

Selection of reliable reference genes for quantitative real-time polymerase chain reaction studies in maize grains

Vanessa Galli · Rafael da Silva Messias ·
Sérgio Delmar dos Anjos e Silva ·
Cesar Valmor Rombaldi

Received: 8 March 2013 / Revised: 25 July 2013 / Accepted: 26 August 2013 / Published online: 8 September 2013
© Springer-Verlag Berlin Heidelberg 2013

Abstract

Key message The stability of candidate reference genes was evaluated in maize landrace varieties and during multiple grain developmental stages to evaluate the expression of carotenoid-related genes by RT-qPCR for application to maize biofortification.

Abstract Vitamin A deficiency affects millions of children worldwide; therefore, increasing the content of vitamin A precursors in maize grains is of interest. The study of the expression of genes involved in the carotenoid biosynthetic pathway in maize grains has provided useful information for metabolic engineering approaches. However, reliable results using real-time quantitative polymerase chain reaction (RT-qPCR) experiments are dependent on the use of the appropriate reference genes. In this study, we utilized geNorm and NormFinder softwares to identify the most stably expressed candidate reference genes in samples from seven stages of grain development and from eight landrace varieties. The results of the analysis

performed using geNorm indicated that tubulin (*TUB*) and actin (*ACT*) were the most suitable reference genes among all experimental conditions, while glyceraldehyde-3-phosphate dehydrogenase gene (*GAPDH*) showed the least stability. The same result was obtained with the NormFinder software. The minimum number of genes required in each experimental condition to normalize the gene expression data was also determined by geNorm. The expression of phytoene synthase gene (*PSY1*), the first enzyme in the carotenoid biosynthetic pathway, was overestimated when the least stable candidate gene (*GAPDH*) was used as the internal control instead of the most stable gene pair (*ACT* + *TUB*), thus highlighting the importance of validating reference genes before conducting a RT-qPCR experiment to obtain accurate results. This study is the first survey of the stability of genes for use as reference genes to normalize RT-qPCR data from maize landraces during multiple stages of grain development.

Keywords Reference genes · *Zea mays* L. · RT-qPCR · GeNorm · NormFinder

Communicated by H. Jones.

V. Galli (✉) · R. da Silva Messias · S. D. dos Anjos e Silva
Empresa Brasileira de Pesquisa Agropecuária de Clima
Temperado, Rodovia BR 396, Km 78 Postal code 403, Pelotas,
RS CEP 96001-970, Brazil
e-mail: vane.galli@yahoo.com.br

V. Galli
Universidade Federal Do Rio Grande Do Sul, Centro de
Biotecnologia, Av. Bento Gonçalves 9500, Campus do Vale,
Porto Alegre, RSCaixa Postal 15005, CEP 91501-970, Brazil

C. V. Rombaldi
Universidade Federal de Pelotas, Faculdade de Agronomia
Eliseu Maciel, Campus Universitário s/n, Pelotas,
RSCaixa Postal 354, CEP 96010-900, Brazil

Introduction

Carotenoids are a class of isoprenoid pigments that provide nutritional and functional values as provitamin A and non-provitamin A compounds (Berardo et al. 2009). Several efforts have been made to specifically address increasing the levels of vitamin A precursors in food (Aluru et al. 2008; Naqvi et al. 2009; Vallabhaneni and Wurtzel 2009) because vitamin A deficiency affects millions of children worldwide (Fraser and Bramley 2004). Maize has been considered an important target for biofortification with vitamin A precursors because it is an important staple food,

especially in developing countries (Berardo et al. 2009). The study of the expression of genes involved in the carotenoid biosynthetic pathway during maize grain development and in different varieties has provided useful information regarding metabolic engineering approaches for the biofortification of the maize crop (Aluru et al. 2008; Naqvi et al. 2009; Vallabhaneni and Wurtzel 2009). One of the most studied genes for this purpose is *PSY1*, which encodes phytoene synthase, the first enzyme in the carotenoid biosynthetic pathway. The PSY enzyme has been considered a rate-controlling step in carotenoid accumulation because the *PSY1* expression profile showed 96 % correlation with the carotenoid content in grains of maize varieties at 20 days after pollination (DAP) (Vallabhaneni and Wurtzel 2009).

Real-time PCR (RT-qPCR) has been widely used to analyze gene expression in different organisms and tissues and under multiple conditions because it allows sensitive, specific, and reproducible quantification of nucleic acids (Bustin 2002; Derveaux et al. 2010). However, substantial variations exist in the quality, stability and input of RNA as well as in the efficiency of the cDNA synthesis step and the polymerase chain reactions (PCR); therefore, reliable results are dependent on the accuracy and precision under the experimental conditions used to perform RT-qPCR (Derveaux et al. 2010; Fleige et al. 2006). In an attempt to improve the quality of expression results, the MIQE guidelines (minimum information for publication of quantitative real-time PCR experiments) have been proposed, which include the standardization of experimental design, the quality of RNA isolation, reverse transcription, the design and optimization of oligonucleotides, and data normalization and analysis (Bustin et al. 2009). The selection of a suitable reference gene to normalize the data is an absolute requirement to minimize non-biological variation between samples and obtain accurate results (Derveaux et al. 2010; Gutierrez et al. 2008). Furthermore, comparisons between different studies may be improved by the use of the same reference gene.

