JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

ARTICLE

pubs.acs.org/JAFC

Selection of Reference Genes for Gene Expression Studies in Zucchini (*Cucurbita pepo*) Using qPCR

Ángeles Obrero,*[,†] Jose V. Die,[‡] Belén Román,[†] Pedro Gómez,[§] Salvador Nadal,[#] and Clara I. González-Verdejo[†]

[†]Área de Mejora y Biotecnología, IFAPA, Centro Alameda del Obispo, Apartado 14080 Córdoba, Spain

[‡]Mejora Genética Vegetal, CSIC, Instituto de Agricultura Sostenible, Apartado 14080 Córdoba, Spain

[§]Área de Mejora y Biotecnología, IFAPA, Centro La Mojonera, Autovía del Mediterráneo sal. 420. E-04745 La Mojonera, Almería, Spain

[#]Área de Producción Agraria, IFAPA, Centro Alameda del Obispo, Apartado 14080 Córdoba, Spain

Supporting Information

ABSTRACT: The zucchini (*Cucurbita pepo*) is an important food crop, the transcriptomics of which are a fundamental tool to accelerate the development of new varieties by breeders. However, the suitability of reference genes for data normalization in zucchini has not yet been studied. The aim of this study was to assess the suitability of 13 genes for their potential use as reference genes in quantitative real-time PCR. Assays were performed on 34 cDNA samples representing plants under different stresses and at different developmental stages. The application of geNorm and NormFinder software revealed that the use of a combination of *UFP*, *EF-1A*, *RPL36aA*, *PP2A*, and *CAC* genes for the different experimental sets was the best strategy for reliable normalization. In contrast, *18S rRNA* and *TUA* were less stable and unsuitable for use as internal controls. These results provide the possibility to allow more accurate use of qPCR in this horticultural crop.

KEYWORDS: *Cucurbita pepo*, qPCR, normalization, reference genes, zucchini

INTRODUCTION

Zucchini, also known as summer squash or courgette, is the edible immature fruit of *Cucurbita pepo* spp. *pepo*. It is a worldwide value crop, being an excellent dietary source of vitamins, minerals, and fiber. There is increasing interest from seed companies and zucchini growers in enhancing this vegetable's quality to address consumer desires. New agro-alimentary technologies and, in particular, the implementation of biotechnological tools in classic breeding programs, allow for the more efficient development of more productive varieties that have greater nutritional value. In this sense, transcriptomic studies are becoming increasingly important, as understanding gene expression patterns is expected to reveal new genes involved in the production of desirable characteristics.

Quantitative real-time PCR (qPCR) has become the most prominent emerging method for the quantification of mRNA levels because of its high sensitivity, good reproducibility, and wide dynamic quantification range.[1,2] To obtain accurate results it is necessary to ensure quality measures and to increase experimental transparency in qPCR analysis. In this sense, the Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) guidelines show basic details that must be taken into account in the publication of qPCR results.[3] Moreover, appropriate normalization strategies are required to control for experimental error. The purpose of normalization is to minimize the impact of nonbiological variation on the results. Several normalization strategies have been proposed, each with its own advantages and disadvantages.[4] The use of one or more reference genes is currently the preferred means of normalization in quantification, with the assumptions that expression levels of these genes should be relatively constant across the tissues and cells tested and should be unaffected by any experimental treatment.[5,6] Selection of appropriate reference genes for data normalization is essential to ensure correct interpretation of results. This is a laborious task, especially in species with limited available sequence information, such as zucchini, but it is an established procedure that is necessary for the accurate and reliable quantification of differentially regulated mRNAs. Many studies to optimize normalization genes for qPCR experiments have been published for humans, animals, yeasts, and bacteria. In recent years, this analysis has also begun in plants,[7−9] but remains unavailable for zucchini.

Previous *Arabidopsis* microarray analyses have shown how novel reference genes may have more stable expression than traditional reference genes.[10] Further works in plants based on these results also showed greater stability of expression in these novel reference genes.[11−14] Ten of the 13 zucchini genes evaluated in this work [18S rRNA (*18S rRNA*), 60S ribosomal protein L36a/L44 (*RPL36aA*), actin (*ACT*), α-tubulin (*TUA*), ubiquitin (*UBI*), glucose-6-phosphate dehydrogenase (*G6PDH*), ubiquitin fusion protein (*UFP*), NAD-dependent malic enzyme (*NAD*), NADP-dependent malic enzyme (*NADP*), and elongation factor-1 α (*EF-1A*)] are classical reference genes that have commonly been used as internal controls for expression studies.[7,15,16] We also included three new candidate genes: helicase (*HELI*),

Received: December 13, 2010
Revised: April 8, 2011
Accepted: April 8, 2011
Published: April 08, 2011

Table 1. Reference Gene Primer Sequences and Amplicon Characteristics

gene symbol	accesion no.	name	forward and reverse primer sequence (5′−3′)	amplicon size (bp)	amplicon T_m (°C)	PCR efficiency (%)	regression coefficient (R^2)
18S rRNA		18S rRNA			83.27	95	0.997
RPL36aA	HM594174	60S ribosomal protein L36a/L44	^{33}GATAGTCTTGCTGCACAGGGAAA ^{98}GGTCTGACCTCCATATCCTGATTG	66	79.79	95	0.998
ACT	HM594170	actin	^{100}CCTCTCAATCCCAAAGCTAACAG ^{190}CGGCCTGGATAGCAACATACA	91	77.10	97	0.999
TUA	HM594172	α-tubulin	^{148}TGATCTCTGCTTGGATCGTATAAGA ^{206}CGAGGAAACCTTGAAGACCAGTA	65	77.04	93	0.998
UBI	HM594175	ubiquitin	^{201}GCAACTGGAAGATGGTAGGACTCT ^{265}CCAAGTGCAATGTCGTTTCCTT	65	76.11	93	0.999
G6PDH	AF260736	glucose-6-phosphate dehydrogenase	^{399}CCCCTGCAACTCGAAGATGT ^{462}ACGATTTCCCACCCTTGCT	64	79.08	94	0.998
UFP	CD726808	ubiquitin fusion protein	^{168}CGGACCAGCAGAGGCTTATC ^{223}GAGAGTTCGCCCATCCTCAA	84	80.27	98	0.998
PP2A	HM594171	protein phosphatase 2A regulatory subunit A	^{98}TGGTAGCATCCTTTCCCAATACA ^{164}CATGCCCGTTCAGCTTTAGC	67	78.89	96	0.999
EF-1A	HO702383	elongation factor-1α	^{82}GCTTGGGTGCTCGACAAACT ^{148}TCCACAGAGCAATGTCAATGG	67	79.03	97	0.996
CAC	HM594173	clathrin adaptor complexes medium subunit	^{148}GGACAAACAGAACCAACCATGA ^{212}GGTTTCCTTTCCGTCACTGTAGA	65	77.28	94	0.998
NAD	AF260732	NAD-dependent malic enzyme	^{352}TTCCAGAGCAAATGGGCATT ^{433}CCGTTCCCTGAACATCATCAT	81	75.70	92	0.999
NADP	AF260735	NADP-dependent malic enzyme	^{400}CCGATCACACTTTCTTGTTCCTT ^{490}GCGTTCGTCTGTTTCGATATTTC	91	79.98	94	0.998
HELI	HM594176	DEAD-box RNA helicase-like protein	^{1}ACACTGGTCCCTCCCACACA ^{60}GCGGGCACTTGGAGATTATC	60	78.18	93	0.995
CAT1	D55645	catalase 1	^{1490}GTCACCCATGAGATCCGCA ^{1551}CCAAGAGACCTATCCGCCTG	62	76.79	92	0.999

