

Research article

Open Access

## Identification of suitable internal control genes for expression studies in *Coffea arabica* under different experimental conditions

Carla F Barsalobres-Cavallari\*<sup>1</sup>, Fábio E Severino<sup>1</sup>, Mirian P Maluf<sup>2</sup> and Ivan G Maia<sup>1</sup>

Address: <sup>1</sup>Laboratório de Biotecnologia e Genética Molecular, Departamento de Genética, Instituto de Biociências, UNESP, Distrito de Rubião Júnior s/n, 18618-000, Botucatu, São Paulo, Brazil and <sup>2</sup>Embrapa/IAC, Centro de Café Alcides Carvalho, Campinas, São Paulo, Brazil

Email: Carla F Barsalobres-Cavallari\* - barsalobres@gmail.com; Fábio E Severino - fabio.bjj@gmail.com; Mirian P Maluf - maluf@iac.sp.gov.br; Ivan G Maia - igmaia@ibb.unesp.br

\* Corresponding author

Published: 6 January 2009

Received: 16 July 2008

BMC Molecular Biology 2009, 10:1 doi:10.1186/1471-2199-10-1

Accepted: 6 January 2009

This article is available from: <http://www.biomedcentral.com/1471-2199/10/1>

© 2009 Barsalobres-Cavallari et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

### Abstract

**Background:** Quantitative data from gene expression experiments are often normalized by transcription levels of reference or housekeeping genes. An inherent assumption for their use is that the expression of these genes is highly uniform in living organisms during various phases of development, in different cell types and under diverse environmental conditions. To date, the validation of reference genes in plants has received very little attention and suitable reference genes have not been defined for a great number of crop species including *Coffea arabica*. The aim of the research reported herein was to compare the relative expression of a set of potential reference genes across different types of tissue/organ samples of coffee. We also validated the expression profiles of the selected reference genes at various stages of development and under a specific biotic stress.

**Results:** The expression levels of five frequently used housekeeping genes (reference genes), namely *alcohol dehydrogenase (adh)*, *14-3-3*, *polyubiquitin (poly)*, *β-actin (actin)* and *glyceraldehyde-3-phosphate dehydrogenase (gapdh)* was assessed by quantitative real-time RT-PCR over a set of five tissue/organ samples (root, stem, leaf, flower, and fruits) of *Coffea arabica* plants. In addition to these commonly used internal controls, three other genes encoding a cysteine proteinase (*cys*), a caffeine synthase (*ccs*) and the 60S ribosomal protein L7 (*rpl7*) were also tested. Their stability and suitability as reference genes were validated by geNorm, NormFinder and BestKeeper programs. The obtained results revealed significantly variable expression levels of all reference genes analyzed, with the exception of *gapdh*, which showed no significant changes in expression among the investigated experimental conditions.

**Conclusion:** Our data suggests that the expression of housekeeping genes is not completely stable in coffee. Based on our results, *gapdh*, followed by *14-3-3* and *rpl7* were found to be homogeneously expressed and are therefore adequate for normalization purposes, showing equivalent transcript levels in different tissue/organ samples. *Gapdh* is therefore the recommended reference gene for measuring gene expression in *Coffea arabica*. Its use will enable more accurate and reliable normalization of tissue/organ-specific gene expression studies in this important cherry crop plant.

## Background

The study of biological regulations is very often correlated to quantification assays. In order to detect differential expression of a gene(s) in distinct biological samples, such as tissue types or under different experimental conditions, the invention of quantitative PCR (qPCR) has transformed the field of gene expression analysis in living organisms [1]. In comparison to classical reverse transcription-polymerase chain reaction (RT-PCR), the main advantages of qPCR are higher sensitivity, specificity and broad quantification range of up to seven orders of magnitude [2-6]. Regardless of being an extremely powerful technique, qPCR has its pitfalls, the most important one being the need of appropriate data normalization with a reference gene [6-14].

According to Andersen et al. [8], accurate data normalization is an absolute requirement for correct measurement of gene expression. Expression of the reference gene used to normalize qPCR analyses should be unaffected throughout many biological contexts; otherwise, it may lead to erroneous results [6-9,15-18]. Until recently, several such genes (*β-actin*, *rRNA*, *β-tubulin*, *alcohol dehydrogenase*, *glyceraldehyde-3-phosphate dehydrogenase*, *14-3-3* and *polyubiquitin*) have been used as internal controls for gene expression analyses under the assumption of stable expression [19-26]. However, several reports have demonstrated that the expression levels of these so-called reference genes differ among different tissue/organ types [10,16,27-33]. Consequently, these genes are unsuitable as transcriptional inner controls, and their use to normalize qPCR data in different tissues may induce significant experimental errors that could result in inappropriate biological data interpretation [9,14,17,21-25,34-39].

Recognizing the importance of reference gene(s) in normalization of qPCR data, various housekeeping genes have been evaluated for stable expression under specific conditions in various organisms. In plants, only a few of them have been investigated in some detail in rice [15,40,41], poplar [36], potato [39], soybean [42,43] and *Arabidopsis thaliana* [30]. So far, suitable internal controls for gene expression studies have not been defined for *Coffea arabica*.

Coffee is an agricultural crop of significant economic importance. *Coffea arabica* L. (arabica type coffee) is typical of the highland growing regions and is responsible for almost 75% of world production [44]. In this study, we report the validation of housekeeping genes to identify the most suitable internal reference gene(s) for normalization of qPCR data obtained among five different tissues/organs (root, stem, leaf, flower, and fruits) of *C. arabica*. To further validate our results, we evaluated the expression levels of our best reference genes at different develop-

mental stages of flowers and cherries and under a specific biotic stress. Following the current literature, five candidate reference genes, namely *alcohol dehydrogenase (adh)*, *polyubiquitin (poly)*, *14-3-3*, *β-actin (actin)* and *glyceraldehyde-3-phosphate dehydrogenase (gapdh)*, were selected. In addition to these commonly used internal controls, three other genes coding for a cysteine proteinase (*cys*), a caffeine synthase (*ccs*) and the 60S ribosomal protein L7 (*rpl7*), respectively, were included in this analysis. These potential reference genes were ranked according to their expression profiles and stability.

## Results and discussion

The expression profile of eight candidate reference genes (*actin*, *adh*, *14-3-3*, *ccs*, *gapdh*, *poly*, *rpl7*, or *cys*) was firstly assessed by qPCR over a panel of five coffee tissue/organ samples (root, stem, leaf, flower, or fruit).