A suitable reference gene should be expressed at a constant level in samples, and its expression is assumed to be unaffected by the experimental conditions. Moreover, the reference gene and the target genes should have similar ranges of expression in the samples (Bustin 2002). Usually, they include genes that encode products with functions in maintaining cell wall structure and primary metabolism such as 18S ribosomal RNA (*18S*), actin (*ACT*), tubulin (*TUB*), ubiquitin (*UBI*) and glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*). Nevertheless, the stability among several of these commonly used reference genes is relative, and no single gene shows constant stable expression under all conditions (Radonic et al. 2004; Czechowski

et al. 2005). Thus, to ensure appropriate normalization of RT-qPCR experiments, the validation of reference genes under specific experimental conditions is of great importance. Therefore, several reference genes were recently evaluated for stable expression under specific conditions in various plant species, including *A. thaliana* (Czechowski et al. 2005), soybean (Kulcheski et al. 2010), rice (Qian-Feng et al. 2010), wheat (Paolacci et al. 2009), tomato (Lovdal and Lillo 2009), and others. Currently, only two recent studies have reported the stability of putative reference genes in maize (Chen et al. 2012; Manoli et al. 2012); however, neither of these studies included samples from maize ear developmental series or from different genotypes, which limits further studies involving the transcriptional profile of carotenoid-related genes on these samples and the biofortification of maize grains through metabolic engineering or conventional breeding.

The goal of this study was to examine the stability of five candidate reference genes (*GAPDH*, *ACT*, *TUB*, *18S* and *UBI*) in grains of eight different maize varieties collected at 22 DAP and in the seven stages of maize grain development to suggest a suitable reference gene for gene expression studies in maize.

Materials and methods

Experimental conditions and sampling

Eight landrace varieties (accessions from Embrapa Temperate Agriculture, Pelotas, Brazil) with large variations in grain color, hardness and shape and the hybrid maize variety 30F53 (Pioneer) were grown in a field in triplicate in 10-m plots in four rows from November 2010 to April 2011. Cobs were covered at the beginning of their formation and were manually pollinated to avoid cross pollination. Grain samples of different landrace varieties were collected at 22 DAP. Grain samples of commercial maize hybrid were collected at 0, 10, 13, 16, 19, 21 and 25 DAP. The samples were immediately frozen in liquid nitrogen and stored at -80°C until analysis.

Total RNA extraction and first strand cDNA synthesis

The total RNA of maize grains was isolated using the CTAB (hexadecyltrimethylammonium bromide) protocol described by Chang et al. (1993) with several modifications, including reduction of the milled sample amount to 100 mg and, consequently, reduction in the volume of extraction buffer (without spermidine) to 1.25 mL and a proportional reduction in the volumes of the other reagents, thus allowing all the steps to be performed in microcen-

trifuge tubes. The samples were incubated with ¼ of 10 M LiCl for 1 h at -70°C instead of an overnight incubation at 4°C ; RNA was further precipitated with sodium acetate and ethanol, and washed with 70 % ethanol to remove residual contaminants. Three replicates of RNA extraction were performed. The RNA quality was evaluated using a 1 % agarose gel after electrophoresis and by spectrometry using the A_{260}/A_{280} and A_{260}/A_{230} ratios. RNA concentration was measured in a Qubit[®] fluorometer (Invitrogen). Total RNA (1 μg) was digested with 1 U DNase I and DNase 1 \times reaction buffer and reverse transcribed using the M-MLV enzyme and oligo-dT primers according to the manufacturer's instructions (Invitrogen).

Primer design

Five candidate reference genes, *GAPDH* (Magneschi et al. 2009), *ACT* (Maroufi et al. 2010; Qian-Feng et al. 2010), *TUB* (Wan et al. 2010; Coker and Davies 2003), *18S* (Jain et al. 2006) and *UBI* (Lei et al. 2011), which have been reported to be good potential candidates in previously published studies, were selected for the present study. Primers for the amplification of maize candidate reference genes and the gene coding for phytoene synthase (*PSY1*) were designed based on the sequences extracted from GenBank using the Vector NTI10 software (Invitrogen) (Table 1). Primers were selected with T_m of 58–62 $^{\circ}\text{C}$ and GC content of 45–55 %, and all of the amplicons were designed to be <150 bp. The specificity of the amplicons was verified by the presence of a single peak in the RT-qPCR melting curve products and a single band of expected size in a 3 % agarose gel after electrophoresis.

Quantitative real-time PCR

The cDNAs were amplified by RT-qPCR in a final volume of 20 μL containing 1 μL cDNA, 10 μL of Platinum Sybr green UDG (Invitrogen), and 2–5 pmol of each primer. Amplification was standardized in a 7,500 Real-time Fast thermal cycler (Applied Biosystems) using the following conditions: 50 $^{\circ}\text{C}$ for 20 s, 95 $^{\circ}\text{C}$ for 10 min followed by 45 cycles of 15 s at 95 $^{\circ}\text{C}$ and 60 s at 60 $^{\circ}\text{C}$. The PCR products for each primer set were subjected to melting curve analysis to verify the presence of primer dimers or non-specific amplicons. The melting curve analysis ranged from 60 to 95 $^{\circ}\text{C}$, with an increase in the temperature stepwise by 1 %. No-template controls and a reverse transcription negative control were included to ensure that no reagent or genomic DNA contamination occurred. The efficiency of the primers was verified in RT-qPCR using serial dilutions of a reference cDNA synthesized using a RNA sample previously quantified as described above.