protein phosphatase 2A (*PP2A*), and clathrin adaptor complexes medium subunit family protein (*CAC*).

The expression patterns of these genes were tested in different tissues and organs (roots, stems, leaves, flowers, and fruits), during different fruit and flower developmental stages and under several abiotic stress conditions (salt, hormonal stress, and low temperature) using the software applications geNorm[17] and NormFinder[18] to calculate the most stably expressed reference genes and to determine the optimal number of reference genes required for reliable normalization of gene expression data.

The use of nonconstant reference genes can have a strong impact on the results of relative expression analysis.[19] Therefore, our aim was to identify internal controls that exhibit highly constant expression throughout the experimental conditions analyzed. Once the genes were identified, we chose a known expression profile of a well-studied gene in Cucurbitaceae, *CAT1*,[39] to test the impact of reference gene selection in the results using different normalization factors based on the more and least stable genes (determined by geNorm and NormFinder). We show that *CAT1* expression levels are significantly dependent on the choice of reference genes and that proper evaluation of reference gene

Table 2. Degenerate Primer Pairs Designed for PCR Fragment Cloning[a]

gene	forward and reverse primers (5′−3′)	T_m (°C)	amplicon size (bp)
Deg-RPL36aA	F: ACN CAR TAY AAR AAR GGN AA R: CAY TGV ARC YTC ARM ACA AT	50	126
Deg-CAC	F: TTY GGG CWA ARA TGT TTG CHC TKG G R: GGM ACC TGR AAY TCC ATY TGR ATH GG	50	382
Deg-PP2A	F: TAA GTC HTG GCG YGT KCG YTA YAT G R: ACC TGR TTB ACT TGR TCA AGY TTG CT	55	213
Deg-ACT	F: GTN ARY AAC TGG GAT GAY ATG G R: ACA ATA CCW GTW GTR CGA CC	56	231
Deg-TUA	F: GAA GAT GCT GCH AAY AAY TTY GC R: GAR AGH ACA CTG TTR TAA GGY TC	56	366
Deg-UBI	F: GAY TAC AAC ATY CAG ARG GAG R: GCR AAR ATC ARC CTC TGC TG	56	422

[a] Sequence and T_m of the primers as well as product length are given.

stability is therefore mandatory before qPCR results in zucchini can be reported.

MATERIALS AND METHODS

Plant Material. Experiments were performed using *C. pepo* L. seeds that were routinely purchased from Nunhems Zaden BV (Haelen, The Netherlands). Seeds were germinated in Petri dishes on wet glass fiber filter paper and covered with aluminum foil, to exclude light, at 20 °C for 7 days. Seedlings were transplanted into perlite-containing pots in trays with Hoagland nutrient solution.[20] Plants were maintained in a greenhouse at IFAPA Center Alameda del Obispo, Córdoba (Spain), from February through March.

Material from different organs and development stages and from the stress treatments was dissected and harvested from at least five different zucchini plants to obtain a pool. The procedure was repeated with five distinct plants to obtain a second biological replicate. All samples were immediately frozen in liquid nitrogen and stored at −80 °C until needed for RNA extraction.

Fruit and Flower Developmental Stages. Two stages of flower development (before anthesis and anthesis) were collected, and three stages of fruit development (fruits of 1−5, 6−10, and 11−15 cm in length) were also included.

Stress Treatments. Salt and hormone treatments were applied to 21-day-old zucchini plants by adding 100 mM NaCl or 5 μM 2, 4-dichlorophenoxy acetic acid (2,4-D) for 24 h, with control plants kept in water for the same length of time. Cold treatment was performed by exposing the zucchini plants to 10 °C for 48 h, whereas control plants were kept at 18−25 °C. For each stress situation, as well as for the control conditions, plant roots, stems, and leaves were sampled after treatments.

Total RNA Isolation. All samples were ground to a fine powder with a mortar and pestle in liquid N_2. Total RNA was extracted using the TRIsure reagent (Bioline, London, U.K.) according to the manufacturer's instructions. RNA concentration and purity were determined with a biophotometer (Eppendorf, Hamburg, Germany), which calculated absorbance at 260 nm. Only RNA samples with 260/280 ratios between 1.9 and 2.1 and 260/230 ratios of >2.0 were used for cDNA synthesis. RNA integrity was assessed by microcapillary electrophoresis with an Experion RNA StdSens Chip and an Experion bioanalyzer (Bio-Rad Laboratories, Hercules, CA), showing rRNA subunits with 18S/25S peaks on the virtual gel and electropherograms. Only RNA samples with electropherograms of high quality and RQI > 8 were used for further study (see Figure S1of the Supporting Information).

cDNA Synthesis. cDNA was synthesized from 1 μg of total RNA for each sample using the QuantiTec Reverse Transcription Kit (Qiagen, Hilden, Germany) with a blend of oligo-dT and random primers according to the manufacturer's instructions. In this kit, genomic DNA is efficiently removed in a single step; nevertheless, we included a negative control to test for contaminating genomic DNA. This control contained all of the reaction components except the reverse transcriptase. Samples without enzyme were checked by qPCR, and only when there was no amplification in minus RT controls after 40 cycles with *ACT* primers were the cDNAs used for further analysis. The cDNA samples used for qPCR assays were diluted 1:5 with nuclease-free water.