### Descriptive analysis of the reference candidate genes

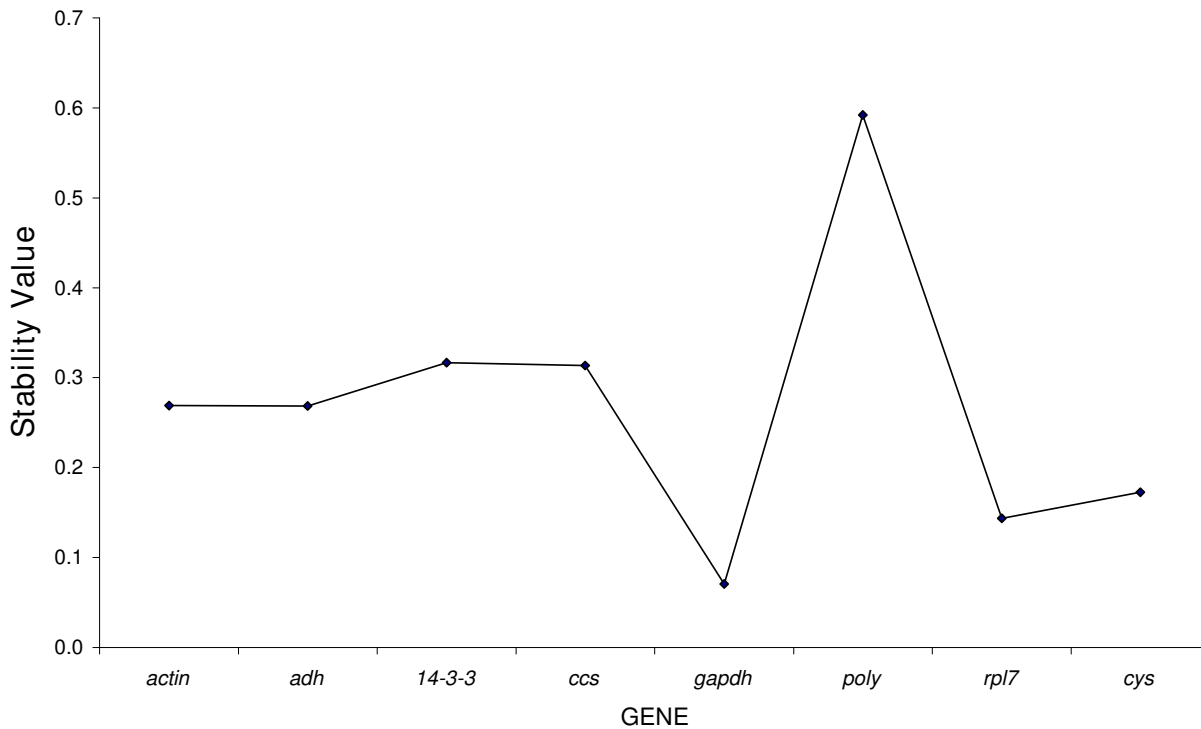
Descriptive statistics of the derived crossing points (CP), based on BestKeeper program [45], were calculated to investigate the variation level of each candidate gene following Pfaffl et al. [46]. According to this analysis (see Table 1), the gene with lowest expression level was *actin*, for which CP values were obtained around cycles 31–34; while the highest was *gapdh*, whose CP values were obtained around cycles 21–23. The expression levels of *14-3-3*, *ccs*, *gapdh*, and *rpl7* presented fluctuations of approximately  $\pm 0.6$  x-fold ( $0.52$  x-fold  $< SD < 0.82$  x-fold), whereas *poly* expression showed higher ranges of CP variation ( $SD = \pm 1.39$  cycles) as well as up- down-regulation ( $\pm 2.09$  x-fold). The coefficient of variation (CV) of the assay was 3.57% (total assay variability), which is within the range (from 3.4% to 11.6%) of previously reported values for qPCR [46].

Descriptive statistics for the expression analyses of each reference candidate gene in the five distinct coffee plant tissue/organ types was also obtained (see Additional File 1). According to this analysis, the gene that exhibit the minor gene expression variation among the analyzed tissues was *gapdh* [ $SD (\pm \text{x-fold}) = 0.04$ ], while the gene with major variation was *ccs* [ $SD (\pm \text{x-fold}) = 0.59$ ]. In addition, *ccs* also presented the highest expression variation in flowers and fruits tissue samples, and therefore it cannot be used as a reference gene. Numerous studies have shown that the expression of housekeeping genes can vary under given situations [28]. This may partly be explained by the fact that housekeeping genes are not only implicated in the basal cell metabolism but also participate in other cellular functions [47,48].

**Table 1: Descriptive statistics and expression level analyses of the tested candidate reference genes based on their crossing point (CP) values**

Factor	<i>actin</i>	<i>adh</i>	<i>14-3-3</i>	<i>ccs</i>	<i>gapdh</i>	<i>poly</i>	<i>rpl7</i>	<i>cys</i>
N	21	21	21	21	21	21	21	21
GM [CP]	32.72	29.7	29.63	26.67	22.86	28.46	30.76	26.67
AM [CP]	32.74	29.71	29.64	26.67	22.88	28.49	30.77	26.7
Min [CP]	31.4	28.64	28.94	25.92	21.74	26.79	30.13	25.24
Max [CP]	34.35	31.12	31.21	27.37	23.77	30.07	31.93	28.44
SD [ $\pm$ CP]	1.23	0.99	0.92	0.58	0.94	1.39	0.69	1.31
CV [%CP]	3.75	3.35	3.09	2.19	4.12	4.89	2.24	4.89
Min [x-fold]	0.41	0.43	0.66	0.53	0.54	0.18	0.68	0.41
Max [x-fold]	3.04	3.1	2.62	1.83	1.64	5.22	2.02	3.04
SD [ $\pm$ x-fold]	1.07	1.08	0.82	0.52	0.52	2.09	0.53	1.08

Abbreviations: N: number of samples; CP: crossing-point; GM [CP]: geometric CP mean; AM [CP]: arithmetic CP mean; Min [CP] and Max [CP]: CP threshold values; SD [ $\pm$  CP]: CP standard deviation; CV [%CP]: variance coefficient expressed as percentage of CP level; Min [x-fold] and Max [x-fold]: threshold expression levels expressed as absolute x-fold over- or under-regulation coefficient; SD [ $\pm$  x-fold]: standard deviation of absolute regulation coefficient.



**Figure 1**  
**Stability of the investigated candidate reference genes.** Stability values of the eight candidate reference genes according to the model-based approach. A lower value of average expression stability indicates more stable expression.

**A model-based approach for estimation of expression variation**

The model-based variance estimation approach, a Visual Basic application for Microsoft Excel (termed NormFinder) [8,49], was used to evaluate expression stability of reference candidate genes. This analysis allowed the ranking of candidate genes since the estimated variation directly indicates the introduced error associated with their use. According to this method, the gene with minimal estimated intra- and intertissue variation was *gapdh* (expression stability = 0.071) while the gene with the maximal variation was *poly* (expression stability = 0.592) (see Figure 1), thus corroborating the results obtained in descriptive analysis.