Table 1 Primer sequences and amplicon characteristics of reference and target genes

Gene	Genbank accession	Forward primer sequence	Reverse primer sequence	Amplicon length	Product T_m ($^{\circ}\text{C}$) ^a	RT-qPCR efficiency ^b	R^2 ^b
TUB	NM_001111988.1	AGAACTGGGACTGCCTCCAAAGG	AGATGAGCAGGGTGCCCATTC	84	82.65	1.90	0.986
ACT	J01238.1	CATGGAGA-ACTGGCATCACACCTT	CTGCGTCATTTTCTCTCTGTGGC	118	84.35	1.93	0.994
GAPDH	NM_001111943.1	ACTGTTTCATGCCATCACTGC	GAGGACAGGAAGCACTTTGC	144	84.18	1.88	0.998
UBI	M73785.1	GTTTAAGCTGCCGATGTCCTG	GACACGACTCATGACACGAACA	89	82.28	1.91	0.999
18S	AH001709.1	AGAAAACGGCTACCACATCCAA	CTACCTCCCCGTGTCAGGATT	69	81.59	1.95	0.996
PSY1	FJ971251.1	GACAGATGAGCTGTAGATGGGC	TCAGAGAGCGGCATCAAGCA	129	83.1	1.93	0.945

^a The melting temperature was calculated by SDS version 1.1 software in a 7500 Real-time Fast thermocycler (Applied Biosystems)

^b The primer efficiency and correlation coefficients (R^2) were determined with RT-qPCR using serial dilutions of a reference cDNA

Data analysis

To estimate the expression stability of the five candidate reference genes, all amplification plots were analyzed with a threshold fluorescence value of 0.2 to obtain amplification cycle (C_q) values using SDS version 1.1 software (Applied Biosystems). The raw C_q data were processed on a linear scale using the ΔC_q method, and expression stability was evaluated using the geNorm (Vandesompele et al. 2002) and NormFinder (Andersen et al. 2004) software packages for Microsoft Excel. The geNorm software calculates the average of the pairwise variation for a candidate reference gene with all other genes tested, expressing the result as the *M* value (Vandesompele et al. 2002). Genes with highly variable results have a high *M* value, which indicates a low stability of expression, and vice versa, when using a cutoff of 1.5. GeNorm also calculates a normalization factor for each sample and suggests the optimal number of reference genes necessary to normalize the experiment. NormFinder software uses an ANOVA-based model to consider intra- and inter-group variations of the candidate reference genes to evaluate the expression stability and provide a direct measure of the variation.

Impact of using inappropriate reference genes on the expression studies of *PSY1*

Expression analysis of the *PSY1* gene was used to verify the impact of the use of inappropriate reference genes on the gene expression analysis. To this end, the most stable

and the most unstable genes as determined by geNorm and NormFinder were used to determine the gene expression levels of *PSY1* using samples from grain development stages and six varieties. The PCR amplification conditions were the same as described above. The relative expression data were calculated according to the $2^{-\Delta\Delta C_t}$ method and were presented as the fold change (Livak and Schmittgen 2001). Samples at 0 DAP and variety V5 were used as the reference samples to calculate the *PSY1* expression levels during different stages of grain development and in the different varieties, respectively. Pearson's test ($P \leq 0.01$) was used, and the analyses were conducted on Systems Analysis Software (SAS).

Results

Specificity of the amplified product

The results of electrophoresis in the agarose gel showed that all genes were amplified with a single band of expected size (Fig. 1a). The RT-qPCR products showed a single peak in the melting curve, confirming the specificity of the amplifications (Fig. 1b). The melting temperatures of all PCR products are shown in Table 1. No primer dimers or other non-specific amplification products were observed. In addition, no RT-qPCR detection signals were observed in the no-template controls and reverse transcription negative control reactions. The efficiency of the PCR analyses varied from 1.88 to 1.95 (Table 1).

