

Selection of Reference Genes for Transcription Studies on Adventitious Root Induction in Olive (*Olea Europaea* L.) Microshoots Considering Co-expression and Average Transcriptional Stability

Carlos Noceda (✉ cana293@yahoo.es)

University of Armed Forces-ESPE <https://orcid.org/0000-0001-6633-2698>

Augusto Peixe

University of Evora

Birgit Arnholt-Schmitt

University of Evora

Research

Keywords: RT-qPCR, reference genes, expression stability, co-regulation, olive, adventitious rooting, alternative oxidase

DOI: <https://doi.org/10.21203/rs.3.rs-60001/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background

Selection of reference genes (RGs) for normalization of PCR-gene expression data includes two crucial steps: determination of the among-sample transcriptionally more stable genes and subsequent choosing of the most suitable genes as internal controls. Both steps can be carried-out through generally accepted strategies each having different strengths and weaknesses. The present study proposes to reinforce normalization of gene expression data by integrating and adding analytical revision at critical steps of those accepted procedures. Especially crucial is to counterbalance a higher representative number of RGs with a correspondent increase in their average transcriptional instability or a generalised co-expression trend among the samples. This methodological study used *in vitro* olive adventitious rooting as an experimental system, since the underlying morphogenetic process -which is common to diverse species- is still not completely understood.

Results

Firstly, RG candidates were ranked according to transcriptional stability following a simple statistical method that reduces biasing effects of concomitant, systematic biological variations associated to experimental conditions, such as the variations caused by gene co-regulation. Those types of systematic co-variation are unconsidered by several popular *ad hoc* informatics programmes. To select the adequate genes among those already ranked, an algorithm of one of the *ad hoc* informatics programmes (GeNorm) was adapted to allow partial automatization of RG selection for any strategy of transcriptional-gene stability ordering. In order to delve into the resulting possible RG sets suitability for inter-assay comparisons and technical-error compensation, separate statistics were formulated. The achieved results were compared with those obtained by standard stability ranking methods. Finally, a double evaluation was performed to accurately contrast two choice RG sets. The whole strategy was applied to a panel considering several independent factors, but the suitability of the obtained putative RG sets was tested for cases restricted to fewer variables. *H2B*, *OUB* and *ACT* are valid for normalization in transcriptional studies on olive microshoot rooting when comparing treatments, time points and assays.

Conclusions

The set of genes identified as internal reference is now available for wider expression studies on any target gene in similar biological systems. The overall methodology aims to constitute a guide for general application.

Background

RNA reverse-transcription (RT) followed by real-time quantitative polymerase chain reaction (qPCR) seems to be the most appropriate technique to study transcript levels of a reduced number of genes in complex "sets of conditions" (panels) [1,2]. The accuracy and biological significance of RT-qPCR assays depend on proper normalization with internal or external controls. When transcripts from genes are used as internal controls, their molecule amounts, absolute or relative, are assumed as representative of the total integer RNA mass in each sample. Consequently, the genes selected for this purpose are named reference genes (RGs). RGs must be transcriptionally stable among samples in terms of accumulation of number of gene transcripts with regard to the total RNA mass per sample. The use of RGs is commonly accepted as the most appropriate normalization methodology in RT-qPCR assays [3], but even the more stably expressed genes among tissues, individuals or groups can be differentially regulated under certain conditions (reviewed in [4]). Thus, it is necessary to critically evaluate the expression stability of potential RGs under different experimental set-ups [5]. This is especially true for complex design panels, such as the case of testing the influence of several factors in gene expression.

To calculate different transcript relative levels, a normalization factor (NF, the geometric mean of the relative levels of the RG transcripts) is applied to the measured value of each target transcript level for each sample. The usual procedure to determine the gene composition of the NF begins ordering a group of genes candidate to be RGs according to their expression stability. Such a ranking can be established through diverse accepted mathematical strategies with different assumptions, strengths and weaknesses each. Consequently, distinct mathematical strategies frequently throws different orders of genes ([5–9], reviewed in [10]). Thus, commonly researchers take on the task of justifying their preferred ranking.

Often, the most transcriptionally stable genes are selected after transcriptional gene-stability ranking. The number of those selected genes is crucial: too few genes may be not representative and too many may include not enough transcriptionally stable genes, thus decreasing average stability of RG set. Then, again, the researcher is the person who has to argue which genes should be chosen. An algorithm exists that automatizes this procedure, so far limited to be used after applying a concrete stability ranking software (GeNorm [11]), and which could lead to the inclusion as RGs of not enough stable genes if additional checking criteria are not applied. An example of this issue is the case showed in the present work, in which we developed a strategy for RG selection for expression studies on olive *in vitro* adventitious rooting.

The differential capacity of distinct olive (*Olea europaea* L.) cultivars to develop adventitious roots has been attributed to anatomical [12,13] as well as physiological, biochemical and genetic [14,15] causes. Although gene expression studies during adventitious root formation may provide significant information on the regulatory control of genes involved in the rooting process, transcriptional modulation has been scarcely studied on explants growing or rooting, especially when *in vitro* conditions are used [9]. In the olive, some enzyme-coding genes related with cell oxidative status were proposed to have a role in adventitious rooting [16,17] through phenylpropanoid and lignin metabolism [18], including those coding alternative oxidase (AOX), a mitochondrial respiratory enzyme involved in stress resistance [19,20]. The supplementation of salicylhydroxamic acid (SHAM), an AOX inhibitor, in rooting inductive conditions, inhibits the adventitious rooting processes in olive semi-hardwood cuttings [21] and *in vitro* cultured microshoots [18].

Based on the use of SHAM, we established olive microshoot *in vitro* experiments for transcriptional studies under adventitious rooting-permissive and inhibitory conditions. To overcome the above-mentioned limitations on RG selection strategies, we developed a comprehensive method by critically integrating

the more robust steps from several strategies, to thus strengthen normalization of gene expression data. The method includes an unbiased transcriptional stability-ordering of genes and the adaptation of the GeNorm selection algorithm to any transcriptional stability ranking. The achieved results were compared with those obtained with the standard stability ranking strategies. In order to delve into the resulting possible RG sets suitability for inter-assay comparisons and technical-error compensation, separate statistics were formulated. Finally, a double-evaluation process was developed to accurately contrast two choice RG sets. The whole strategy was applied to an experimental panel considering several independent factors -such as treatment (BA with or without SHAM, control) time of culture and assay (repeat with same explant type proceeding from different biological material)- but the suitability of putative RG sets was tested for cases restricted to fewer variables.

The overall methodology was developed in a case-specific study, but constitutes a guide for general application. A set of three RGs was identified as internal reference and is now available for wider expression studies on any target gene in similar systems.

Methods

Analysis of SHAM effects on adventitious rooting

Explants from a single clone of *Olea europaea* L. cv. 'Galega vulgar', pre-cultured *in vitro* according to [22], were used in all trials. Indole-3-butyric acid (IBA; Sigma-Aldrich, St. Louis, MO, USA) was used as the root promoting auxin at 14.7 mM. Since an involvement of AOX in adventitious rooting has been supposed [21], SHAM (Sigma-Aldrich, St Louis, MI, USA), a potential AOX inhibitor, was used to provide a restrictive treatment for adventitious root formation. Fresh SHAM was prepared in dimethyl sulfoxide (DMSO, Fluka, France), and the final concentrations into the IBA solution were SHAM 100 mM, DMSO 26 % (v/v). Used in this concentration, DMSO has no inhibitory effect on rooting [18]. Microshoots with 3-5 nodes, keeping the four full expanded apical leaves were used for the rooting trials. The basal parts of the explants (approx. 1 cm) were dipped for 10 s into the IBA solution (with and without SHAM). After that, the explants were *in vitro* cultured on a rooting medium devoid of growth regulators as proposed in [22]. A control without any dipping treatment was also used. Thus, three treatments were established in the final subculture: rooting medium without immersion of basal part of explants (negative control), with initial immersion in IBA (rooting) or initial immersion in IBA+SHAM (rooting inhibition). Six assays each with these three treatments were performed. The derived *in vitro* grown plantlets were used in all subsequent analyses. Visible root formation was recorded at 22 and 28 days after transfer into rooting medium.

Biological material for transcript quantification analyses

Three assays were selected for molecular analyses, on the basis of differential rooting capacity for the treatment with IBA and without SHAM: assays I, II and III, ordered according to rooting capacity (Table 1). This differential behaviour confers more robustness when obtaining conclusions about factors influencing adventitious root induction processes. Thus, in spite that assay is a fixed factor, results would be also valid when considering it as a random factor, and it can function as biological replicate when considering only other factors.

Table 1. Rooting percentages¹ for each treatment and assays ordered by rooting capacity

22 days after induction			28 days after induction			Assay (rooting capacity)
Control	IBA	IBA+SHAM	Control	IBA	IBA+SHAM	
0 (50)	60 (50)	0 (50)	-	-	-	I
0 (25)	38 (40)	0 (40)	-	-	-	II
0 (25)	0 (50)	0 (25)	0 (25)	32 (25)	0 (25)	III

¹In parenthesis, number of total microshoots.

For RT-qPCR analyses on gene expression, relative transcript accumulation was measured at five different time points (0 h, 4 h, 24 h, 48 h and 96 h) after microshoots inoculation in the rooting media. These time points were selected after inspection of the RT-semi-quantitative PCR results on *AOX* transcript accumulation under the same conditions here assessed along a longer period of time (data not shown). For real time whole quantitative analyses, a trifactorial, complete panel was designed: 5 time points x 3 treatments x 3 assays. For a bifactorial panel, the assays were considered as biological replicates, as above explained. In all cases each biological replicate consisted of a bulk of eight basal portions of microshoots, cut from the half of the first basal visible complete internode. All samples were immediately frozen in liquid nitrogen and stored at -80 °C until required. Two technical replicates (RT-level) were performed.

Nucleic acid extraction

RNeasy Plant Mini Kit (QIAGEN, Hilden, Germany) including DNase (RNase-free DNase set, Hilden, Germany) application, and DNeasy Plant Mini Kit (QIAGEN, Hilden, Germany) were used to extract RNA and DNA, respectively, from microshoot bottoms. Final elutions were made in 50 µl nucleases free water. DNA and RNA integrity, so as RNA contamination with DNA was checked carrying out electrophoresis in 1 % (w/v) SeaKem LE Agarose (Lonza, Rockland, ME, USA) gel (50 mV for 45 min and revealing with ethidium bromide solution (2 ng mL⁻¹) using the Gene Flash Bio Imaging system (Syngene, Cambridge,

UK). After extraction, RNA samples were aliquoted to be used once each. Both DNA and RNA concentrations were determined using NanoDrop 2000C (Thermo Fisher Scientific Inc., USA).

Reverse transcription of mRNA

Total RNA concentration was adjusted to 100 ng/uL after dilution, measurement and re-adjustment. Previous heating of RNA at 60 °C for 5 min, to eliminate tertiary structures, was tested and did not affect later reverse transcription plus amplification, thus it was always performed. Three hundred ng of total mRNA were 1st-strand reverse-transcribed using the Maxima[®] Reverse Transcriptase (Thermo Fisher Scientific Inc.®, Waltham, MA) in a total volume of 20 µl according to the manufacturer's instructions. The Supermix contained only oligo dT primer to decrease variability factors affecting RT efficiency. The reaction mixture was incubated at 25°C for 5 min for priming, then at 42°C for 30 min for reverse transcription, and finally at 85°C for 5 min for reverse transcriptase inactivation. The complementary DNA (cDNA) was stored at -20°C until further use. cDNA samples were aliquoted to be used always fresh in later PCRs.

