

Silver, N., Best, S., Jiang, J., Thein, S.L., 2006. Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR. BMC Mol. Biol. 7, 33.

Suda, T., Katoh, M., Hiratsuka, M., Fujiwara, M., Irizawa, Y., Oshimura, M., 2003. Use of real-time RT-PCR for the detection of allelic expression of an imprinted gene. Int. J. Mol. Med. 12, 243–246.

Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Geurts, R., Leunissen, J.A., 2007. Primer3Plus, an enhanced web interface to Primer3. Nucleic Acids Res. 35, W71–W74.

Valasek, M.A., Repa, J.J., 2005. The power of real-time PCR. Adv. Physiol. Educ. 29, 151–159.

van Doorn, R., Szemes, M., Bonants, P., Kowalchuk, G.A., Salles, J.F., Ortenberg, E., Schoen, C.D., 2007. Quantitative multiplex detection of plant pathogens using a novel ligation probe-based system coupled with universal, high-throughput real-time PCR on OpenArrays. BMC Genomics 8, 276.

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

Villasenor, T., Brom, S., Davalos, A., Lozano, L., Romero, D., Los Santos, A.G., 2011. Housekeeping genes essential for pantothenate biosynthesis are plasmid-encoded in *Rhizobium etli* and *Rhizobium leguminosarum*. BMC Microbiol. 11, 66.

Wang, H.L., Chen, J., Tian, Q., Wang, S., Xia, X., Yin, W., 2014. Identification and validation of reference genes for *Populus euphratica* gene expression analysis during abiotic stresses by quantitative real-time PCR. Physiol. Plant. 152, 529–545.

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