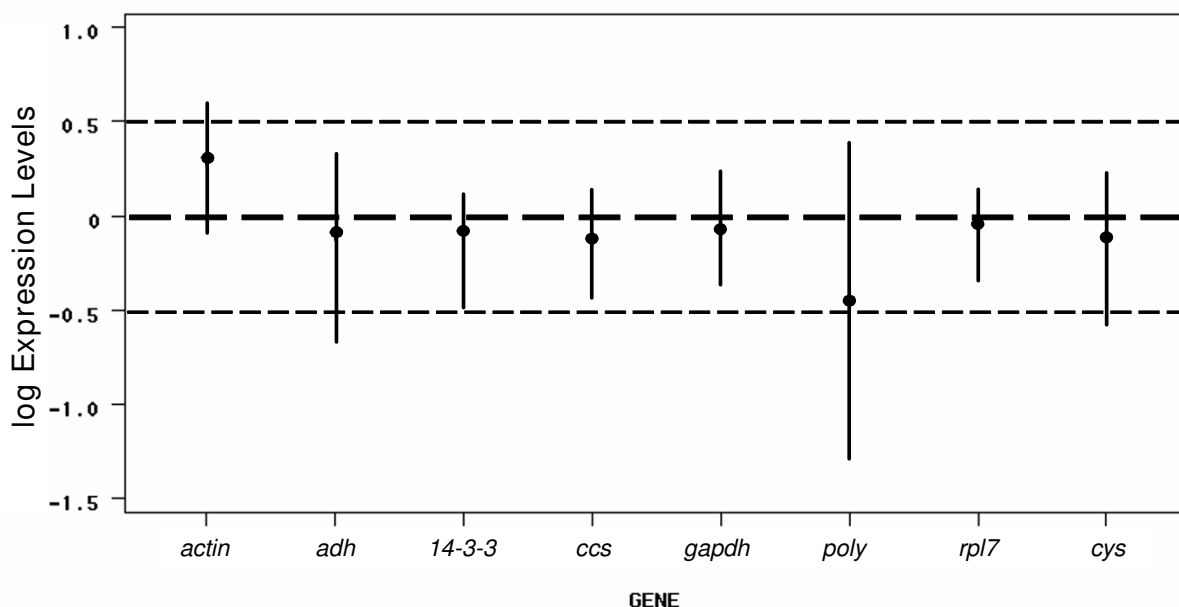
**Ranking the candidate reference genes**

The relationship between the stability value and the intra- and intertissue expression variations is present in Figure 2. This figure clearly demonstrates the distinct specificities of the investigated genes. According to this analysis, the best candidate gene should present the minimal combined inter- and intra-tissue expression variation. Consistent with the descriptive analysis (see Table 1 and Additional File 1) and the model-based variance estimation approach (see Figure 1), *gapdh* showed not only the highest expres-

sion levels but was also the most stable gene studied. In Figure 2, it can be observed that almost all genes presented average of log expression levels near 0 (the thick dashed line). Log difference >0 (as observed for *actin*) or <0 (as observed for *poly*) implies that variability in expression levels is significant, so the gene could be incorrectly used as reference gene for normalization. In this context, the *poly* gene presents the highest intertissue expression variation (SD = ± 0.65; see Figure 2). Thus, among the tested genes, *gapdh*, followed by *rpl7* and *14-3-3*, showed the most stable expression over the investigated panel of five distinct coffee tissue/organ types.

The candidate genes were also ranked according to their M values using the geNorm program. The average expression stabilities (M values) of all tested genes were lower than 1.5, with *14-3-3* and *actin* showing the most stable expression (data not shown). Although *actin* gene has shown highest stability following geNorm analysis, this gene presented the lowest expression profile according the Best-Keeper analysis. Corroborating the previous analysis, *poly* remained the least stable gene.

As a whole, our analysis indicates that housekeeping genes are differently regulated in different tissues/organs



**Figure 2**  
**Gene expression differences among the candidate reference genes.** The log-transformed gene expression levels are represented by black circles. The intertissue variation is indicated by vertical bars that give a confidence interval for the difference. The two thin dashed lines represent the maximal standard deviation of the reference candidate genes, with a log expression levels difference between 0.5 and -0.5.

of coffee plants and may exhibit variable expression patterns. The observed differences in gene expression ratios along a comprehensive panel of tissues/organs are consistent with the data presented by Barber et al. [1], Iskandar et al. [50] and Jain et al. [15]. Our results also provide evidences that normalizations to the expression level of a single gene in samples from distinct tissue types may induce to errors, thus corroborating previous studies [29,51,52].

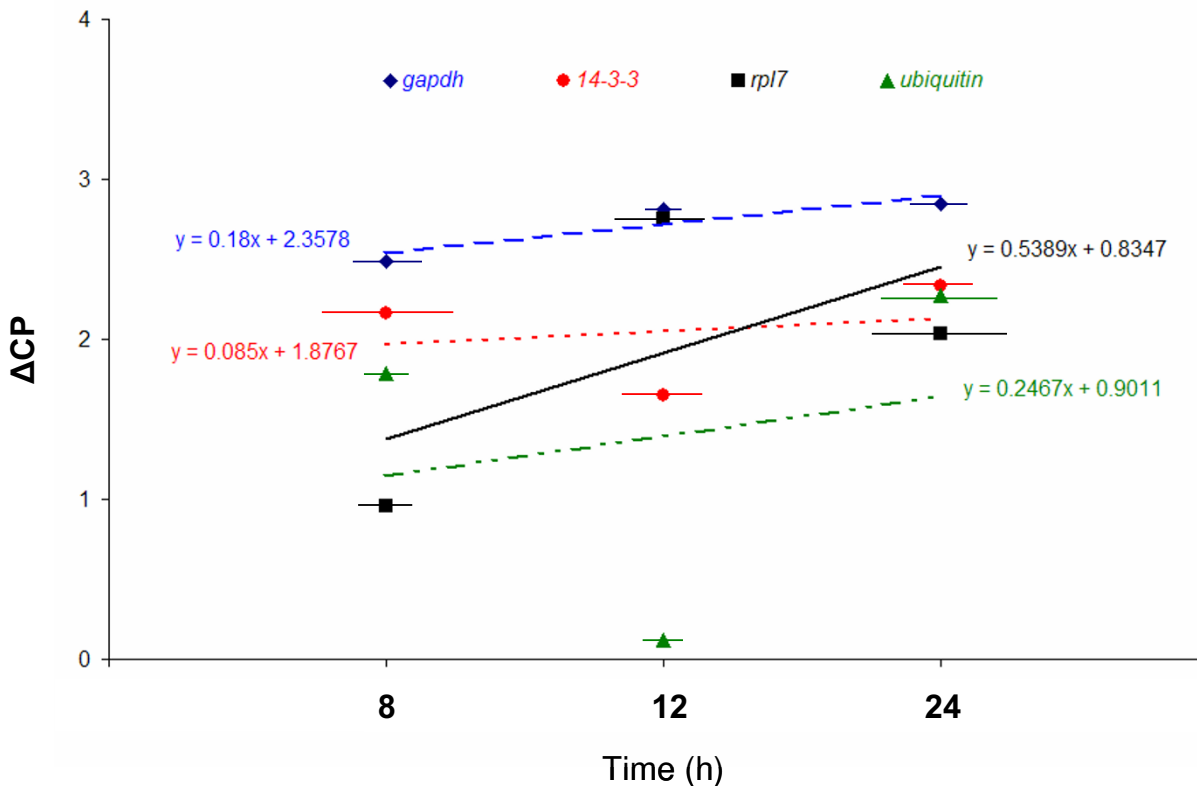
**Comparison of the identified best reference genes to published data**

In order to validate our potential candidate reference genes (*gapdh*, *rpl7* and *14-3-3*), the expression stability of these genes under the influence of a specific biotic stress was investigated. In this case, the obtained results were

compared to those dealing with similar coffee gene expression analysis but using *ubiquitin* as a reference gene for normalization [53-56].

