Australian Journal of Crop Science AJCS
AJCS 9(7):646-655 (2015) ISSN:1835-2707

Identification of stable reference genes for quantitative PCR in jute under different experimental conditions: An essential assessment for gene expression analysis

Ahlan Sabah Ferdous[†], Md. Tariqul Islam[†], Salma Sultana Alam, Haseena Khan[*]

Molecular Biology lab, Department of Biochemistry and Molecular Biology, University of Dhaka-1000, Bangladesh
[†]These authors contributed equally.

***Corresponding author: haseena@du.ac.bd**

Abstract

A reputation for accuracy and rapidity has made Quantitative reverse transcription PCR (qRT-PCR) a popular method for gene expression analysis. However, the precision of qRT-PCR critically relies on stable control genes for data normalization. To obtain consistent data, validation of internal controls is vital. We examined seven candidate genes of jute (*ACT*, *ELF*, *TUB*, *UBC*, *U6*, *G3PDH*, and *18S*) under three abiotic (salinity, dehydration, and low temperature) and one biotic (fungal infection) stress conditions. Their stabilities were evaluated using four distinct algorithms (comparative deltaCT, geNorm, NormFinder, and BestKeeper) all integrated in a web-based tool, RefFinder. Our data suggest that *ELF* and *UBC* are most homogenously expressed under all three abiotic stress conditions, whereas *TUB* is most stable under fungal infestation. Contrasting expression was observed for *G3PDH* and *ELF*, the former being unsuitable for all abiotic conditions but was ranked second highest for biotic. The reverse was true for *ELF*. For validation of our data, expression of a cold responsive gene, *CBF-1* was assessed under low temperature using the most and least stable reference genes. Striking differences in *CBF-1* expression illustrate the accuracy of the reference genes. This set of genes will be useful for quantitative analyses of genes in jute.

Keywords: Reference gene, qRT-PCR, jute, biotic stress, abiotic stress.
Abbreviation: ACT_actin, ELF_elongation factor-alpha, TUB_beta-tubulin, UBC_ubiquitin conjugating enzyme, U6_U6-snRNA, G3PDH_glyceraldehydes-3-phosphate dehydrogenase, 18S_18S rRNA, and CBF-1_C-repeat binding factor-1.

Introduction

With the advancement of molecular biology-based research, analysis of gene expression has become progressively more important in furthering the understanding of pathways that underlie developmental and cellular processes. To assess changes in gene expression, reverse transcription followed by quantitative polymerase chain reaction (qRT-PCR) is one of the most widely used techniques in recent years (Tong et al., 2009). Popularity of qRT-PCR is burgeoning owing to its high sensitivity and specificity, speed, ease of use, and requirement of a relatively low amount of RNA (Paolacci et al., 2009). The qRT-PCR method reflects the need to accurately quantify mRNA levels in the fields of molecular medicine, biotechnology, microbiology, and molecular diagnostics (Bustin and Dorudi, 1998) as well as in plants (Gachon etal., 2004). Nevertheless, to ensure reproducible and accurate measurements of transcript abundance, this sensitive technique should be performed by following some golden rules (Udvardi et al., 2008). The central problem in exact gene transcription analysis is at the same time one of the benefits of highly accurate qRT-PCR: the precise determination of amplifiable template nucleic acid present in the reaction (Radonić et al., 2004; Guénin et al., 2009). Quantification of gene expression is affected by several factors, including experimental sources of variation (de Almeida et al., 2010). Several experimental sources of variation exist in qRT-PCR, such as sample-to-sample variation in RNA integrity, variability of extraction protocols that may co-purify inhibitors, differences in reverse efficiency, and the amount of cDNA template used in each reaction (Huggett et al., 2005; Libault et al., 2008; de Almeida et al., 2010). Consequently, normalization of the qRT-PCR data is essential to minimize experimental deviations or errors that inevitably occur during sample preparation procedures and data analysis, all of which make quantitation of gene transcripts unreliable (Hong et al., 2008). There are many strategies that can be chosen for normalization; these include normalization to sample size, total RNA, and the popular practice of measuring an internal reference or housekeeping gene (Huggett et al., 2005). Normalization to total RNA content poses a number of problems because it is difficult to quantify small amounts of RNA, and variation in RT and PCR reaction efficiencies are not accounted for by this method (Brunner et al., 2004). Similarly, normalization to an external RNA standard is challenging because of the instability of RNA (Brunner et al., 2004). A number of strategies have been proposed to normalize qRT-PCR data; however, normalization remains one of the most important challenges concerning this technique (Huggett et al., 2005). The most commonly applied approach for normalization for qRT-PCR is the use of one or more reference gene(s), which should be expressed at a constant level across various conditions, and its expression is assumed to be unaffected by experimental parameters (Radonićet al., 2004; Huggett et al., 2005). Normalization of the expression of a target gene against a stably expressed internal gene can compensate for most variations and result in a relative quantification of levels of gene expression across samples. Moreover, correct and accurate sample

normalization is crucial to reveal small but significant differences in expression when comparing samples collected under different conditions (Guénin et al., 2009). Therefore, the accuracy of the results obtained by qRT-PCR strongly depends on the choice of one or preferably more reference genes that are stably expressed across all samples (Gutierrez et al., 2008). Cellular homeostasis genes, more commonly known as housekeeping genes, include actin (ACT), ubiquitin (UBQ)glucose-6-phosphate dehydrogenase (GAPDH), ribosomal genes (18S, 28S rRNA genes), cycophilin (Cyp), elongation factor alpha (ELF alpha), tubulin (TUB); they are involved in basic and ubiquitous cellular processes (Reddy et al., 2013) and are used as reference genes for gene expression studies in many plant species (Kim et al., 2003; Lee et al., 2010; Mallona et al., 2010). However, recent studies indicate that the traditional housekeeping genes are not always stably expressed when tested in other species or in a wider range of experimental treatments. This means that the most stable reference gene(s) should be identified for a specific species under study or in a new experimental set-up (Iskandar et al., 2004). With a goal of aiding this normalization process, a number of statistical analysis strategies, such as geNorm (Vandesompele et al., 2002), NormFinder (Andersen et al., 2004), and BestKeeper (Pfaffl et al., 2004), have been proposed for the evaluation of reference gene expression stability and selection of the most suitable reference gene under different experimental conditions (Jain et al., 2006). Until recently, several stable reference genes have been reported under different conditions, including for arabidopsis (Czechowski et al., 2005), tobacco (Schmidt and Delaney, 2010), rice (Kim et al., 2003; Jain et al., 2006), soybean (Libault et al., 2008; Hu et al., 2009), tomato (Expósito-Rodríguez et al., 2008), and wheat (Paolacci et al., 2009). For gene expression studies in jute, stability of the reference genes needs to be verified prior to use in qRT-PCR. The direct use of traditional and recently proposed novel candidate reference genes to non-model plants, such as jute, is hampered by the limited availability of genomic sequences. Thus, a series of candidate reference genes were selected for which such sequence information was available for other plant species. Jute is an annual dicotyledonous fiber-yielding plant with immense economic importance. It is a natural fiber, well known for its high tensile strength and biodegradability (Ahmed et al., 2011). In terms of usage, production, and global consumption, it is second only to cotton (Roy et al., 2006). Vital physiological processes of jute, such as germination, growth, development, reproduction, senescence and death, are greatly influenced by environmental conditions (Sharmin et al., 2011). Some other key factors affecting the production of jute fall into two broad categories, namely abiotic and biotic stresses. In the former category, the major challenges include low yield under unfavorable growth conditions, such as salinity, drought, flood, or cold, and the latter includes susceptibility to insect pests, fungal and viral diseases, etc. (Keka et al., 2008). In this study, seven potential reference genesactin (ACT), elongation factor alpha (ELF), glyceraldehydes-3-phosphate dehydrogenase (G3PDH), beta-tubulin (TUB), ubiquitin conjugating enzyme (UBC), U6 snRNA (U6), and 18S rRNA (18S)—were selected and evaluated in search of a suitable reference gene to be used for quantitative data analysis in jute under four different environmental conditions, including three abiotic stressors (salinity, dehydration, and low temperature) and fungal infestation as a biotic stressor. Estimation of their stability was calculated by four different popular algorithms (comparative deltaCT method, geNorm, NormFinder, and BestKeeper) all integrated together in a single web tool, RefFinder under the Cotton EST Database (Xie et al., 2011).