Selection of candidate genes as internal reference genes

Candidate genes to be used as internal RGs for mRNA differential accumulation studies are usually cell housekeeping genes, which are supposed to be stably expressed in most cases. Thus, seven candidate housekeeping genes (CHGs) were selected on the basis of 1) their average transcript accumulation and measured transcript accumulation stability during rooting induction and other plant physiological processes ([9] and references therein), and 2) their differential encoding of molecular components representing a cross-section of functional diversity in plant cell physiology. This last measure minimizes the likelihood of a putative co-regulation effect among genes that may respond in parallel to particular experimental conditions. Such precautions are a prerequisite for one of the statistical procedures here used (the geNorm –gNo-algorithm) to identify stably accumulated gene transcripts belonging to several CHGs [23]. After determination of the CHGs, these were searched in NCBI databases of olive. When not found, the ortholog sequences of *Arabidopsis thaliana* were searched and submitted to a BLAST to find the homologous olive sequences, either in genes or in cDNA. The characteristics of the selected CHGs and correspondent amplicons are summarized in [Table 2](#).

Table 2. Characteristics of the obtained amplicons of the tested candidate housekeeping genes.

Name	Abbreviation	Role	Blasted sequence (NCBI)	Blast identities and gaps	Found <i>O. eu</i> used for primer design
Beta-actin	<i>ACT</i>	Microfilament component	AF545569.1 (act1)	act1 mRNA, partial CDS	<u>TTGCTCTC</u>
Elongation factor 1-alpha	<i>EF</i>	Translational elongation	AM946404.1	Elongation factor partial gene, exons 1-2	<u>TTTTGAGG</u>
Glyceraldehyde-3-phosphate dehydrogenase	<i>GADPH</i>	Glycolysis enzyme	NM_106601.3 (A. thaliana GAPCP1 mRNA, complete CDS)	447/553 (81 %), 0/553 (0 %)	FL684222.1
Histone H2B	<i>H2B</i>	Chromatin structure	NC_003076.8 At5g59910 (A. thaliana HTB4)	328/408 (80 %), 3/408 (1%)	G0244518.1
Small heat shock protein 18.3	<i>Hsp</i>	Stress response	FN554869.1	mRNA for putative class I hsp18.3, cv. Cellina di Nardo	<u>ACTTGGCACCGCAT</u>
Polyubiquitin	<i>OUB</i>	Protein degradation	AF429430.1	OUB2 mRNA, complete CDS	<u>AGGCATCC</u>
Alpha-tubulin	<i>TUA</i>	Microtubule structure	EF506517.1 (O. europaea putative alpha	237/282 (84 %),	G0245051.1

Design and testing of primers

See Additional file 1

Cloning of amplified fragments and verification of obtained amplicons

PCR fragments generated from each gene amplification were purified using the GFX PCR DNA and Gel Purification kit (GE Healthcare, Little Chalfont, England). The addition of a single deoxyadenosine 5'-monophosphate to the 3'-end of amplicons required for the cloning procedure was made by adding 0.1 U μL^{-1} of Taq polymerase (Promega, Madison, WI), 1 \times manufacturer supplied $(\text{NH}_4)_2\text{SO}_4$ buffer, 2.5 mM MgCl_2 and 0.2 mM dATP (Fermentas, Ontario, Canada). The final mix was incubated for 30 min at 72 °C in a 2720 Thermalcycler (Applied Biosystems, Foster City, CA). The amplicons were cloned into a pGEM[®]-T Easy System I vector (Promega, Madison, WI) and subsequently used for the JM109 *Escherichia coli* competent cells transformation (Promega, Madison, WI). Bacterial handling was followed according to [24]. Plasmid DNA of white bacterial colonies was extracted using the GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific Inc., Lithuania), characterized by the restriction enzyme EcoRI and visualized after electrophoresis (100 mV, 30 min) in SeaKem LE Agarose (Lonza, Rockland, ME, USA) 1.4 %, as described above for checking nucleic acid integrity. Recombinant clones were quantified by NanoDrop 2000C (Thermo Fisher Scientific Inc., USA) and prepared for commercial sequencing (Macrogen[®], Amsterdam, Netherlands). The obtained sequences were compared with original sequence to corroborate amplification of the selected region, by using the online tool GeneBee.

Quantitative polymerase chain reaction (qPCR)

Two-step RT-qPCR setting was carried out. Amplification reactions with all primer pairs were performed with sample maximization design of the plate [25]. Aliquots of the same cDNA sample were used with all primer sets for qPCR. Reactions were done in a 18 μL volume containing 600 nM (tested before for optimal concentration) of each primer, 1.8 μL of 9 template cDNA:41 water (v:v) (what is equivalent to a total of 4.86 ng of input RNA), Maxima SYBR Green qPCR Master Mix (2X) and 10 nM ROX as reference dye (both last products from Thermo Fisher Scientific Inc.[®], Waltham, MA) in 96-well microtitre plates (Applied Biosystems, China) which were spun in a microplate centrifuge (VWR) to avoid bubbles. qPCR step was performed on the ABI Prism 7500 Sequence Detection System (Applied Biosystems, China) using the parameters recommended by the manufacturer (50 °C for 2 min, 95 °C for 10 min and 40 cycles of 95 °C for 15 s and 60 °C for 1 min; maximum ramp rate). To verify that the used primer pair produced only a single product, a dissociation protocol was added after thermocycling, determining dissociation of the PCR products from 65 °C to 95 °C increasing the temperature stepwise by 0.5 °C every 10 s. Furthermore, PCR products were run in a gel to confirm a single band. Baseline and quantification cycle (C_q, defined as the number of cycles needed for the fluorescence signal to reach a specific threshold of detection), so as specificity of the amplifications were automatically determined using the 7500 Software v. 2.0.5 (Applied Biosystems 2010 Life Technology Corporation). The assay included two no-template controls without RNA or cDNA. No-reverse transcription controls (RNA from reverse transcription without reverse transcriptase) were run for each sample for the gene with lowest measured transcript accumulation. Although 8 samples per group are advisable for RG selection using NormFinder (NFI) software [26], this programme was run with 6 (time \times treatment). This number of samples was considered enough since: i) it is constant for a considerable number groups (all groups) and, ii) the number of CHG to test is higher than the minimum advisable (five) [26]. Less robust was the inter-group NFI analyses when taking into account also the assay inside of each treatment \times time cell.

Additional test for gDNA contamination

gDNA from two microshoot pools of the olive clone under evaluation was extracted employing DNeasy Plant Mini Kit (QIAGEN, Hilden, Germany) following manufacturer instructions. Additionally to no-reverse transcription control, the possible gDNA contamination in cDNA samples was made by checking the absence of gDNA bands of *AOX1a* (higher than the correspondent cDNA bands because of an intron) in the RT-PCR reactions (data not shown).

Calculation of PCR efficiency

To compare different qPCR runs, performed on different plates, all runs were adjusted to the threshold of C_q 0.1. C_q is, as it is deduced from its definition, inversely correlated to the input amount of total RNA [27]. Raw data, including the melting and amplification curves obtained by the Applied Biosystems software, were extracted to Microsoft[®] Excel files and then loaded for further data analysis.

The PCR efficiency of each primer pair was obtained by using LinRegPCR [28]. This program uses amplification data captured during the exponential phase of each PCR reaction after reconstructing baselines, a source of variation of the observed efficiency [28]. From a total of 78 amplification plots (i.e. two technical replicates of three biological replicates –assays– of a total of 13 different experimental conditions) per CHG, those that passed all quality tests of the program were used for calculations.

To overcome the problems derived from downstream mathematical proceedings, that by default do not take into account PCR efficiencies, corrected quantification cycle –C_q' = C_q log₂(efficiency)– values were obtained using calculated efficiencies. Then, relative contents for each amplicon were calculated

via the ΔCq method using one replicate of time 0 as reference.

Data analysis

A method for RG selection using estimators of coefficients of total variation (CV) and inter-group variation (F) of gene transcript accumulation values [5] was carried out (for simplicity, here named CV/F method). For a formulation of the mentioned relative variations, see [5]. To compare and eventually support CV/F results with different algorithms for the selection of the most transcriptionally stable reference genes, RefFinder (RFi) (<http://www.leonxie.com/referencegene.php>), a web-based comprehensive tool, was applied. To calculate values on gene transcript accumulation stability, RFI uses the currently available algorithms BestKeeper [29] –regarding Cq standard deviation (BKS)–, comparative ΔCq (CoD) [30], geNorm (gNo) [23] and NormFinder (NFi) [26] –only ungrouped samples–, assigning an appropriate weight to each one and ordering the CHG according to the output gene transcript accumulation stability measure. Each individual algorithm implemented in RFI also generates a ranking of genes. gNo, f.e., via a stepwise exclusion of the least stable gene, creates a stability ranking. RFI also calculates the BestKeeper Pearson coefficient of correlation (BKR) [29], although it is not used in later RFI weighting of stability parameters. Independent NFi software, additionally used in this work, uses an ANOVA-based model to estimate intra- and inter-group variation, and combines these estimates to provide a direct measure of the variation in transcript accumulation for each gene. To estimate differences among normalization factors (NFs, which equals the geometric means of the RG relative transcript accumulations for each replicate) the methodology of gNo for pairwise comparisons [23] was applied. For NF evaluation tests, two- and three-way ANOVA models with Bonferroni post-hoc adjustments ($P < 0.05$) were performed, after data transformation to better keep the model assumptions, by loading data into the statistical programme IBM SPSS (v21, SPSS Inc., USA). For further revision of the results, inter-replicate variations (estimated as relative variation within biological samples) and inter-assay variations (estimated as ratio of the relative variations within bifactorial groups –time x treatment– and within trifactorial groups –biological samples) were calculated, so as overall and individual CHG inter-replicate mean distances.

Results And Discussion

The justification and results of every step of the procedure for selection of RGs is discussed as follows.

Step 1. Selection of genes as candidates for reference genes and amplification tests

Genes were selected according to procedures specified in Material and Methods. Single amplicons for each CHG were obtained whose specificity was confirmed by the observation of single-peak melting curves of the qPCR products or by the presence of a single band of expected size for each primer pair in agarose gel electrophoresis after PCRs employing either gDNA (data not shown) or cDNA as templates [see Additional file 2]). No primer dimers or other products resulting from non-specific amplification were found.

Step 2. Calculation of Cq's

Calculated amplification efficiencies ranged from 1.855 to 1.935 (Table 2), values appropriated to be subsequently utilised in Cq's calculation. The different abundance of each RG transcript affect the normalized results [31], but in the present case no considerable differences in transcript accumulation levels were observed. Furthermore, suitable RGs should be equivalent in transcript abundance to that of the target gene (TG), whose Cqs should be between 15 and 30 [26], limits within which the Cqs obtained in this analysis fits (data not shown; See Fig. 1 for a visual inspection of Cq's).

Fig. 1. Corrected quantification cycle (Cq) values for the seven candidate housekeeping genes. Mean of biological groups is represented, with upper semi-bars indicating Cq' SD for biological groups ($n=6=3$ assays x 2 RT-level replicates), and bottom semi-bars indicating Cq' mean SD for samples (3 samples with 2 RT-level replicates each)

Step 3. Estimation of transcript level variations:

The theoretically optimal way to identify the more transcriptionally stable candidate genes is trough estimation of both overall and inter-group variations in transcript levels under the employed experimental conditions [5]. With this purpose, two consecutive procedures were performed, as follows:

Displaying of Cq' means and standard deviation for experimental groups.