The comparison was conducted by linear regressions analyzes of the CP difference ( $\Delta CP$ ) obtained from the assayed expression levels of the tested genes in leaves of *C. arabica* inoculated, or not (control plants), with *Hemileia vastatrix*. The average CP ( $N = 3$ ) was calculated for each gene and the  $\Delta CP$  ( $CP_{\text{inoculated leaf}} - CP_{\text{non-inoculated leaf}}$ ) was determined for each time-point (8, 12 and 24 h after fungus inoculation).

As it can be observed in Figure 3, the regression lines for *14-3-3* and *gapdh* have slopes close to zero (*14-3-3* = 0.085; *gapdh* = 0.18), indicating similar expression levels



**Figure 3**

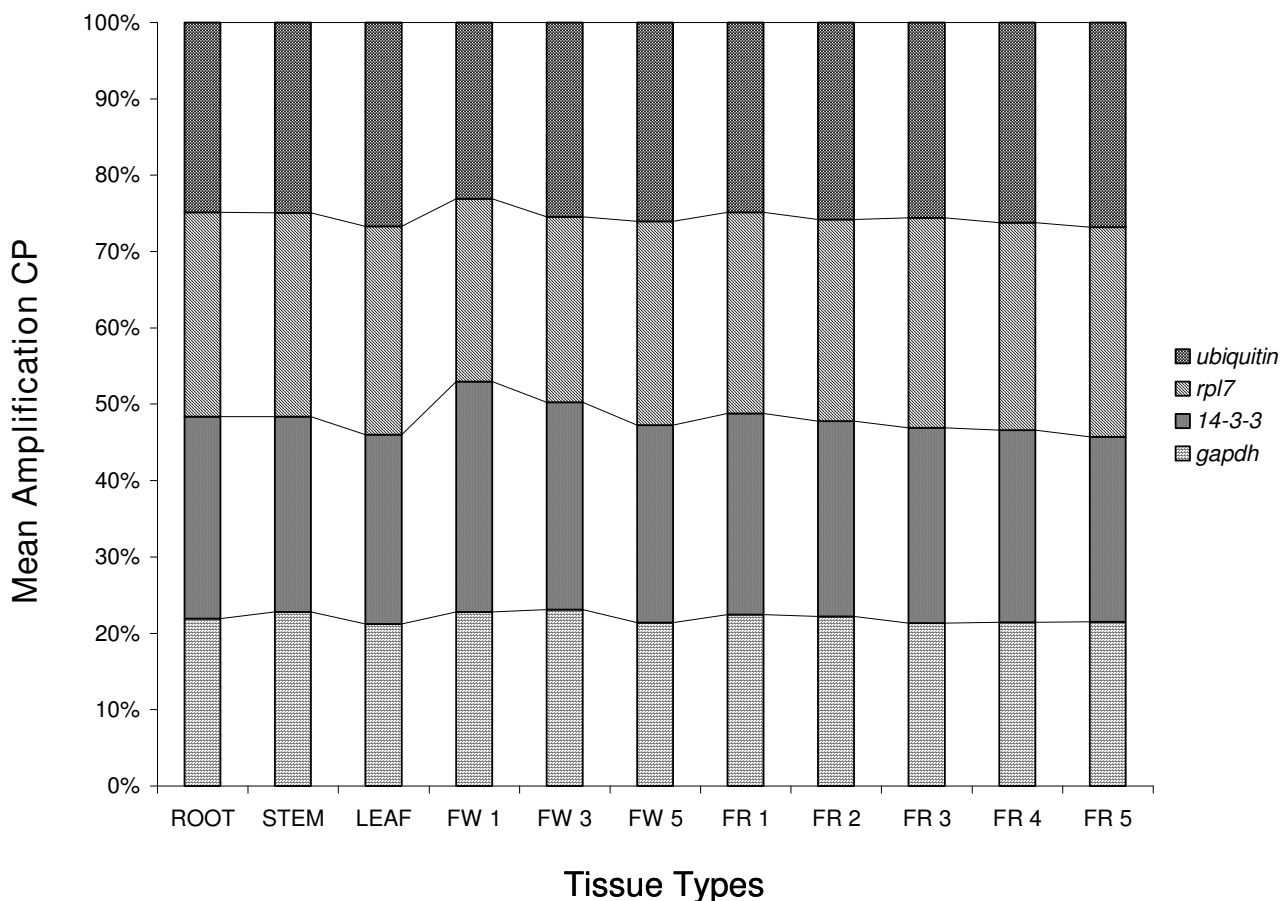
**Evaluation of the expression of selected reference genes during fungus infection.** The expression of selected reference genes (*gapdh*, *rpl7* and *14-3-3*) and of a commonly used coffee normalization gene (*ubiquitin*) was monitored in leaves of *C. arabica* var. Mundo Novo inoculated with *Hemileia vastatrix*. The crossing point (CP) difference ( $\Delta CP = CP_{\text{inoculated leaf}} - CP_{\text{non-inoculated leaf}}$ ) was calculated for each time-point (8, 12 and 24 h after challenge by the rust fungus) to investigate the expression levels of each reference gene. The standard error of the triplicates for each time-point is indicated by horizontal bars.

in inoculated and non-inoculated leaves, and reinforcing their use as effective normalization genes [1,57,58]. In contrast, the slope value for *rpl7* was significantly different from zero and higher than the one obtained for *ubiquitin* (*ubiquitin* = 0.2467; *rpl7* = 0.5389), thus limiting its use as a normalization factor under biotic stress condition.

**Validation of data results in different developmental stages of flowers and coffee cherries**

An additional validation step of the expression levels of *gapdh*, *14-3-3*, *rpl7* and *ubiquitin* was performed using unpooled tissue samples from flower and cherry developmental series. The employed sample set is given in the Additional File 2.

In this assay, the highly expressed gene was *gapdh* ( $\overline{CP} = 21$  cycles) followed by *ubiquitin* ( $\overline{CP} = 24$  cycles), while *14-3-3* and *rpl7* presented the same mean CP (25 cycles). The comparison of gene contributions is present in Figure 4. As already observed, *gapdh* showed the greatest stability in expression among all coffee tissue/organ samples analyzed, while expression of *rpl7* and *14-3-3* varied the most, especially in flowers at stage 1 of development (see Figure 4 and Additional File 2). As mentioned earlier, *ubiquitin* was used as a standard reference gene for comparison.