Results

Selection of reference genes and testing of primer efficiency

A total of seven candidate reference genes and, to validate their stability, a low temperature responsive gene were retrieved from closely related species (data not shown) by using the most conserved region among them for designing degenerate primers. Such primers were then used to amplify the desired regions. Gel extraction and sequencing were performed next to confirm their presence and to design jute-specific primers for qRT-PCR. Additionally, a primer pair was designed by targeting exon-exon junctions and used to compare amplicons from cDNA templates with the amplicons derived from genomic DNA. Because introns are eliminated during mRNA processing, amplicon sizes were much smaller than that obtained from genomic DNA (S-1). This comparison indicates that the cDNAs were free of genomic DNA contamination, which was further validated by no amplification in NRT (no RT) controls during real time PCR. Both agarose gel electrophoresis (Fig 1) and single peaks in real time PCR amplifications (S-2) confirmed the specificity of the primers, and their efficiencies were substantiated with linear regression (R^2) and efficiency calculation (E values) (Table 1, S-3). All the primers met the criteria with R^2 values greater than 0.980 and efficiency ranging from 94%–101%.

Expression analysis of the selected reference genes

Expression patterns of the selected reference genes were analyzed at the seedling stage under different abiotic (low temperature, dehydration, and salinity) and biotic (fungal) stresses at five different time intervals from 0 to 48 hours of stress induction. An equal amount of RNA followed by an equal amount of cDNA from biological duplicates and technical triplicates were used for amplification by quantitative PCR in order to avoid any experimental deviations (S-4). Values of quantification cycle (Cq) across all samples used to generate a relative amplification distribution curve (Fig 2) revealed that ELF and UBC are the most stable genes under different abiotic stress conditions, with Cq values ranging from 16.077–26.308 and 19.163–26.666, respectively. Although U6 has a relatively stable expression, it is dispersed critically by the minimum and maximum Cq values. The quantification cycle also infers the differences in the levels of transcripts, suggesting that expressions of ELF and 18S are relatively high with lower Cq value (average 20.13124 and 21.70161, respectively) amid all genes, whereas ACT and TUB have the lowest expression level. The wide ranges of Cq values even in the most stable genes confer discrepancy in expression according to the variation of stress conditions. Thus, it indicates that no single gene is stable throughout the stresses, suggesting an urgency to validate the apt reference gene(s) under different experimental conditions.

Data analysis for stability of the candidate reference genes

Evaluation of gene expression stability and selection of suitable reference genes for the normalization purpose was performed using RefFinder, an online tool integrated with

Table 1. Details of jute-specific primer sequences used for validation.

Gene symbol	Primers sequence(5'-3')	Tm (°C)	Primer size(bp)	Amplicon size(bp)	Amplification efficiency (%)	R^2
ACT	TGGTATTGGATGTTGGAGAT	57	20	83	94.6	0.9953
	GGTATTGACTTAATGCTGCT		20			
ELF	CTTGAGGCTCTTGACCAGAT	61	20	105	100.1	0.9967
	AACAGGGACAGTTCCGATAC		20			
TUB	CTCCAACCGGCTTGAAGATG	63	20	139	96.5	0.9872
	CCTCTCCAGTGTACCAATGC		20			
UBC	GGTGTTGCTCTCTATCTGCT	61.3	20	135	100.2	0.9874
	CATGGCATACTTCTGGGTCCA		21			
U6	GGACCATTTCTCGATTTGTACGTG	62	24	60	101.3	0.9981
	TACAGATAAGATTAGCATGGCCCC		24			
18S	ACTACGTCCCTGCCCTTTGTA	62	21	170	97.4	0.9834
	CACCTACGGAAACCTTGTTACG		24			
G3PDH	CACTTGAAGGGTGGTGCTAAG	62	21	146	95.4	0.99
	GGAGCAAGGCAGTTAGTAGTG		21			
CBF1	GGTCAGCTTGTTTGAACTTC	59	20	70	97.78	0.989
	TCCTTTGGATCACTAGAAGC		20			

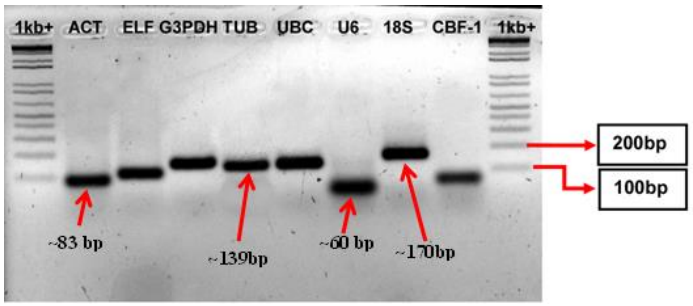


Fig 1. Confirmation of the presence of candidate reference genes in jute (1kb+: 1Kb+ ladder, ACT: actin, ELF: elongation factor alpha, G3PDH: glyceraldehyde 3 phosphate dehydrogenase, TUB: tubulin, UBC: ubiquitin conjugating enzyme, U6: U6 snRNA, 18S: 18S rRNA, CBF-1: (C-repeat binding factor-1).

