

## Reference genes for quantitative real-time polymerase chain reaction studies in soybean plants under hypoxic conditions

T.J. Nakayama<sup>1,2</sup>, F.A. Rodrigues<sup>1</sup>, N. Neumaier<sup>1</sup>, F.C. Marcelino-Guimarães<sup>1</sup>, J.R.B. Farias<sup>1</sup>, M.C.N. de Oliveira<sup>1</sup>, A. Borém<sup>3</sup>, A.C.B. de Oliveira<sup>4</sup>, B.M. Emygdio<sup>4</sup> and A.L. Nepomuceno<sup>1-5</sup>

<sup>1</sup>Laboratório de Biotecnologia Vegetal, Embrapa Soja, Londrina, PR, Brasil

<sup>2</sup>Departamento de Biologia Geral, Universidade Estadual de Londrina, Londrina, PR, Brasil

<sup>3</sup>Departamento de Fitotecnia, Universidade Federal de Viçosa, Viçosa, MG, Brasil

<sup>4</sup>Embrapa Clima Temperado, Pelotas, RS, Brasil

<sup>5</sup>Plant Gene Expression Center, Embrapa Labex USA Plant Biotechnology, Albany, CA, USA

Corresponding authors: A.L. Nepomuceno  
E-mail: alexandre.nepomuceno@embrapa.br

Genet. Mol. Res. 13 (1): 860-871 (2014)

Received March 22, 2013

Accepted August 18, 2013

Published February 13, 2014

DOI <http://dx.doi.org/10.4238/2014.February.13.4>

**ABSTRACT.** Quantitative real-time polymerase chain reaction (RT-qPCR) is a powerful tool used to measure gene expression. However, because of its high sensitivity, the method is strongly influenced by the quality and concentration of the template cDNA and by the amplification efficiency. Relative quantification is an effective strategy for correcting random and systematic errors by using the expression level of reference gene(s) to normalize the expression level of the genes of interest. To identify soybean reference genes for use in studies of flooding stress, we compared 5 candidate reference genes (CRGs) with the NormFinder and GeNorm programs to select the best internal control. The expression

stability of the CRGs was evaluated in root tissues from soybean plants subjected to hypoxic conditions. Elongation factor 1-beta and actin-11 were identified as the most appropriate genes for RT-qPCR normalization by both the NormFinder and GeNorm analyses. The expression profiles of the genes for alcohol dehydrogenase 1, sucrose synthase 4, and ascorbate peroxidase 2 were analyzed by comparing different normalizing combinations (including no normalization) of the selected reference genes. Here, we have identified potential genes for use as references for RT-qPCR normalization in experiments with soybean roots growing in O<sub>2</sub>-depleted environments, such as flooding-stressed plants.

**Key words:** Endogenous genes; Internal control genes; Flooding; Housekeeping genes; Gene expression

## INTRODUCTION

Oxygen availability is one of the primary forces shaping the evolution of living organisms. Therefore, Mustroph et al. (2010) investigated evolutionarily conserved and species specific responses to oxygen availability. Specifically, they evaluated the transcriptomic reconfiguration of different organisms from the Plantae, Animalia, Fungi, and Bacteria kingdoms under low-oxygen conditions. Responses associated with glycolysis, fermentation, alternative respiration, metabolite transport, reactive oxygen species detoxification, chaperone activity, and ribosome biogenesis were all conserved among these organisms, whereas changes in signaling and transcriptional regulation were found to be species specific (Mustroph et al., 2010).

Quantitative real-time polymerase chain reaction (RT-qPCR) is often used to measure the relative expression of genes that have been previously identified by global gene-expression assays under different stress conditions. Such experiments have been used to examine the effects of flooding in *Arabidopsis* (Liu et al., 2005), rice (Lasanthi-Kudahettige et al., 2007), maize (Zou et al., 2010), *Populus* (Kreuzwieser et al., 2009), and cotton (Christianson et al., 2010). Compared with other local methods of analyzing transcript abundance, such as Northern blotting (Böhm-Hofstätter et al., 2010) and semiquantitative (Guénin et al., 2009) or competitive PCR (Zentilin and Giacca, 2007), RT-qPCR is less labor intensive, does not require manipulation after PCR amplification, and accurately quantifies gene expression over a wide detection range (Nolan et al., 2006). However, because of its high sensitivity, this method is strongly influenced by the quality and concentration of the initial nucleic acid samples and the amplification efficiency (Nolan et al., 2006). A typical strategy to correct for analytical variation is to normalize the expression of the target gene using 1 or more reference genes (Vandesompele et al., 2002). Although these reference genes require stable expression in general, they can nonetheless vary across developmental stages, tissues, or experimental treatments (Jian et al., 2008; Hu et al., 2009). Therefore, to avoid biasing the normalized expression of the target gene because of gene-specific variation in the reference gene, it is necessary to validate the best candidate reference genes (CRGs) for each experimental treatment (Xu et al., 2011). To our knowledge, this study is the first to evaluate the expression stability of CRGs in root tissues from hypoxia-stressed soybean plants. We also analyzed the relative expression of 3 target genes after normalizing with the 2 best CRGs.