Cloning Partial Sequences. *C. pepo* nucleotide sequences for *G6PDH*, *UFP*, *NAD*, and *NADP* were obtained from the GenBank database (see Table 1 for accession numbers). For *18S rRNA* gene amplification, the QuantumRNA 18S Internal Standards kit (Ambion, Austin, TX) was used. Sequences for *ACT*, *TUA*, *UBI*, *PP2A*, *RPL36aA*, *CAC*, *EF-1A*, and *HELI* genes were not available from the NCBI public database. Specific PCR fragments for *ACT*, *TUA*, *UBI*, *PP2A*, *RPL36aA*, and *CAC* genes were cloned using degenerate primers. *ACT*, *TUA*, and *UBI* fragments were amplified using primers previously described[21] with new PCR conditions adapted to *C. pepo* material (Table 2). For *RPL36aA*, *CAC*, and *PP2A* genes, degenerate primers were designed on the basis of highly conserved regions of proteins within orthologous genes (Table 2). *EF-1A* and *HELI* fragments were cloned using *Pisum sativum* primers previously used to select appropriate reference genes in pea.[22] Amplification of gene fragments was performed using PCR with cDNA as a template. The following PCR conditions were used: an initial denaturation at 94 °C for 35 s, 40 cycles of denaturation at 94 °C for 35 s, annealing at 50−56 °C (according to the gene) for 35 s, and extension at 72 °C for 1 min, with a final extension of 7 min at 72 °C. Taq DNA polymerase (Bioline, London, U.K.) was used in the reactions. PCR products for *ACT*, *TUA*, *UBI*, *PP2A*, *RPL36aA*, *CAC*, *EF-1A*, and *HELI* were separated by electrophoresis, purified using a Favorgen GEL/PCR purification kit (Favorgen Biotech Corp., Kaohsiung, Taiwan), cloned

into the pGEM-T vector (Promega, Madison, WI), and sequenced (STAB-VIDA, Oeiras, Portugal). The sequence of the amplification product for each primer pair was compared in GenBank (NCBI) using the BLASTX algorithm and searches for homologous genes in other organisms. All sequences showed 76−100% identity, and the resulting sequence data were deposited in the GenBank database under accession no. HM594170, HM594172, HM594175, HM594171, HM594174, HM594173, HO702383, and HM594176, respectively.

Primer Design and qPCR Conditions. Specific primer pairs for qPCR amplification were designed to amplify products of 50−150 bp, with an optimal primer melting temperature (T_m) of 60 °C and GC contents between 35 and 65%. For *18S rRNA* gene amplification, the QuantumRNA 18S Internal Standards kit (Ambion, Austin, TX) was used (Table 1).

MFOLD software v3.2 (http://mfold.bioinfo.rpi.edu/cgi-bin/dna-form1.cgi) was subsequently used to evaluate the possible formation of secondary structures at the sites of primer binding, using the default settings with 50 mM Na^+, 3 mM Mg^{2+}, and an annealing temperature of 60 °C.[23]

qPCR reactions were performed in a 96-well plate with an Mx3000P Real-Time PCR System (Stratagene, La Jolla, CA), using SYBR Green detection. Reactions were performed in a total volume of 10 μL, containing 0.2 μL of 50× SYBR Green Solution, 5 μL of 2× SensiMix (dT) (Quantace, London, U.K.), 1 μL of cDNA (corresponding to ∼10 ng of total RNA), and 0.4 μL (200 nM) of each gene-specific primer. Except for *RPL36aA*, the following standard thermal profile was used for all PCR reactions: polymerase activation (95 °C for 10 min) and amplification and quantification cycles repeated 40 times (95 °C for 1 min, 60 °C for 1 min). For *RPL36aA*, primer annealing at 65 °C was necessary to generate a sharp peak in melting curve analysis. Finally, a dissociation analysis of the PCR products was performed by running a gradient from 60 to 95 °C to confirm the presence of a single product in each reaction and the specificity of the qPCR. Two biological replicates were analyzed using Mx3000P analysis software v4.1 (Stratagene). All amplification plots were analyzed with an R_n threshold of 0.03 to obtain quantification cycle (C_q) values for each gene−cDNA combination. The PCR efficiency (E) was estimated using LinReg software with data obtained from the exponential phase of each individual amplification plot and the equation $(1 + E) = 10^{slope}$.[24] Table 1 shows the primer sequences, amplicon sizes, melting temperatures, and PCR efficiencies.

Statistical Analyses. Gene expression levels were calculated for all individual reference genes based on the number of C_q. To analyze the stability of expression and to identify the most suitable reference genes, we used the freely available statistical algorithms geNorm v3.5n (http://medgen.ugent.be/∼jvdesomp/genorm/) and NormFinder (http://www.mdl.dk/publicationsnormfinder.htm) according to the authors' recommendations. For both programs, C_q values were converted to linear scale expression quantities using the formula $Q = E^{-\Delta Cq}$, where E is the efficiency and ΔC_q is the difference between the C_q value of the target sample and the value of the sample with the lowest C_q. geNorm calculates a gene expression normalization factor for each tissue sample based on the geometric mean of a user-defined number of reference genes. The expression stability measure (M) is calculated as the mean of the pairwise variation of a gene compared to that of all other genes. The genes with the lowest M values were considered the most stable ones, whereas the highest M value indicated the least stable expression. The NormFinder program is another Visual Basic application tool for Microsoft Excel, used for the determination of stability of expression of reference genes. This approach is based on a variance estimation, and it focuses on finding smaller intra- and intergroup expression variation and combining both results in a stability value. To assess both intra- and intergroup variations, four subgroups were established as geNorm analysis (salt stress, hormonal stress, cold stress, and fruit and flower developmental stages). The mean values of the two biological replicates were calculated for the analysis. Normalized ratios of the target gene *CAT1* (GenBank accession no. GI: 862451) in leaves and roots under saline conditions as well as in control samples were calculated according to the geNorm manual.[25] Primers used for *CAT1* amplification are shown in Table 1.

RESULTS

Performance of Amplification Primers. qPCR was conducted on the 34 samples with 13 primer pairs. To optimize PCR conditions, the secondary structures of the sequences were checked with nucleic acid-folding software (Figure S3 of the Supporting Information). All primers that generated amplicons possessing secondary structures with significant negative ΔG values (free energy in kcal/mol) at the primer binding sites were excluded from further analysis. The specificity of the amplifications was confirmed by the single-peak melting curves of the PCR products, with no visible primer−dimer formation (Figure 1 and Figure S2 of the Supporting Information). PCR efficiencies varied from 92 to 98%, and correlation coefficients (R^2) ranged between 0.995 and 0.999 (Table 1).

Expression Profiles of Reference Genes. qPCR analysis methods were used to measure the transcription levels of 13 candidate reference genes involved in different cellular functions, from 34 different cDNA samples. The expression levels were presented as quantification cycles (C_q) (Figure 2), and these values were compared. The lowest mean C_q value was for *18S rRNA* (15.18), and the highest value was for *HELI* (27.40); therefore, *18S rRNA* transcript levels were about 4770-fold more abundant than those of helicase. These results indicate that the tested genes show a wide range of expression levels.