**Figure 4**  
**Validation of the selected reference genes in samples from flowers and cherries at different developmental stages.** Comparison of gene contributions, by mean amplification crossing points (CP) represented in percentage, in each coffee tissue/organ type. The investigated tissue/organ sample set was: root, stem, leaf, three different stages of flower development (FW 1, FW 3 and FW 5) and five different kinds of coffee cherries (FR 1, FR 2, FR 3, FR 4 and FR 5).

### General remarks about the selected reference genes

According to these results, the gene encoding glyceraldehyde-3-phosphate dehydrogenase (GAPDH), an enzyme of glycolysis [59], outperformed all other reference genes tested and should therefore be considered a suitable reference gene for expression studies in arabica coffee plants. This observation corroborates the quantification of *gapdh* expression in different tissues of sugarcane [50] and *Eucalyptus* [20]. In contrast, several reports in human and animal systems have suggested that this reference gene has limitations for its use as internal control due to its marked variability of expression among tissue types [1,16,51,60].

The assayed 14-3-3 gene also showed a stable expression (see Figure 2) that was supported by the descriptive analysis (see Table 1 and also Additional File 1) and confirmed in the biotic stress assay (see Figure 3). However, its expression presented some variation among different stages of coffee flower development (see Figure 4). Papini-Terzi et al. [19] recommended 14-3-3 as a suitable reference gene for expression normalization in a wide range of tissue samples of sugarcane.

The gene encoding the transcription regulator and structural constituent of the 60S subunit of the cytosolic ribosome (*rpl7*) could also be used as an internal control in gene expression studies in *C. arabica*, due to its stability (see Figure 2) and acceptable variation among tissue/organ types (see Table 1). Our results are in agreement with previous published data for this gene since small variation among tissue types was detected by descriptive analysis (see Additional File 1). Nevertheless, in leaves of *C. arabica* inoculated with *H. vastatrix*, it was observed that the expression ratio of *rpl7* (expressed by  $\Delta$ CP) was not constant and the absolute value of *rpl7* linear regression slope was superior to that observed for a commonly used coffee normalization gene (*ubiquitin*) (see Figure 3). In addition, this gene, like 14-3-3, presented a variable expression level among different stages of coffee flower development (see Figure 4). In sugarcane, the relative expression of *rpl35-4*, a gene coding for the ribosomal L35-4 60S protein, was also reported to be stable [61]. These authors estimated the sugarcane leaf transcriptome using Serial Analysis of Gene Expression (SAGE) and reported that tag associated with the *rpl35-4* transcript presented minimum variation among the analyzed SAGE libraries.

The remaining tested genes showed to be unsuitable as internal controls for normalization purposes in *C. arabica*.

### Conclusion

This study provides the most extensive collection of arabica coffee tissue/organ mRNA expression data for eight reference candidate genes. Our analysis evidenced stable

levels of *gapdh*, 14-3-3 and *rpl7* mRNA in different *Coffea arabica* tissue/organ types. Consequently, these genes can be used for accurate and reliable normalization in future gene expression studies in coffee (e.g., they can be used as a reference for a target gene in a specific tissue or experimental condition). In this respect, we suggest *gapdh* as the most relevant reference gene for accurate normalization purposes in *C. arabica*, showing almost constant expression levels in the investigated experimental set-up.

Moreover, we have shown that depending on the reference inner control gene, the within-tissue variation of mRNA expression levels is generally small, whereas among tissues/organs the variation can be substantial. This indicates that normalizations to a single gene across different tissue types are unwise. Since the variation observed between normal tissues of different types may in part be due to the different metabolic demands of those tissues, comparisons within a tissue type between normal and diseased states are similarly unwise.

### Methods

#### Plant material and growth conditions

Freshly harvested roots, stems, and leaves were obtained from ten 4 month-old coffee plants (*Coffea arabica* var. Mundo Novo IAC 388-17-1) grown under greenhouse conditions (28 °C, 60% RH) in Campinas, São Paulo, Brazil. Flower and fruit samples, at different developmental stages, were collected from 4–5 year-old plants of var. Mundo Novo grown under field conditions at Botucatu and Campinas, São Paulo, Brazil (see Additional File 2). After harvesting, fresh tissue samples were frozen immediately in liquid nitrogen until RNA extraction.

#### Biotic stress assay

For the biotic stress assay, equally-aged sets of *Coffea arabica* var. Mundo Novo plants were kept in growth chamber (16 h/8 h light/dark; 23 °C; 70% RH) for at least one week, before being inoculated with the coffee leaf-rust fungus *Hemileia vastatrix* Berk and Br. race II, that elicits a compatible reaction in coffee. The urediniospores (100 mg) were harvested in a *C. arabica* field in Campinas, São Paulo, Brazil, and diluted in 10 ml of sterile water under dark conditions.

Leaves from the second pair of plagiotropic shoots from the apex of 4 month-old coffee plants were inoculated with an aqueous suspension of fresh urediniospores (10 mg/ml). To allow spore germination, the inoculated leaves were covered with a wet black plastic film for 24 h. Inoculated leaves were not detached from the plants. Mock-inoculated controls as well as non-inoculated controls were performed. The biological samples were obtained from three independent experiments.

Leaves were randomly sampled at different time-points after inoculation: 0, 8, 12, and 24 h, and immediately deep-frozen. To confirm the infection by the leaf-rust fungus, some inoculated leaves were maintained in plants.

### RNA isolation and quality controls

Tissue samples of 2.0 to 2.5 mg were weighed and ground to fine powder in liquid nitrogen using a pre-cooled mortar and pestle. Total RNA from the majority of the samples was extracted using TRIzol reagent (Invitrogen) according to manufacturer's instructions. Alternatively, total RNA from seeds was isolated by lithium chloride (LiCl) method, according to Mason and Schmidt [62]. Only RNA samples with 260/280 ratio between 1.9 and 2.1 and 260/230 ratio greater than 2.0 were used for subsequent analyses. The integrity of the RNA samples was also assessed on 1.0% agarose/formaldehyde gel electrophoresis.

### Reverse transcription

Five micrograms of total RNA were treated with DNase I (Promega) and an aliquot of 500 ng of the treated RNA was reverse-transcribed using SuperScript First-Strand Synthesis System for RT-PCR (Invitrogen). Both were used following the manufacturer's instructions. The cDNA sample concentration was determined using the NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies).

### Primer design

Primers were designed for *Coffea* orthologs of commonly used housekeeping genes representing distinct functional classes, identified by BLAST searches in the Brazilian Coff

fee EST database [63,64] as well as in the public coffee EST database (at the SOL site hosted by Cornell University [65]). For primer design, the Primer Express 2.0 software (PE Applied Biosystems, USA) with default parameters was employed. The accession numbers, gene description, primers sequences and amplicon lengths are shown in Table 2. All primer pairs produced a single product and amplified the target transcript with equal efficiency over a 1000-fold range of input material.

