Validation of reference genes for accurate normalization by quantitative polymerase chain reaction in sugarcane drought stress studies using two cultivars

Dennis Crystian[1]   Jackeline Terto[1]   José Vieira Silva[2]   Cícero Almeida[1*]

[1]Laboratório de Recursos Genéticos, Universidade Federal de Alagoas (UFAL), Campus Arapiraca, Arapiraca, 57300-970, AL, Brasil. E-mail: cicero@arapiraca.ufal.br. *Corresponding author.
[2]Laboratório de Fisiologia, Universidade Federal de Alagoas (UFAL), Campus Arapiraca, Arapiraca, AL, Brasil.

ABSTRACT: *The aim of this study was to analyze the expression of putative reference genes in sugarcane under drought stress. The varieties RB72454 and RB72910 were cultivated and the treatments control and drought stress applied to 135-day-old plants grown under field conditions. The stress level of the plants was measured by rate of photosynthesis, transpiration, and stomatal conductance. For each biological replicate, expression analyses were conducted using quantitative polymerase chain reaction for the genes α-tubulin, β-tubulin, β-actin, cyclophilin, eukaryotic elongation factor 1, glyceraldehyde 3-phosphate dehydrogenase (GAPDH), histone H3 and ubiquitin. Stability was evaluated using the geNorm, NormFinder, and BestKeeper software packages. Among the candidate genes, GAPDH was identified as the most stable in all software, indicating its suitability for gene expression studies in sugarcane undergoing drought stress; the gene β-actin was the second most stable. These findings suggest using GAPDH and β-actin for normalization in relative gene expression in sugarcane.*
Key words: *gene expression, drought stress, reference gen, Saccharum.*

Validação de genes de referência para normalização de reação em cadeia de polimerase quantitativa em estudos de estresse hídrico em cana-de-açúcar

RESUMO: *O objetivo do estudo foi analisar a expressão de genes de referência em cana-de-açúcar sob estresse* hídrico. *As cultivares RB72454 e RB72910 foram cultivadas em campo e os tratamentos controle e estresse foram aplicados após 135 dias. O estresse das plantas foi mensurado utilizando as variáveis fisiológicas: taxa fotossintética, transpiração e condutância estomática. Para cada repetição biológica, a expressão gênica foi conduzida usando PCR tempo real para os genes α-tubulina, β-tubulina, β-actina, ciclofilina, fator 1 de alongamento eucariótico, gliceraldeído 3-fosfato desidrogenase (GAPDH), histona H3 e ubiquitina. A estabilidade dos genes foi avaliada usando os softwares geNorm,NormFinder e BestKeeper. Entre os genes candidatos, GAPDH foi identificado como o mais estável nos três métodos de análises, indicando ser estável para estudos de expressão gênica em cana-de-açúcar sob estresse* hídrico; *seguindo o gene β-actina como o segundo mais estável. Os resultados concluem que os genes GAPDH e* β-actina *são indicados para normalização em expressão gênica relativa, associada com estresse hídrico em cana-de-açúcar.*
Palavras-chave: *expressão gênica, estresse hídrico, genes de referência, Saccharum.*

INTRODUCTION

Among the approaches available for gene expression studies, the quantitative polymerase chain reaction (qPCR) is extensively used. The qPCR is a fast, easy for analyzing and understanding gene expression in living organisms by measuring the expression of target genes in different samples. In qPCR analysis, some factors are important: RNA integrity, number of samples, tissue type, environmental conditions where the organism develops, and cell activity (KOZERA & RAPACZ, 2013). In relative gene expression analysis, it is essential to use a gene that is highly stable in diverse conditions of biotic and abiotic stress, different body tissues, and in various stages of development. These genes are known as reference genes (KOZERA

& RAPACZ, 2013) and are generally recruited from those involved in primary metabolism or basic activities and the components of cellular structure. Generally, an adequate reference gene for gene expression analysis should be defined as a gene which is stably expressed among the different samples and unaffected by any experimental treatment. Accurate normalization is prerequisite for consistent qPCR results and the inappropriate reference genes can impose serious impacts on results obtained (KOZERA & RAPACZ, 2013). To analyze these reference genes, statistical methodologies have been developed to infer on gene stability: *geNorm* (VANDESOMPELE et al. 2002), *NormFinder* (ANDERSEN et al. 2004), and *BestKeeper* (PFAFFL et al. 2004) are most commonly used.