Data Analysis. *geNorm.* Tested genes were ranked according to stability values (M) and were analyzed for total samples and across four subsets (Figure 3). The 13 genes' M values in the different experimental conditions were <1.4, which is below the geNorm threshold of 1.5. When all 34 tissue samples were considered together, *UFP* and *EF-1A* were selected as having the most stable expression levels, whereas *18S rRNA* was expressed most variably (Figure 3A). The lowest values of M for fruit and flower developmental stages were for *RPL36aA* and *UFP*, whereas the highest value was for *TUA* (Figure 3B). For treatments with salt and cold stress, expression levels of *PP2A* and *CAC* were the most stable (Figure 3C,E), whereas *PP2A* and *EF-1A* were the most stable under hormonal stress (Figure 3D). For the three treatments (salt, hormonal, and cold), the highest M value was observed for *18S rRNA*.

It is crucial to determine the optimal numbers of genes that are necessary for accurate normalization in each experiment. geNorm uses the pairwise variation V_n/V_{n+1} to define the optimal number of reference genes. Vandesompele et al.[17] established a threshold of 0.15, below which the addition of more genes is not necessary. Panels G, H, and I of Figure 3 show that the inclusion of a second gene did not contribute significantly to the variation of the normalization factor, as the value of $V_{2/3}$ was <0.15. Nevertheless, a third gene could be included in salt and hormonal stresses ($V_{3/4} < 0.15$) because the use of at least three reference genes is recommended by some authors.[17,26] In panels F and J of Figure 3, the values $V_{3/4} > 0.15$ and $V_{4/5} < 0.15$ indicated that the use of the four most stable genes was necessary for normalization in total samples and cold stress analyses.

NormFinder. The stabilities of the 13 genes were also analyzed using the NormFinder program. In Figure 4, the lowest stability

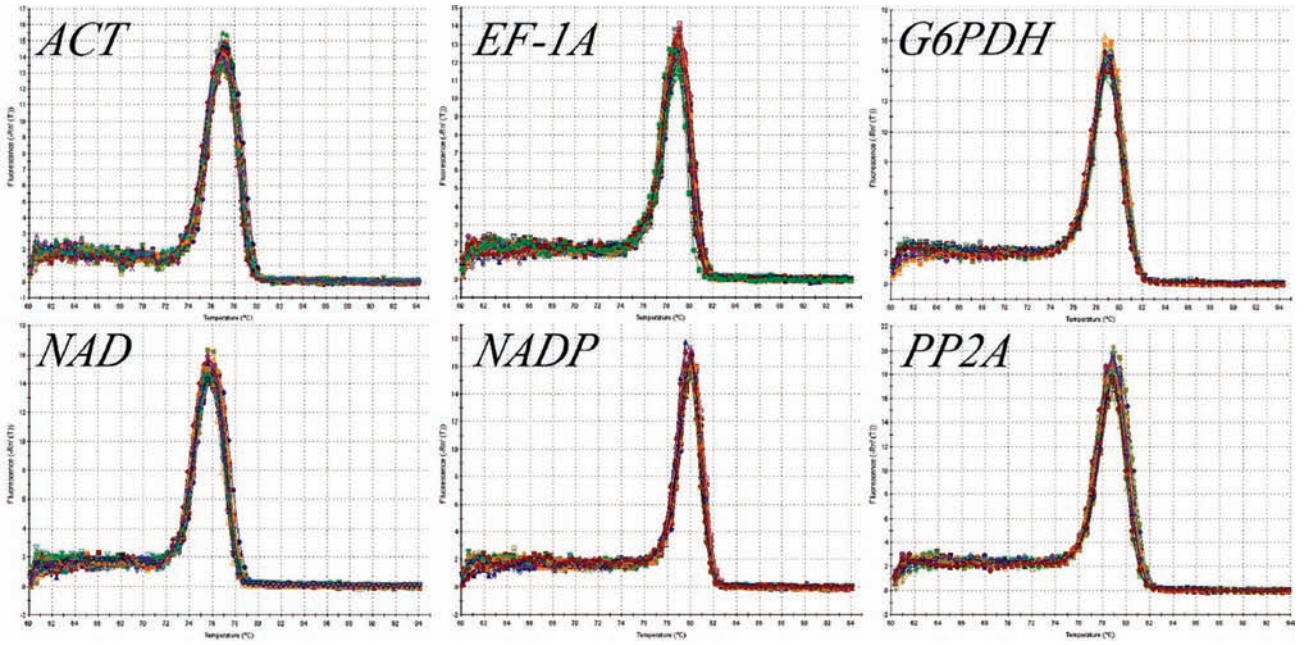

Figure 1. Dissociation curves of six representative reference genes, showing single peaks.

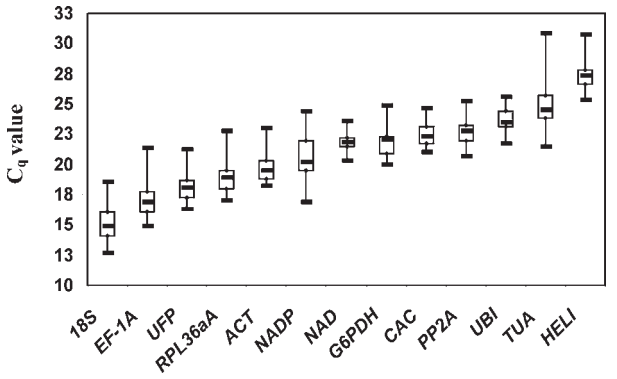

Figure 2. qPCR C_q values for reference genes in all samples. Boxes indicate the first and third quartiles. The vertical lines (whiskers) represent the maximum and minimum values. The central line represents the median value.

value indicates the most stably expressed gene. The results obtained for NormFinder analysis were similar to those from geNorm analysis. When all samples were evaluated, four subgroups were established, taking into account the intra- and intergroup variation (fruit and flower development stages, salt stress, hormonal stress, and cold stress). In this case, *UFP* and *ACT* genes showed the most stable expression levels, whereas *18S rRNA* was the least stable (Figure 4A). Among the subsets, *PP2A* and *CAC* were identified, as was the case in geNorm, as the two most stable genes in salt, hormonal, and cold treatment. *18S rRNA* was also found by both algorithms to be the least stable gene in the three treatments (Figures 4C,D,F). Fruit and flower developmental stages showed slight differences between geNorm and NormFinder algorithms; *PP2A* and *ACT* were found to be the most stable genes, whereas both analyses ranked *TUA* as the least stable gene (Figure 4B).