The sample maximization design here used for plate arrangement avoids inter-run variability among samples, frequently underestimated [25]. Cq' values for the amplicons of the seven CHGs were depicted (Fig. 1) to allow a visual evaluation of the transcript level stabilities and common trends among genes along the different factors (time, treatment). *Hsp* resulted as the most unstable CHG, with a general opposite trend to *ACT*, *EF*, *GADPH*, *H2B* and *TUA*, all of them genes showing a high degree of co-expression, *v.g.*, co-regulation, in spite of they are functionally unrelated in a direct way. Co-regulation may be induced by stress or multi-stress [32,33], and could bias RG selection. In spite of their co-regulation, those genes show an opposite trend to the remaining and also stable gene, *OUB*, a behaviour that can be associated to its pre-proteolytic function, i.e., it is expressed when protein replacement is needed. The observed intra-group (time x treatment) variations were moderate except for *Hsp*, to what contributes a heterogeneous group-dependent variability among biological samples, since its intra-group inter-replicate variabilities were low (Fig. 1).

Calculation of transcriptional gene stability values

CV and F statistics estimates, respectively, the overall and inter-group variation of transcript levels of single CHGs without counterbalancing with the rest of CHGs. Thus, CV and F estimators provide stability rankings not so biased by concomitant, systematic biological variations that may be associated to an experimental condition, leading to a group of genes to follow a similar expression trend -such as the variation due to co-regulation. Thus, CV and F provide in most cases a good classification of the stability of CHGs. It is worth to point out that the purpose of CV and F statistics is the generation of a transcriptional

gene stability ranking rather than gene transcriptional stability values, and not at all the obtaining of significant differences. Hence, inference statistics prerequisites such as normal distribution are not mandatory for the pursued goal. F was calculated for bi- (time x treatment) and trifactorial (time x treatment x assay) cases (thus F was named as F2 and F3, respectively).

Step 4. Ranking of candidate housekeeping genes according to their stability

The more transcriptionally stable genes should have a more or less parallel classification in both CV and F descriptive parameters [5]. In the CV/F method, preference is given to CV (overall variation) ranking [5], thereby overcoming possible large discrepancies caused by opposite and extreme values of CV and F. The rankings of CHGs according to CV/F method -so as used software algorithms- are shown in Table 3.

All replicates						Time x Treatment			Time x Treatment x Assay		
RFi algorithms						CV	F2	independent NFi	F3	independent NFi	
BK _r ²	BKS	CoD	GNo	NFi	RFi						
All tested housekeeping genes	EF (0.67)	H2B (0.34)	H2B (1.07)	EF/GADPH	OUB (0.24)	H2B (1.57)	H2B (0.29)	<i>Hsp</i> (4.84)	<u>OUB (0.24)</u>	H2B (11.66)	<u>OUB (0.2)</u>
ACT (0.57)	OUB (0.38)	ACT (1.12)	(0.59)	H2B (0.32)	OUB (2.45)	OUB (0.34)	H2B (7.70)	<u>H2B (0.26)</u>	OUB (14.59)	<u>H2B (0.28)</u>	
OUB (0.55)	ACT (0.58)	OUB (1.14)	H2B (0.69)	ACT (0.38)	EF (2.83)	ACT(0.43)	OUB(8.49)	ACT (0.31)	ACT (26.05)	ACT (0.33)	
GADPH (0.54)	EF (0.68)	EF (1.16)	TUA (0.74)	EF (0.66)	ACT (3.08)	EF (0.56)	ACT(10.79)	EF (0.45)	TUA (33.03)	EF (0.46)	
TUA (0.43)	GADPH (0.76)	GADPH (1.17)	ACT (0.79)	GADPH (0.76)	GADPH (3.34)	GADPH(0.63)	EF (13.53)	GADPH (0.49)	EF (34.86)	GADPH (0.5)	
H2B (0.38)	TUA (0.80)	TUA (1.23)	OUB (0.82)	TUA (0.88)	TUA (5.42)	TUA (0.67)	TUA (17.89)	TUA (0.58)	GADPH (38.6)	TUA (0.56)	
<i>Hsp</i> (0.30)	<i>Hsp</i> (2.01)	<i>Hsp</i> (2.79)	<i>Hsp</i> (1.38)	<i>Hsp</i> (2.74)	<i>Hsp</i> (7.00)	<i>Hsp</i> (3.04)	GADPH (19.02)	<i>Hsp</i> (1.43)	<i>Hsp</i> (1237.12)	<i>Hsp</i> (1.37)	
Best 6 tested housekeeping genes	EF (0.86)	H2B (0.34)	H2B (0.74)	EF/GADPH (0.58)	H2B (0.42)	H2B (1.32)	H2B (0.29)	H2B (7.70)	H2B (0.21)	H2B (11.66)	H2B (0.2)
GADPH (0.85)	OUB (0.38)	GADPH (0.80)	GADPH (0.57)	GADPH (2.11)	OUB (0.34)	OUB (8.49)	ACT (0.28)	OUB (14.59)	ACT (0.30)		
TUA (0.80)	ACT (0.58)	EF (0.82)	H2B (0.70)	EF (0.58)	EF (2.45)	ACT (0.43)	ACT (10.79)	EF (0.31)	ACT (26.05)	EF (0.32)	
H2B (0.64)	EF (0.68)	ACT (0.82)	TUA (0.74)	ACT (0.59)	ACT (3.94)	EF (0.56)	EF (13.53)	<u>GADPH</u> (0.32)	TUA (33.03)	<u>GADPH</u> (0.32)	
ACT (0.60)	GADPH (0.76)	TUA (0.86)	ACT (0.79)	TUA (0.64)	OUB (4.56)	GADPH (0.63)	TUA (17.89)	<u>OUB</u> (0.34)	EF (34.86)	<u>OUB</u> (0.36)	
OUB (0.37)	TUA (0.80)	OUB (0.88)	OUB (0.82)	OUB (0.70)	TUA (4.95)	TUA (0.67)	GADPH (19.02)	TUA (0.37)	GADPH (38.6)	TUA (0.37)	

Table 3. Transcriptional stability rankings for transcript accumulation of candidate housekeeping genes for the whole panel of experimental conditions¹.

¹According to BestKeeper Pearson coefficient of correlation (BK_r), and the stability parameters for BestKeeper SD (BKS), comparative ΔCt method (CoD), GeNorm (GNo), NormFinder (NFi), RefFinder (RFi) and CV/F method. A lower value for stability parameters indicates higher transcript accumulation stability. F and NF were calculated for different sample groups. Underlayed are the genes contributing to the 2-genes normalization factor with maximum stability according to NFi independent software. In bold, genes selected as reference genes (see Fig. 2). In cursive, the discarded CHG *Hsp*.

²All tested genes with P<0.01 for Pearson correlation coefficient for the correspondent BestKeeper normalization factor, according to Pfaffl et al (2004).

Step 5. Discarding of transcriptionally unstable genes

As it is also evident from visual distribution of CV and F parameters (Fig. 1), *Hsp* resulted in the more transcriptionally unstable CHG for all methods used. Moreover, in Table 3, *Hsp* ranked as the less stable CHG for CV and F3, and its first position in F2 is an effect of an intolerably high intra-group *Hsp* instability.

Consequently, *Hsp* was discarded and removed for subsequent re-rankings.

Step 6. Re-rankings of remaining genes

For CV/F method, the exclusion of *Hsp* does not change stability values and the order of the rest of CHGs. Thus, omitting *Hsp*, the order *H2B*, *OUB*, *ACT* and *EF* is clear for CV/F method, and, taking into account the preference given to CV ranking over that from F, those genes may be followed by *GADPH* and, finally, *TUA*. *GADPH* showed the lowest inter-group stability according to both calculated F rankings, and had a relatively low score according to CV, being ranked as 5th for this parameter.

Step 7. Ranking with software methods as support and comparison with CV/F ranking

Selection of functionally distant CHGs, such as those from the group here tested (Table 2) and some previously considered as suitable RGs for plant *in vitro* growth and rooting studies ([9] and references therein), diminish the risk that most of them be transcriptionally affected in a similar way by the experimental conditions. In spite of this, it has been advised to evaluate each individual CHG stability by other means previously to the software usage [5]. CV has been used as a main parameter for normalization [34], frequently complementing normalization software [5,35,36].

Software methods employed in the present study (except BKS algorithm implemented in RFi) use statistics that relate transcript levels among all used CHGs, and consequently rely under the hard to support assumption that a majority of the evaluated CHGs do not show real biological concomitant, systematic variations on their transcript accumulation among sample groups, i.e., the measured variation is assumed to be due to technical error for most of the tested genes. In other words, all those programmes are sensitive to concomitant variations such as co-regulation [37], especially CoD [38] and gNo [39]. Furthermore, NFi software, although less affected by co-regulation, does not account for inter-group systematic errors associated to sample preparation [38] – in contrast to F-. On the other hand, BestKeeper requirements are too strict with the inter-sample variance [40]. Moreover, rankings offered by all those methods are frequently conflicting ([5–9], reviewed in [10]). The RFi programme uses the geometric mean of the four statistical approaches (BKS in the case of BestKeeper, and overall variation for NFi) to rank the CHGs. RFi weighting may help to overcome some pitfalls, but it is obviously also influenced by the systematic variations affecting the integrated algorithms.

In spite of these commented pitfalls, with the exception of gNo, which algorithm more strongly ranks according to similar gene transcript accumulation trends –likely a biasing effect-, the rest of used software agreed with CV/F method in including *H2B*, *OUB* and *ACT* into the three or four firstly ranked CHGs, although without concordance in the exact order positions (Table 3). : *H2B* ranked from 1st to 2nd position, *OUB* also (except for CoD method, where it ranked third), and *ACT* ranked between 2nd and 4th positions. The remaining CHGs, in a sort of parallelism to CV/F rank, usually oscillated between positions 3th and 5th (*EF*) or 4th and 6th (*GADPH* and *TUA*), except in the case of gNo software, which considered *EF* and *GADPH* as the optimal option.

The removal of the highly unstable *Hsp* affected the ranking of methods for which transcript accumulation or Cq' values of each CHG algorithmically interact with those from the others. Table 3 shows that *H2B* is still well ranked for all algorithms. Nonetheless, in those algorithms into RFi where the gene ranking is dependent on interactions among genes (CoD, gNo, and NFi), *ACT* and *OUB* pass to be into the last three positions. The removal of *Hsp*, which strongly affected average trends (see changes in BKr values), changed the rankings into CoD and gNo, two algorithms highly biased by co-regulation, a fact to consider especially in cases of plant multi-stress induction (such as the present one) even in cases of normally non-regulated genes [32,33]. The NFi integrated in RFi is also sensible to inter-sample systematic variations: in the present case, it is sensible to co-regulation. The use of the remaining six CHGs for subsequent estimations is into the limit of the advised number of CHGs to be used in ranking algorithm NFi [26].

After *Hsp* removal, all methods showed *H2B* as the most stable CHG (or 3rd most stable for gNo) (Table 3). In other words, *H2B* is the only selected CHG common for all methods. *H2B* was previously recommended to be used as RG due to its high stability in *in vitro* rooted *Eucalyptus globulus* microcuttings [9]. *ACT* usually ranks between 2nd and 4th positions. This ranking reinforces its 2nd place for CV/F. *OUB*, 3rd for CV/F ranking, ranks in the last two positions for those algorithms more susceptible to co-regulation effects. Since this is due to its transcript accumulation trend opposite to the rest of the five remaining CHGs, in a way its behaviour compensates the opposite trend of both *H2B* and *ACT*, stabilizing the correspondent NF, as happens when the selection of two genes is configured for independent NFi. It should be noticed that when the selection of only two genes is switched for this software, *OUB* –together with *H2B*– is suggested to compose the NF.

Selection of RGs composing NFs is performed after establishment of a stability ranking of CHGs. In spite of the commented preference for CV/F method in order to establish a stability ranking for CHGs, it is worth to compare if NFs obtained from each method can be similarly valid. There is a wide consensus in accepting different criteria for establishment of a proper NF after CHG ranking, but they have not been applied together in a systematic way. According to the 2nd round classifications (after discarding *Hsp*) of the prompted CHGs, such criteria were adapted for an optimal selection of RGs into each ranking methodology, as follows in next Step.