## MATERIAL AND METHODS

### Plant growth and experimental design

Soybean seeds from cultivar BR 4 were germinated on moist filter paper and maintained in a growth chamber at  $25^{\circ} \pm 1^{\circ}\text{C}$  and approximately 90% relative humidity. The experiment was carried out in a greenhouse using a hydroponic system. Plantlets were transferred to polystyrene supports, which were placed on blue plastic boxes (59 x 38 x 16 cm) containing 36 L Hoagland solution (Hoagland and Arnon, 1950). Plantlets were grown under natural daylight at  $27^{\circ} \pm 3^{\circ}\text{C}$  (light) and  $24^{\circ} \pm 2^{\circ}\text{C}$  (dark), with humidity ranging from  $64 \pm 9$  (light) to  $71 \pm 5\%$  (dark) until the V1 stage (Fehr et al., 1971). The nutrient solution was kept at pH 6.6 and constantly aerated with atmospheric air. Plantlets were placed in a randomized block design comprising 2 oxygen conditions and 3 treatment sampling times, each with 3 biological replicates (4 plantlets per replicate).

At the V1 stage, the plantlets were randomly divided into 2 groups: plants under normoxia (the control condition with plantlets under constant atmospheric aeration) and plantlets under hypoxia (the nutrient solution was constantly bubbled with nitrogen gas). The oxygen concentration was determined before the start of the experiment. The boxes were sealed to avoid reoxygenation of the nutrient solutions. The oxygen dissolved in the nutrient solution was measured using a portable oxymeter (HI 9146, Hanna). Normoxic conditions contained oxygen concentrations between 7.1 and 7.8 mg/L, whereas in the hypoxic group, the oxygen concentration was between 0.1 and 0.3 mg/L (Figure S1). Plantlets were maintained under normoxic (control plants) or hypoxic conditions for 0.5, 4, and 28 h. At each time point, root tissues were collected and immediately frozen in liquid nitrogen before being stored at  $-80^{\circ}\text{C}$ .

### Total RNA extraction and cDNA synthesis

Total RNA was extracted from root tissues using Trizol reagent (Invitrogen) according to manufacturer instructions. RNA concentration and purity were measured using a spectrophotometer (NanoDrop, ND-1000), and the integrity of the molecules was analyzed on 1% agarose gels stained with ethidium bromide. After DNase (Invitrogen) treatment, 5  $\mu\text{g}$  total RNA was reverse transcribed into cDNA using the SuperScript First Strand Synthesis System from an RT-PCR Kit (Invitrogen). The total RNA samples were primed with 500 ng oligo(dT)12-18 and 50 ng random hexamers following the protocol recommended by the manufacturer. Genomic DNA contamination was not observed after DNase treatment. Primers spanning an intron for  $\beta$ -actin amplification were used as controls (data not shown).

### RT-qPCR

The elongation factor 1- $\beta$  (*ELF1B*), alpha tubulin 5 (*TUA*), alpha tubulin 2 (*TUA2*), actin 11 (*ACTB*), and 18S rRNA (*18S*) genes were chosen as CRGs (Table 1). Primer sequences of the CRGs were obtained from Byfield et al. (2006), Jian et al. (2008), and Stolf-Moreira et al. (2011) studies. The target genes, alcohol dehydrogenase 1 (*ADH*), sucrose synthase 4 (*SUS4*), and ascorbate peroxidase 2 (*APX2*), were selected because of their responsiveness to flooding stress (Liu et al., 2005; Bieniawska et al., 2007; Shi et al., 2008). To identify the orthologous genes in soybeans, the amino acid sequences of the *Arabidopsis thaliana* genes were obtained from The Arabidopsis Information Resource (TAIR) database (Swarbreck et al., 2008) at

<http://arabidopsis.org> and searched using the BLASTp tool (Ye et al., 2006) in the Phytozome database (Schmutz et al., 2010). The selected soybean sequences were used to query the TAIR database using BLASTp to ensure that the most similar sequences were chosen.

**Table 1.** Candidate reference and target genes analyzed in soybean roots by RT-qPCR.

Name	Gene locus	<i>Arabidopsis</i> ortholog locus	<i>Arabidopsis</i> locus description	Function
Candidate reference genes				
<i>ELF1B</i>	Glyma02g44460	AT5G19510	Eukaryotic elongation factor 1-beta	Translational elongation
<i>TUA</i>	Glyma05g29000	AT5G19780	Alpha tubulin 5	Structural constituent of cytoskeleton
<i>TUA2</i>	Glyma20g27280	AT1G04820	Alpha tubulin 2	Structural constituent of cytoskeleton
<i>ACTB</i>	Glyma15g05570	AT5G19510	Actin 11	Structural constituent of cytoskeleton
<i>18S</i>	M.L. <sup>A</sup>			Translation
Target genes				
<i>ADH</i>	Glyma04g41990	AT1G77120	Alcohol dehydrogenase 1	Fermentation
	Glyma06g12780	AT1G77120		
<i>SUS4</i>	Glyma13g17420	AT3G43190	Sucrose synthase 4	Transferase activity
<i>APX2</i>	Glyma12g07780	AT1G07890	Ascorbate peroxidase 2	Hydrogen peroxide scavenging

<sup>A</sup>Multiple loci; 18S ribosomal RNA (accession No. X02623.1).

A global alignment of the different isoforms was performed using Clustal W2 to identify unique sites for primer design (Larkin et al., 2007). Primers were designed for the 3' untranslated region with Primer Express 3.0 (Applied Biosystems). Putative primer dimers and hairpin structures were analyzed using the program Vector NTI 11.0 (Invitrogen). Primers were analyzed with BLAST, using the nucleotide sequence against the Phytozome database (Schmutz et al., 2010) to assess their specificity for the selected gene. A standard curve was created from serial dilutions of a cDNA pool, which were analyzed in duplicate. The amplification reaction efficiency was calculated using the formula  $E = 10^{-1/slope}$ , and the best primer concentrations were determined for each gene amplification (Table 2).