### Quantitative PCR

The PCR mixture contained 5 µl of a 1:10 dilution of the synthesized cDNA, primers to a final concentration of 700 nM each, 17.5 µl of the SYBR Green PCR Master Mix (Applied Biosystems, USA) and PCR-grade water up to a total volume of 35 µl. The mixes were homogenized and split in three samples of 10 µl, thus each gene reaction was performed in triplicate. PCR reactions in the absence of template were also performed as negative controls for each primer pair. An equimolar pool of cDNA samples of five coffee tissue/organ types (root, stem, leaf, flower, and fruit) was prepared to be used as a common reference for all qPCR. The quantitative PCRs were performed employing the ABI Prism 7300 Sequence Detection System (PE Applied Biosystems, USA). All PCR reactions were performed under the following conditions: 2 min at 50°C, 10 min at 95°C, and 45 cycles of 15 s at 95°C and 1 min at 65°C in 96-well optical reaction plates (Applied Biosystems, USA). Confirmation of amplicon specificity was based on the dissociation curve at the end of each run and by product visualization after electrophoresis on an 8% polyacrylamide gel.

**Table 2: Candidate reference genes and primer sequences used for quantitative PCR analysis**

Gene name	Source gene <sup>a</sup>	Gene description	Primer sequence <sup>b</sup>	Amplicon length (bp)
<i>poly</i>	<b>SGN-U347154</b>	hexameric polyubiquitin	5' CGCTGACTACAATATCCAAAAGGA 3' 5' CTGCATTCCACCCCTCAGA 3'	67
<i>adh</i>	<b>SGN-U350348</b>	alcohol dehydrogenase class III	5' CCTCAAGCCGGCGAAGT 3' 5' CTGTATGGCAGAGGGCAGTGT 3'	55
<i>actin</i>	<b>SGN-U353034</b>	actin 7	5' AATTGTCCGTGACATCAAGGAA 3' 5' TGAGCTGCTCTTGGCTGTTTC 3'	82
<i>gapdh</i>	<b>SGN-U347734</b>	glyceraldehyde-3-phosphate dehydrogenase	5' TTGAAGGGCGGTGCAA 3' 5' AACATGGGTGCATCCTTGCT 3'	59
<i>rpl7</i>	<b>SGN-U351477</b>	60S ribosomal protein L7	5' CATTTCGAGGTATCAATGCTATGCA 3' 5' TGTCTCAGGCGCAGAAGCT 3'	66
<i>ccs</i>	<b>SGN-U350284</b>	caffeine synthase I	5' CAATGCCCGGCTCTTTCTAC 3' 5' GTAACAAGAGTGTAATAATGCATGGA 3'	68
<i>cys</i>	<b>SGN-U352616</b>	cysteine proteinase	5' GCGATCGCTACCGTCCAA 3' 5' CTTTTTCTCTCCAGTCAATGGAGTT 3'	63
<i>l4-3-3</i>	<b>SGN-U356404</b>	l4-3-3 protein	5' TTGTGCTCTTTAGCTTCCAAACG 3' 5' CTTACGAGACATATTGTCTTACTCAA 3'	75

<sup>a</sup> Unigene accession number according to the SOL Genomics Network [65]

<sup>b</sup> Forward (upper line) and reverse (lower line) primer sequences



### Analysis of candidate reference genes

To estimate the expression variation level of the eight candidate genes (*actin*, *adh*, *ccs*, *14-3-3*, *gapdh*, *poly*, *rpl7*, or *cys*) over a five coffee tissue/organ sample set (root, stem, leaf, flower, or fruit), the BestKeeper descriptive statistical method [45,66] was applied.

To access the levels of gene expression for each gene in the different coffee tissue types, the method described by Ramakers et al. [67] with modifications was used. Optic data were exported from 7300 Real-Time PCR System (PE Applied Biosystems, USA) into MS Excel. Four cycles at the exponential phase, near and including the crossing point (CP), were used. The fluorescence data were logarithmically transformed, and pasted into statistical software package (SAS version 8e, SAS Institute, Cary, NC, USA) for linear regression analysis, including determination of intercepts, slopes (x), PCR efficiency ( $E = 10^{\text{slope}}$ ) and their respective standard errors and correlation coefficients ( $R^2$ ). The gene expression values (x-fold) were obtained according to a mathematical model proposed by Pfaffl [68]:  $x\text{-fold} = E_{\text{reference gene}}^{\Delta CP}$ , where  $\Delta CP = CP_{\text{pool of tissues}} - CP_{\text{tissue sample}}$ .

Knowing the expression levels, the stability value was assessed utilizing NormFinder [8,49] and geNorm [17,69] programs.

### Abbreviations

qPCR: Quantitative PCR; *poly*: hexameric polyubiquitin; *adh*: alcohol dehydrogenase class III; *actin*: actin 7; *rpl7*: 60S ribosomal protein L7; *ccs*: caffeine synthase 1; *gapdh*: glyceraldehyde-3-phosphate dehydrogenase; *cys*: cysteine proteinase; CP: crossing point.

### Authors' contributions

CFB-C and IGM conceived and designed this study. FES performed the biological assays. CFB-C and FES carried out the molecular genetic studies, participated in the qPCR experiments (acquisition, analysis and interpretation of data). MPM and IGM contributed with reagents/materials/analysis tools. IGM coordinated the study. CFB-C wrote the manuscript. All authors contributed, read, corrected and approved the final manuscript. The authors declare no conflict of interest in this work.

### Additional material

#### Additional file 1

Results from BestKeeper descriptive statistical analysis. The data provided represent the descriptive statistics, based on crossing point (CP) values, for the expression analyses of the candidate reference genes in the five distinct coffee plant tissue/organ set.

Click here for file

[http://www.biomedcentral.com/content/supplementary/1471-2199-10-1-S1.doc]

#### Additional file 2

Coffee tissue/organ sample set (*Coffea arabica* var. *Mundo Novo* – IAC 388-17-1) used in the present study. Freshly harvested roots, stems, and leaves were obtained from 4 month-old coffee plants grown under greenhouse conditions (28°C, 60% RH) in Campinas, São Paulo, Brazil. Flower and fruit samples, at different developmental stages, were collected from 4–5 year-old field grown plants in Botucatu and in Campinas, São Paulo, Brazil. After harvesting, fresh tissue samples were frozen immediately in liquid nitrogen until RNA extraction.

Click here for file

[http://www.biomedcentral.com/content/supplementary/1471-2199-10-1-S2.doc]

### Acknowledgements

The authors thank Adalgisa Soares de Oliveira, Carla Cristina da Silva and Marcos Brandalise (EMBRAPA/IAC) for providing part of the plant materials used in this study. CFB-C was recipient of a fellowship from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brazil. FES was recipient of a fellowship from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Brazil. This research was sponsored by Consórcio Brasileiro de Pesquisa e Desenvolvimento do Café (CBP&D-Café), Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA) and Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP).