Step 8. Selection of RG sets

Calculation of pairwise variations between possible NFs

A proper NF should be i) as stable among samples as possible as long as ii) its variations mainly reflect the real technical errors in order these errors be compensated. The first criterion, NF stability, is usually procured through selection of a number of the most stable CHGs –i.e., RGs, which compose the NF. According to this, six NFs were calculated for each gene ranking method, being each NF composed from the most stable n CHGs, being n=1 to 6, and ordered according to n (Fig. 2).

The replicate pairwise variation (V) [23] indicates the variability between NF pairs. If V is low, both NFs are considered equivalent for normalization. If the compared NFs are composed each by the n and the $n+1$ most stable CHGs, and V is low, both NFs may be valid for normalization. If the pairwise comparisons are made from $n=1$ to the total number of CHGs, the first low V found is considered to correspond to the optimal NF pair, and usually the NF composed with the minor number of genes is selected. A V threshold value of 0.15 is accepted as low enough, but this consensus should not be strict [11], especially when considering complex designs of various factors as in case presented in this article. Also, a trend of changing V values when adding new genes for the calculation of the NF is recognized to be equally informative [31,41,42], and in fact values higher than 0.15 have been accepted [32,43,44]. Concretely, the best option for a decreasing V trend are the NFs corresponding to the V lowest value. Moreover, if after a decreasing trend, even to less than 0.15, V starts to grow again by the addition to the NF of a worse ranked CHG, more instability may be being introduced, and consequently more error. Thus, the inclusion of such a CHG as RG is not advisable.

The methodology for V calculation has been used so far only for comparisons of pairs of NFs composed by n and $n+1$ CHGs as ranked according to the CHG stability obtained with gNo , since this algorithm is implemented in that software [23]. In the present work, such an algorithm was adapted to determine V between any pair of possible NFs, hereby obtaining left graphics of Fig. 2.

Fig. 2. Pairwise variation (left) and descriptive parameters (right) for the normalization factors. Normalization factors are composed from the indicated candidate housekeeping genes, determined according to the indicated methods. Descriptive parameters are CV and F for bifactorial – F_2 – and trifactorial – F_3 – cases. Discontinuous lines indicate the maximum value of the parameter into the candidate housekeeping genes contributing to the indicated normalization factors. Arrows indicate the selected NFs after the determination of representative candidate NFs (left) and the assessment of congruent stability (right) (see text, Step 8).

Determination of representative candidate NFs

Normalization against more than one gene is *a priori* recommended [45]. At least three RGs to calculate NF are recommendable [46,47]. Three non-physiologically related genes are more representative, as long as NF is stable.

Assessment of congruent stability and selection of NFs for each gene ranking method

Selected NF among those constructed for each ranking method should be as stable as possible as long as the previous criteria are kept. Maximum or high NF stability underlies the method proposed in [48] for choosing NFs, but this may overlook CHG representation. Thus, the overall CHG representation in the candidate NFs and the congruency of these NFs –according to the previous criteria of stability and compensation of errors– were tested by assessing the stability of the candidate NFs regarding discarded NFs and individual CHGs, a matter that tends to be ignored. To inspect the suitability of the candidate NFs, the next rules were introduced:

1. i) Ideally, important variability parameters (CV and F) for NFs should be lower than those of the CHG –composing the NF– having the maximum values for such parameters (see Fig. 2 right). Otherwise, this worse ranked CHG may be unnecessarily contributing to additional NF instability.
2. ii) In this regard, as in general, when selecting NFs the ranking provided by overall variability (CV) should ideally have priority to that given by inter-group variability (F).
- iii) F would be more representative of the whole experimental panel if it spans a higher number of factors: in the present case, F_3 should have more priority than F_2 .
1. iv) Economical criteria may be taken into account additionally when two or more NFs may be similarly valid according to all above criteria, consequently selecting the RG set with the minimum number of CHGs.

In Table 3 (genes in bold) and Fig. 2, the proposed and selected RGs (genes which contribute to the NFs) for different algorithms are represented after discarding *Hsp* as highly unstable gene and applying the above described criteria. In all cases those rules brought NFs composed from three RGs. Results show three possible NFs, depending on the ranking method: *H2B*, *OUB* and *ACT* (for CV/F and BKS methods); *H2B*, *ACT* and *EF* (for independent NF algorithms), and *H2B*, *EF* and *GADPH* (for the rest of the methods). There is a consensus for the selection of these three CHGs amongst RFi and their implemented methods, except for BKS. This apparent congruence can be an effect of the bias by co-regulation. In the case of independent NF algorithm, which takes into account biological groups, the *GADPH*^{2nd} position resulting from those algorithms is substituted by *ACT*, but this can be still an effect of the influence of co-regulation. BKS and CV/F methods, not affected by the commented pitfalls, change the *ACT* of the selected NF of independent NF algorithm by *OUB*. The different trend of this gene with respect to the majority of CHGs increases the stability of the resulting NF. Only *H2B*, the most stable CHG for all methods except for gNo (where ranked 3rd), constituted a consensus among the selected NFs.

In order to test the possible validity of any of the three selected NFs, V values between them were calculated. V values were always much higher than 0.15 (0.27 was the lowest one –data not shown), thus there being important significant differences between the correspondent NFs. The commented deficiencies of algorithms suggest the CV/F method as a reasonable solution, and its use at least as a first approach. Only BKS, unaffected by co-regulation since uses a similar way to CV/F , had analogous results to CV classification. If being not too strict with high standard deviation for BestKeeper, BKS classification could be a good combination with F to classify stable genes. In fact, according to [49], the BestKeeper approach may be useful to narrow down a search if no specific genes are known to be plausible candidates.

Step 9. Error compensation versus stability: inspection of the quality of the selected NFs

The quality of NFs may be partially assessed by inspecting: i) in what measure they compensate the technical errors of transcript levels measurement, and ii) if they can also be valid for inter-assay (biological replicates in the present case) comparisons.

The total variation of the measurement of the transcriptional levels comprises the real biological variation plus technical errors. These errors are systematic when they are due to factors of experimental design (e.g. in the case of samples from a given experimental condition which are all more prone to RNA degradation), and may increase from RNA isolation with associated effects like carried over effects (such as analytical processing of experimental groups at different times or in different conditions). Carried over effects are minimized when using a sample maximization design [25] –as in the present case–, e.g., run-to-run variation. On the other hand, non-systematic, i.e., non-experiment-associated technical errors should weigh proportionally less among experimental groups (formed by, for example, different treatments or time points) than among biological, equivalent replicates. In the present case, since assay is a factor fixed by rooting potential, the assay-associated systematic variation (and consequently the included systematic error) should weigh proportionally more in the overall inter-assay variation than in assays for which the rooting potential would have been similar. Thus, the inter-assay systematic errors should be compensated as possible by the selected NF avoiding the overcompensation, i.e., the compensation of the real biological inter-assay variation. This convenience is kept for the NFs selected in the analyses (Fig. 2) for all the indicated CHG ranking methods, since *inter-assay variation* of these NFs is moderate or low with respect to those non-selected and the individual CHGs (Fig. 3). What is more, except for the independent NF_i, the inter-assay variation of the selected NFs is similar to that of the CHG, among those composing the NF, with maximal inter-assay variation. Thus, selection criterion i for Step 8 (Assessment of congruent stability) is also kept for inter-assay case. These considerations can lead to the conclusion that the selected NFs, for both bi- and trifactorial cases, are also valid for inter-assay comparisons.

Fig. 3. Error compensation and stability of normalization factors. Mean distances between replicates (MD), and compensations (standard deviations of mean distances) (SDC) (graphics on the left), and inter-assay and inter-replicate variations (graphics on the right) for the normalization factors composed from the indicated candidate housekeeping genes, obtained according to the indicated methods. MD and SDC are shown in Cq's. Discontinuous lines indicate the maximum value of the indicated parameter into the candidate housekeeping genes contributing to the indicated normalization factors.

Obviously, NFs should also compensate non-systematic analytical errors. The *inter-replicate variation* (Fig. 3) virtually only accounts for the technical error accumulated during RT-replicates preparation and later associated effects, i.e., a non-systematic, non-factor associated error. Inter-replicate variation tends to decrease asymptotically with the number of RGs to a theoretically precise average value, as can be confirmed in Fig. 3. All genes are in reality absolutely stable for technical replicates, and consequently inter-replicate error would be optimally compensated by a NF composed from a large number of RGs. Nevertheless, the contribution to the NFs of such a large number of RGs would compromise other requirements such as previously indicated criteria on this regard, so as a good compensation for systematic errors. Thus, a moderate NF inter-replicate variation value, of course lower than the NF-composing CHG with maximum value for this parameter, would be an acceptable deal. This is the case of selected NFs. This means a partial compensation of inter-replicate error, without compromising systematic error compensations and a reasonable stabilization of real transcript levels represented by the NF.

The *inter-replicate mean distance* of the NFs composed from the all seven CHGs ($Ct=0.0042$) specifically accounts of average error due to RT-replicate preparation. In exchange, the *individual CHG inter-replicate mean-distances* include all non-systematic technical errors, carried over or not. Selected NFs are again well ranked for this descriptor, having an intermediate-high value (0.0072 corrected cycles), and then overcompensating RT average error (0.0042 corrected cycles) and thereby accounting of later errors too (see replicate compensation ranking, given by standard deviations of the individual CHG inter-replicate mean distances) (Fig. 3).

Step 10. Determination of the optimal normalization factor for the complete bi- and trifactorial panels

For CV/F method, NF3, composed from *H2B*, *OUB* and *ACT*, is a good compromise to reduce V between two consecutive CHG groups (Fig. 2), having a major representation than only two CHGs and keeping the stability conditions stated above. *OUB* has a slight opposite trend to both *H2B* and *ACT* transcript accumulations, then compensating bifactorial (time x treatment) and trifactorial (time x treatment x assay) inter-group variations for NF3. NF3 offers low inter-group (bi- or trifactorial) variation (Fig. 2 right) and intermediate inter-replicate variation (Fig. 3 right), thus being reasonably stable for group-associated systematic variations without compromising the compensation of non-systematic technical errors. The inclusion of additional genes introduces more instability (Fig. 2 right). Furthermore, the addition of one or two worse ranked genes is not advisable, since they rarely are selected when testing experimental conditions separately (see next Step). As a general conclusion, this three-gene NF represents the optimal combination of CHGs, amongst those tested, for the general panel of assayed experimental conditions. Slight instability among groups may account for systematic variations.

Step 11. Normalization factors for more specific experimental conditions

Additionally, the validity of NF3 for specific treatments inside of a time and along time inside of a treatment was checked. In these cases, the experimental condition groups consisted of the different levels of time or treatment (thus giving a one factor design) or in combination with the assays (thus giving a two factor design). Single CHGs were subjected to graphic inspection and ranked for each time or treatment according to the same algorithms used for the complete bi- or trifactorial panels (Table 4). After inspection of the graphics (Fig. 1), *Hsp*, due to its high instability, was removed from further consideration for IBA, IBA+SHAM, 4 h and 24 h. The genes ranked in last position for F were not removed as always ranked in the first four positions according to CV (Table 4). After determination of a CHG ranking according to CV/F method – the method discussed above to be considered in general more suitable - the before described RG selection criteria were followed (Fig. 4).