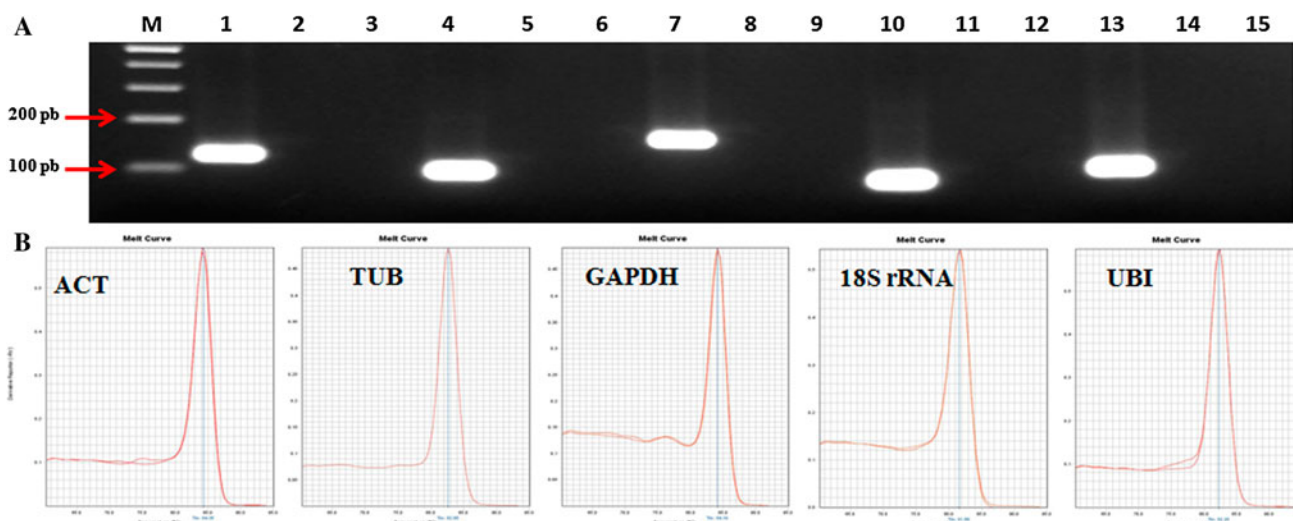


Fig. 1 Specificity of qRT-PCR amplification products. **a** Agarose gel (3 %) showing amplification of a specific PCR product of the expected size of five reference genes in maize grains. Line 1, *ACT*; Line 4, *TUB*; Line 7, *GAPDH*; Line 10, *18S*; Line 13, *UBI*, Lines 2, 5,

8, 11, 14, 9, reverse transcription negative control; Lines 3, 6, 9, 12, 15, no-template control; *M* 1 kb plus (Invitrogen). **b** RT-qPCR melting curve analysis calculated by SDS version 1.1 software in a 7500 Real-time Fast thermocycler (Applied Biosystems)

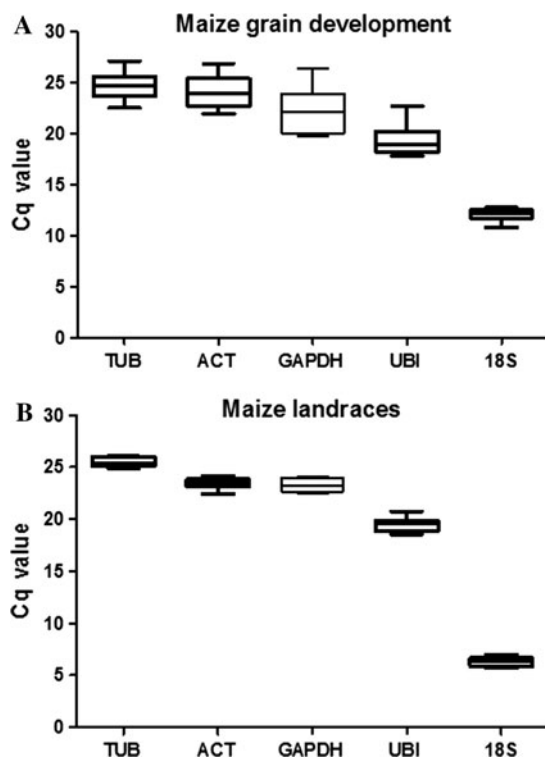


Fig. 2 RT-qPCR quantification cycle (Cq) values for the candidate reference genes. Cq data obtained from maize grains from different developmental stages (**a**) and from maize landraces at 22 DAP (**b**). The line crossing the box represents the median. The box indicates the 25 and 75th percentiles. Whiskers represent the maximum and minimum values

Expression profiling of candidate reference genes

The Cq values for the genes studied showed a wide range from 10 to 27 in the maize grain developmental stage samples and from 6 to 26 in the samples from maize landrace varieties, and the majority of these values were between 23 and 26 (Fig. 2). The expression level of the individual reference candidate genes was similar among grains of different maize landraces at the same developmental stage; however, the expression level was affected in grains at different developmental stages (especially *GAPDH*). The gene encoding 18S rRNA was highly expressed compared to the other genes, reaching threshold fluorescence after only 6 amplification cycles.

Expression stability of candidate reference genes

The expression stability of candidate reference genes was evaluated in the geNorm software, which calculates the average pairwise variation of a particular control gene with all other control genes. The program recommends using candidate genes with an M value below the threshold of 1.5. As shown in Table 2, all candidate genes showed

Table 2 The expression stability values (M) of the candidate reference genes estimated by geNorm algorithm

Gene name	Maize grain development	Maize landraces	All samples
TUB	1.043	0.466	0.755
ACT	1.162	0.472	0.799
UBI	1.322	0.558	0.901
18S	1.471	0.586	0.983
GAPDH	1.499	0.774	1.075

M values lower than 1.5. According to the M value, the most stable gene in the samples from maize grain development and maize landraces is *TUB* ($M = 1.043$ and 0.466 , respectively), followed by *ACT* ($M = 1.162$ and 0.472 , respectively). *GAPDH* showed the highest M value ($M = 1.499$ in the samples from maize grain development and 0.774 from maize landraces), which implies the lowest stability in both evaluated conditions, indicating that it was not suitable for expression analysis under these experimental conditions. The same result was observed when all samples were considered in the analysis; therefore, *GAPDH* was excluded by the software for further analysis regarding the optimal number of reference genes to normalize expression data using these maize samples.