Sugarcane has great economic importance because it is the raw material for sugar, first and second generation ethanol, and electricity production. In Brazil, it has been cultivated in different environments, with varying soil types, temperature, and precipitation. Among abiotic factors, water availability has been most important for sugarcane cultivation and has motivated a number of studies on gene expression (RODRIGUES et al. 2009, 2011; OLIVEIRA et al. 2014; LING et al. 2014). The main studies of gene expression associated to drought stress have been conducted in root and leaf tissues to identify physiological processes related to drought tolerance (LING et al. 2014).

Studies conducted on sugarcane to determine reference genes have been conducted in plants grown in pots (SILVA et al. 2014), in vitro with *Saccharum officinarum* (LING et al. 2014), and in the field (ISKANDAR et al. 2004). Nevertheless, only SILVA et al. (2014) determined stability associated with drought stress in plants in glasshouse pots, analyzing gene expression in the roots, and GUO et al. (2014) analyzed gene expression in seedlings by inducing drought stress with polyethylene glycol 8000 (PEG). Neither study included gene stability analysis in field-grown adult plants. In sugarcane, the most commonly used reference genes are: 25S ribosomal ribonucleic acid (rRNA), glyceraldehyde 3-phosphate dehydrogenase (GAPDH), β-actin, γ-actin, α-tubulin, β-tubulin, and ubiquitin, among others. In this study, eight genes: ubiquitin (Ubi), α-tubulin (α –tub), β-tubulin (β-tub), β-actin (β-act), GAPDH, histone H3 (ht3), eukaryotic elongation factor 1α (eEF1) and cyclophilin (Cyc) were evaluated as reference genes for drought stress in sugarcane leaves under field conditions of drought stress (SILVA et al. 2014; GUO et al. 2014; RODRIGUES et al. 2011; ANDRADE et al. 2017).

MATERIALS AND METHODS

Drought stress field experiment

The experimental design was a split-plot in strips with four replications, and the main plot (Varieties) arranged in a randomized block design. The experiment was conducted using the varieties RB72454 and RB72910 and two water conditions. Each subplot consisted of a single, two-meter-long row and, borders were established using RB72910. Plants were grown under favorable water conditions for 135 days when the treatments were established. Treatments were defined as follows: moisture close to field capacity for the test without stress, with 80–100% available water (RB72454-control and RB72910-control) and suppression of the water by stopping irrigation (RB72454-stressed and RB72910-

stressed) for testing under severe stress. Suppression of the water is necessary to analyzer the patterns of the gene expression under drought as described by Silva et al. (2012). After 15 days of stress, leaves were collected from the RB72910 and the RB72454 varieties exposed to the water availability and drought stress regimes. In each treatment combination, three biological replicates were collected for total RNA extraction and individual analysis.

To assess the state of drought stress in the plants subjected to the treatments, the following physiological analyses were performed after 15 days of stress, using two repeats for each replication: rate of photosynthesis (A), transpiration (E), and stomatal conductance (g_s), using an Infrared Gas Analyzer (IRGA) platform, with air flow of $300mLmin^{-1}$ and a $995mmol^2s^{-1}$ light source, photosynthesis photon flux density (PPFD) of $958mol\ m^{-2}\ s^{-1}$ and vapor saturation deficit (VPD) of 4,072kPa (ADC, LCi model, Hoddesdon,UK) according to the user manual. Analyses of variance were applied to the variables and means were compared using Tukey's test at a 5% significance level, using R software (TEAM,2012). The gas exchange measurements have been showed a positive association with the drought tolerance in different species(HURA et al. 2007; KARIMI et al. 2011, ASHRAF & HARRIS 2013).