most popular major statistical programs (geNorm, NormFinder, Bestkeeper, and deltaCT method). Samples were grouped according to the stress conditions, and stable reference genes were determined individually by analyzing the Cq values from that group of data produced by performing qRT-PCR. Based on the rankings from each program, RefFinder assigns an appropriate weight to an individual gene and calculates the geometric mean of their weights for the overall final ranking (http://www.leonxie.com/referencegene.php?type=reference) (Table 2). Under a low temperature stress condition, recommended comprehensive ranking determined UBC as the most stable gene and ELF as the second most, whereas U6 was the least stable. In contrast, under dehydration and salinity stress conditions, the most stable gene was ELF, and G3PDH was the least stable. It was found that UBC was the second most stable gene under both dehydration and salinity stresses. Thus, if the two most stable genes are used for expression analyses, in the case of dehydration, salinity, and low temperature treatment, ELF and UBC should be used. Interestingly, under fungal stress, TUB and G3PDH were the most stable ones, and ELF was least stable. This implies the significance of studying the stability of reference genes under stress-specific conditions. Based on RefFinder analyses, the overall stability of the genes under different stress conditions is shown in Fig 3.

Determination of the optimal number of reference gene(s) for normalization

Pairwise variation among the candidate genes is calculated in order to determine the number of genes to be included to produce valid data. High pairwise variation indicates the need to take into account more reference gene to determine a corroborative result. Ranking of stability recommends the consecutive order of genes according to the consistency of their expression. However, it is important to determine the optimal number of reference gene(s) for normalization; high pairwise variation among the genes may cause an impaired expression profile if an optimal number of reference genes is not included. Based on the formula described by geNorm (Vandesompele et al., 2002), pairwise variations of seven candidate genes under each stress condition were analyzed. For the two most stable genes, the normalization factor (NF_n), which is the geometric mean of those genes, was calculated, and then stepwise inclusion of other control genes were made until the n + 1^{th} gene had no significant contribution to the newly calculated normalization factor (NF_{n+1}). To determine the possible need or utility of including more than two genes, pairwise variation ($V_{n/n+1}$) was calculated between two sequential normalization factors (NF_n and NF_{n+1}) for all genes under different stress conditions (Fig 4). In this study pair wise variation was calculated with the ranking of genes recommended by comprehensive ranking. Vandesompele et al. (2002) suggested that pair wise variation of more than 0.15 needs to be added to this value between two sequential genes for normalization. Our study found $V_{2/3}$ for all of the four stress conditions to be <0.15, implying that no further inclusion of genes was needed.

Validation of the selected reference genes

Expression pattern of a transcription factor, CBF1/DREB1b (CRT binding factor or DRE binding protein), involved in

Table 2. Ranking of candidate genes according to their expression stability (DC: deltaCT, BK: BestKeeper, NF: NormFinder, GN: geNorm, RCR: recommended comprehensive ranking).

Stressors	Low temperature					Dehydration				
Rank	DC	BK	NF	GN	RCR	DC	BK	NF	GN	RCR
1	UBC	UBC	ELF	UBC | ELF	UBC	UBC	ELF	UBC	ELF TUB	ELF
2	ELF	ELF	UBC		ELF	ELF	U6	ELF		UBC
3	TUB	TUB	TUB	TUB	TUB	ACT	UBC	ACT	UBC	TUB
4	18S	18S	18S	ACT	18S	U6	TUB	U6	ACT	U6
5	ACT	ACT	ACT	G3PDH	ACT	TUB	ACT	TUB	U6	ACT
6	G3PDH	G3PDH	G3PDH	18S	G3PDH	18S	18S	18S	18S	18S
7	U6	U6	U6	U6	U6	G3PDH	G3PDH	G3PDH	G3PDH	G3PDH
Stressors			Salinity					Fungal		
Rank	DC	BK	NF	GN	RCR	DC	BK	NF	GN	RCR
1	ELF	ELF	UBC	ELF | UBC	ELF	TUB	TUB	TUB	U6| G3PDH	TUB
2	UBC	TUB	ELF		UBC	18S	G3PDH	18S		G3PDH
3	TUB	UBC	TUB	TUB	TUB	G3PDH	U6	UBC	18S	18S
4	ACT	U6	ACT	ACT	ACT	UBC	18S	G3PDH	TUB	U6
5	U6	ACT	U6	U6	U6	U6	ACT	ACT	UBC	UBC
6	18S	18S	18S	18S	18S	ACT	UBC	U6	ACT	ACT
7	G3PDH	G3PDH	G3PDH	G3PDH	G3PDH	ELF	ELF	ELF	ELF	ELF

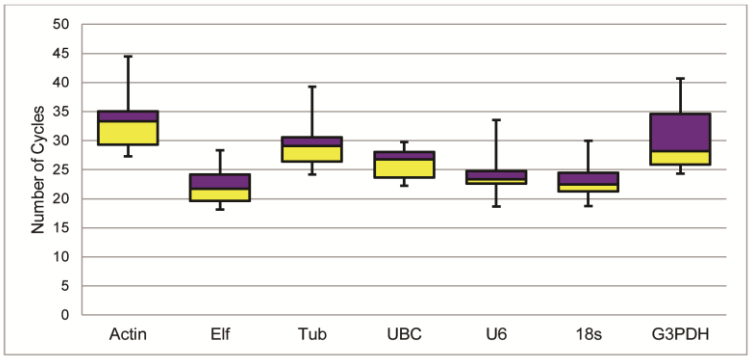


Fig 2. Relative amplification distribution curve of the selected reference genes based on Cq values across all control and stress-induced samples. The line across the box depicts the median. Yellow and violet portions indicate 25th and 75th percentiles, respectively. Whiskers represent maximum and minimum values.

cold acclimation pathways (Jaglo-Ottosen et al., 1998) was taken into account for validation of the stability of the selected reference genes under a low temperature condition. Pairwise variation analysis confirmed that using the most stable pair of genes would be satisfactory for use in normalization. For low temperature, UBC and ELF were the most stable, and U6 and G3PDH were the least stable. Relative expression of CBF-1 was analyzed by normalizing with both pairs of genes to investigate any contradictions in the pattern of expression (Fig 5). In addition, low pairwise variation among the genes recommends studying the expression of CBF-1 by normalizing with a single most stable gene because a steady result with a single reference gene could minimize the cost and time (Fig 5A). Interestingly, it was found that relative expression of the desired gene was almost same when normalized with both the stable pair and single most stable control. Thus, in case of jute, use of a single control will not hinder analyses of expression under low temperature stress conditions. However, when the relative expression was analyzed by normalizing with the two least stable pairs of reference genes (Fig 5B), there was a drastic change in the expression pattern. Therefore, selection and validation of candidate reference genes under different stress conditions are essential.