RT-qPCRs were performed using 6.9 ng cDNA, 60-300 nM of each forward and reverse primer (Table 2), 6.5  $\mu$ L SYBR Green Master Mix (Invitrogen), and ultrapure water to a final volume of 12.5  $\mu$ L. Amplification reactions were performed in a 7300 Real Time Systems thermocycler (Applied Biosystems) under the following conditions: 50°C for 2 min; 95°C for 10 min; followed by 40 cycles of 95°C for 15 s and 60°C for 1 min. The threshold and the baseline were adjusted manually for each gene using the 7300 System SDS RQ Study software (Applied Biosystems). The expression levels of the CRGs and the target genes were analyzed from the technical triplicates of each of the 3 biological replications. In order to confirm that the samples were not contaminated, a control sample from the RT-qPCR and samples with no template were also assayed. Dissociation curves and electrophoresis analyses confirmed the amplification specificity and lack of contamination (data not shown).

### Analysis of gene expression stability

The logarithm threshold cycle (Ct) values for the CRGs were converted to nonnormalized relative quantities ( $Q$ ; linear scale). Here,  $Q = E^{\Delta Ct}$ , where  $E$  is the amplification efficiency and  $\Delta Ct$  is the lowest Ct from the data set minus the sample Ct. The relative quantities were analyzed using the NormFinder v0.953 (Andersen et al., 2004) and GeNorm v3.5 (Vandesompele et al., 2002) softwares to rank the genes according to their expression stability. The NormFinder algorithm is based

on an analysis of variance model. It estimates the variation in the expression of the CRGs within and among the sample sets and generates a value for the expression stability of each CRG. The best reference gene has the lowest stability value (lowest variation in gene expression). GeNorm calculates a value (M value) that represents the average pairwise variations between a given gene and all other control genes. After successive stepwise exclusions of the genes with the highest M values (unstable gene expression), the gene pair with the most stable expression (lowest M values) was obtained.

GeNorm was also used in this study to analyze the optimal number of reference genes to normalize the expression of target genes. The estimation is based on the variation between 2 normalization factors (the geometric mean value of the best reference genes), with stepwise inclusions of less stable reference genes ( $V_{n/n+1}$ ). According to Vandesompele et al. (2002), a cut-off value of 0.150 is appropriate. Below this value, the inclusion of another reference gene has no significant effect.

### Estimating the expression and significance analyses

The relative expression and statistical significance of the CRGs and target genes were analyzed with the REST 2009 v2.0.13 software (Pfaffl et al., 2002). The REST algorithm allows the input of different amplification efficiencies for the reference and target genes and also accepts multiple reference genes, which increases the accuracy of the results. The statistical significance of the expression levels was evaluated through randomization (10,000 interactions) and bootstrapping the data.

## RESULTS

### Expression level and stability of the CRGs

The amplification efficiency ( $E$ ) and squared correlation coefficient ( $R^2$ ) of the primers were determined using a standard curve made from a serially diluted cDNA pool (Table 2). The  $E$  of the CRGs and the target genes ranged from 87.3 (*TUA*) to 97.6% (*18S*) and from 90.0 (*ADH*) to 99.9% (*SUS4*) (Table 2). An  $E$  equal to 100% indicates amplicon duplication after each cycle of PCR amplification during the exponential phase. The  $R^2$  was higher than 0.992 for all primers, except for *ADH* ( $R^2 = 0.980$ ) (Table 2). An  $R^2$  of 1 represents an optimal adjustment in Ct values in relation to the cDNA quantities for each sample (serial dilution).

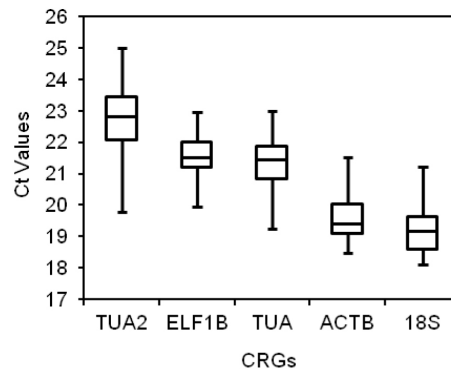
**Table 2.** Primers for the candidate reference and target genes.