RNA extraction and cDNA synthesis

Total RNA was extracted from three biological replicates of each genotype. From each plant evaluated, about 15cm were harvested from the middle part of the leaves and kept in liquid nitrogen until they were macerated using a mortar and pestle. Total RNA was extracted using the *PureLink*® kit (Life Technologies Corporation), following the manufacturer's instructions. Consequently, three biological replicates were obtained for each of the following treatment combinations: RB72910-Control, Stressed-RB72910, RB72454-Control, and RB72454-Stressed. Next, RNA quality was checked with 1% agarose gel.

Synthesis of the first cDNA strand was performed using 2μg of total RNA, lX First-Strand and 40 units of avian myeloblastosis virus reverse transcriptase (20units/mL). The reaction was incubated at 42° Celsius for 1h and 30min. The quality of the synthesis was checked with 1% agarose gel.

Synthesis of primers

Eight reference genes were selected from the published literature and by searching the NCBI database for available expressed sequence tags to develop oligos (Table 1). For the genes with no primers,

Table 1 - Selected candidate reference gene primers and their parameters.

Access number[a]	Gene Name	Sequence for forward (F) and reverse (R) primers	Amplicon length (bp)	Primer efficiency
AF093505	Ubi	F:AGCCTCACGGGACTCTTTCTC R:TGGATGGATGAGGCAAGGA	73	1.92
CA116806	ht3[*]	F:CGCACACGCACACTGAAAG R:CGGCCGGAGCAACTTCTT	81	1.92
CA148161	β-act	F:CCACACTGTGCCCATTTACG R:ACGGCCAGCAAGATCCAA	70	2.03
CA222437	β-tub[*]	F:GGAGGAGTACCCTGACAGAATGA R:CAGTATCGGAAACCTTTGGTGAT	114	2.03
CA254672	GAPDH	F:TCACCCTCAAGGACTCCAAGA R:GGGTTCCTGATGCCAAAGAC	70	2.07
EF581011	eFF1	F:GGCGCCGTAACCGTATCAT R:ACCCATGGCTGAAGGGAAAT	97	1.93
GQ246462	Cyc	F:GCCCTCCCCTCCGTCAT R:TCACCATCATCTTATTCAGATCGAA	79	2.01
JQ230105	α-tub	F:GCACCATGCCCAGTGATACC R:GAACCGGTCTCGCTGAAGAA	73	2.03

[a]Gene access number in the NCBI database.
[*]Developed by RODRIGUES et al. (2009).

oligo sequences were determined using the *StepOne Real-Time PCR System* primer design software (Applied Biosystems Inc., Foster City, CA, USA).

Quantitative polymerase chain reaction (qPCR) was performed using a *Thermo Scientific SYBR® Green Master Mix* in a ABI *StepOne Real-Time PCR System* thermocycler. Each gene was analyzed in three repeats for each biological sample. Each reaction was composed of 20ng of cDNA, 25nM of primer, and 12.5µL of *SYBR® Green Master Mix* with the volume topped up to 25µL. Amplification was done at 95°C for 10min., 40 cycles at 95°C for 15sec., 60°C for 1min., and 95°C for 15sec.

Statistical analysis of gene normalization

The stability of gene candidates for normalization in sugarcane under drought stress and non-stress control regimes was assessed using *geNorm* (VANDESOMPELE et al., 2002), *NormFinder* (ANDERSEN et al. 2004), and *BestKeeper* (PFAFFL et al. 2004) software. The analyses were based on the cycle threshold (C_q) results available from the qPCR and were carried out following the corresponding user guides. A standard curve was drafted for each gene using a sequence of tenfold dilution from the cDNA of cultivar RB72910 in the control water regime; the resulting regression coefficient of the slope provided an estimate of the efficiency primers amplification.

Analyses of variance were applied to Cq values and means were compared using Tukey's test at a 5% significance level, using R software (TEAM, 2012).