Choice of Reference Genes Affects Normalization. To test the impact of reference gene selection and to evaluate the suitability of the reference genes selected in this study, the differential expression of *CAT1* was calculated in leaves and roots subjected to salt stress as well as in control samples. On the basis of our data, three normalization factors (NF) were derived by calculating the geometric averages of the following gene combinations: NF_1, two top-ranking genes selected by geNorm and NormFinder in salt stress (*PP2A, CAC*); NF_2, three top-ranking genes identified by geNorm and NormFinder in salt stress (*EF1A, PP2A, CAC*); and NF_3, the least stably expressed gene according to both geNorm and NormFinder in salt stress (*18S rRNA*). Subsequently, we investigated whether these strategies resulted in significantly different expression results. Figure 5A shows relative expression levels of *CAT1* after normalization with the three normalization factors. When *CAT1* transcript quantities were normalized with the least stable gene (*18S rRNA*), the standard error increased. Figure 5B shows the mean coefficients of variation (CV) for the three normalization factors; NF_3 exhibited the majority of the variation (CV 75%). Normalization of *CAT1* expression with NF_1 (CV 38%) for salt treatment resulted in <1.2-fold enrichment of transcription compared to normalization with NF_2 (CV 47%). On the contrary, when normalization was performed with NF_3 (CV 75%), nearly 4.5-fold enrichment of transcription was obtained, compared to normalization with NF1.

DISCUSSION

Many factors in qPCR may affect the accuracy of the results and the reliability of the conclusions, including the selection of the reference genes. Normalizing to reference gene expression is a popular method for internally controlling for error and for eliminating nonbiological variation in qPCR. To our knowledge, there is no information in the literature regarding the choice of reference genes for gene expression studies in zucchini. Compared with other species of Cucurbitaceae, zucchini is a species with limited genetic information, hence the importance and difficulty of this study. In this work, we conducted a systematic study of the expression stability of 13 reference genes in zucchini,

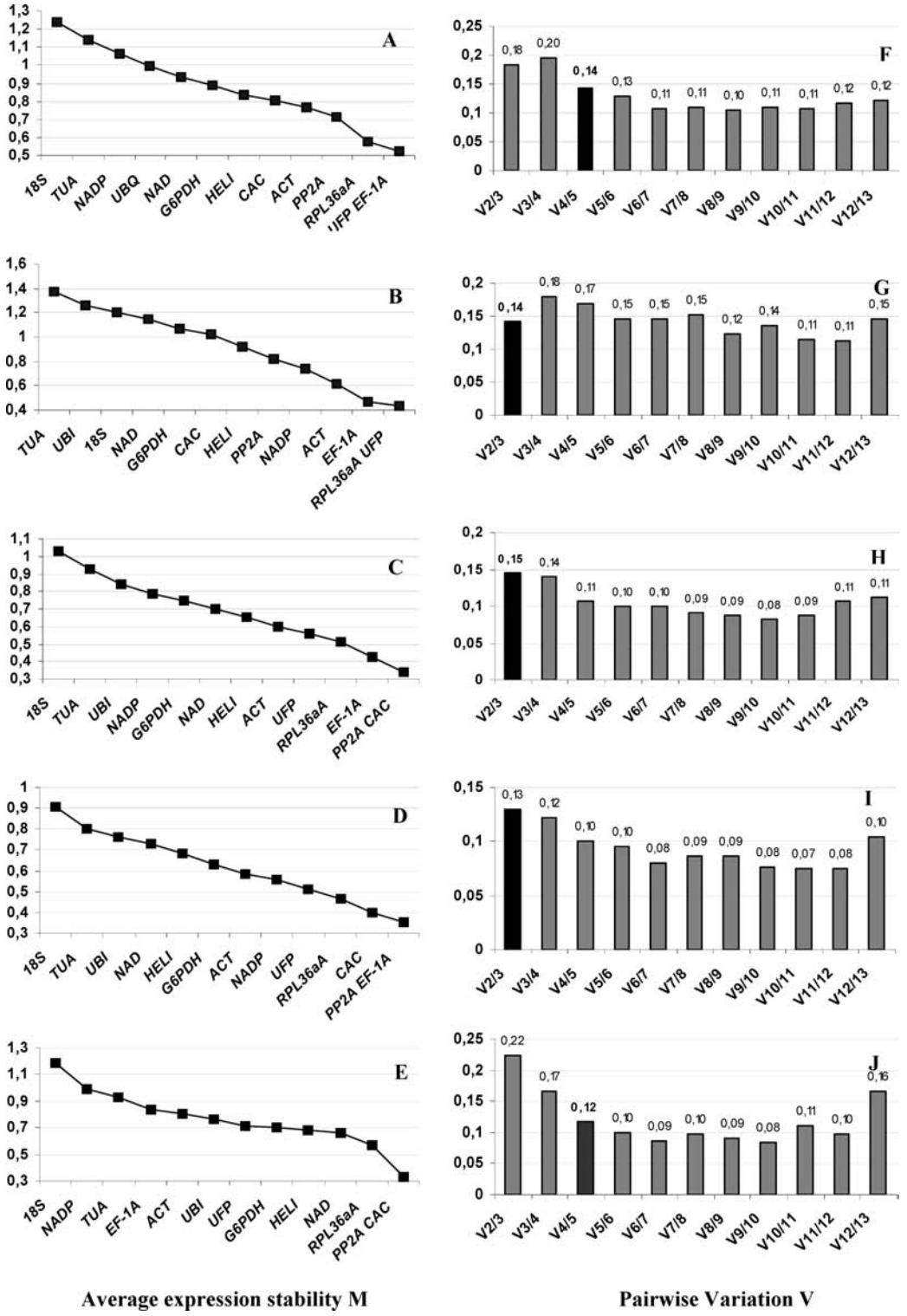

Figure 3. Expression stability and ranking of reference genes as calculated by geNorm in all samples (A), developmental stages in fruit and flower (B), salt treatment (C), hormonal treatment (D), and cold treatment (E). A lower value of average expression stability, *M*, indicates more stable expression. Pairwise variations calculated by geNorm to determine the minimum numbers of reference genes for accurate normalization in all samples (F), developmental stages in fruit and flower (G), salt treatment (H), hormonal treatment (I), and cold treatment (J).

along a series composed of 34 samples from different tissues/organs subjected to different stresses and at development stages.

Before starting qPCR quantification, we paid careful attention to preparative steps, which were performed in compliance with MIQE guidelines. Among these steps, it is considered critically important to select RNA samples with high RQI values. It is known that tissue handling in clinical applications must be carefully controlled to preserve the purity and integrity of the RNA.