Table 4. Stability rankings for transcript accumulation of candidate housekeeping genes for the indicated experimental conditions¹

All replicates (Treatment)						Time		Time x Assay		
Level	BKS	CoD	GNo	NFi	RFi	CV	F	NFi	F	NFi
Control Treatment	H2B	ACT	ACT/OUB	H2B	ACT (1.57)	H2B (0.18)	Hsp (3.85)	OUB (0,22)	H2B (6,6)	<u>H2B</u> (0,29)
	OUB	H2B	ACT	H2B (1.57)	OUB (0.26)	H2B (4.78)	<u>H2B</u> (0,26)	OUB (13,55)	OUB (0,29)	
	ACT	OUB	H2B	OUB	OUB (2.06)	ACT (0.36)	OUB (5.86)	<u>ACT</u> (0,32)	EF (24,01)	<u>ACT</u> (0,3)
	EF	EF	EF	EF	EF (4.00)	EF (0.41)	GADPH (7.27)	GADPH (0,47)	GADPH (27,98)	GADPH (0,5)
	GADPH	GADPH	GADPH	GADPH	GADPH (5.00)	GADPH (0.42)	EF (8.93)	EF (0,5)	TUA (40,48)	EF (0,52)
	TUA	TUA	TUA	TUA	TUA (6.00)	<u>TUA</u> (0.54)	<u>TUA</u> (9.63)	TUA (0,64)	ACT (42,01)	TUA (0,62)
	Hsp	Hsp	Hsp	Hsp	Hsp (7.00)	Hsp (1.32)	<u>ACT</u> (22.53)	Hsp (1,01)	Hsp (442,64)	Hsp (1,08)
IBA	OUB	OUB	EF GADPH	OUB	OUB (1.32)	H2B (0.32)	OUB (5.51)	<u>OUB</u> (0,39)	OUB (13,16)	<u>OUB</u> (0,43)
	H2B	<u>GADPH</u>	<u>GADPH</u>	<u>GADPH</u> (2.21)	OUB (0.34)	H2B (7.13)	H2B (0,44)	H2B (14,69)	H2B (0,44)	
	ACT	H2B	OUB	H2B	H2B (2.91)	ACT (0.49)	ACT (12.54)	TUA (0,47)	TUA (29,5)	TUA (0,46)
	EF	TUA	H2B	TUA	EF (3.16)	TUA (0.62)	Hsp (12.54)	EF (0,48)	ACT (34,59)	EF (0,51)
	TUA	EF	TUA	EF	TUA (4.47)	<u>EF</u> (0.64)	<u>EF</u> (16.73)	ACT (0,49)	EF (53,02)	<u>GADPH</u> (0,52)
	GADPH	ACT	ACT	ACT	ACT (5.05)	<u>GADPH</u> (0.67)	<u>GADPH</u> (20.47)	<u>GADPH</u> (0,49)	GADPH (53,55)	ACT (0,53)
	Hsp (2.27)	TUA (24.77)	Hsp (2664,02)							
IBA + SHAM	H2B	H2B	H2B TUA	H2B	H2B (1.00)	H2B (0.25)	Hsp (4.04)	<u>ACT</u> (0,31)	H2B (8,68)	<u>H2B</u> (0,33)
	OUB	ACT	ACT	TUA (2.45)	OUB (0.32)	ACT (7.51)	<u>H2B</u> (0,34)	ACT (13,65)	<u>ACT</u> (0,34)	
	ACT	TUA	ACT	TUA	ACT (2.45)	ACT (0.35)	TUA (8.69)	TUA (0,5)	OUB (14,68)	TUA (0,51)
	TUA	GADPH	GADPH	GADPH	GADPH (4.23)	GADPH (0.42)	H2B (12.06)	GADPH (0,53)	GADPH (14,69)	GADPH (0,51)
	GADPH	EF	EF	EF	OUB (4.56)	<u>EF</u> (0.42)	EF (13.78)	EF (0,63)	EF (17,63)	EF (0,62)
	EF	OUB	OUB	OUB	EF (5.23)	TUA (0.53)	OUB (15.85)	OUB (0,66)	TUA (19,87)	OUB (0,67)
	Hsp (1.97)	GADPH (22.17)	Hsp (1107,17)							
All replicates (Time)						Treatment		Treatment x Assay		
Level	BKS	CoD	GNo	NFi	RFi	CV	F	NFi	F	NFi
4 h	<u>GADPH</u>	<u>GADPH</u>	EF GADPH	<u>GADPH</u>	<u>GADPH</u> (1.00)	H2B (0.16)	Hsp (3.19)	<u>H2B</u> (0,2)	TUA (4,11)	H2B (0,18)
	OUB	EF	TUA	EF (2.06)	TUA (0.18)	TUA (3.92)	<u>GADPH</u> (0,2)	OUB (15,79)	<u>GADPH</u> (0,22)	
	EF	TUA	TUA	EF	TUA (2.91)	OUB (0.22)	EF (5.28)	TUA (0,2)	GADPH (20,15)	EF (0,3)
	TUA	H2B	H2B	H2B	H2B (4.43)	GADPH (0.25)	H2B (5.51)	EF (0,28)	EF (21,49)	<u>TUA</u> (0,3)
	ACT	ACT	ACT	ACT	OUB (4.56)	<u>EF</u> (0.37)	<u>ACT</u> (8.01)	ACT (0,33)	H2B (24,26)	ACT (0,35)
	H2B	OUB	OUB	OUB	ACT (5.00)	<u>ACT</u> (0.39)	OUB (12.15)	OUB (0,45)	ACT (38,25)	OUB (0,45)

	Hsp (1.53)	GADPH (13.54)	Hsp (1408,8)							
1 d	H2B	EF	EF OUB	EF	EF (1.41)	ACT (0.15)	ACT (1,22)	<u>EE</u> (0,15)	ACT (1,58)	<u>EE</u> (0,22)
	ACT	OUB	OUB	OUB (1.86)	H2B (0.2)	H2B (2,85)	<u>OUB</u> (0,17)	H2B (1,98)	OUB (0,23)	
	OUB	ACT	ACT	ACT	ACT (2.71)	OUB (0.27)	OUB (3,15)	H2B (0,25)	GADPH (8,59)	<u>H2B</u> (0,27)
	EF	TUA	TUA	TUA	H2B (3.34)	EF (0.33)	TUA(7,6)	ACT (0,3)	EF (10,25)	ACT (0,32)
	GADPH	H2B	H2B	H2B	TUA (4.43)	GADPH (0.39)	EF(10,14)	TUA (0,31)	OUB (11,74)	TUA (0,37)
	TUA	GADPH	GADPH	GADPH	GADPH (5.73)	TUA (0.52)	GADPH(34,04)	GADPH (0,34)	TUA (20,03)	GADPH (0,39)
	Hsp (1.19)	Hsp(145,76)	Hsp (41,6)							
2d	H2B	H2B	ACT/OUB	H2B	H2B (1.32)	H2B (0.37)	EF (6.19)	OUB (0,19)	OUB (18,46)	<u>H2B</u> (0,23)
	OUB	OUB	OUB	OUB (1.68)	OUB (0.49)	H2B (7.79)	<u>H2B</u> (0,22)	H2B (28,88)	<u>OUB</u> (0,26)	
	ACT	ACT	H2B	ACT	ACT (2.28)	ACT (0.52)	OUB (9.76)	<u>EE</u> (0,23)	TUA (33,74)	ACT (0,29)
	GADPH	GADPH	GADPH	GADPH	GADPH (4)	EF (0.53)	ACT (13.31)	ACT (0,26)	ACT (37,23)	EF (0,35)
	TUA	EF	EF	EF	EF (5.23)	TUA (0.59)	GADPH (17.13)	GADPH (0,43)	EF (40,52)	GADPH (0,37)
	EF	TUA	TUA	TUA	TUA (5.73)	GADPH (0.61)	TUA (19.76)	TUA (0,49)	GADPH (126,8)	TUA (0,46)
	Hsp	Hsp	Hsp	Hsp	Hsp (7.00)	Hsp (0.7)	Hsp (29.99)	Hsp (1,09)	Hsp (291,95)	Hsp (1,11)
4d	H2B	H2B	ACT/H2B	H2B	H2B (1.00)	EF (0.16)	GADPH (1.21)	<u>ACT</u> (0,12)	OUB (4,12)	<u>H2B</u> (0,14)
	EF	ACT	ACT	ACT (2.00)	H2B (0.19)	H2B (2.16)	H2B (0,14)	H2B (5,36)	<u>EE</u> (0,21)	
	OUB	EF	EF	EF	EF (2.71)	OUB (0.24)	ACT (2.29)	<u>EE</u> (0,15)	EF (5,66)	ACT (0,25)
	ACT	GADPH	GADPH	GADPH	GADPH (4.23)	ACT (0.24)	TUA (2.68)	GADPH (0,21)	ACT (6,45)	GADPH (0,35)
	GADPH	OUB	OUB	OUB	OUB (4.4)	GADPH (0.32)	Hsp (5.65)	OUB (0,32)	GADPH (8,63)	OUB (0,4)
	TUA	TUA	TUA	TUA	TUA (6.00)	TUA (0.35)	EF (6.92)	TUA (0,34)	TUA (9,82)	TUA (0,44)
	Hsp	Hsp	Hsp	Hsp	Hsp (7.00)	Hsp (0.48)	OUB (6.99)	Hsp (0,37)	Hsp (33,83)	Hsp (0,5)

¹According to stability parameters for BestKeeper SD (BKS), comparative ΔC_t method (CoD), GeNorm (GNo), NormFinder (NFi), RefFinder (RFi) and CV/F method. A lower value indicates higher transcript accumulation stability. F and NFi were calculated for different sample groups. Underlayed are the genes contributing to the 2-genes normalization factor with maximum stability according to NFi independent software. In bold, genes selected as reference genes (see Fig. 4).

Fig. 4. Pairwise variation (left) and descriptive parameters (right) for the normalization factors. Normalization factors are composed from the indicated candidate housekeeping genes, determined for CV ranking. Descriptive parameters are CV and F for bifactorial –F2– and trifactorial –F3– cases. Discontinuous lines indicate the maximum value of the parameter into the CHGs contributing to the indicated normalization factors. Arrows indicates the selected normalization factors after the determination of representative candidate NFs (left) and the assessment of congruent stability (right) (see text, Step 8).

To see a detailed selection of RGs for groups see Additional file 3. Summarizing (Fig. 4), except for 96 h, where the combination *H2B*, *ACT* and *EF* is the optimal, *H2B* and *OUB* (NF2) belong to all selected NFs for the cases of specific experimental conditions (one given treatment with time point as a factor or a given time point with treatment as a factor), whatever assay is included or not as a second factor. In some cases, the addition of *EF* (Control treatment) or *TUA* (4 h) to NF2 forms the optimal set for normalization. In most of the cases (IBA, IBA+SHAM, 24 h and 48 h), the addition of *ACT* to NF2 (and then constituting

NF3) improves normalization according to the above pointed criteria. As previously indicated, for the bi- and trifactorial whole panels (treatment x time and treatment x time x assay, respectively), NF3 (*H2B*, *OUB* and *ACT*) is proposed as the optimal NF.

For more specific cases (a given treatment at a given time point, or even biological or technical replicates in the same experimental conditions of the same assay), checking of the variability of CHGs among biological or technical replicates would be required, following a similar methodology to that exposed. With the present transcript accumulation dataset, the total number of replicates inside of a given experimental condition (6 = 3 assays x 2 technical replicates) may be too low to perform a consistent test. By this reason, if a NF should be estimated from the present datasets, the results obtained for the level-associated bifactorial cases (treatment level x assay or time level x assay) can be more valid. In spite of this, the decomposition of bi- and trifactorial whole panels for separate normalization and statistical comparison into the more simple level-associated mono- or bifactorial panels, respectively, is not advisable if a considerable loss of statistical power may be involved. Anyway, in such cases, CV and F values for NF3 are always low, which support the general use of NF3 as the most practical NF for most situations. Nevertheless, the obtained results coincide with those of other reports where differences in RG stability depend on characteristics of biological material or environmental conditions ([50] and references therein). The complexity level of an experimental panel diminishes normalization accuracy, as can be expected *a priori*.