RESULTS

Physiological analysis of the plants
To analyze the physiological status of the plants under the respective treatment conditions, variables related to gas exchange were measured. Statistical analysis revealed that the rate of photosynthesis (A), transpiration (E), and stomatal conductance (g_s) showed differences between the non-stress control and drought stress regimes (Table 2). For the variable rate of photosynthesis (A), the genotypes showed mean values of 22.14m showe$_2$ m^{-2}s^{-1} and 5.32 mand 5.$_2$ m^{-2}s^{-1} for the control and drought stress regimes, respectively. Stomatal conductance (g_s) showed a sharp decrease from 0.084mol H_2O m^{-2}s^{-1} to 0.021mol H_2O m^{-2}s^{-1} under control and drought stress regimes, respectively. Mean values for leaf transpiration (E) ranged from 3.43mmol H_2O m^{-2}s^{-1} in the non-stress control water regime to 1.08mmol H_2O m^{-2}s^{-1} in the stress regime (Figure 1).

Real-time qPCR analyses
Gene expression was analyzed using the qPCR method based on *SYBR Green* detection.

Table 2 - Analysis of variance of the physiological variables: A (photosynthetic rate); E (transpiration); g_s (stomatal conductance).

	df	A	E	g_s
		--------------------------------------Mean Square--------------------------------------		
Treatments (T)	1	1954.44[**]	44.97[**]	0.0395[**]
Blocks	3	40.11	1.51	0.001
Residual (a)	3	32.48	0.73	0.0009
Genotypes (G)	1	0.23	0.033	1.1×10^{-6}
T x G	1	35.51	0.323	0.00027
Repeats/Blocks	4	15.23	0.659	0.00037
Residual (b)	18	11.63	0.24	0.00026
CV (a)%		39.7	36.8	51.4
CV (b)%		23.8	21.1	27.7

[**]Significant at the 1% significance level.

Analysis of variance was performed to verify the C_q variations in the factors analyzed (water regime, genotypes, and genes as a factor on Analysis of variance); analysis of variance revealed an interaction between genes and water regime, indicating that genes have expression patterns that depend on growing conditions (Table 3). Dissociation curve analysis evidenced the generation of single products-specific melting temperature. The C_q varied according to the gene analyzed: for example, GAPDH and $CycC_q \approx 21$; whereas, α-tub and $UbiC_q \approx 29$, showing a wide variation in gene expression (Figure 2).

Stability analysis using the *geNorm, NormFinder,* and *BestKeeper* methods indicated that GAPDH is the most suitable reference gene for drought stress studies in sugarcane; however, the β-act gene showed satisfactory levels of stability (Table 4). Conversely, the β-tub gene presented greater instability with high rates of variation between samples and, therefore, is not recommended for qPCR under drought stress conditions (Table 4). It should be noted that using *geNorm* methodology, the GAPDH, β-act, and Cyc genes had stability indices M<1.5, suggesting good stability between samples.

Table 3 - Analysis of variance of C_q (Quantification cycle).

Sources of variation	df	-------------Mean Square-------------
Treatments (T)	1	4.00
Biological Replications	2	10.1
Residual (a)	2	1.5
Genes (G)	7	353.85[**]
Genotypes (Gn)	1	17.81[**]
T x G	7	10.36[**]
T x Gn	1	2.08
G x Gn	7	1.65
T x G x Gn	7	0.62
Replicates/T x G x Gn	64	5.56
Residual (b)	188	2.5
CV (a)		6.2%
CV (b)		4.7%

[**]Significant at the 1% significance level.

Table 4 - Stability expressions of candidate reference genes in sugarcane under water stress.

Genes	----------------geNorm--------------		----------------NormFinder---------------		-------------BestKeeper------------	
	M index	Ranking	Stability	Ranking	R	Ranking
GAPDH	1.337	1	0.262	1	0.684	1
β-act	1.372	2	0.288	2	0.675	2
Cyc	1.402	3	0.448	4	0.471	6
Ubi	1.502	4	0.473	5	0.655	4
eFF1	1.656	5	0.317	3	0.675	2
ht3	2.209	6	0.772	8	0.451	7
α-tub	2.228	7	0.642	6	0.531	4
β-tub	2.785	8	0.745	7	0.314	8