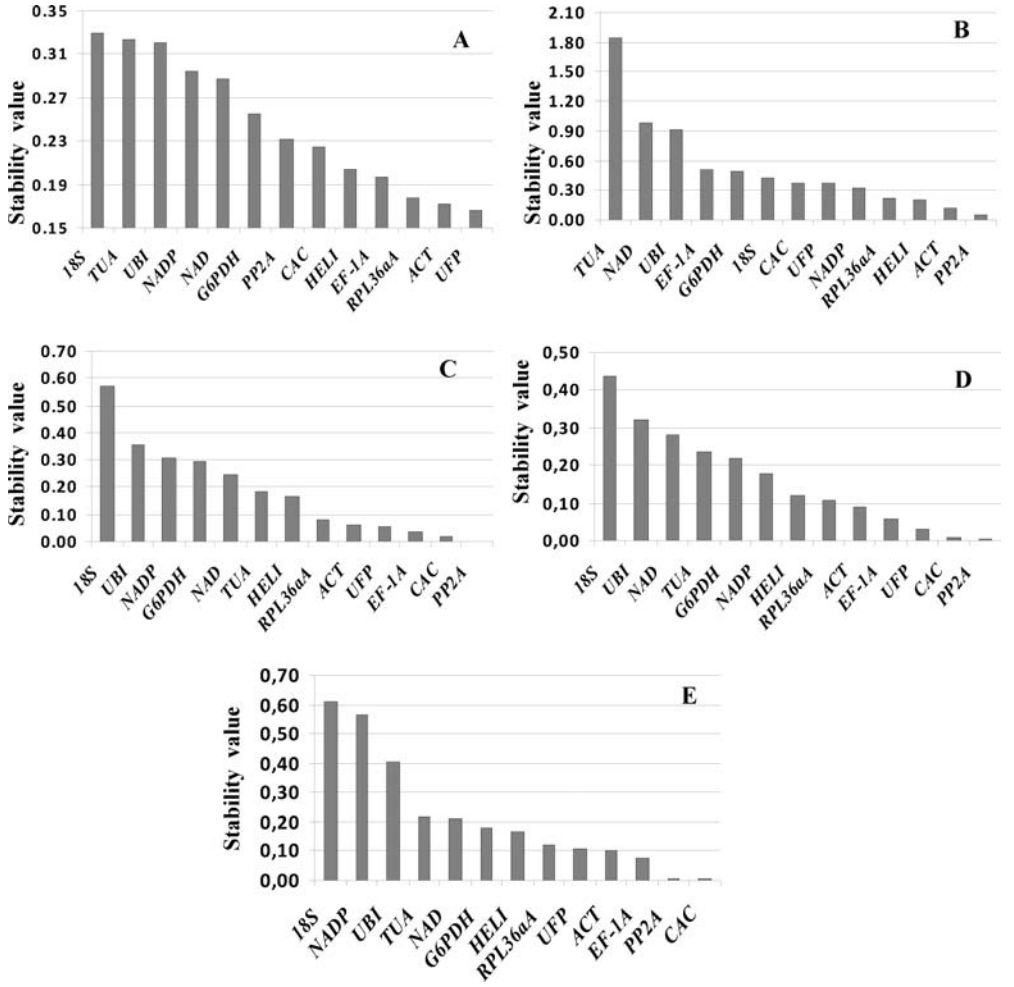

Figure 4. NormFinder analysis for the 13 reference genes. A lower stability value indicates more stable expression in all samples (A), developmental stages in fruit and flower (B), salt treatment (C), hormonal treatment (D), and cold treatment (E).

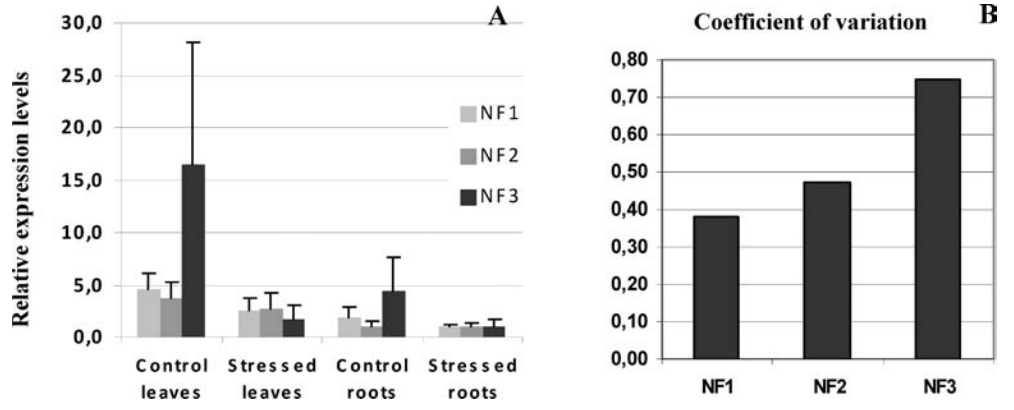

Figure 5. (A) Relative expression levels of *CAT1* in zucchini leaves and roots under salt treatment. *CAT1* expression was normalized with three normalization factors: NF_1 (*CAC*/*PP2A*), NF_2 (*CAC*/*PP2A*/*EF-1A*), and NF_3 (*18S*). Normalized values of *CAT1* relative expression are given as averages. Bars indicate standard errors. (B) Variation in *CAT1* after normalization. Values are given as averages of three different coefficients of variation.

Nevertheless, this step receives less attention in plants. Imbeaud et al.[27] reported that working with low-quality RNA may strongly compromise the results of downstream applications, which are often labor-intensive, time-consuming, and highly expensive. For this reason, we performed careful RNA analysis by using microcapillary electrophoresis with Experion before carrying out gene expression studies. Another important factor that affects the accuracy of gene expression analysis is DNA contamination. A minus RT control demonstrated the absence of contaminating genomic DNA, avoiding overestimation of the amount of RNA

present. Primer pairs for each gene were also optimized obtaining correct efficiencies because it is known that the presence of hairpins overlapping primer annealing sites and forming secondary structures in the amplicon have negative impacts on PCR efficiency.[28] Taking into account the fact that all qPCR assays were performed with proper controls and with the checklist discussed above, we were confident that our gene expression results were accurate and reliable.

Once the quality parameters were confirmed, we studied the expression patterns of genes being evaluated for their suitability as reference genes, using two different algorithms implemented in the programs geNorm and NormFinder. These statistical models were chosen because they take PCR efficiency into account, they are freely available, and they can analyze >10 genes. In addition, geNorm supplies additional information regarding the minimal number of genes necessary for valid normalization, and NormFinder is not significantly affected by the coregulation of candidate genes.