Step 12. Evaluation of NF3 by comparison with NF2

It was argued above (Steps 10 and 11) the proposition of NF3 as the best possible option for normalization of the whole panel, and also for several more concrete experimental conditions here assayed. In spite of this, NF3 may be sub-optimal for the rest of the cases and could be used as a less accurate solution. Nevertheless, according to the results on overall and intergroup variability for restricted panels (Fig. 4), NF2 (which is like NF3 without including the less stable gene, *ACT*), although less representative, could constitute also a good option, especially for more concrete experimental situations, such as monofactorial case IBA+SHAM (Step 11). NF3 is evaluated below by comparison with NF2 when normalizing a hypothetical highly stable gene and a real TG with low stability.

NF3/NF2 comparison by normalizing a stable gene

The theoretical results on an ideally completely stable TG (without any variation) regarding NF3 normalization were compared with those that would be obtained by normalising with NF2 on the same TG. The results are equivalent to normalize NF3 against NF2, and provide accurate differences between both NFs. Fig. 5 shows the deviations for such normalization after equalizing both NFs for time 0 (which constitutes the same situation for all treatments). Differences in corrected cycles (C_t') were only between 0.10 and 0.55, which are in the ranges of the intra-group standard deviations even of the more stable genes (Fig. 1), thus indicating that normalization of a real stable gene (which would show intra-group standard deviations) with NF2 would easily trough non-significant differences regarding NF3. In other words, in the practice, NF2 or NF3 could be used for less stable genes normalization without obtaining too much differences. In fact, V for both NFs is about the threshold value of 0.15 accepted as low enough.

Fig. 5. Estimated marginal means of an ideally stable gene according to NF3 when NF2 is applied. NF3: normalization factor composed from *H2B*, *OUB* and *ACT*. NF2: normalization factor composed from *H2B* and *OUB*.

In spite of it, since all differences gave higher transcript levels than time 0, NF2 normalization may lead to overestimation of relative transcript levels obtained after NF3 normalization (Table 5). According to the above justifications for adding a third RG, it is supposed that the introduction of *ACT* in NF2 to compose NF3 corrects not only punctual non-systematic technical errors but also some systematic, group associated errors. Among these, can be detected a decrease of "NF2- transcript accumulation" 4 days after the experiment initiation or in samples from not-SHAM treated explants (with stress palliated by AOX) 4 h after the experiment initiation.

NF3/NF2 comparison by normalizing an unstable gene

According to the previous conclusions, to evaluate the applicability of NF2 or NF3, significant differences in transcript levels of a real unstable gene (*Hsp*) were determined according to both NFs. As shown in Fig. 1 and reflected in CV and F2 values (Table 3), *Hsp* displayed high bifactorial intragroup variation but low bifactorial intergroup variation. Thus, despite showing instability, *Hsp* resulted not highly modulated by different conditions in the system. Assumptions to perform ANOVAs were tested using Fligner-Killeen test so as graphics of residuals and observed and predicted values. The differences found when normalizing against NF2 and NF3 obtained after Bonferroni transformation are in table and graphics in Additional file 4. The use of NF3 renders lower calculated transcript levels than the use of NF2, as explained above, and thereby generated slight small number of significant differences among groups, i.e., does not throw some differences generated by F2. Thus, in both bifactorial (time levels inside of treatment) –11 for NF2 and 10 for NF3– and trifactorial –6 for NF2 and 5 for NF3– cases. Bifactorial (treatment levels inside of time) case showed a total of 5 significant differences for both NFs. Thus, globally, NF3 presented slightly more conservative results than NF2. Summarizing, only inside of IBA+SHAM treatment, NF2 generated significant differences between 1d and 2d that NF3 did not do ($P = 0.062$), and the same happens in BA for assays I and III.

Combined evaluation.

After applying both comparisons, it can be concluded that the more TG-transcript level variations appear (higher intra-group standard deviations and more evident inter-group variation), the less discrepancies about significant differences in gene transcript accumulation are generated by the application of different NFs. Thus, NF2, here considered as sub-optimal, can generate the same significant differences than NF3 when the TG is unstable enough, including when its expression depends on experimental conditions.

Conclusions

Selection of appropriate RGs is crucial for the validity of transcription studies. The CV/F method (Steps 3 to 7) demonstrates to be a friendly and useful gene stability-ranking procedure lacking the inaccuracy generated by the influence of biological systematic errors such as gene co-regulation. This inaccuracy may affect in different ways the most popular algorithms applied for such a purpose, leading to possible misinterpretations. Nevertheless, the additional use of popular software is advisable as complementary support of core methods. This is generally accepted state-of-the-art knowledge. However, combining the showed criteria for selection of a proper number of stable RGs provides more accuracy to normalization. The nucleus of the integrated procedure is the selection of the RG sets (Step 8): i) calculation of pairwise variations between any pair of possible proper NFs by implementing a method until now used only when gNo software have been run, ii) determination of representative candidate NFs, and iii) assessment of a congruent stability of candidate NF, a fact usually underconsidered.

The quality of the resulting NF may be additionally validated (Step 9) by inspecting its suitability for comparisons among non-considered factors or biological or technical replicates, and also checking in what extension such NF compensate errors of transcript level measurement. For those purposes, special statistical estimators were formulated for the case in study.

The NF composed from *H2B*, *OUB* and *ACT* provides a valid normalization for TGs in studies on olive microshoot adventitious rooting when comparing treatments, time points and assays (Steps 10 and 11).

Finally, double evaluation (Step 12) against both a theoretically highly stable gene and a real gene with relatively high instability provided information about the suitability of possible alternative NFs. The validity of the use of a sub-optimal NF depends on the variability of the studied TG. The more stable a TG is, the more transcript accumulation differences the application of different NFs will bring.

Abbreviations

AOX

Alternative oxidase

BK_r

BestKeeper Pearson coefficient of correlation

BKS

BestKeeper regarding standard deviation

CHG

Candidate housekeeping gene

CoD

Comparative ΔCq

Cq

Quantification cycle

Cq'

Corrected quantification cycle

CV

Coefficient of total variation (for data on relative transcript accumulation)

DMSO

Dimethyl sulfoxide

F

Inter-group variation (for data on relative transcript accumulation)

gNo:

geNorm

IBA

Indole-3-butyric acid

NF

Normalization factor

NFi

NormFinder

qPCR

Quantitative polymerase chain reaction

RFi

RefFinder

RG

Reference gene

RT

Reverse transcription

SHAM

Salicylhydroxamic acid

TG

Target gene

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

All data generated or analysed during this study are included in this published article [and its supplementary information files].

Competing interests

The authors declare that they have no competing interests

Funding

Research was financially supported by FEDER Funds through the Operational Programme for Competitiveness Factors (COMPETE) and National Funds through Foundation for Science and Technology (FCT) under the project PTDC/AGR-AAM/103377/2008, which also financed the Postdoc scholarship to Carlos Noceda. The State University of Milagro (Ecuador) financed in 2016 the project "Involvement of oxidative stress-regulator genes in olive adventitious rooting", that provided time and resources for processing and writing.

Authors' contributions

BA and AP developed the main idea and designed the biological experiment. CN carried out the experiment and developed the analysis protocol and data processing, and was a major contributor in writing the manuscript. All authors made major arrangements in the original draft and read and approved the final manuscript.

Acknowledgments

The authors gratefully acknowledge Nuno Ribeiro and Paulo Infante, from the University of Évora, for their considerable help with statistics analyses, and also to Virginia Sobral from the same University for technical assistance provided during the experiments. José Flores, from the Polytechnic Higher School of the Litoral (ESPOL) is acknowledged for his help in editing the manuscript. Prometeo Project from the Secretariat for Higher Education, Science, Technology and Innovation (SENESCYT) of the Republic of Ecuador is also acknowledged for allowing Carlos Noceda the writing of this article, so as the international

scientific network BIOALI-CYTED, which contributed to present insights and to establish FunCROP contacts. Special thanks are given to Nadine Fladd, from the Writing and Communication Centre at University of Waterloo, who helped us with the Grammar of the draft.

References

1. Bustin SA. Quantitative real-time RT-PCR - a perspective. *J Mol Endocrinol* [Internet]. 2005;34(3):597–601. Available from: <http://jme.endocrinology-journals.org/cgi/doi/10.1677/jme.1.01755>
2. Wong ML, Medrano JF. Real-time PCR for mRNA quantification. *Biotechniques* [Internet]. 2005;39(1):75–85. Available from: <http://www.biotechniques.com/article/05396BM03>
3. Huggett J, Dheda K, Bustin S, Zumla A. Real-time RT-PCR normalisation; strategies and considerations. *Genes Immun* [Internet]. 2005;6(4):279–84. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/15815687>
4. Dundas J, Ling M. Reference genes for measuring mRNA expression. Vol. 131, *Theory in Biosciences*. 2012. p. 215–23.
5. Kortner TM, Valen EC, Kortner H, Marjara IS, Krogdahl shild, Bakke AM. Candidate reference genes for quantitative real-time PCR (qPCR) assays during development of a diet-related enteropathy in Atlantic salmon (*Salmo salar* L.) and the potential pitfalls of uncritical use of normalization software tools. *Aquaculture*. 2011;318(3–4):355–63.
6. Brunner AM, Yakovlev I a, Strauss SH. Validating internal controls for quantitative plant gene expression studies. *BMC Plant Biol* [Internet]. 2004;4:14. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=515301&tool=pmcentrez&rendertype=abstract>
7. Langnaese K, John R, Schweizer H, Ebmeyer U, Keilhoff G. Selection of reference genes for quantitative real-time PCR in a rat asphyxial cardiac arrest model. *BMC Mol Biol*. 2008;9(1):53.
8. Spinsanti G, Panti C, Bucalossi D, Marsili L, Casini S, Frati F, et al. Selection of reliable reference genes for qRT-PCR studies on cetacean fibroblast cultures exposed to OCs, PBDEs, and 17 β -estradiol. *Aquat Toxicol*. 2008;87(3):178–86.
9. de Almeida MR, Ruedell CM, Ricachenevsky FK, Sperotto R a, Pasquali G, Fett-Neto AG. Reference gene selection for quantitative reverse transcription-polymerase chain reaction normalization during in vitro adventitious rooting in *Eucalyptus globulus* Labill. *BMC Mol Biol*. 2010;11:73.
10. Kozera B, Rapacz M. Reference genes in real-time PCR. *J Appl Genet*. 2013;54:391–406; 2013;54:391–406.
11. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, et al. *GeNorm software manual*. 2007.
12. Salama M, Zahran M, Hassan M. Comparing the rooting ability of some olive cultivars propagated by leafy cuttings under mist. *Ann Agric Sci, Fac Agric Ain Shams Univ*. 1987;32:577–90.
13. Qrunfleh MM, Rushdi Y, Musmar T, Rushdi L. Root formation in cuttings of “Nabali” olives treated with uniconazole and indole butyric acid. *Dirasat*. 1994;21(6):71–9.
14. Fabbri A. The effect of various anatomical characteristics on the rooting of cuttings in olives, cv. Frangivento. *Riv Della Ortoflorofruttic Ital*. 1980;64(4):325–35.
15. Bakr E, Selim H, Nour G, Gabr M. Developmental anatomy of adventitious roots on stem cuttings of “Wetaken” olive cultivar. *Egypt J Hort*. 1977;4(1):91–7.
16. Santos-Macedo E, Vieira C, Carrizo D, Porfirio S, Hegewald H, Arnholdt-Schmitt B, et al. Adventitious root formation in olive (*Olea europaea* L.) microshoots: Anatomical evaluation and associated biochemical changes in peroxidase and polyphenol oxidase activities. *J Hortic Sci Biotechnol*. 2013;88(1):53–9.
17. Porfirio S, Calado ML, Noceda C, Cabrita MJ, da Silva MG, Azadi P, et al. Tracking biochemical changes during adventitious root formation in olive (*Olea europaea* L.). *Sci Hortic (Amsterdam)*. Elsevier; 2016;204:41–53.
18. Santos Macedo E, Sircar D, Cardoso HG, Peixe A, Arnholdt-Schmitt B. Involvement of alternative oxidase (AOX) in adventitious rooting of *Olea europaea* L. microshoots is linked to adaptive phenylpropanoid and lignin metabolism. *Plant Cell Rep*. 2012;31(9):1581–90.
19. Arnholdt-Schmitt B, Costa J, de Melo D. Alternative respiration: less energy-efficient but highly effective. *Trends Plant Sci*. 2006;6(11):281–7.
20. Arnholdt-Schmitt B, Santos Macedo E, Peixe A, Cardoso H, Cordeiro A. AOX—A potential functional marker for efficient rooting in olive shoot cuttings. In: *Proceedings of Second International Seminar Olivebioteq, Marsala Mazara del Vallo, Italy*. 2006. p. 249–54.
21. Santos-Macedo E, Cardoso H, Hernández A, Peixe A, Polidoros A, Ferreira A, et al. Physiologic responses and gene diversity indicate olive alternative oxidase as a potential source for markers involved in efficient adventitious root induction. 2009.
22. Peixe A, Serras M, Campos C, Zavattieri M, Dias M. Adventitious rooting in olive (*Olea europaea* L.) cuttings: a histological evaluation. *Rev Ciências Agrárias (Sociedade Ciências Agrárias Port. Revista de Ciências Agrárias (Sociedade de Ciências Agrárias de Portugal)*; 2007;30(1):476–82.
23. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* [Internet]. 2002;3(7):RESEARCH0034. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=126239&tool=pmcentrez&rendertype=abstract>
24. Sambrook J, Fritsch E, Maniatis T. *Molecular cloning*. 1989.
25. Hellemans J, Mortier G, De Paepe A, Speleman F, Vandesompele J. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol* [Internet]. 2007;8(2):R19. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1852402&tool=pmcentrez&rendertype=abstract>
26. Andersen C, Jensen J, Ørntoft T. 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(15):5245–50.
27. Walker N. A technique whose time has come. 2002.