DISCUSSION

An ideal reference gene has stable expression, regardless of experimental conditions applied to the organism. However, a number of studies have shown that expression of reference genes could considerably vary according to experimental conditions (THELLIN et al., 1999) because of their involvement in various cellular functions (SINGH & GREE 1993; ISHITANI et al. 1996). In this context, it is critical to determine the stability of genes to be used as reference genes in qPCR for the methodological development of gene expression studies (KOZERA & RAPACZ, 2013). Variation in reference genes depends on the biotic and abiotic conditions where the organism develops (THELLIN et al., 1999). In plants, drought tolerance has been shown to be an extremely complex abiotic condition featuring interactions associated with cell maintenance, production, and degradation of

proteins and changes in metabolism, as reviewed by SHANKER et al. (2014). From this perspective, the present study contributed by determining genes that exhibit stability in field-grown sugarcane genotypes under drought stress and control regimes. It should be noted that the physiological status of the plants was accurately monitored in this study to verify actual drought stress. The variables rates of photosynthesis (A), transpiration (E), and stomatal conductance (g_s) were analyzed to show the physiological responses of the plants to drought stress.

Many genes have been evaluated as candidates for stability under drought stress; the strawberry gene DNA binding protein (DBP), for example, was most stable under drought conditions (GALLI et al. 2015), the tonoplast intrinsic protein TIP41 gene of *Pennisetum glaucum* (L.) showed the highest stability in another study (SAHA & BLUMWALD, 2014), and the glucosidase II a-subunit gene of *Populus euphratica* was most stable under
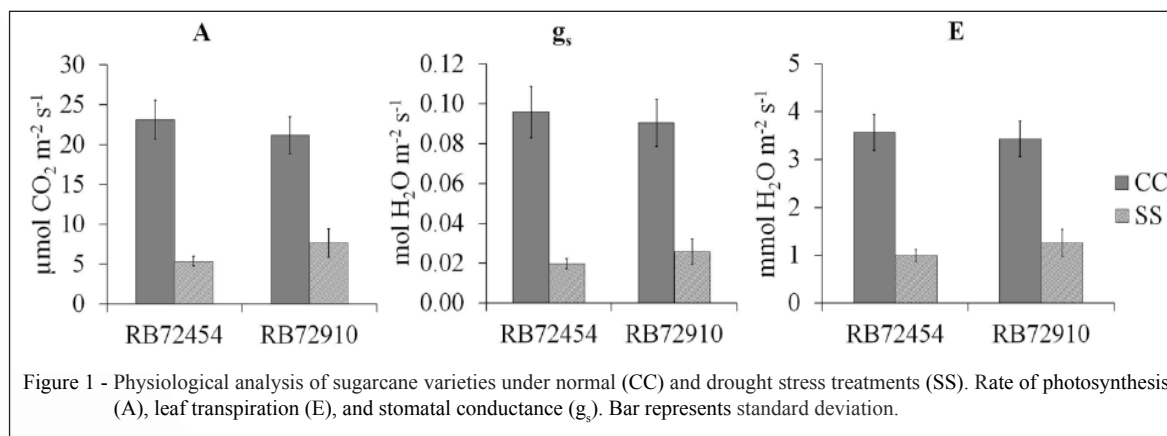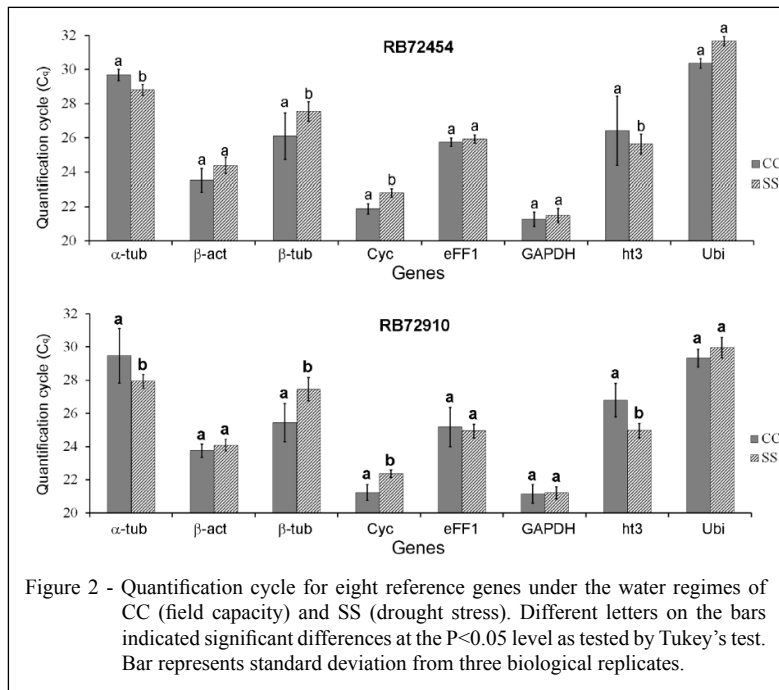