The Normfinder software was also applied to evaluate the stability of the candidate reference genes. According to this approach, *TUB*, followed by *ACT*, showed the best stability using both experimental conditions (Table 3). *GAPDH* showed the least stability among candidate reference genes evaluated; therefore, NormFinder provided the same result as geNorm.

Determination of the optimal number of reference genes

GeNorm also calculated the pairwise variation (V_n/V_{n+1}) between the normalization factors for each sample, suggesting the optimal number of reference genes necessary to normalize the experiment. Vandesompele et al. (2002) suggest that a stepwise inclusion of genes until the $(n + 1)$ th gene has no significant contribution to the calculated normalization factor, and the software suggests a cutoff threshold of $V = 0.15$. According to these criteria, the use of two reference genes was sufficient to normalize the results of gene expression using maize landraces because the value of $V_{2/3}$ was 0.112 and the $V_{3/4}$ was 0.121 (Fig. 3d). The $V_{2/3}$ value in the samples from maize grain development was 0.36, which dropped to 0.267 in the $V_{4/5}$, and therefore was still higher than the suggested cutoff value of 0.15. However, 0.15 is not an absolute cutoff value but rather a suggested

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

Rank	Maize grain development		Maize landraces		All samples	
	Gene name	Stability value	Gene name	Stability value	Gene name	Stability value
1	TUB	0.051	TUB	0.130	TUB	0.173
2	ACT	0.449	ACT	0.158	ACT	0.323
3	UBI	0.655	UBI	0.159	UBI	0.428
4	18S	0.858	18S	0.340	18S	0.605
5	GAPDH	0.879	GAPDH	0.463	GAPDH	0.660

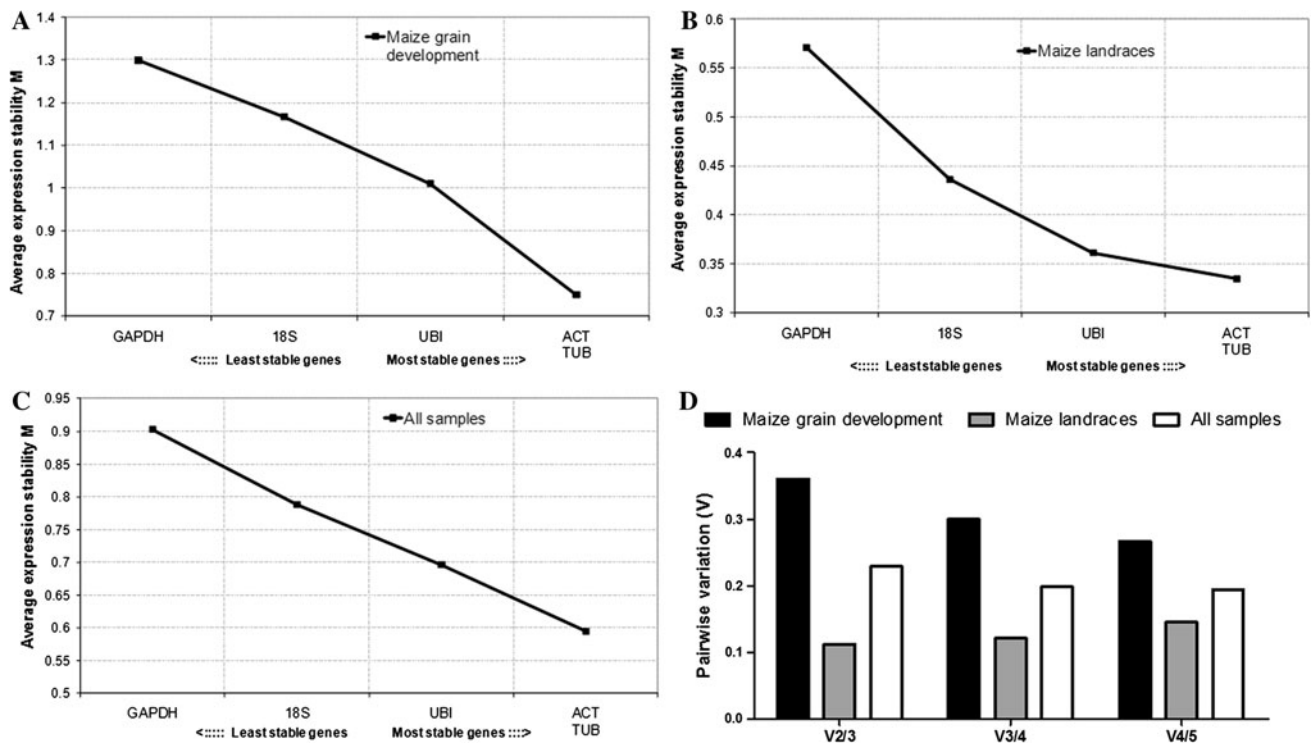


Fig. 3 Average expression stability (M value) and pairwise variation (V) analysis of the five candidate reference genes using the geNorm software. Expression stability was evaluated in samples from maize grain developmental stages (a), maize landraces (b) and all samples (c). The most stable reference genes were measured during stepwise exclusion of the least stable reference genes. A lower average