Name	Forward primer sequence	Reverse primer sequence	nM	Amplicon length (bp)	E (%)	$R^2$
<i>ELF1B</i> <sup>A</sup>	5'-GTTGAAAAGCCAGGGGACA-3'	5'-TCTTACCCCTTGAGCGTGG-3'	60	118	95.5	0.998
<i>TUA</i> <sup>A</sup>	5'-AGGTCGGAAACTCCTGCTGG-3'	5'-AAGGTGTTGAAGGCGTCGTG-3'	300	159	87.3	0.999
<i>TUA2</i> <sup>A</sup>	5'-CCTCGTTCGAATTCGCTTTTIG-3'	5'-CAACTGTCTTGTCGCTTGCCAT-3'	100	161	95.6	0.997
<i>ACTB</i> <sup>B</sup>	5'-GAGCTATGAATTGCCTGATGG-3'	5'-CGTTTCATGAATTCAGTAGC-3'	60	118	94.4	1.000
<i>18S</i> <sup>C</sup>	5'-AAACGGCTACCACATCCAAG-3'	5'-CCTTCAATGGATCCATCGTTA-3'	100	155	97.6	0.993
<i>ADH</i>	5'-CCTGTGAGGGACTGAGTTCTCTTTC-3'	5'-ACAGCAGCAAAGGTCCTATCATTCT-3'	300	148	90.0	0.980
<i>SUS4</i>	5'-TGTTGTTGCATGATTGGAICTTG-3'	5'-CACGGCTAAAATTGAATTGATGG-3'	60	80	99.9	0.999
<i>APX2</i>	5'-CACGGTGCCATAATATTCTCTC-3'	5'-CAACCCAACTCCAATCATCAC-3'	300	134	91.7	0.995

Primers obtained from: A (Jian et al. 2008), B (Byfield et al. 2006), C (Stolf-Moreira et al. 2011). Primer concentration (nM), amplicon length (bp), amplification efficiency (E), and correlation coefficient ( $R^2$ ) are shown for each gene. Primer concentration was determined by optimizing the amplification efficiency reaction, which is calculated as  $E = 10^{-1/slope}$ , where *slope* is the angular coefficient of the linear regression model fitted over Ct values versus log input of cDNA, measured in duplicate.



The transcript abundance of each CRG was obtained from the average of the technical triplicates amplified for each biological sample (composed of 4 plantlets collected in bulk) and is shown as the Ct value (Figure 1). Across all 18 biological samples, the Ct values ranged from 18.11 (*18S*) to 25.02 (*TUA2*). The *18S* gene showed the highest transcript level, with a median Ct of 19.15, while the *TUA2* gene presented the lowest, with a median Ct of 22.81. The *ELF1B* gene showed the smallest amplitude of variation in Ct values (3.02 cycles) and the lowest variation in relation to the median (an amplitude of 0.82 cycles in the distribution of 50% of the values relative to the median). The second ranked gene was *ACTB* (3.06; 0.92 cycles), whereas *TUA2* presented the largest amplitude of variation in Ct (5.24 cycles) and dispersion relative to the median (1.39 cycles) (Figure 1). For *18S* and *TUA*, the amplitudes of variation detected were 3.11 and 3.73 cycles, respectively, and the dispersion of the Ct values relative to the median were 1.03 and 1.01, respectively (Figure 1).



**Figure 1.** Variation in the transcript abundance of the candidate reference genes. The box plots represent the Ct variation measured for the *TUA2*, *ELF1B*, *TUA*, *ACTB*, and *18S* genes.

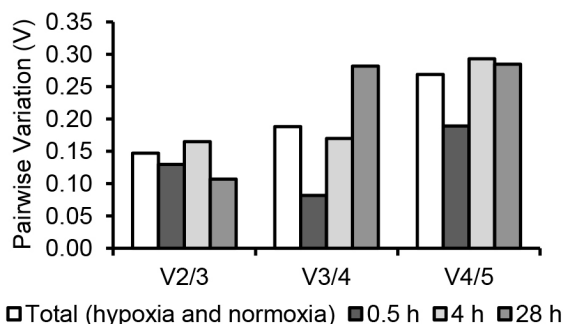
The expression stability of the CRGs was analyzed with the NormFinder and GeNorm software using the following data sets: 1) all samples under both hypoxia and normoxia (subsets 1A, 1B, and 1C represent samples under hypoxia and normoxia for 0.5, 4, and 28 h, respectively), 2) all samples under hypoxia (subsets 2A, 2B, and 2C represent hypoxic samples of 0.5 vs 4 h, 0.5 vs 28 h, and 4 vs 28 h, respectively), 3) all samples under normoxia (subsets 3A, 3B, and 3C represent normoxic samples of 0.5 vs 4 h, 0.5 vs 28 h, and 4 vs 28 h, respectively). For the CRG expression analysis, normalization was not used and was performed only for subsets 1A, 1B, 1C, 2A, 2B, 2C, 3A, 3B, and 3C because REST provides significance tests between 2 sample sets only. Using GeNorm, in all 12 data sets, the least stable and the 3 most stable CRGs had M values below 1.0 and 0.5, respectively (data not shown). *ELF1B* and *ACTB* were the most frequently ranked CRGs that showed higher expression stability among the sample sets (Table 3). In contrast, *18S* was the least stable gene, except in subset 3B (0.5 vs 28 h under normoxia) as analyzed by NormFinder. The rankings of some CRGs contradicted the REST test (Table 3; [Table S1](#)). In the subsets 2A (0.5 vs 4 h under hypoxia) and 2C (4 vs 28 h under hypoxia), significant variation was detected for all genes, except for the least stable predicted gene (*18S*). This result is the opposite of what was expected, as a significant and greater variation existed only for the less stable predicted CRGs in subsets 1B and 1C (4 and 28 h under hypoxia vs normoxia, respectively). In addition, some CRGs were differentially expressed even under normoxia [subsets 3B (0.5 vs 28 h under normoxia) and 3C (4 vs 28 h under normoxia)].