Figure 1 - Physiological analysis of sugarcane varieties under normal (CC) and drought stress treatments (SS). Rate of photosynthesis (A), leaf transpiration (E), and stomatal conductance (g_s). Bar represents standard deviation.

Figure 2 - Quantification cycle for eight reference genes under the water regimes of CC (field capacity) and SS (drought stress). Different letters on the bars indicated significant differences at the P<0.05 level as tested by Tukey's test. Bar represents standard deviation from three biological replicates.

drought stress (WANG et al. 2014). These results suggest the importance of validating specific reference genes for different treatments and organisms.

Reference stability genes for quantitative polymerase chain reaction in sugarcane were described by ISKANDAR et al. (2004), indicating GAPDH as the reference gene recommended for gene expression studies, with great stability in various tissues and genotypes. LING et al. (2014) also suggested that the GAPDH gene was most stable in sugarcane tissues grown in test tubes, analyzing various tissues in abiotic conditions of abscisic acid, methyl jasmonate, salicylic acid, hydrogen peroxide, sodium chloride, polyethylene glycol, copper chloride, and cadmium chloride. Reference genes for tolerance to drought stress were only reported in the root tissue of young plants grown in a greenhouse, indicating histone H1, α-tub, and GAPDH as the most stable genes (SILVA et al. 2014). GUO et al. (2014) analyzed the stability of genes under drought stress induced by polyethylene glycol 8000 in seedlings; the results indicating GAPDH and eFF1 as the most stable.

In this study, conducted on 150-day-old field grown plants, analysis with the *geNorm, NormFinder,* and *BestKeeper* software tools suggested GAPDH was the gene with greatest stability for gene expression studies under drought stress, confirming the findings of LING et al.(2014). Therefore, GAPDH appears to be a high-stability gene in abiotic conditions and an excellent reference for gene expression studies in sugarcane. The gene with the second-highest stability was β-act, contrary to the findings of SILVA et al. (2014) and GUO et al. (2014), who recommended α-tub and eEF1, respectively. The β-act has been used as a reference gene in gene expression studies of sugarcane under drought stress (RODRIGUES et al. 2011). Among the genes analyzed in this study, α-tub and β-tub presented the least stability and; therefore, are not recommended as reference genes in studies of drought stress in field plants. It can be concluded that the GAPDH and β-act genes are the most stable, when analyzed in different field-grown sugarcane genotypes under drought stress conditions and are; therefore, recommended as reference genes for normalization in gene expression studies.

ACKNOWLEDGEMENTS

DECLARATION OF CONFLICTING INTERESTS

The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the

collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

ETHICAL APPROVAL

This article does not contain any studies with human participants or animals performed by any of the authors.

AUTHORS' CONTRIBUTIONS

DC and JT performed the experiments. JVS and CA analyzed the data and led the writing. All authors conceived the ideas and approved the final version of the manuscript.

REFERENCES

ANDRADE, L.M. et al. Reference genes for normalization of qPCR assays in sugarcane plants under water deficit. **Plant Methods**, v.13, p.28, 2017. doi: 10.1186/s13007-017-0178-2.