expression stability M value indicates more stable expression. **d** Pairwise variation (V) calculated by geNorm to determine the minimum number of reference genes for accurate normalization in samples from maize grain development, from maize landraces and from all samples

value for the study performed by Vandesompele et al. (2002); therefore, the use of four reference genes is suggested to normalize samples for maize grain development. The results confirmed that no single reference gene had a constant expression in the samples evaluated. GeNorm establishes a rank order of gene stability via stepwise exclusion of the least stable gene. According to this software, *TUB* and *ACT* was the best combination to use in the normalization of RT-qPCR data for all conditions evaluated (Fig. 3a–c).

Expression analysis of *PSY1* with the most stable/unstable reference genes

The relative expression of *PSY1* in seven stages of maize grain development and from six maize landraces was used to detect the effect of using different reference genes in the data normalization (Fig. 4). The most stable (*TUB* + *ACT*) and the least stable (*GAPDH*) reference genes, according to geNorm and NormFinder, were used as internal controls. *PSY1* was expressed at a higher level at the end of maize

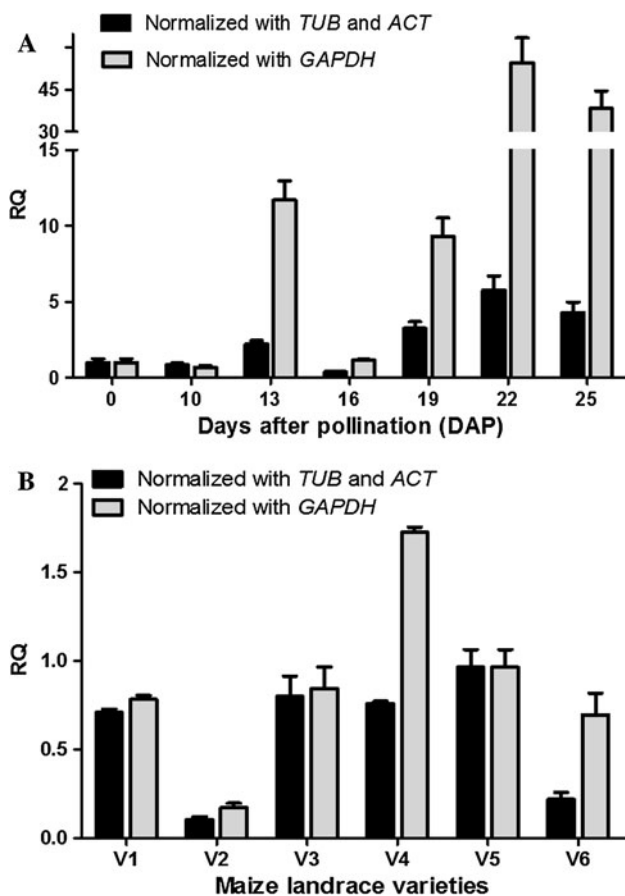


Fig. 4 Expression profile of *PSY1* during maize grain development (a) and in different varieties (b). The best stable combination of reference genes (*TUB* + *ACT*) and the least stable (*GAPDH*) were used to normalize the expression data. A grain sample at 0 DAP (a) and variety V5 (b) were used as reference samples. The error bar shows the mean standard error calculated from three biological replicates

grain development (from 19 to 25 DAP), and its expression varied among the varieties. However, *PSY1* was shown to have a higher expression level at 13, 19, 22 and 25 DAP (Fig. 4a) and in varieties V4 and V6 (Fig. 4b) when using the least stable reference *GAPDH* compared to when *TUB* + *ACT* were used as internal controls. Thus, the use of unsuitable references leads to differences in the relative expression profile. These results further confirmed the importance of validating reference genes prior to experimental applications.

Discussion

The use of non-validated reference genes for qPCR data normalization may generate unreliable results (Derveaux et al. 2010; Bustin et al. 2009; Radonic et al. 2004), which implies that studies regarding the expression stability of candidate reference genes must be performed to select the

most suitable reference genes to normalize the RT-qPCR data in each species and under different conditions. The present study tested the stability of five candidate reference genes in grain samples from eight maize varieties and seven developmental stages. As far as we know, this study is the first survey on the stability of genes to use as reference genes to normalize RT-qPCR data from maize landraces and grain developmental stages.