**Table 3.** Expression stability of candidate reference genes.

	NormFinder	GeNorm	NormFinder	GeNorm	NormFinder	GeNorm	NormFinder	GeNorm
Hypoxia and Normoxia								
Rank	(1) Total		(1A) 0.5 h		(1B) 4 h		(1C) 28 h	
1	<i>ELF1B</i>	<i>ELF1B ACTB</i>	<i>ACTB</i>	<i>ACTB ELF1B</i>	<i>TUA</i>	<i>ELF1B ACTB</i>	<i>ELF1B</i>	<i>ELF1B TUA</i>
2	<i>ACTB</i>		<i>ELF1B</i>		<i>ACTB</i>		<i>TUA</i>	
3	<i>TUA</i>	<i>TUA</i>	<i>TUA2</i>	<i>TUA2</i>	<i>ELF1B</i>	<i>TUA</i>	<i>ACTB</i>	<i>ACTB</i>
4	<i>TUA2</i>	<i>TUA2</i>	<i>TUA</i>	<i>TUA</i>	<i>TUA2</i>	<i>TUA2</i>	<i>TUA2***</i>	<i>TUA2***</i>
5	<i>18S</i>	<i>18S</i>	<i>18S</i>	<i>18S</i>	<i>18S***</i>	<i>18S***</i>	<i>18S***</i>	<i>18S***</i>
Hypoxia								
	(2) Total		(2A) 0.5 and 4 h		(2B) 0.5 and 28 h		(2C) 4 and 28 h	
1	<i>ELF1B</i>	<i>ELF1B ACTB</i>	<i>ELF1B*</i>	<i>ELF1B* TUA*</i>	<i>ACTB</i>	<i>ACTB ELF1B</i>	<i>ELF1B***</i>	<i>ELF1B*** TUA2***</i>
2	<i>ACTB</i>		<i>TUA*</i>		<i>ELF1B</i>		<i>TUA2***</i>	
3	<i>TUA</i>	<i>TUA</i>	<i>ACTB***</i>	<i>ACTB***</i>	<i>TUA</i>	<i>TUA</i>	<i>ACTB***</i>	<i>ACTB***</i>
4	<i>TUA2</i>	<i>TUA2</i>	<i>TUA2***</i>	<i>TUA2***</i>	<i>TUA2***</i>	<i>TUA2***</i>	<i>TUA*</i>	<i>TUA*</i>
5	<i>18S</i>	<i>18S</i>	<i>18S</i>	<i>18S</i>	<i>18S</i>	<i>18S</i>	<i>18S</i>	<i>18S</i>
Normoxia								
	(3) Total		(3A) 0.5 and 4 h		(3B) 0.5 and 28 h		(3C) 4 and 28 h	
1	<i>ACTB</i>	<i>ACTB ELF1B</i>	<i>TUA2</i>	<i>ELF1B ACTB</i>	<i>ELF1B***</i>	<i>ELF1B*** ACTB</i>	<i>ACTB</i>	<i>ACTB ELF1B</i>
2	<i>ELF1B</i>		<i>ACTB</i>		<i>TUA***</i>		<i>ELF1B</i>	
3	<i>TUA</i>	<i>TUA</i>	<i>TUA</i>	<i>TUA2</i>	<i>ACTB</i>	<i>TUA***</i>	<i>TUA***</i>	<i>TUA***</i>
4	<i>TUA2</i>	<i>TUA2</i>	<i>ELF1B</i>	<i>TUA</i>	<i>18S</i>	<i>TUA2*</i>	<i>TUA2***</i>	<i>TUA2***</i>
5	<i>18S</i>	<i>18S</i>	<i>18S</i>	<i>18S</i>	<i>TUA2*</i>	<i>18S</i>	<i>18S</i>	<i>18S</i>

\* $P < 0.05$ , \*\*\* $P < 0.001$  indicate significant differences, as determined using the REST (2009) software, between treatments (hypoxia/normoxia) at each time point (1A, 1B, and 1C) and between different time points for each treatment (2A, 2B, 2C, 3A, 3B, and 3C). The expression stability was analyzed by NormFinder and GeNorm, and the significance test performed with REST (2009).

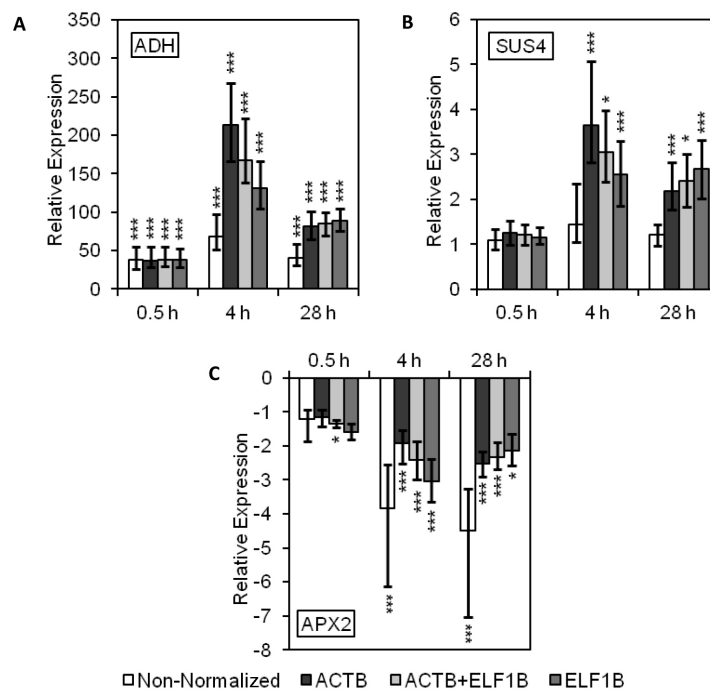
The appropriate number of reference genes for normalization was estimated with GeNorm by using pairwise variation ( $V$ ) between 2 normalization factors,  $NF_n$  and  $NF_{n+1}$ . The threshold was set to 0.150 because no additional reference genes were required below this value (Vandesompele et al., 2002). As shown in Figure 2, two genes were appropriate for normalizing the relative expression of the total sample set ( $V_{2/3} = 0.147$ ). The addition of a third gene increased  $V$  above the threshold value ( $V_{3/4} = 0.188$ ). The use of 2 reference genes was justifiable because, when analyzing the subsets 1B and 1C (samples of 4 and 28 h, respectively), the addition of a third gene did not reduce the  $V$  of subset 1B to a value below the threshold ( $V_{2/3} = 0.165$ ;  $V_{3/4} = 0.170$ ), and it increased the  $V$  of subset 1C from a  $V_{2/3}$  of 0.107 to a  $V_{3/4}$  of 0.285 (Figure 2).