Although the two programs are based on different statistical algorithms, they showed several common features, and only slight differences in the rankings of the best scored genes were found. For example, *UPF*, *RPL36aA*, and *EF-1A* were among the four most stable genes all samples, but *ACT* was ranked second by NormFinder and fourth by geNorm. In contrast, both programs determined that *PP2A* and *CAC* were ranked highest for different abiotic stresses. *PP2A* and *CAC* were found to be the most stably expressed genes in *Arabidopsis* under abiotic stresses,[10] and *CAC* also showed high stability of expression in tomato, coffee, and cucumber studies,[11,29,30] indicating that novel reference genes are very often more stable in various experimental conditions than the commonly used internal controls. The largest differences between both algorithms were found when we analyzed flower and fruit developmental stages. geNorm ranked *RPL36aA* and *UFP* as the most stable genes, whereas NormFinder ranked *ACT* and *PP2A* highest. These discrepancies could be due to the C_q values from flower samples, which were the most variable among repetitions in our analysis compared with other tissues. The differences found in the results between geNorm and NormFinder methods were expected given that the programs are based on distinct statistical algorithms. Some previous publications that identified reference genes for plant research have also found discrepancies between both algorithms,[11,13,29,31−33] and there is not a consensus regarding which is the best method. Although we have found differences between both algorithms in this results set, the two most stable genes for geNorm (*RPL36aA* and *UFP*) were within the six most stable genes for NormFinder (stability value < 0.37), and the two most stable genes for NormFinder (*PP2A* and *ACT*) were also found within the six most stable genes for geNorm (stability value M < 0.82). geNorm indicated that studies involving cold stress and total samples required at least four control genes for reliable and accurate normalization, whereas two control genes were enough for experiments within different development stages or salt and hormonal stresses. However, a third gene could be included as is recommended by Vandesompele et al.[17]

In some others plants studied, *18S rRNA* and *TUA* were identified as the most stable expressed reference genes.[7,34−36] However, in the present analysis, *18S rRNA* and *TUA* showed considerable variation in different experimental conditions by the two algorithms and were always classified among the least reliable control genes for the four subsets and total samples. The poor stability of these genes was also found by several other authors.[30−33,37,38]

The expression of catalase (*CAT1*) was analyzed to verify the utility of the most stable reference genes as internal controls. Catalase is one of several cellular antioxidant defenses that are involved in detoxification of active oxygen species that are generated by various environmental stresses. In higher plants, several different catalase subunits are encoded by small multigene families, and differential patterns of expression have been reported in different of plant tissues and developmental stages.[39,40] Esaka et al.[39] demonstrated by Northern analysis differential expression of *CAT1* in various tissues of pumpkin. In the present experiment, roots and leaves were included to determine the different expression patterns by using qPCR technology. The authors considered cold and hormonal stress in their study, so it was our aim to complete this information by considering the salt stress. *CAT1* expression was found to be influenced more by tissue type than by salt stress, with mRNA levels higher in leaves than in roots.

To check the impact of including a third gene in the NF, we performed normalization for *CAT1* gene expression data by using two (NF_1) or three (NF_2) reference genes. In this particular case, normalization of *CAT1* with three genes did not result in significant differences in *CAT1* relative expression. Because the same results were obtained with two and three reference genes, we concluded that the inclusion of a third gene was not necessary; normalization with *PP2A* and *CAC* was sufficient for accurate normalization. Nevertheless, the number of reference genes for normalization must be carefully considered for each particular experimental situation and should reflect a balance between economic adequacy, experimental rigor, and practical considerations, with the aim of minimizing variation in NF and obtaining accurate results. Normalization with *18S rRNA* (NF_3) indicated significantly higher expression levels than normalization with a combination of several stably expressed genes, resulting in an overestimation of the fold changes in gene expression across control leaves and control roots. This indicates that normalizing with an improper reference gene greatly affects the apparent expression pattern of the target gene.

In conclusion, this is the first detailed study on the evaluation of selected reference genes in zucchini. By a combination of two software programs for data analysis, this study showed the genes *UFP*, *EF-1A*, and *RPL36aA* to be the most stable expressed reference genes in the whole sample set and the genes *PP2A* and *CAC* to be the most stably expressed under abiotic stresses. In the context of different development stages *RPL36aA* and *UFP* were identified by geNorm and *PP2A* and *ACT* by NormFinder. Because the two programs showed high stability value for the four genes, these genes can be used for normalization in fruit and flower developmental stages.

Our identification and validation of suitable zucchini control genes will facilitate future developmental transcriptomic studies in this economically important plant. With the genome sequence of other Cucurbitaceae available, it seems possible to transfer this information to zucchini, for which transcriptomic approaches are still lacking.

■ ASSOCIATED CONTENT

Supporting Information. Figure S1. Example of good-quality RNA, analyzed using Experion RNA StdSens Chip: (A) electropherogram of high-quality RNA for roots (sample 1, RQI = 10); (B) virtual gel for 12 samples (L = ladder). Samples 3 and 10 showed bands corresponding to plastids. Figure S2.

Dissociation curves of 7 reference genes, showing single peaks. Figure S3. Theoretical two-dimensional secondary structures of the amplicons generated by six different primers:(A) no secondary structures were present in the regions where the primers annealed; (B) the secondary structure between the primer annealing sites had a very small negative ΔG value and hence did not influence the amplification efficiency. Primers are indicated by the arrows. This material is available free via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Phone: +34-957-516061. Fax: +34-957-016043. E-mail: angeles.obrero@juntadeandalucia.es.

Funding Sources

Financial support was provided by the Spanish Project INIA-RTA2008-00035-C02-01, FEDER and FSE funds (EU). AOC acknowledges a predoctoral fellowship from INIA. JVD is researcher funded by the 'Juan de la Cierva' programme. CIGV has a contract from INIA (subprogram DOC INIA-CCAA). The authors thank the Ministerio de Ciencia e Innovación from Spain and FSE funds.

REFERENCES

(1) Bustin, S. A. Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *J. Mol. Endocrinol.* **2000**, *25*, 169–193.

(2) Pfaffl, M. W.; Hageleit, M. Validities of mRNA quantification using recombinant RNA and recombinant DNA external calibration curves in real-time RT-PCR. *Biotechnol. Lett.* **2001**, *23*, 275–282.

(3) Bustin, S. A.; Benes, V.; Garson, J. A.; Hellemans, J.; Huguett, J.; Kubista, M.; Mueller, R.; Nolan, T.; Pfaffal, M. W.; Shipley, G. L.; Vandesompele, J.; Wittwer, C. T. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem.* **2009**, *55*, 611–622.