Assessing the stability of a specific gene expression is challenging because it has to be performed without using any other references. In an attempt to solve this problem, several statistical algorithms have been developed to evaluate the stability of candidate reference genes, such as NormFinder (Andersen et al. 2004), geNorm (Vandesompele et al. 2002), the ‘Stability index’ (Brunner et al. 2004), the DCt approach (Livak and Schmittgen 2001), and BestKeeper (Pfaffl et al. 2004). In this study, we used geNorm and NormFinder. GeNorm software is one of the most commonly used algorithms because of its robustness and convenience. However, if several of the candidate reference genes are co-regulated, they will occupy closed positions in the ranking provided by the analysis because geNorm software assumes that none of the genes being analyzed are co-regulated (Vandesompele et al. 2002). According to geNorm, *TUB* followed by *ACT* was the most stable reference gene to normalize the expression of genes in maize samples in the present study. The genes are involved in cell structure maintenance and may be regulated by similar conditions; therefore, NormFinder was also used to assess whether this result reflects the co-regulation of these genes. NormFinder ranks the set of candidate reference genes according to the least of their estimated intra- and inter-group variations; therefore, this software is not affected by co-regulation (Andersen et al. 2004). Using this software, the same results as geNorm were obtained. Interestingly, the same genes were considered the most stable under all experimental conditions, corroborating that *TUB* and *ACT* are suitable to normalize the expression data using maize grains. *ACT* also had the highest expression stability across leaf and root tissues in chicory (Maroufi et al. 2010), banana (Lei et al. 2011) and during rice grain development (Qian-Feng et al. 2010), and *TUB* was considered appropriate to normalize expression in different cucumber sample pools (Wan et al. 2010) and tomato (Coker and Davies 2003). However, Manoli et al. (2012) demonstrated that the gene expression stability of *ACT*, *TUB* and *18S* rRNA was significantly lower than most other genes evaluated using the roots, leaves and stems of maize under different stress conditions. Furthermore, Chen et al. (2012) showed that during the priming and germination of maize seeds, actin depolymerizing factor and ubiquitin were the most stable genes, as calculated by geNorm, and that *ACT*, *18S*, α *TUB* and β *TUB* were less

stable, confirming that no single gene should be used as a reference gene under all conditions and with all tissues from the same species.

One of the most commonly used genes as an internal control in RT-qPCR studies is *GAPDH* (Magneschi et al. 2009). In the present study, *GAPDH* was the gene with the lowest stability value because it was the lowest ranked gene averaged across all the samples and is therefore not considered suitable to use as a reference gene in maize samples. Similarly, the *M* value of *18S* was lower than 1.5 under all evaluated conditions, but it was the third gene in the geNorm rank for maize grain development and the fourth gene for maize landraces, suggesting that it is not the most appropriate for use under these conditions. Furthermore, the *C_q* values of *18S* were much lower than the other candidate genes; therefore, it is not adequate to normalize the expression of genes with low expression levels, such as those from the carotenoid metabolic pathway, as suggested by Bustin (2002).

According to geNorm, in this study, the use of two genes was sufficient to normalize gene expression when comparing different varieties using grains in the same developmental stage (22 DAP). In contrast, the use of four genes to normalize transcript levels during the development of the grain (from 0 to 25 DAP) did not reach the cutoff value of 0.15 suggested by the geNorm software. These results suggest that the reference genes are differentially affected among samples under different experimental conditions and further confirm the importance of validating the reference gene stability for each experimental setting.

To evaluate the effect of using the appropriate reference genes to normalize transcript levels, the expression profiles of *PSYI* were assessed at different stages of maize grain development and in different varieties. The relative transcript abundance was overestimated in three of the seven stages of grain development and in some of the varieties evaluated when the most variable reference gene *GAPDH* was used as an internal control (Fig. 4) instead of the most stable genes (*ACT* + *TUB*). The expression of *GAPDH* most likely varies because it may not only play a role in the glycolytic pathway but may also participate in other cellular processes (Muñoz-Bertomeu et al. 2010). These results suggest that the use of inappropriate reference genes may introduce bias in the analysis and affect the interpretation of the RT-qPCR data. For example, if the result of gene expression obtained using *GAPDH* in Fig. 4b was taken into account, the V4 maize variety would be considered a promising target for studies involving the biofortification of maize grains with carotenoids because the expression of *PSYI* is highly correlated with the content of carotenoids in maize grains (Vallabhaneni and Wurtzel 2009).

In conclusion, *TUB* and *ACT* were the most stable genes according to geNorm and NormFinder in all maize grain

samples tested and are thus considered reliable reference genes; however, *GAPDH* should be avoided because it performed poorly in both statistical packages. The use of appropriate reference genes significantly improved the accuracy of the *PSYI* expression profile determination, which emphasizes the need to accurately validate candidate internal control genes before use in RT-qPCR studies. The results obtained in the present study will provide information to further evaluate the expression of carotenoid-related genes in maize grains to biofortify this crop through metabolic engineering.

Acknowledgments The authors gratefully acknowledge the support of Embrapa Temperate Agriculture, CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) and FAPEG (Fundação de Apoio a Pesquisa e Desenvolvimento Agropecuário Edmundo Gastal).

Conflict of interest There are no conflicts of interests to declare.