Discussion

Rapid progression of functional genomic studies involves rigorous experimentation on levels of gene expression. Such analysis is of vital importance to comprehend the cellular and molecular processes in any biological entity. However, quantification of gene expression is affected by several factors, such as the quantity of the initial material, quality of the RNA, efficiency of cDNA synthesis, primer performance, and methods to be used for statistical analysis (Maroufi et al., 2010). An ideal reference gene should be representative of overall expression across all possible tissues (cells) and experimental conditions (Guénin et al., 2009).

However, such genes may not exist because plant growth is affected by the environment; differences in results would arise with the use of different reference genes, and inaccurate assessment of gene expression could be obtained if suitable reference genes are not used (Ma et al., 2013).

Table 3. Reference genes selected in this study with their functions.

Gene name	Gene symbol	Function
Actin	*Act*	Structural constituent of cytoskeleton
Elongation factor-alpha	*Elf*	Translation factor activity, nucleic acid binding
Glyceraldehyde 3 Phosphate dehydrogenase	*G3PDH*	Involved in glycolysis process.
Beta-tubulin	*Tub*	Structural constituent of cytoskeleton, microtubule based processes.
Ubiquitin conjugating enzyme	*Ubc*	Performs the second step in the ubiquitination reaction that targets a protein for degradation via the proteasome.
U6-snRNA	*U6*	Associated with spliceosome, processing of pre-mRNA
18S rRNA	*18S*	Involved in translation process.
C-repeat binding factor-1	*CBF-1*	Transcription factor in cold acclimation pathway

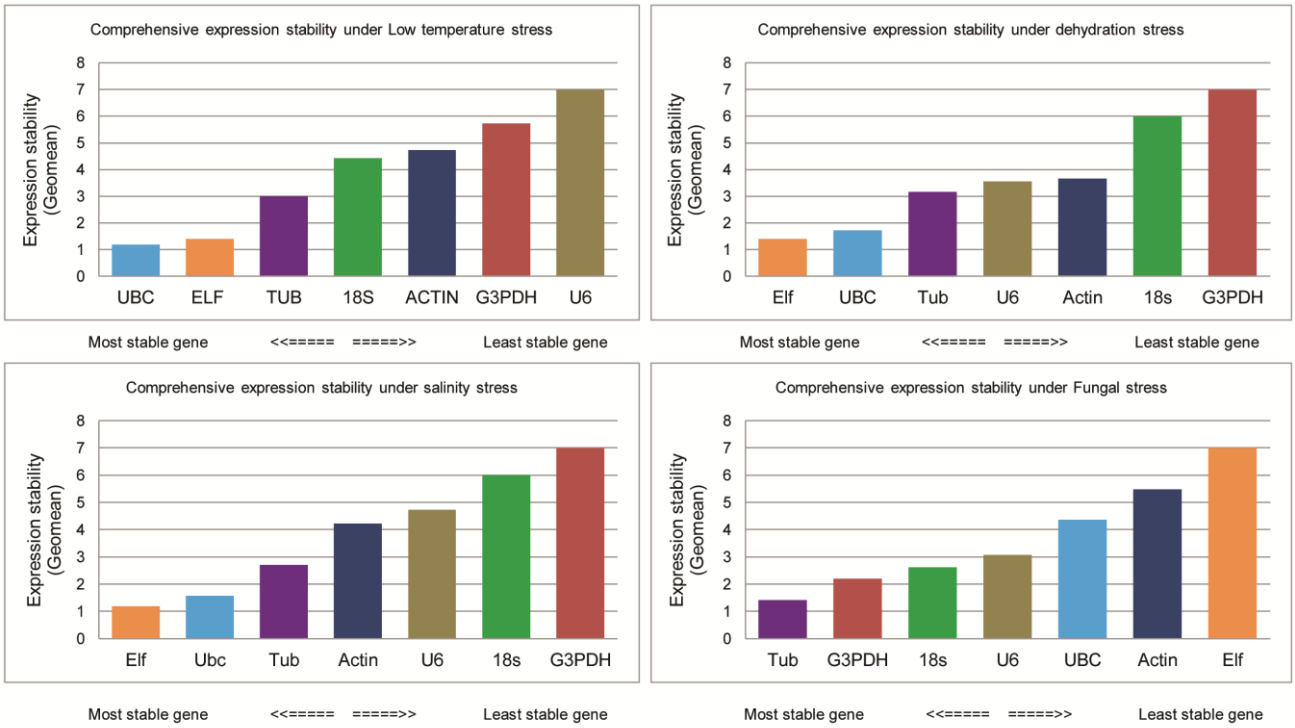
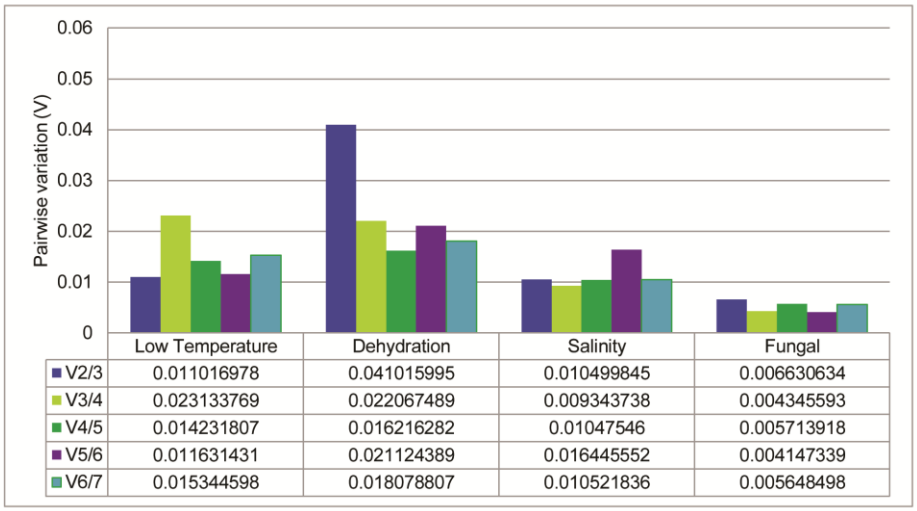


Fig 3. Expression stability of the reference genes based on comprehensive ranking under different stress conditions.