**Figure 2.** Pairwise variation ( $V$ ) of the candidate reference genes in the total sample set and in the 0.5-, 4-, and 28-h subsets under hypoxia and normoxia.

### Expression of the target genes

The *ELF1B* and *ACTB* reference genes, which ranked as the most stably expressed genes in this study, were used to evaluate the expression of the target genes *ADH*, *SUS4*, and *APX2*. The transcript abundance at the 3 hypoxia time points was compared with the respective normoxia time points (e.g., 4 h hypoxic vs 4 h normoxic treatments). This procedure, which differs from some previous RT-qPCR studies that used only a 0 h time point as the control for gene expression, facilitates the specification of the effects of the experimental treatment on target gene expression. Comparing stressed and unstressed samples at the same time point can remove putative additive effects, such as gene-intrinsic effects (e.g., circadian rhythm), differences in developmental stages among individuals, or any unknown variation between the time points. Normalization enabled us to observe an increase in the expression of *ADH* and *SUS4*, as well as a reduction in the standard error bars for the downregulated gene *APX2* (Figure 3). Common normalization effects occurred between up- and downregulated genes. A statistically significant difference was detected in *SUS4* expression at 4 and 28 h and in *APX2* expression at 0.5 h only after normalization. *SUS4* was differentially expressed when the reference genes *ACTB* or *ELF1B* were used for normalization, as well as when the NF (*ACTB* + *ELF1B*) was used. However, *APX2* expression was significant only when expression was normalized using the NF.



**Figure 3.** Relative quantification of the *ADH* (A), *SUS4* (B), and *APX2* (C) genes. Raw data were normalized using the *ELF1B* and *ACTB* reference genes. The transcript abundance of the target genes from plants subjected to hypoxic conditions for different periods of time was compared with the respective controls (normoxic condition). Relative gene expression: randomization value ( $N = 10,000$ ), error bars, \* $P < 0.05$ , \*\*\* $P < 0.001$ .



## DISCUSSION

RT-qPCR has been widely used to measure gene expression. However, because of its high sensitivity, cDNA quality and concentration must be accurately evaluated to reduce experimental error. Random and systematic errors can be reduced by using reference genes that are stably expressed across experimental treatments. This is important since reference genes are subjected to the same errors as target genes during all stages of cDNA preparation (Huggett et al., 2005). As shown by Caradec et al. (2010) and our study, it is necessary to evaluate CRGs to avoid biased results, which might lead to misinterpretation.

Tucker et al. (2011), who sought to understand which regulatory elements of cell separation associated with aerenchyma development were shared with those involved in the formation of the cyst nematode syncytium, compared the expression of 3 reference genes in soybean root tissues under flooding. Likewise, Narsai et al. (2010) analyzed a dataset from rice transcriptomes under different experimental conditions, including embryo and coleoptile tissues subjected to low oxygen conditions, in order to identify reference genes for use in biotic- and abiotic-stressed rice plants. Nevertheless, we did not find any published reports that systematically analyzed the expression stability of CRGs in plants under hypoxia. To determine the genes most appropriate for accurately normalizing relative gene expression that could be used in forthcoming studies of flooding-stressed soybean plants, 5 CRGs were evaluated in soybean roots under hypoxic conditions at 3 different periods of time.

The CRGs were chosen based on studies of gene expression under hypoxic conditions, such as the *18S* gene analyzed in *Arabidopsis* (Liu et al., 2005) and soybeans (Komatsu et al., 2010); a number of genes validated as soybean reference genes, including *ELF1B*, *TUA*, and *TUA2* in different tissues at different developmental stages (Jian et al., 2008); and *ACTB* in drought-stressed soybean roots and leaves (Stolf-Moreira et al., 2011).