28. Ruijter JM, Ramakers C, Hoogaars WMH, Karlen Y, Bakker O, van den hoff MJB, et al. Amplification efficiency: Linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Res.* 2009;37(6).
29. Pfaffl MW, Tichopad A, Prgomet C, Neuvians TP. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper–Excel-based tool using pair-wise correlations. *Biotechnol Lett [Internet].* 2004;26(6):509–15. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/15127793>
30. Silver N, Best S, Jiang J, Thein SL. Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR. *BMC Mol Biol [Internet].* 2006;7(1):33. Available from: <http://www.biomedcentral.com/1471-2199/7/33>
31. 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.
32. Goulao L, Fortunato A, Ramalho J. Selection of reference genes for normalizing quantitative real-time PCR gene expression data with multiple variables in *Coffea* spp. *Plant Mol Biol Rep.* 2012;30(3):741–59.
33. Le DT, Aldrich DL, Valliyodan B, Watanabe Y, van Ha C, Nishiyama R, et al. Evaluation of Candidate Reference Genes for Normalization of Quantitative RT-PCR in Soybean Tissues under Various Abiotic Stress Conditions. *PLoS One.* 2012;7(9).
34. de Jonge HJM, Fehrmann RSN, de Bont ESJM, Hofstra RMW, Gerbens F, Kamps WA, et al. Evidence based selection of housekeeping genes. *PLoS One.* 2007;2(9).
35. Boda E, Pini A, Hoxha E, Parolisi R, Tempia F. Selection of reference genes for quantitative real-time RT-PCR studies in mouse brain. *J Mol Neurosci [Internet].* 2009;37(3):238–53. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/18607772>
36. Zampieri M, Ciccarone F, Guastafierro T, Bacalini M, Calabrese R, Moreno-Villanueva M, et al. Validation of suitable internal control genes for expression studies in aging. *Mech Ageing Dev.* 2010;131(2):79–85.
37. Khanlou KM, van Bockstaele E. A critique of widely used normalization software tools and an alternative method to identify reliable reference genes in red clover (*Trifolium pratense* L.). *Planta.* 2012;236(5):1381–93.
38. Zsóri K, Muszbek L, Csiki Z, Shemirani A. Validation of reference genes for the determination of platelet transcript level in healthy individuals and in patients with the history of myocardial infarction. *Int J Mol Sci.* 2013;14(2):3456–66.
39. de Santis C, Smith-Keune C, Jerry DR. Normalizing RT-qPCR Data: Are We Getting the Right Answers? An Appraisal of Normalization Approaches and Internal Reference Genes from a Case Study in the Finfish *Lates calcarifer*. *Mar Biotechnol.* 2011;13(2):170–80.
40. Axtner J, Sommer S. Validation of internal reference genes for quantitative real-time PCR in a non-model organism, the yellow-necked mouse, *Apodemus flavicollis*. *BMC Res Notes.* 2009;2:264.
41. Reid KE, Olsson N, Schlosser J, Peng F, Lund ST. An optimized grapevine RNA isolation procedure and statistical determination of reference genes for real-time RT-PCR during berry development. *BMC Plant Biol [Internet].* 2006;6:27. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17105665>
42. Hong SY, Seo PJ, Yang MS, Xiang F, Park CM. Exploring valid reference genes for gene expression studies in *Brachypodium distachyon* by real-time PCR. *BMC Plant Biol [Internet].* 2008;8:112. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/18992143>
43. Maroufi A, Van Bockstaele E, De Loose M. Validation of reference genes for gene expression analysis in chicory (*Cichorium intybus*) using quantitative real-time PCR. *BMC Mol Biol.* 2010;11:15.
44. Marum L, Miguel A, Ricardo CP, Miguel C. Reference gene selection for quantitative real-time PCR normalization in *quercus suber*. *PLoS One.* 2012;7(4).
45. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, et al. The MIQE guidelines: Minimum Information for publication of quantitative real-time PCR experiments. *Clin Chem.* 2009;55(4):611–22.
46. Bustin S a, Beaulieu J-F, Huggett J, Jaggi R, Kibenge FSB, Olsvik P a, et al. MIQE précis: Practical implementation of minimum standard guidelines for fluorescence-based quantitative real-time PCR experiments. *BMC Mol Biol.* 2010;11:74.
47. Derveaux S, Vandesompele J, Hellemans J. How to do successful gene expression analysis using real-time PCR. *Methods.* 2010;50(4):227–30.
48. Chervoneva I, Li Y, Schulz S, Croker S, Wilson C, Waldman SA, et al. Selection of optimal reference genes for normalization in quantitative RT-PCR. *BMC Bioinformatics [Internet].* 2010;11:253. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20470420%5Cnhttp://www.ncbi.nlm.nih.gov/pmc/articles/PMC2889935/pdf/1471-2105-11-253.pdf>
49. Lyng MB, Laenkholm A-V, Pallisgaard N, Ditzel HJ. Identification of genes for normalization of real-time RT-PCR data in breast carcinomas. *BMC Cancer [Internet].* 2008;8:20. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2248196&tool=pmcentrez&rendertype=abstract>
50. Klie M, Debener T. Identification of superior reference genes for data normalisation of expression studies via quantitative PCR in hybrid roses (*Rosa hybrida*). *BMC Res Notes [Internet].* BioMed Central Ltd; 2011;4(1):518. Available from: <http://www.biomedcentral.com/1756-0500/4/518>

Figures

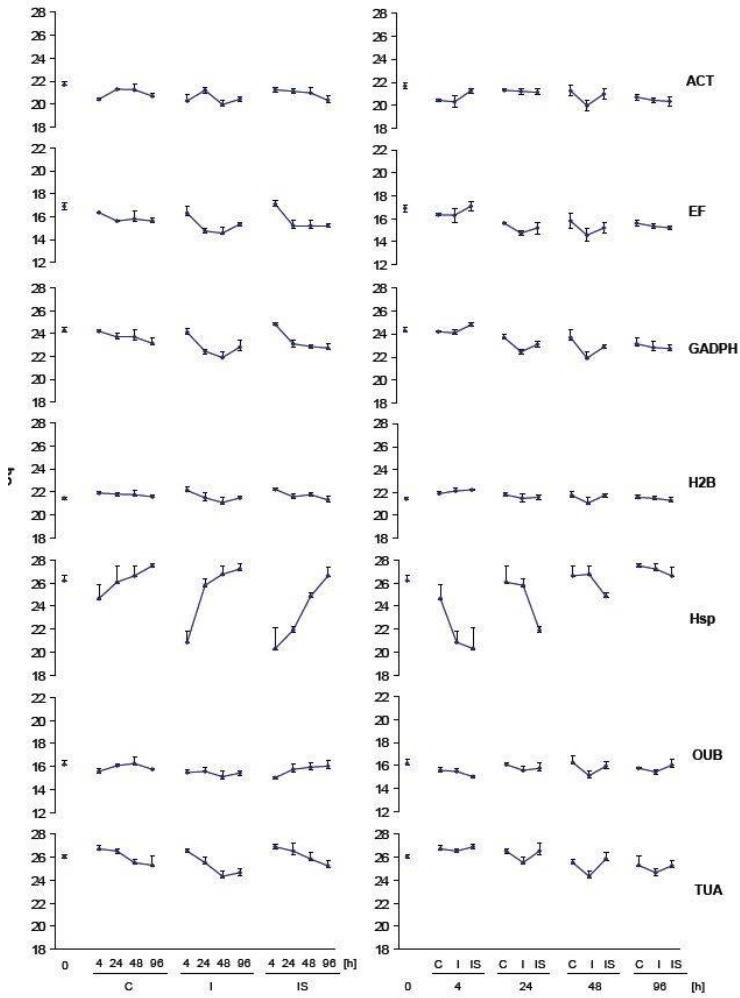


Figure 1
 Corrected quantification cycle (Cq') values for the seven candidate housekeeping genes. Mean of biological groups is represented, with upper semi-bars indicating Cq' SD for biological groups (n=6=3 assays x 2 RT-level replicates), and bottom semi-bars indicating Cq' mean SD for samples (3 samples with 2 RT-level replicates each)