For example in *Caragana intermedia*, Zhu et al. (2013) studied the expression of DREB1 (also known as CBF-1, a transcription factor in cold acclimation pathways) at low temperature. When the pattern of expression was calculated using the evaluated stable genes, Elf-alpha and SAND, and least stable gene, UNK2, a wide variation in the expression levels was observed (Zhu et al., 2013). Thus, there is a need to select appropriate reference genes to be used as internal controls for each experimental set up in a specific species. A suitable reference gene for one species might be unstable for another. For example, ACT performed as a highly reliable reference gene when evaluated in chicory (Maroufi et al., 2010), litchi (Zhong et al., 2011), and peanut (Chi et al., 2012) but was found to be unstable in peach (Tong et al., 2009) and papaya (Zhu et al., 2012). The current study undertook experiments to evaluate the expression stability of seven potential candidate reference genes (ACT, ELF, G3PDH, TUB, UBC, U6 and 18S) for their use in quantitative analyses of gene expression in jute. Consistency of expression of these genes under four stress conditions (low temperature, dehydration, salinity, and fungal infestation) were analyzed by four different algorithms (deltaCt method, geNorm, NormFinder, and BestKeeper) integrated in a single web tool, RefFinder under the Cotton EST Database (Xie et al., 2011). Our study identified ELF as the most stable and UBC as the second best under salinity stress, UBC as most stable and ELF as second most stable when exposed to a low temperature condition, ELF as the top most and UBC as next best stable under dehydration treatment, and TUB as most stable followed by G3PDH as second most stable when exposed to fungal stress; these are represented in a recommended comprehensive ranking by RefFinder (Table 2). Overall, ELF, UBC, and TUB appeared to be expressed stably under different stress conditions. U6, ACT, 18S rRNA, and G3PDH were relatively unstable in abiotic experimental conditions with the former three showing an irregular expression pattern under biotic stress as well. It is to be noted that in the case of abiotic stressors (salinity, dehydration, and low temperature) UBC and ELF appeared to be more stable, whereas for fungal stress they ranked lower in the different algorithms. The recommended comprehensive ranking showed ELF in the 7th position and UBC in the 5th position among the seven genes tested under the fungal stress condition. Meanwhile, G3PDH, which ranked the lowest under salinity and dehydration conditions and second lowest in low temperature condition, was shown to be more stable under biotic stress condition. This has also been found by other researchers; for example in



	Low Temperature	Dehydration	Salinity	Fungal
V2/3	0.011016978	0.041015995	0.010499845	0.006630634
V3/4	0.023133769	0.022067489	0.009343738	0.004345593
V4/5	0.014231807	0.016216282	0.01047546	0.005713918
V5/6	0.011631431	0.021124389	0.016445552	0.004147339
V6/7	0.015344598	0.018078807	0.010521836	0.005648498

Fig 4. Pairwise variations among the candidate genes under different stress conditions.

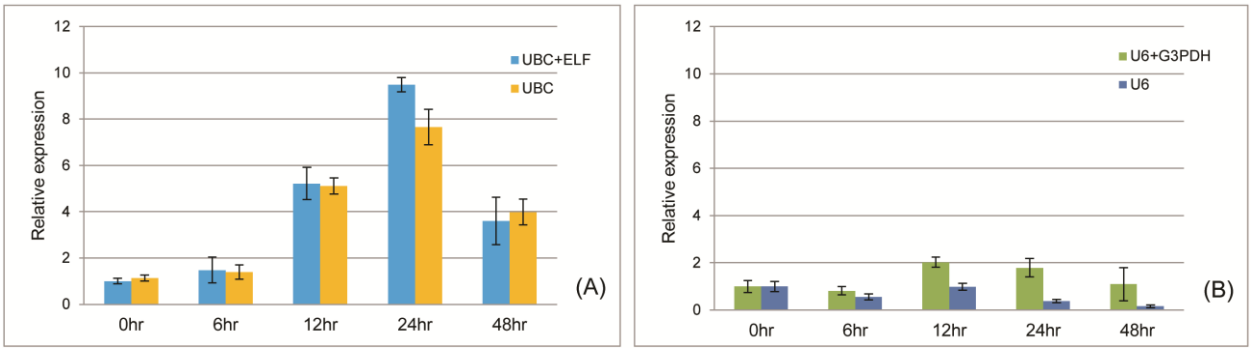


Fig 5. Relative expression pattern of CBF-1 obtained by normalizing with (a) most stable pair (UBC + ELF) and (b) least stable pair (U6 + G3PDH) under low temperature stress condition.

common bean (Borges et al., 2012), TUB proved to be highly stable under biotic stress conditions, whereas under abiotic stress it ranked as an unsuitable reference gene for normalization. ELF has been reported as a suitable reference gene in *Caragana intermedia* (Zhu et al., 2013), barley (Janská et al., 2013), *Litsea cubeba* (Lin et al., 2013), and rice (Jain et al., 2006). UBC was suggested as a reliable internal control in rose (Klie and Debener, 2011) and *Brachypodium* (Hong et al., 2008). Tubulin and G3PDH were reported as stable internal controls in cereals (Jarošová and Kundu, 2010) under biotic stress conditions. Contradiction in the stability of expression of these genes in biotic and abiotic stress conditions indicates that although these genes are called "housekeeping" for their role in basal cellular functions (Wei et al., 2013), they invariably change their expression in response to experimental conditions (Hu et al., 2009). Therefore, using unevaluated reference genes by assuming their consistent expression might prove wrong in actual experimentations. This is further corroborated by the validation process. Validation of selected reference genes in the present study was performed by analyzing the expression of the CBF-1 (also called DREB1b) gene under a low temperature condition. CBF-1 is a member of CBF/DREB 1 (CRT binding factor or DRE binding protein) regulon (Gilmour et al., 2004). Expression of jute CBF-1, when normalized with the most stable genes under a low temperature condition, showed highest expression at 24 hrs, followed by a rapid decline in its expression (Fig 5a). In contrast, when the least stable gene was used, the level of expression was found anomalous (Fig 5b). This emphasizes the importance of selecting and validating reference gene(s) before studying gene expression in any species and in any experimental condition. Data analyses showed that the order of most stable genes were not the same in all four algorithms (Table 2). This incongruity entails differences in calculation strategies that are followed in the statistical analysis by these computational programs. Nonetheless, when compared, the three highest ranked genes in one algorithm were generally present in the relatively high positions in other algorithms. Primer specificity is determined by analyzing melting curves; however, a single peak or the absence of a replicon in a no template control (NTC) do not necessarily mean a single product, and agarose gel electrophoresis is also necessary to confirm the specificity (Zhu et al., 2012). The threshold of pairwise variation as described by Vandesompele et al. (2002) is 0.15 for inclusion of more reference genes in the normalization process. In this study, values were found to be much lower, with 0.06 as the highest, suggesting that more than two most stable genes was not necessary to be included. Initial expression of CBF-1 was analyzed by normalizing with the two most stable genes (UBC and ELF) under a low temperature condition. However, when the most stable reference gene (UBC) was used, the pattern of expression was found to be the same as that obtained when two controls were used. The threshold-based choice of the number of normalizers does not need to be an exclusive criterion. A researcher may consider other factors, such as cost and accuracy, and can therefore either assume a large number of

genes in order to increase data accuracy or fewer genes when analysis of many treatments and tissues are necessary (Hu et al., 2009). Quantitation of gene expression in jute under the conditions studied here can therefore be performed using only the most stable gene as an internal control.