Compared with other studies of expression stability, such as adventitious rooting in *Eucalyptus globulus* (de Almeida et al., 2010) and *Populus* (Xu et al., 2011), in the different soybean tissues (Jian et al., 2008; Hu et al., 2009) and drought-stressed soybean roots and leaves (Stolf-Moreira et al., 2011), *18S* transcripts were the most abundant (de Almeida et al., 2010; Stolf-Moreira et al., 2011). *TUA2* [called TUB by Jian et al. (2008)] and *ELF1B* exhibited the largest (Jian et al., 2008) and smallest (Hu et al., 2009) variations in Ct values, respectively. Here, *ELF1B* and *ACTB*, followed by *TUA*, were ranked as the most stable reference genes according to our GeNorm and NormFinder analyses, whereas the *18S* gene was found to be the most unstable. Greater expression stability has already been demonstrated for the plant genes *ELF1B* (Jian et al., 2008; Hu et al., 2009), *ACTB* (Stolf-Moreira et al., 2011), and *TUA* (Jian et al., 2008; de Almeida et al., 2010). Previous plant gene expression studies have classified the *18S* gene among the most stable (Stolf-Moreira et al., 2011; Xu et al., 2011), whereas in other studies, it is ranked as the least stable (de Almeida et al., 2010), which is consistent with our findings.

The present study was designed using a dataset composed of 3 levels (total of samples, 2 oxygen availability conditions, and 3 time points). Therefore, NormFinder was the most suitable program for evaluating CRGs because it calculates both intragroup and intergroup variations. GeNorm does not take intergroup variation into account, but instead reflects the overall variation observed in the sample set. NormFinder and GeNorm performed similarly with our data (Table 3), which has also been observed in other studies (Hu et al., 2009; Rytönen et al., 2010). Although not observed in the hypoxia vs normoxia comparisons (1A, 1B, and 1C), the REST test

contradicted the CRG ranking when comparing time points, primarily in the 2A and 2C hypoxic subsets (Table 3, [Table S1](#)). This result demonstrates the limitations of NormFinder for estimating variation between certain groups. Rytönen et al. (2010) also observed discrepancies between NormFinder and significance testing with respect to the analysis of variation among groups.

The two best CRGs identified in this study (*ELF1B* and *ACTB*) were used to normalize the target genes *ADH*, *SUS4*, and *APX2* (Figure 3). The *ADH* gene is often used to validate experimental treatments of hypoxia (Liu et al., 2005; Kreuzwieser et al., 2009; Tucker et al., 2011). Our results showed that *ADH* was highly expressed at all time points under hypoxic conditions, regardless of the normalization reaction (Figure 3). However, only after normalization with the reference gene *ACTB* or the NF (geometric mean of the expression values of *ACTB* plus *ELF1B*) we were able to observe that *ADH* was induced at a lower level at 0.5 compared to 28 h, and its highest expression was observed at 4 h (Figure 3). A similar expression pattern was verified in *Arabidopsis* plants under hypoxic conditions (Liu et al., 2005). The gene *SUS4* was found to be differentially expressed under hypoxic conditions only after normalization. This result further indicates that normalization is crucial in RT-qPCR experiments. Without normalization, we would not have been able to detect the induction of *SUS4* triggered by the hypoxia treatment, as has been observed in the roots of other species, such as *Populus* (Kreuzwieser et al., 2009), *Arabidopsis* (Baud et al., 2009), maize (Zou et al., 2010), and cotton (Christianson et al., 2010). Similarly, only after normalization with specific endogenous control genes we found that *APX2* expression was downregulated at 0.5 h under hypoxia, a decrease that was maintained throughout the entire experiment (Figure 3). This finding corroborates the results of Shi et al. (2008), who observed a decrease in the *APX2* transcript and protein levels in soybean roots after 3 days under hypoxic conditions.

To our knowledge, this study is the first to evaluate soybean reference genes for use as internal controls in RT-qPCR studies while quantifying transcripts expressed in soybean root tissues under low oxygen conditions. Of the genes studied here, *ELF1B* and *ACTB* were found to be the best for RT-qPCR normalization. Using these genes, we were able to better estimate the relative expressions of *ADH*, *SUS4*, and *APX2*, which are known markers of hypoxia. *ELF1B* and *ACTB* are potential reference genes that should be validated in other plants subjected to hypoxic stress. Although tools for the analysis of expression stability are useful for ranking CRGs, attention should be paid to intergroup variation by using standalone tools to analyze the statistical significance of gene expression data.

## ACKNOWLEDGMENTS

This research is part of T.J. Nakayama MSc thesis from Pós-Graduação em Genética e Biologia Molecular (Universidade Estadual de Londrina, Brazil); T.J. Nakayama was supported by a fellowship from CAPES. The authors are also grateful to the Embrapa fund for research support.

## [Supplementary material](#)

## REFERENCES

- Andersen CL, Jensen JL and Orntoft TF (2004). Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res.* 64: 5245-5250.