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

(5) Nolan, T.; Hands, R. E.; Bustin, S. A. Quantification of mRNA using real-time RT-PCR. *Nat. Protoc.* **2006**, *1*, 1559–1582.

(6) VanGuilder, H. D.; Vrana, K. E.; Freeman, W. M. Twenty-five years of quantitative PCR for gene expression analysis. *BioTechniques* **2008**, *44*, 619–26.

(7) Brunner, A. M.; Yakovlev, I. A.; Strauss, S. H. Validating internal controls for quantitative plant gene expression studies. *BMC Plant Biol.* **2004**, *4*, 14.

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

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

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

(11) Expósito-Rodríguez, M.; Borges, A. A.; Borges-Pérez, A.; Pérez, J. A. Selection of internal control genes for quantitative real-time RT-PCR studies during tomato development process. *BMC Plant Biol.* **2008**, *8*, 131.

(12) Remans, T.; Smeets, K.; Opdenakker, K.; Mathijsen, D.; Vangronsveld, J.; Cuypers, A. Normalisation of real-time RT-PCR gene expression measurements in *Arabidopsis thaliana* exposed to increased metal concentrations. *Planta* **2008**, *227*, 1343–1349.

(13) Paolacci, A. R.; Tanzarella, O. A.; Porceddu, E.; Ciaffi, M. Identification and validation of reference genes for quantitative RT-PCR normalization in wheat. *BMC Mol. Biol.* **2009**, *10*, 11.

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

(15) Radonic, A.; Thulke, S.; Mackay, I. M.; Landt, O.; Siegert, W.; Nitsche, A. Guideline to reference gene selection for quantitative real-time PCR. *Biochem. Biophys. Res. Commun.* **2004**, *313*, 856–862.

(16) Nicot, N.; Hausman, J. F.; Hoffmann, L.; Evers, D. Housekeeping gene selection for real time RT-PCR normalization in potato during biotic and abiotic stress. *J. Exp. Bot.* **2005**, *56*, 2907–2914.

(17) Vandesompele, J.; De Preter, K.; Pattyn, F.; Poppe, B.; Van Roy, N.; De Paepe, A.; Speleman, F. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* **2002**, *3*, RESEACH0034.

(18) 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 Res.* **2004**, *64*, 5245–5250.

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

(20) Hoagland, H. D.; Arnon, D. I. The water-culture method for growing plants without soil. *Calif. Agric. Exp. Stn. Circ.* **1950**, *347*, 1–32.

(21) González-Verdejo, C. I.; Die, J. V.; Nadal, S.; Jiménez-Marín, A.; Moreno, M. T.; Román, B. Selection of housekeeping genes for normalization by real-time RT-PCR: analysis of Or-MYB1 gene expression in Orobanche ramosa development. *Anal. Biochem.* **2008**, *379*, 38–43.

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

(23) Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Bioinformatics* **2003**, *31*, 3406–3415.

(24) Ramakers, C.; Ruijter, J. M.; Deprez, R. H.; Moorman, A. F. Moorman, Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neurosci. Lett.* **2003**, *339*, 62–66.

(25) geNorm software manual update. http://medgen.ugent.be/~jvdesomp/ genorm/.

(26) Derveaux, S.; Vandesompele, J.; Hellemans, J. How to do successful gene expression analysis using real-time PCR. *Methods* **2010**, *50*, 227–230.

(27) Imbeaud, S.; Graudens, E.; Boulanger, V.; Barlet, X.; Zaborski, P.; Eveno, E.; Mueller, O.; Schroeder, A.; Auffray, C. Towards standardization of RNA quality assessment using user-independent classifiers of microcapillary electrophoresis traces. *Nucleic Acids Res.* **2005** *33*, e56.

(28) Hoebeeck, J.; van der, L. R.; Poope, B.; De, S. E.; Yigit, N.; Claes, K.; Zewald, R.; de Jong, G. J.; De, P. A.; Speleman, F.; Vandesompele, J. Rapid detection of VHL exon deletion using real-time quantitative PCR. *Lab. Invest.* **2005**, *85*, 24–33.

(29) Cruz, F.; Kalaoun, S.; Nobile, P.; Colombo, C.; Almeida, J.; Barros, L.; Romano, E.; Grossi-de-Sá, M. F.; Vaslin, M.; Alves-Ferreira, M. Evaluation of coffee reference genes for relative expression studies by quantitative real time RT-PCR. *Mol. Breed.* **2009**, *23*, 607–616.

(30) Migocka, A.; Papierniak, M. Identification of suitable reference genes for studying gene expression in cucumber plants subjected to abiotic stress and growth regulators. *Mol. Breed.* **2010**, doi: 10.1007/s11032-010-9487-0.

(31) Wan, H.; Zhao, Z.; Qian, C.; Sui, Y.; Malik, A. A; Chen, J. Selection of appropriate reference genes for gene expression studies by quantitative real-time polymerase chain reaction in cucumber. *Anal. Biochem.* **2010**, *399*, 257–261.

(32) Lin, Y. L.; Lai, Z. X. Reference gene selection for qPCR analysis during somatic embryogenesis in longan tree. *Plant Sci.* **2010**, *178*, 359–365.

(33) Huis, R.; Hawkins, S.; Neutelings, G. Selection of reference genes for quantitative gene expression normalization in flax (*Linum usitatissimum* L). *BMC Plant Biol.* **2010**, *10*, 71.

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

(35) Jarosová, J.; Kundu, J. K. Validation of reference genes as internal control for studying viral infections in cereals by quantitative real-time RT-PCR. *BMC Plant Biol.* **2010**, *10*, 146.

(36) Maroufi, A.; Bockstaele, E. V.; Loose, M. D. Validation of reference genes for gene expression analysis in chicory (*Cichorium intybus*) using quantitative real-time PCR. *BMC Mol. Biol.* **2010**, *11*, 15.

(37) Cortleven, A.; Remans, T.; Brenner, W. G.; Valcke, R. Selection of plastid- and nuclear-encoded reference genes to study the effect of altered endogenous cytokinin content on photosynthesis genes in *Nicotiana tabacum*. *Photosynth. Res.* **2009**, *102*, 21–29.

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

(39) Esaka, M.; Yamada, N.; Kitabayashi, M.; Setoguchi, Y.; Tsugeki, R.; Kondo, M.; Nishimura, M. cDNA cloning and differential gene expression of three catalases in pumpkin. *Plant Mol. Biol.* **1997**, *33*, 141–155.

(40) Robertson, C. Regulation of catalases in *Arabidopsis*. *Free Radical Biol. Med.* **1997**, *23*, 489–496.