### References

- Barber RD, Harmer DW, Coleman RA, Clark BJ: **GAPDH as a housekeeping gene: analysis of GAPDH mRNA expression in a panel of 72 human tissues.** *Physiol Genomics* 2005, **21(3)**:389-395.
- Gachon C, Mingam A, Charrier B: **Real-time PCR: what relevance to plant studies?** *J Exp Bot* 2004, **55(402)**:1445-1454.
- Bustin SA: **Real-time, fluorescence-based quantitative PCR: a snapshot of current procedures and preferences.** *Expert Rev Mol Diagn* 2005, **5(4)**:493-498.
- Ginzinger DG: **Gene quantification using real-time quantitative PCR: an emerging technology hits the mainstream.** *Exp Hematol* 2002, **30(6)**:503-512.
- Kubista M, Andrade JM, Bengtsson M, Forootan A, Jonak J, Lind K, Sindelka R, Sjoberg R, Sjogreen B, Strombom L: **The real-time polymerase chain reaction.** *Mol Aspects Med* 2006, **27**:95-125.
- Bustin SA, Nolan T: **Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction.** *J Biomol Tech* 2004, **15(3)**:155-166.
- Bustin SA: **Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems.** *J Mol Endocrinol* 2002, **29**:23-39.
- Andersen CL, Jensen JL, Orntoft TF: **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)**:5245-5250.
- Huggett J, Dheda K, Bustin S, Zumla A: **Real-time RT-PCR normalisation; strategies and considerations.** *Genes Immun* 2005, **6(4)**:279-284.
- Schmittgen TD, Zakrajsek BA: **Effect of experimental treatment on housekeeping gene expression: validation by real-time, quantitative RT-PCR.** *J Biochem Biophys Methods* 2000, **46(1-2)**:69-81.
- Gorzelnik K, Janke J, Engeli S, Sharma AM: **Validation of endogenous controls for gene expression studies in human adipocytes and preadipocytes.** *Horm Metab Res* 2001, **33(10)**:625-627.
- von Smolinski D, Leverkushoehe I, von Samson-Himmelstjerna G, Gruber AD: **Impact of formalin-fixation and paraffin-embedding on the ratio between mRNA copy numbers of differently expressed genes.** *Histochem Cell Biol* 2005, **124(2)**:177-188.