- Baud S, Vaultie MN and Rochat C (2009). Structure and expression profile of the sucrose synthase multigene family in *Arabidopsis*. *Exp. Biol.* 55: 397-409.
- Bieniawska Z, Paul Barratt DH, Garlick AP, Thole V, et al. (2007). Analysis of the sucrose synthase gene family in *Arabidopsis*. *Plant J.* 49: 810-828.
- Böhm-Hofstätter H, Tschernutter M and Kunert R (2010). Comparison of hybridization methods and real-time PCR: their value in animal cell line characterization. *Appl. Microbiol. Biotechnol.* 87: 419-425.
- Byfield GE, Xue H and Upchurch RG (2006). Two genes from soybean encoding soluble D9 stearyl-ACP desaturases. *Crop Sci.* 46: 840-846.
- Caradec J, Sirab N, Keumeugni C, Moutereau S, et al. (2010). 'Desperate house genes': the dramatic example of hypoxia. *Br. J. Cancer* 102: 1037-1043.
- Christianson JA, Llewellyn DJ, Dennis ES and Wilson IW (2010). Global gene expression responses to waterlogging in roots and leaves of cotton (*Gossypium hirsutum* L.). *Plant Cell Physiol.* 51: 21-37.
- de Almeida MR, Ruedell CM, Ricachenevsky FK, Sperotto RA, et al. (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: 73.
- Fehr WR, Caviness CE, Burmood DT and Pennington JS (1971). Stage of development description for soybeans [*Glycine max* (L.) Merrill]. *Crop Sci.* 11: 929-931.
- Guénin S, Mauriat M, Pelloux J, Van Wuytswinkel O, et al. (2009). Normalization of qRT-PCR data: the necessity of adopting a systematic, experimental conditions-specific, validation of references. *J. Exp. Bot.* 60: 487-493.
- Hoagland DR and Arnon DI (1950). The water-culture method for growing for plants without soil. *AES Circ.* 347: 1-32.
- Hu R, Fan C, Li H, Zhang Q, et al. (2009). Evaluation of putative reference genes for gene expression normalization in soybean by quantitative real-time RT-PCR. *BMC Mol. Biol.* 10: 93.
- Huggett J, Dheda K, Bustin S and Zumla A (2005). Real-time RT-PCR normalisation; strategies and considerations. *Genes Immun.* 6: 279-284.
- Jian B, Liu B, Bi Y, Hou W, et al. (2008). Validation of internal control for gene expression study in soybean by quantitative real-time PCR. *BMC Mol. Biol.* 9: 59.
- Komatsu S, Kobayashi Y, Nishizawa K, Nanjo Y, et al. (2010). Comparative proteomics analysis of differentially expressed proteins in soybean cell wall during flooding stress. *Amino. Acids* 39: 1435-1449.
- Kreuzwieser J, Hauberg J, Howell KA, Carroll A, et al. (2009). Differential response of gray poplar leaves and roots underpins stress adaptation during hypoxia. *Plant Physiol.* 149: 461-473.
- Larkin MA, Blackshields G, Brown NP, Chenna R, et al. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947-2948.
- Lasanthi-Kudahettige R, Magneschi L, Loreti E, Gonzali S, et al. (2007). Transcript profiling of the anoxic rice coleoptile. *Plant Physiol.* 144: 218-231.
- Liu F, Vantoi T, Moy LP, Bock G, et al. (2005). Global transcription profiling reveals comprehensive insights into hypoxic response in *Arabidopsis*. *Plant Physiol.* 137: 1115-1129.
- Mustroph A, Lee SC, Oosumi T, Zanetti ME, et al. (2010). Cross-kingdom comparison of transcriptomic adjustments to low-oxygen stress highlights conserved and plant-specific responses. *Plant Physiol.* 152: 1484-1500.
- Narsari R, Ivanova A, Ng S and Whelan J (2010). Defining reference genes in *Oryza sativa* using organ, development, biotic and abiotic transcriptome datasets. *BMC Plant Biol.* 10: 56.
- Nolan T, Hands RE and Bustin SA (2006). Quantification of mRNA using real-time RT-PCR. *Nat. Protocol.* 1: 1559-1582.
- Pfaffl MW, Horgan GW and Dempfle L (2002). Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Res.* 30: e36.
- Rytkönen KT, Renshaw GM, Ashton KJ, Williams-Pritchard G, et al. (2010). Elasmobranch qPCR reference genes: a case study of hypoxia preconditioned epaulette sharks. *BMC Mol. Biol.* 11: 27.
- Schmutz J, Cannon SB, Schlueter J, Ma J, et al. (2010). Genome sequence of the palaeopolyploid soybean. *Nature* 463: 178-183.
- Shi F, Yamamoto R, Shimamura S, Hiraga S, et al. (2008). Cytosolic ascorbate peroxidase 2 (cAPX 2) is involved in the soybean response to flooding. *Phytochemistry* 69: 1295-1303.
- Stolf-Moreira R, Lemos EGM, Abdelnoor RV and Beneventi MA (2011). Identification of reference genes for expression analysis by real - time quantitative PCR in drought-stressed soybean. *Pesq. Agropec. Bras.* 46: 58-65.
- Swarbreck D, Wilks C, Lamesch P, Berardini TZ, et al. (2008). The *Arabidopsis* Information Resource (TAIR): gene structure and function annotation. *Nucleic Acids Res.* 36: D1009-D1014.
- Tucker ML, Murphy CA and Yang R (2011). Gene expression profiling and shared promoter motif for cell wall-modifying proteins expressed in soybean cyst nematode-infected roots. *Plant Physiol.* 156: 319-329.

- Vandesompele J, De PK, Pattyn F, Poppe B, et al. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* 3: RESEARCH0034.
- Xu M, Zhang B, Su X, Zhang S, et al. (2011). Reference gene selection for quantitative real-time polymerase chain reaction in *Populus*. *Anal. Biochem.* 408: 337-339.
- Ye J, McGinnis S and Madden TL (2006). BLAST: improvements for better sequence analysis. *Nucleic Acids Res.* 34: W6-W9.
- Zentilin L and Giacca M (2007). Competitive PCR for precise nucleic acid quantification. *Nat. Protoc.* 2: 2092-2104.
- Zou X, Jiang Y, Liu L, Zhang Z, et al. (2010). Identification of transcriptome induced in roots of maize seedlings at the late stage of waterlogging. *BMC Plant Biol.* 10: 189.