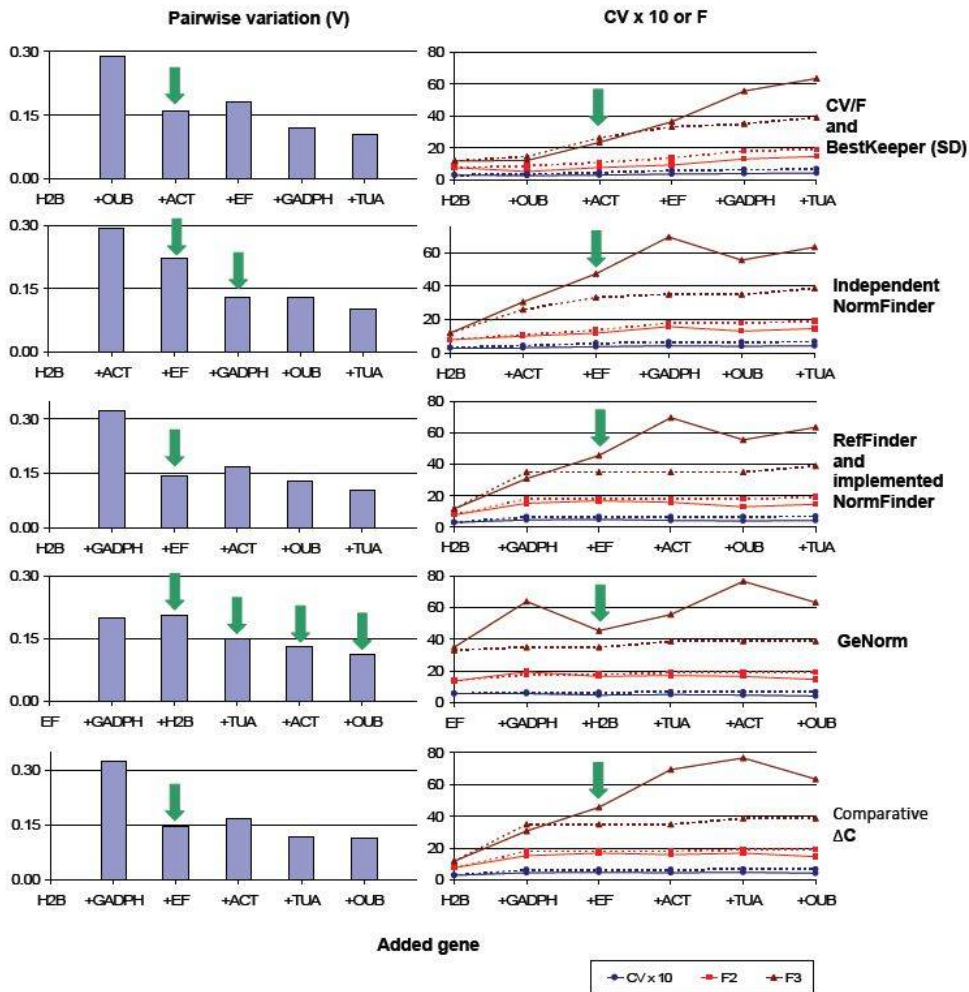


Figure 2

Pairwise variation (left) and descriptive parameters (right) for the normalization factors. Normalization factors are composed from the indicated candidate housekeeping genes, determined according to the indicated methods. Descriptive parameters are CV and F for bifactorial –F2– and trifactorial –F3– cases. Discontinuous lines indicate the maximum value of the parameter into the candidate housekeeping genes contributing to the indicated normalization factors. Arrows indicate the selected NFs after the determination of representative candidate NFs (left) and the assessment of congruent stability (right) (see text, Step 8).

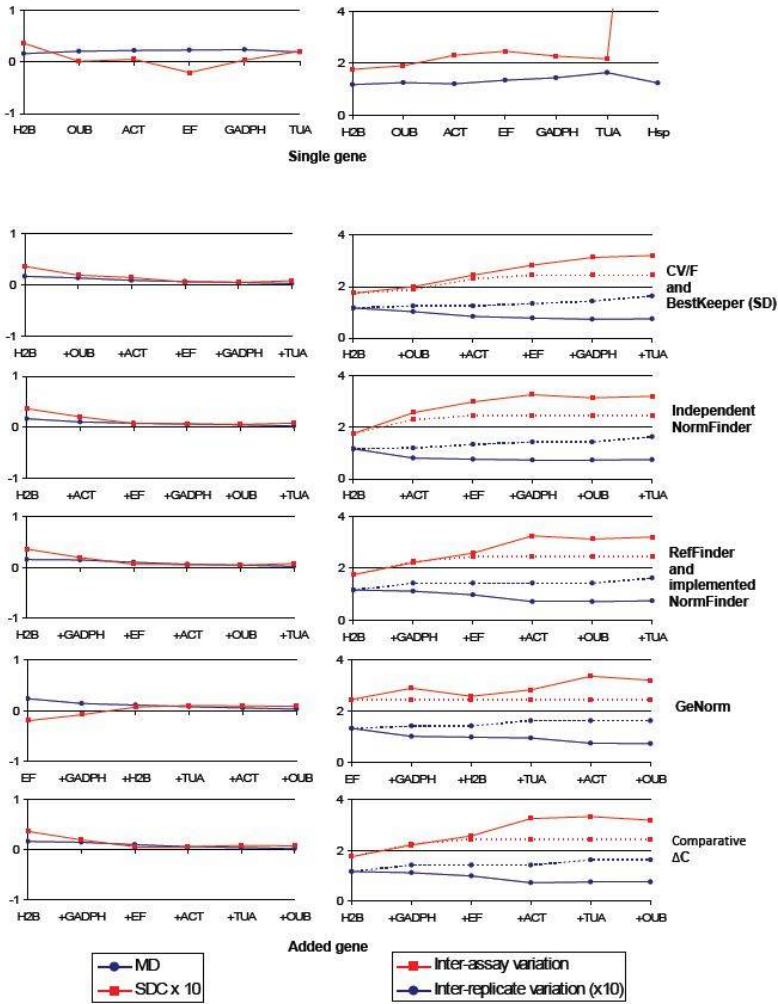


Figure 3 Error compensation and stability of normalization factors. Mean distances between replicates (MD), and compensations (standard deviations of mean distances) (SDC) (graphics on the left), and inter-assay and inter-replicate variations (graphics on the right) for the normalization factors composed from the indicated candidate housekeeping genes, obtained according to the indicated methods. MD and SDC are shown in Cq's. Discontinuous lines indicate the maximum value of the indicated parameter into the candidate housekeeping genes contributing to the indicated normalization factors.

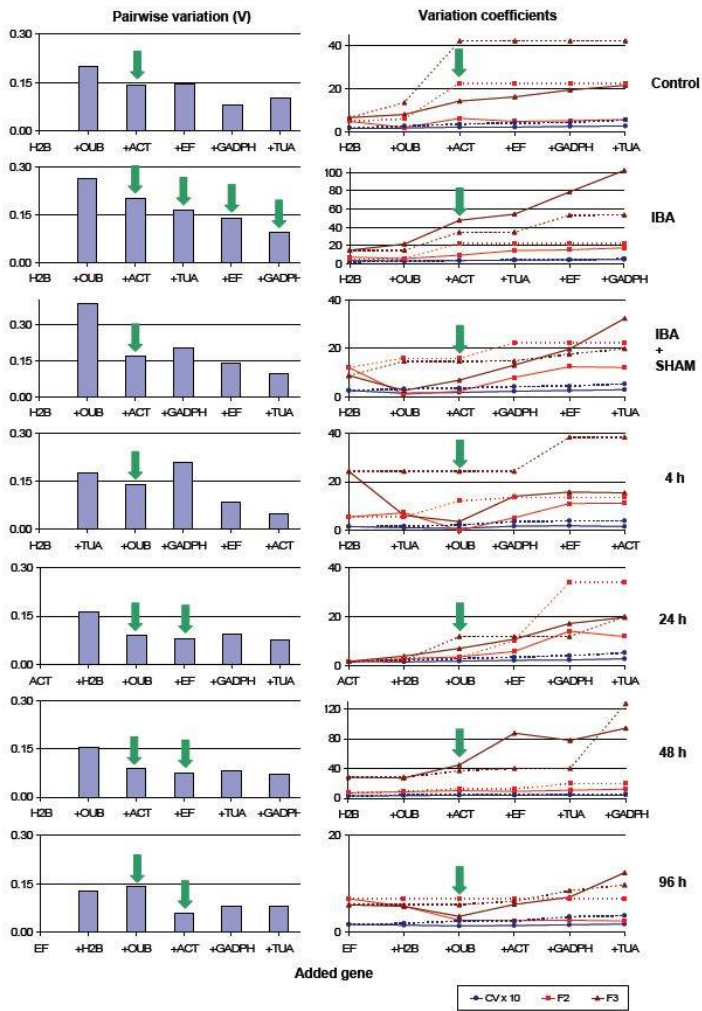


Figure 4 Pairwise variation (left) and descriptive parameters (right) for the normalization factors. Normalization factors are composed from the indicated candidate housekeeping genes, determined for CV ranking. Descriptive parameters are CV and F for bifactorial –F2– and trifactorial –F3– cases. Discontinuous lines indicate the maximum value of the parameter into the CHGs contributing to the indicated normalization factors. Arrows indicates the selected normalization factors after the determination of representative candidate NFs (left) and the assessment of congruent stability (right) (see text, Step 8).

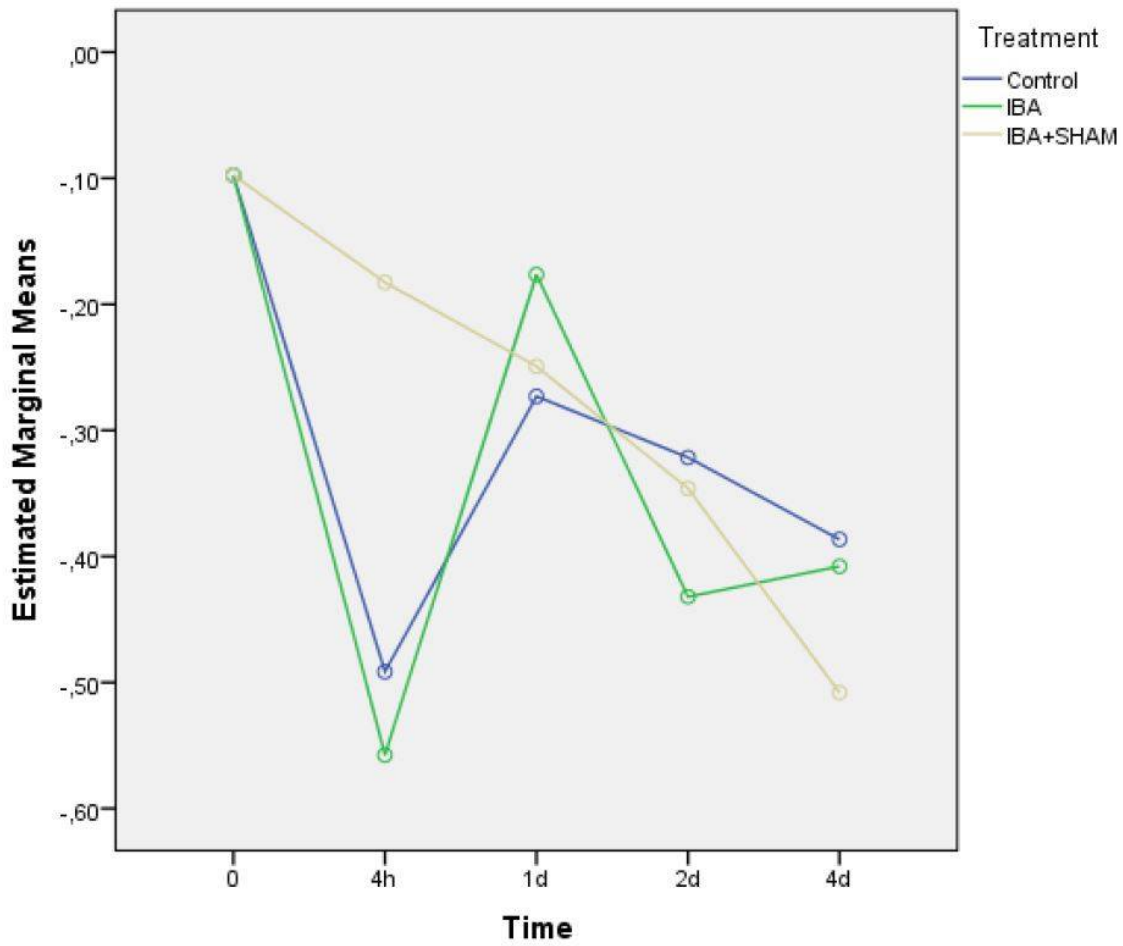


Figure 5

Estimated marginal means of an ideally stable gene according to NF3 when NF2 is applied. NF3: normalization factor composed from H2B, OUB and ACT. NF2: normalization factor composed from H2B and OUB.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Affiliations.docx](#)
- [Additionalfile4.pdf](#)
- [Additionalfile3.pdf](#)
- [Additionalfile2.pdf](#)
- [Additionalfile1.pdf](#)