References

- Aluru M, Xu Y, Guo R, Wang Z, Li S, White W, Wang C, Rodermeil S (2008) Generation of transgenic maize with enhanced provitamin A content. *J Exp Bot* 59:3551–3562
- Andersen CL, Jensen JL, 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
- Berardo N, Mazzinelli G, Valoti P, Lagana P, Redaelli R (2009) Characterization of maize germplasm for the chemical composition of the grain. *Agr Food Chem* 57:2378–2384
- Brunner AM, Yakovlev IA, Strauss SH (2004) Validating internal controls for quantitative plant gene expression studies. *BMC Plant Biol* 4:14
- Bustin SA (2002) Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. *J Mol Endocrinol* 29:23–39
- Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT (2009) The MIQE guidelines: minimum information for publication of quantitative realtime PCR experiments. *Clin Chem* 55:611–622
- Chang S, Puryear J, Cairney J (1993) A simple and efficient method for isolating RNA from pine trees. *Plant Mol Biol Rep* 11
- Chen K, Fessehaie A, Arora R (2012) Selection of reference genes for normalizing gene expression during seed priming and germination using qPCR in *Zea mays* and *Spinacia oleracea*. *Plant Mol Biol Rep* 30:478–487
- Coker JS, Davies E (2003) Selection of candidate housekeeping controls in tomato plants using EST data. *Biotechniques* 35:740–748
- Czechowski T, Stitt M, Altmann T, Udvardi MK, Scheible WR (2005) Genome-wide identification and testing of superior reference genes for transcript normalization in *Arabidopsis*. *Plant Physiol* 139:5–17
- Derveaux S, Vandesompele J, Hellemans J (2010) How to do successful gene expression analysis using real-time PCR. *Methods* 50:227–230

- Fleige S, Walf V, Huch S, Prgomet C, Sehm J, Pfaffl MW (2006) Comparison of relative mRNA quantification models and the impact of RNA integrity in quantitative real-time RT-PCR. *Biotechnol Lett* 28:1601–1613
- Fraser PD, Bramley PM (2004) The biosynthesis and nutritional uses of carotenoids. *Prog Lipid Res* 43:228–265
- Gutierrez L, Mauriat M, Pelloux J, Bellini C, Wuytswinkel UV (2008) Towards a systematic validation of references in real time RT-PCR. *Plant Cell* 20:1734–1735
- Jain M, Nijhawan A, Tyagi AK, Khurana JP (2006) Validation of housekeeping genes as internal control for studying gene expression in rice by quantitative real-time PCR. *BBRC* 343:646–651
- Kulcheski FR, Marcelino-Guimaraes FC, Nepomuceno AL, Abdelnoor RV, Margis R (2010) The use of microRNAs as reference genes for quantitative polymerase chain reaction in soybean. *Anal Biochem* 406:185–192
- Lei C, Hai-ying Z, Jian-fei K, Jian-guo L, Wang-jin L, Jian-ye C (2011) Validation of reference genes for RT-qPCR studies of gene expression in banana fruit under different experimental conditions. *Planta* 234:377–390
- Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2-Delta DeltaCT method. *Methods* 25:402–408
- Lovdal T, Lillo C (2009) Reference gene selection for quantitative real-time PCR normalization in tomato subjected to nitrogen, cold, and light stress. *Anal Biochem* 387:238–242
- Magneschi L, Kudahettige RL, Alpi A, Perata P (2009) Expansin gene expression and anoxic coleoptiles elongation in rice cultivars. *J Plant Physiol* 166:1576–1580
- Manoli A, Sturaro A, Trevisan S, Quaggiotti S, Nonis A (2012) Evaluation of candidate reference genes for qPCR in maize. *J Plant Physiol* 169:807–815
- Maroufi A, Van Bockstaele E, De Loose M (2010) Validation of reference genes for gene expression analysis in chicory (*Cichorium intybus*) using quantitative real-time PCR. *BMC Mol Biol* 11:15
- Muñoz-Bertomeu J, Cascales-Miñana B, Alaiz M, Segura J, Ros R (2010) A critical role of plastidial glycolytic glyceraldehyde-3-phosphate dehydrogenase in the control of plant metabolism and development. *Plant Signal Behav* 5:67–69
- Naqvi S, Zhu C, Farre G, Ramessar K, Bassie L, Breitenbach J, Conesa DP, Ros G, Sandmann G, Capell T, Christou P (2009) Transgenic multivitamin corn through biofortification of endosperm with three vitamins representing three distinct metabolic pathways. *PNAS* 106:7762–7767
- Paolacci AR, Tanzarella OA, Porceddu E, Ciaffi M (2009) Identification and validation of reference genes for quantitative RT-PCR normalization in wheat. *BMC Mol Biol* 10:11
- Pfaffl MW, Tichopad A, Prgomet C, 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* 26:509–515
- Qian-Feng L, Samuel S, Ding-Yang Y, Heng-Xiu Y, Ming-Hong G, Qiao-Quan L (2010) Validation of candidate reference genes for the accurate normalization of real-time quantitative RT-PCR data in rice during seed development. *Plant Mol Biol Rep* 28:49–57
- Radonic A, Thulke S, Mackay IM, Landt O, Siegert W, Nitsche A (2004) Guideline to reference gene selection for quantitative real-time PCR. *Biochem Bioph Res Co* 313:856–862
- Vallabhaneni R, Wurtzel ET (2009) Timing and biosynthetic potential for carotenoid accumulation in genetically diverse germplasm of maize. *Plant Physiol* 150:562–572
- Vandesompele J., De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 3 RESEARCH0034
- Wan H, Zhao Z, Qian C, Sui Y, Malik AA, Chen J (2010) Selection of appropriate reference genes for gene expression studies by quantitative real-time polymerase chain reaction in cucumber. *Anal Biochem* 399:257–261