Materials and Methods

Plant materials and experimental conditions

Seeds of farmer-popular O-9897 variety of tossa jute (*Corchorus olitorius*) were collected from Bangladesh Jute Research Institute (BJRI). Seeds were allowed to grow under controlled conditions (maintained in a growth chamber) for 3 days in petri dishes. Seedlings were then subjected to three different abiotic stressors (salinity, low temperature, and dehydration) and an abiotic stressor (fungal infestation) individually. Stressed seedlings were collected after 0, 6, 12, 24, and 48 hrs of stress application. These samples were immediately snap-frozen in liquid nitrogen and then kept at −80 °C for subsequent use (Lin et al., 2013).

For salinity stress, seedlings were treated with a 300 mM sodium chloride solution on the 4thday of germination and collected at the different time intervals stated previously. For low temperature stress, seedlings were kept at 4 °C and collected at the desired intervals. Dehydration stress was simulated by treating the 4-day old seedlings with 150 mM mannitol solution and collected at specific time intervals of stress induction. For fungal stress, 1% solution of *Macrophomina phaseolina* was sprayed on the 4-day old seedlings and collected at different time intervals. Two independent biological replicates were collected for each of the samples.

Candidate gene confirmation

A set of seven potential reference genes were selected (Table 3) based on literature reviews: ACT, ELF-α, G3PDH, TUB, UBC, U6-snRNA (U6), and 18S. Validation of stability of suitable reference genes was performed by expression profiling of transcription factor CBF-1 under low temperature stress. Because jute genome data are not publicly available, degenerate primers were at first designed from the sequences of gene homologs of other plant species: *Gossipium*, *Vitis venifera*, *Ricinus*, and *Glycine max*. Primers were designed using PRIMER 3.0 (http://frodo.wi.mit.edu/) and analyzed by the OligoAnalyzer (http://www.idtdna.com/analyzer/Applications/OligoAnalyz er/) tool of IDT, with the aim of amplifying only the coding region. Primers were tested for expected product size by performing PCR using Mastercycler gradient (Eppendorf, Germany). The presence of a specific DNA band on an agarose gel following gel electrophoresis was the preliminary indication of the presence of these genes in jute. For further confirmation, the bands were extracted from the gel and sequenced. The sequences were then used against NCBI Blastn database (Altschul et al., 1990) with optimization to "somewhat similar sequences".

Designing primers and checking their efficiency

Upon sequencing, gene-specific primers were designed for real time PCR from the sequences obtained from jute DNA. Primers were designed by evaluating the following criteria: product size (< 150 bp), GC content 40%–60%, primer size range 20–24 bp, and Tm range (> 55°C). These primers were tested for their amplification specificity by PCR followed by gel electrophoresis in 2% agarose gels and also by a single sharp peak in Tm calling curves during real time PCR. The amplification efficiency of seven pairs of candidate reference gene primers and one pair of primers for the validation was calculated by performing qRT-PCR using 10-fold serially diluted cDNA (1, 10, 10^2, 10^3, 10^4 × dilution) each in triplicate using the following formula: $E = (10^{-1}/\text{slope-1}) \times 100$. In addition, a standard curve was drawn, from which the R^2 values (linear regression) were determined (Table 1).

RNA isolation and purity check

The plant samples were ground to a fine paste with liquid nitrogen, and total RNA was isolated using TRIzol reagent (Invitrogen, USA) following the manufacturer's instructions. The quality of the purified RNA was checked by measuring the concentration of samples by a NanoDrop 1000 (Thermo Scientific). RNA samples with 260/280 wavelength ratio between 1.8 and 2.1 and 260/230 wavelength ratio greater than 2.0 were considered as good quality samples. The quality of the RNA was further checked by electrophoresis through 1.3% agarose gels in which sharp bands for 28S and 18S rRNA confirmed good quality.

First strand cDNA synthesis

RNA isolation and cDNA synthesis from all different samples were performed for two biological replicates. Four thousand nanograms of total RNA were used to perform RT reactions with Superscript III first strand synthesis system (Invitrogen), according to manufacturer's instructions, which were further modified for jute. For a reaction volume of 20 µL, 1.0 µL of 10 mM dNTPs and 1.0 µL of oligo dT primer were at first taken together with an appropriate amount of RNA and adjusted amount of DEPC-treated H_2O followed by heating the mixture for 5 min at 65°C then transferring the mixture immediately to ice. Kept on ice for approximately 2 min, 4 µL of 5× FS buffer, 2 µL of 0.1 M dTT, 1.0 µL of M-MLV Superscript III RT (200 U/µL), and 1.0 µL RNaseOUT^{TM} (40 U/µL) were added to the reaction mix. The RT reaction was carried out in a Mastercycler (Eppendorf, Germany), following a pulse RT cycle starting from incubation at 50 min at 50°C and then 5 min at 85°C to inactivate the RT enzyme. This step was followed by another incubation step of 20 min at 37°C after adding 1.0 µL RNAse H to degrade the cDNA-RNA hybrid. An additional pair of primers was used to check for DNA contamination in cDNA samples by targeting flanking regions (i.e., introns) within the coding sequences. All calculations also included 10% overage to adjust pipetting errors.

qRT-PCR conditions

Real time PCR was performed in a 32-well plate Roche LightCyclerNano System with Roche SYBR Green MasterI (Roche Diagnostics, Germany). The PCR reaction volume was 7.5 µL, containing an equal amount of cDNA for each sample, 0.1875 µL of each primer (forward and reverse), and 3.75 µL of SYBR Green MasterI. Thermo cycling conditions were set as an initial polymerase activation step for 600 s at 95°C, followed by 45 cycles of 10 s at 95°C for template denaturation, 10 s at the optimum temperature for each gene-specific primer for annealing, and 15 s at 72°C for extension and fluorescence measurement. Afterwards, a dissociation protocol with a gradient from 50°C to 95°C was used for each primer pair to verify the specificity of the qRT-PCR and absence of primer dimers. In addition, each PCR reaction included a NRT control to check for potential genomic DNA

contamination. Reagent contamination was also detected by a reaction mix without template. All samples were amplified in technical triplicates and biological duplicates (Jain et al., 2006), and the mean of Cq (quantification cycle) value of each sample was used hour-wise for qRT-PCR data analysis.