ANDERSEN, C.L.; JENSEN, J.L.; ORNTOFT, T.F. Normalization of real-time quantitative reverse transcription-PCR data: A model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. **Cancer Research**, v.64, p.5245-5250, 2004.

ASHRAF, M.; HARRIS, P.J.C. Photosynthesis under stressful environments: An overview. **Photosynthetica**, v.51, p.163–190. 2013.doi: 10.1007/s11099-013-0021-6.

GALLI, V. et al. Validation of reference genes for accurate normalization of gene expression for real time-quantitative PCR in strawberry fruits using different cultivars and osmotic stresses. **Gene**, v.554, p.205-214, 2015.

GUO, J. et al. The choice of reference genes for assessing gene expression in sugarcane under salinity and drought stresses. **Scientific Reports**, v.4, 2014.

HURA T. et al. Physiological and biochemical tools useful in drought-tolerance detection in genotypes of winter triticale: Accumulation of ferulic acid correlates with drought tolerance. **Ann. Bot**., v.100, p.767–775, 2007. doi: 10.1093/aob/mcm162.

ISHITANI R. et al. Evidence that glyceraldehyde-3-phosphate dehydrogenase is involved in age-induced apoptosis in mature cerebellar neurons in culture. **Journal of Neurochemistry**, v.66, p.928-935, 1996.

ISKANDAR, H.M. et al. Comparison of reference genes for quantitative real-time polymerase chain reaction analysis of gene expression. **Plant Molecular Biology Reporter**, v.22, p.325-337, 2004.

KARIMI, S. et al. Gas-exchange response of almond genotypes to water stress. **Photosynthetica**, v.53, p.29–34, 2001. doi: 10.1007/s11099-015-0076-7.

KOZERA, B.; RAPACZ, M. Reference genes in real-time PCR. Journal of Applied Genetics, v.54, p.391-406, 2013.

LING, H. et al. Comprehensive selection of reference genes for gene expression normalization in sugarcane by real time quantitative RT-PCR. **Plos One**, v.9, 2014.

OLIVEIRA, S.R.L. et al. Validation of novel reference genes for reverse transcription quantitative real-time PCR in drought-stressed sugarcane. **Scientific World Journal**, 2014.

PFAFFL, M.W. et al. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper - Excel-based tool using pair-wise correlations. **Biotechnology Letters**, v.26, p.509-515, 2004.

RODRIGUES, F.A.; LAIA, M.L.; ZINGARETTI, S.M. Analysis of gene expression profiles under water stress in tolerant and sensitive sugarcane plants. **Plant Sci**, v.176, p.286-302, 2009.

RODRIGUES, F.A. et al. Sugarcane genes differentially expressed during water deficit. **Biologia Plantarum**, v.55, p.43-53, 2011.

SAHA, P.; BLUMWALD, E. Assessing reference genes for accurate transcript normalization using quantitative real-time PCR in pearl millet pennisetum glaucum (L.) R. Br. **Plos One**, v.9, 2014.

SHANKER, A.K. et al. Drought stress responses in crops. Functional & Integrative Genomics, v.14, p.11-22, 2014.

SILVA, P.P. et al. Path analysis for selection of drought tolerant sugarcane genotypes through physiological components. **Industrial Crops and Products**, v.37, p.11-19, 2012.

SINGH, R.; GREEN, M.R. Sequence-specific binding of transfer-RNA by glyceraldehyde-3-phosphate dehydrogenase. **Science**, v.259, p.365-368, 1993.

TEAM, R.C. R: A language and environment for statistical computing. **R Foundation for Statistical Computing**, Vienna, Austria, 2012.

THELLIN, O. et al. Housekeeping genes as internal standards: use and limits. **Journal of Biotechnology**, v.75, p.291-295, 1999.

VANDESOMPELE, J. et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. **Genome Biology**, v.3, 2002.

WANG, H.L. et al. Identification and validation of reference genes for Populus euphratica gene expression analysis during abiotic stresses by quantitative real-time PCR. **Physiologia Plantarum**, v.152, p.529-545, 2014.