Data analysis

The level of expression for the seven candidate reference genes were determined by the cycle number needed for the amplification-related fluorescence to reach a definite threshold level of detection called quantification cycle (Cq). For estimation of reference gene stability, RefFinder (http://www.leonxie.com/referencegene.php) was used; it compiles major statistical programs (geNorm, NormFinder, BestKeeper, and deltaCt method) (Vandesompele et al., 2002; Andersen et al., 2004; Pfaffl et al., 2004; Silver et al., 2006) for evaluation of reference gene expression stability and selection of the most suitable reference gene under different experimental conditions. RefFinder can be fed the raw Cq data, which are then used according to the necessity of the different programs integrated within it. Samples were grouped according to the stressors applied (salinity, low temperature, dehydration, and fungal infestation), and data analysis was performed separately for different groups. Stable reference genes were determined for each environmental stress condition individually by analyzing the Cq value from that group of data produced by performing qRT-PCR. RefFinder gives a unified comprehensive ranking of the candidate genes according to their stability calculated by the four different methods.

The geNorm, Normfinder, and BestKeeper are excel-based tools. NormFinder and geNorm use relative expression values as input data, whereas BestKeeper and the comparative ΔCT method use Cq values directly. The geNorm software provides the two most stable reference genes or a combination of multiple stable genes by calculating a gene expression normalization factor (M value) based on the geometric mean of a number of candidate reference genes (Vandesompele et al., 2002). NormFinder measures the optimal reference gene among a group of candidate genes on the basis of their expression stability in a sample set or specific experimental designs (Andersen et al., 2004). This algorithm evaluates the overall expression variation of the candidate reference genes and the variation between subgroups of samples (Chao et al., 2012). BestKeeper determines the best reference genes using pairwise correlation analysis of candidate reference genes (Pfaffl et al., 2004). It uses standard deviation, percent covariance, and power of the candidates as indicators to determine the best reference gene (Chao et al., 2012). The comparative ΔCT method evaluates the most stable reference genes by comparing relative expression of "pairs of genes" within each sample (Silver et al., 2006; Chao et al., 2012). This method measures the stability of a gene by the mean of standard deviation values derived from comparison between a reference gene and other candidate reference genes (Chao et al., 2012). Requirement of the optimal number of reference genes to be used for normalization was calculated using the geometric mean of the most stable genes and consequently performing a consecutive pairwise variation calculation as suggested in geNorm algorithms (Vandesompele et al., 2002). For expression analysis of CBF-1 (using the most and least stable reference genes) under a low temperature stress condition, calculation was performed by using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001).

Conclusion

To the best of our knowledge, this set of experiments is the first attempt to validate candidate reference genes in jute for normalization of gene expression analysis using qRT-PCR. In this study, reference genes evaluated can be considered useful for future assessments of gene expression analysis when studying expression under stress conditions in jute. Moreover, this study provides useful guidelines for reference gene selection for researchers working on jute.

Acknowledgements

This experiment was funded by Bangladesh Academy of Science and US Department of Agriculture (BAS-USDA PALS program). The authors acknowledge BJRI (Bangladesh Jute Research Institute) for providing seeds, kind assistance from Suprovath Kumar Sarker, as well as technical support from Kamal Hossain.

References

Achard P, Gong F, Cheminant S, Alioua M, Hedden P and Genschik P (2008) The cold-inducible CBF1 factor–dependent signaling pathway modulates the accumulation of the growth-repressing DELLA proteins via its effect on gibberellin metabolism. Plant Cell. 20(8): 2117-2129.

Ahmed S, Shafiuddin M, Azam MS, Islam MS, Ghosh A and Khan H (2011) Identification and characterization of jute LTR retrotransposons. Mob Genet Elements. 1(1): 18-28.

Altschul SF, Gish W, Miller W, Myers EW and Lipman DJ (1990) Basic local alignment search tool. J Mol Biol. 215(3): 403-410.

Andersen CL, Jensen JL and Ørntoft 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 Res. 64(15): 5245-5250.

Borges A, Tsai SM and Caldas DGG (2012) Validation of reference genes for RT-qPCR normalization in common bean during biotic and abiotic stresses. Plant Cell R. 31(5): 827-838.

Brunner AM, Yakovlev IA and Strauss SH (2004) Validating internal controls for quantitative plant gene expression studies. BMC Plant Biol. 4(1): 14.

Bustin S and Dorudi S (1998) Molecular assessment of tumour stage and disease recurrence using PCR-based assays. Mol Med Today. 4(9): 389-396.

Chao WS, Doğramaci M, Foley ME, Horvath DP and Anderson JV (2012) Selection and validation of endogenous reference genes for qRT-PCR analysis in leafy spurge (*Euphorbia esula*). PLoS One. 7(8): e42839.

Chi X, Hu R, Yang Q, Zhang X, Pan L, Chen N, Chen M, Yang Z, Wang T and He Y (2012) Validation of reference genes for gene expression studies in peanut by quantitative real-time RT-PCR. Mol Genet Genomics. 287(2): 167-176.

Czechowski T, Stitt M, Altmann T, Udvardi MK and Scheible W-R (2005) Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. Plant Physiol. 139(1): 5-17.

de Almeida MR, Ruedell CM, Ricachenevsky FK, Sperotto RA, Pasquali G and Fett-Neto AG (2010) Reference gene selection for quantitative reverse transcription-polymerase chain reaction normalization during in vitro adventitious rooting in *Eucalyptus globulus* Labill. BMC Mol Biol 11(1): 73.

Expósito-Rodríguez M, Borges AA, Borges-Pérez A and Pérez JA (2008) Selection of internal control genes for quantitative real-time RT-PCR studies during tomato development process. BMC Plant Biol. 8(1): 131.

Gachon C, Mingam A and Charrier B (2004) Real-time PCR: what relevance to plant studies? J Exp Bot. 55(402): 1445-1454.

Gilmour SJ, Fowler SG and Thomashow MF (2004) Arabidopsis transcriptional activators CBF1, CBF2, and CBF3 have matching functional activities. Plant Mol Biol. 54(5): 767-781.

Guénin 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. J Exp Bot. 60(2): 487-493.

Gutierrez L, Mauriat M, Pelloux J, Bellini C and Van Wuytswinkel O (2008) Towards a systematic validation of references in real-time RT-PCR. Plant Cell. 20(7): 1734-1735.

Hong S-Y, Seo PJ, Yang M-S, Xiang F and Park C-M (2008) Exploring valid reference genes for gene expression studies in *Brachypodium distachyon* by real-time PCR. BMC Plant Biol. 8(1): 112.

Hu R, Fan C, Li H, Zhang Q and Fu Y-F (2009) Evaluation of putative reference genes for gene expression normalization in soybean by quantitative real-time RT-PCR. BMC Mol Biol. 10(1): 93.

Huggett J, Dheda K, Bustin S and Zumla A (2005) Real-time RT-PCR normalisation; strategies and considerations. Genes Immun. 6(4): 279-284.

Iskandar HM, Simpson RS, Casu RE, Bonnett GD, Maclean DJ and Manners JM (2004) Comparison of reference genes for quantitative real-time polymerase chain reaction analysis of gene expression in sugarcane. Plant Mol Biol Rep. 22(4): 325-337.

Jaglo-Ottosen KR, Gilmour SJ, Zarka DG, Schabenberger O and Thomashow MF (1998) Arabidopsis CBF1 overexpression induces COR genes and enhances freezing tolerance. Science. 280(5360): 104-106.

Jain M, Nijhawan A, Tyagi AK and Khurana JP (2006) Validation of housekeeping genes as internal control for studying gene expression in rice by quantitative real-time PCR. Biochem Biophys Res Commun. 345(2): 646-651.

Janská A, Hodek J, Svoboda P, Zámečník J, Prášil IT, Vlasáková E, Milella L and Ovesná J (2013) The choice of reference gene set for assessing gene expression in barley (*Hordeum vulgare* L.) under low temperature and drought stress. Mol Genet Genomics. 288(11): 639-649.

Jarošová J and Kundu JK (2010) Validation of reference genes as internal control for studying viral infections in cereals by quantitative real-time RT-PCR. BMC Plant Biol. 10(1): 146.

Keka SI, Shamsuzzaman M, Pahloan MU, Pervin S, Rahman MM and Khan H (2008) Identifying simple sequence repeat (SSR) marker linked to mite tolerance in jute species. Bangladesh J Botany. 37(2): 161-171.

Kim B-R, Nam H-Y, Kim S-U, Kim S-I and Chang Y-J (2003) Normalization of reverse transcription quantitative-PCR with housekeeping genes in rice. Biotechnol Lett. 25(21): 1869-1872.

Klie M and Debener T (2011) Identification of superior reference genes for data normalisation of expression studies via quantitative PCR in hybrid roses (*Rosa hybrida*). BMC Res Notes. 4(1): 518.

Lee JM, Roche JR, Donaghy DJ, Thrush A and Sathish P (2010) Validation of reference genes for quantitative RT-PCR studies of gene expression in perennial ryegrass (*Lolium perenne* L.). BMC Mol Biol. 11(1): 8.

Libault M, Thibivilliers S, Bilgin D, Radwan O, Benitez M, Clough S and Stacey G (2008) Identification of four soybean reference genes for gene expression normalization. Plant Genome. 1(1): 44-54.

Lin L, Han X, Chen Y, Wu Q and Wang Y (2013) Identification of appropriate reference genes for normalizing transcript expression by quantitative real-time PCR in *Litsea cubeba*. Mol Genet Genomics. 288(12): 727-737.

Livak KJ and Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods. 25(4): 402-408.

Ma S, Niu H, Liu C, Zhang J, Hou C and Wang D (2013) Expression stabilities of candidate reference genes for qRT-PCR under different stress conditions in soybean. PLoS One. 8(10): e75271.

Mallona I, Lischewski S, Weiss J, Hause B and Egea-Cortines M (2010) Validation of reference genes for quantitative real-time PCR during leaf and flower development in Petunia hybrida. BMC Plant Biol. 10(1): 4.

Maroufi A, Van Bockstaele E and De Loose M (2010) Validation of reference genes for gene expression analysis in chicory (*Cichorium intybus*) using quantitative real-time PCR. BMC Mol Biol 11(1): 15.

Paolacci AR, Tanzarella OA, Porceddu E and Ciaffi M (2009) Identification and validation of reference genes for quantitative RT-PCR normalization in wheat. BMC Mol Biol. 10(1): 11.

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. Biotechnol Lett. 26(6): 509-515.

Radonić A, Thulke S, Mackay IM, Landt O, Siegert W and Nitsche A (2004) Guideline to reference gene selection for quantitative real-time PCR. Biochem. Biophys Res. Commun. 313(4): 856-862.

Reddy DS, Bhatnagar-Mathur P, Cindhuri KS and Sharma KK (2013) Evaluation and Validation of Reference Genes for Normalization of Quantitative Real-Time PCR Based Gene Expression Studies in Peanut. PLoS One. 8(10): e78555.

Riechmann JL and Meyerowitz EM (1998) The AP2/EREBP family of plant transcription factors. Biol Chem. 379: 633-646.

Roy A, Bandyopadhyay A, Mahapatra A, Ghosh S, Singh N, Bansal K, Koundal K and Mohapatra T (2006) Evaluation of genetic diversity in jute (*Corchorus* species) using STMS, ISSR and RAPD markers. Plant breeding. 125(3): 292-297.

Schmidt GW and Delaney SK (2010) Stable internal reference genes for normalization of real-time RT-PCR in tobacco (*Nicotiana tabacum*) during development and abiotic stress. Mol Genet Genomics. 283(3): 233-241.

Sharmin S, Moosa MM, Islam S, Kabir I, Akter A and Khan H (2011) Identification of a novel dehydration responsive transcript from tossa jute (*Corcohrus olitorius* L.). Cell Mol Biol. 9(1).

Silver N, Best S, Jiang J and Thein SL (2006) Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR. BMC Mol Biol. 7(1): 33.

Tong Z, Gao Z, Wang F, Zhou J and Zhang Z (2009) Selection of reliable reference genes for gene expression studies in peach using real-time PCR. BMC Mol Biol. 10(1): 71.

Udvardi MK, Czechowski T and Scheible W-R (2008) Eleven golden rules of quantitative RT-PCR. Plant Cell. 20(7): 1736-1737.

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 Biol. 3(7): research0034.

Wei L, Miao H, Zhao R, Han X, Zhang T and Zhang H (2013) Identification and testing of reference genes for Sesame gene expression analysis by quantitative real-time PCR. Planta. 237(3): 873-889.

Xie F, Sun G, Stiller JW and Zhang B (2011) Genome-wide functional analysis of the cotton transcriptome by creating an integrated EST database. PLoS One. 6(11): e26980.

Zhong H-Y, Chen J-W, Li C-Q, Chen L, Wu J-Y, Chen J-Y, Lu W-J and Li J-G (2011) Selection of reliable reference genes for expression studies by reverse transcription quantitative real-time PCR in litchi under different experimental conditions. Plant Cell Rep. 30(4): 641-653.

Zhu J, Zhang L, Li W, Han S, Yang W and Qi L (2013) Reference gene selection for quantitative real-time PCR normalization in *Caragana intermedia* under different abiotic stress conditions. PLoS One. 8(1): e53196.

Zhu X, Li X, Chen W, Chen J, Lu W, Chen L and Fu D (2012) Evaluation of new reference genes in papaya for accurate transcript normalization under different experimental conditions. PLoS One. 7(8): e44405.