13. Bustin SA, Benes V, Nolan T, Pfaffl MW: **Quantitative real-time RT-PCR-a perspective.** *J Mol Endocrinol* 2005, **34**:597-601.
14. Gutierrez L, Mauriat M, Guenin S, Pelloux J, Lefebvre JF, Louvet R, Rusterucci C, Moritz T, Guerinéau 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.
15. Jain M, Nijhawan A, Tyagi AK, Khurana JP: **Validation of housekeeping genes as internal control for studying gene expression in rice by quantitative real-time PCR.** *Biochemical and Biophysical Research Communications* 2006, **345**:646-651.
16. Suzuki T, Higgins PJ, Crawford DR: **Control selection for RNA quantitation.** *Biotechniques* 2000, **29**(2):332-337.
17. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paeppe A, Speleman F: **Accurate normalization of Real-Time quantitative RT-PCR by geometric averaging of multiple internal control genes.** *Genome Biol* 2002, **3**(7):34.
18. Tricarico C, Pinzani P, Bianchi S, Paglierani M, Distante V, Pazzagli M, Bustin SA, Orlando C: **Quantitative real-time reverse transcription polymerase chain reaction: normalization to rRNA or single housekeeping genes is inappropriate for human tissue biopsies.** *Anal Biochem* 2002, **309**(2):293-300.
19. Papini-Terzi FS, Rocha FR, Vêncio RZ, Oliveira KC, Felix Jde M, Vicentini R, Rocha Cde S, Simões AC, Ulian EC, di Mauro SM, da Silva AM, Pereira CA, Menossi M, Souza GM: **Transcription profiling of signal transduction-related genes in sugarcane tissues.** *DNA Res* 2005, **12**(1):27-38.
20. Vicentini R, Sasaki FT, Gimenes MA, Maia IG, Menossi M: **In silico evaluation of the Eucalyptus transcriptome.** *Genetics and Molecular Biology* 2005, **28**:487-495.
21. Ohl F, Jung M, Radonic A, Sachs M, Loening SA, Jung K: **Identification and validation of suitable endogenous reference genes for gene expression studies of human bladder cancer.** *J Urol* 2006, **175**(5):1915-1920.
22. Pombo-Suarez M, Calaza M, Gomez-Reino JJ, Gonzalez A: **Reference genes for normalization of gene expression studies in human osteoarthritic articular cartilage.** *BMC Mol Biol* 2008, **9**:17.
23. Infante C, Matsuoka M, Asensio E, Canavate J, Reith M, Manchado M: **Selection of housekeeping genes for gene expression studies in larvae from flatfish using real-time PCR.** *BMC Mol Biol* 2008, **9**:28.
24. Bonfeld B, Elfving B, Wegener G: **Reference genes for normalization: A study of rat brain tissue.** *Synapse* 2008, **62**(4):302-309.
25. 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.
26. Perez S, Royo L, Astudillo A, Escudero D, Alvarez F, Rodriguez A, Gomez E, Otero J: **Identifying the most suitable endogenous control for determining gene expression in hearts from organ donors.** *BMC Molecular Biology* 2007, **8**(1):14.
27. Foss DL, Baarsch MJ, Murtaugh MP: **Regulation of hypoxanthine phosphoribosyltransferase, glyceraldehyde-3-phosphate dehydrogenase and beta-actin mRNA expression in porcine immune cells and tissues.** *Anim Biotechnol* 1998, **9**(1):67-78.
28. Thellin O, Zorzi W, Lakaye B, De Borman B, Coumans B, Hennen G, Grisar T, Igout A, Heinen E: **Housekeeping genes as internal standards: use and limits.** *J Biotechnol* 1999, **75**(2-3):291-295.
29. Lee PD, Sladek R, Greenwood CM, Hudson TJ: **Control genes and variability: absence of ubiquitous reference transcripts in diverse mammalian expression studies.** *Genome Res* 2002, **12**(2):292-297.
30. Czechowski T, Stitt M, Altmann T, Udvardi MK, Scheible WR: **Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis.** *Plant physiology* 2005, **139**:5-17.
31. Aoki T, Naka H, Katagiri T, Hirono I: **Cloning and characterization of glyceraldehyde-3-phosphate dehydrogenase cDNA of Japanese flounder *Paralichthys olivaceus*.** *Fisheries Sci* 2000, **66**:737-742.
32. Olsvik PA, Lie KK, Jordal AEO, Nilsen TO, Hordvik I: **Evaluation of potential reference genes in real-time RT-PCR studies of Atlantic salmon.** *BMC Mol Biol* 2005, **6**:21.
33. Zhang Z, Hu J: **Development and validation of endogenous reference genes for expression profiling of medaka (*Oryzias latipes*) exposed to endocrine disrupting chemicals by quantitative real-time RT-PCR.** *Toxicol Sci* 2007, **95**(2):356-368.
34. Dheda K, Huggett J, Chang J, Kim L, Bustin S, Johnson M, Rook G, Zumla A: **The implications of using an inappropriate reference gene for real-time reverse transcription PCR data normalization.** *Anal Biochem* 2005, **344**:141-143.
35. Sturzenbaum SR, Kille P: **Control genes in quantitative molecular biological techniques: the variability of invariance.** *Comp Biochem Physiol B Biochem Mol Biol* 2001, **130**(3):281-289.
36. Brunner AM, Yakovlev IA, Strauss SH: **Validating internal controls for quantitative plant gene expression studies.** *BMC Plant Biol* 2004, **4**:14.
37. Dheda K, Huggett JF, Bustin SA, Johnson MA, Rook G, Zumla A: **Validation of housekeeping genes for normalizing RNA expression in real-time PCR.** *Biotechniques* 2004, **37**(1):112-119.
38. Radonic A, Thulke S, Mackay IM, Landt O, Siegert W, Nitsche A: **Guideline to reference gene selection for quantitative real-time PCR.** *Biochem Biophys Res Commun* 2004, **313**(4):856-862.
39. 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**(41):2907-2914.
40. Ding J, Jia J, Yang L, Wen H, Zhang C, Liu W, Zhang D: **Validation of a rice specific gene, sucrose phosphate synthase, used as the endogenous reference gene for qualitative and real-time quantitative PCR detection of transgenes.** *J Agric Food Chem* 2004, **52**(11):3372-3377.
41. Kim BR, Nam HY, Kim SU, Kim SI, Chang YJ: **Normalization of reverse transcription quantitative-PCR with housekeeping genes in rice.** *Biotechnol Lett* 2003, **25**(21):1869-1872.
42. Jian B, Liu B, Bi Y, Hou W, Wu C, Han T: **Validation of internal control for gene expression study in soybean by quantitative real-time PCR.** *BMC Molecular Biology* 2008, **9**(1):59.
43. Libault M, Thibivilliers S, Bilgin D, Radwan O, Benitez M, Clough S, Stacey G: **Identification of four soybean reference genes for gene expression normalization.** *The Plant Genome* 2008, **1**:44-54.
44. De Los Santos-Briones C, Hernández-Sotomayor SMT: **Coffee biotechnology.** *Brazilian Journal of Plant Physiology* 2006, **18**:217-227.
45. **BestKeeper Software** [[http://www.gene-quantification.de/best\\_keeper.html](http://www.gene-quantification.de/best_keeper.html)]
46. 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* 2004, **26**(6):509-515.
47. Singh R, Green MR: **Sequence-specific binding of transfer RNA by glyceraldehyde-3-phosphate dehydrogenase.** *Science* 1993, **259**(5093):365-368.
48. Ishitani R, Sunaga K, Hirano A, Saunders P, Katsube N, Chuang DM: **Evidence that glyceraldehyde-3-phosphate dehydrogenase is involved in age-induced apoptosis in mature cerebellar neurons in culture.** *J Neurochem* 1996, **66**(3):928-935.
49. **NormFinder Software** [<http://www.mdl.dk/publication/snornfinder.htm>]
50. Iskandar HM, Simpson RS, Casu RE, Bonnett GD, Maclean DJ, Manners JM: **Comparison of reference genes for quantitative real-time polymerase chain reaction analysis of gene expression.** *Plant Mol Biol Rep* 2004, **22**:325-337.
51. Glare EM, Divjak M, Bailey MJ, Walters EH: **beta-Actin and GAPDH housekeeping gene expression in asthmatic airways is variable and not suitable for normalising mRNA levels.** *Thorax* 2002, **57**(9):765-770.
52. Savli H, Karadenizli A, Kolayli F, Gundes S, Ozbek U, Vahaboglu H: **Expression stability of six housekeeping genes: A proposal for resistance gene quantification studies of *Pseudomonas aeruginosa* by real-time quantitative RT-PCR.** *J Med Microbiol* 2003, **52**(Pt 5):403-408.
53. Fernandez D, Santos P, Agostini C, Bon M-C, Petitot A-S, C Silva M, Guerra-Guimaraes L, Ribeiro ANA, Argout X, Nicole M: **Coffee (*Coffea arabica* L.) genes early expressed during infection by the rust fungus (*Hemileia vastatrix*).** *Molecular Plant Pathology* 2004, **5**(6):527-536.
54. Ganesh D, Petitot A-S, Silva MC, Alary R, Lecouls A-C, Fernandez D: **Monitoring of the early molecular resistance responses of coffee (*Coffea arabica* L.) to the rust fungus (*Hemileia vastatrix*) using real-time quantitative RT-PCR.** *Plant Science* 2006, **170**(6):1045-1051.

55. Silva MdC, Várzea V, Guerra-Guimarães L, Azinheira HG, Fernandez D, Petitot A-S, Bertrand B, Lashermes P, Nicole M: **Coffee resistance to the main diseases: leaf rust and coffee berry disease.** *Brazilian Journal of Plant Physiology* 2006, **18**:119-147.
56. Petitot A-S, Lecouls A-C, Fernandez D: **Sub-genomic origin and regulation patterns of a duplicated WRKY gene in the allotetraploid species Coffea arabica.** *Tree Genetics & Genomes* 2007, **4**:379-390.
57. Livak KJ, Schmittgen TD: **Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta C_t$  method.** *Methods* 2001, **25**:402-408.
58. Liu W, Saint DA: **A new quantitative method of real time reverse transcription polymerase chain reaction assay based on simulation of polymerase chain reaction kinetics.** *Anal Biochem* 2002, **302**:52-59.
59. Giulietti A, Overbergh L, Valckx D, Decallonne B, Bouillon R, Mathieu C: **An overview of real time quantitative PCR: Applications to quantify cytokine gene expression.** *Methods* 2001, **25**:386-401.
60. Ke L, Chen Z, Yung W: **A reliability test of standard-based quantitative PCR: Exogenous vs. endogenous standards.** *Mol Cell Probes* 2000, **14**:127-135.
61. Calsa T Jr, Figueira A: **Serial analysis of gene expression in sugarcane (*Saccharum* spp.) leaves revealed alternative C4 metabolism and putative antisense transcripts.** *Plant Mol Biol* 2007, **63**(6):745-762.
62. Mason MG, Schmidt S: **Rapid isolation of total RNA and genomic DNA from Hakea actities.** *Funct Plant Biol* 2002, **29**:1013-1016.
63. Vieira LGE, Andrade AC, Colombo CA, Moraes AHdA, Metha Â, Oliveira ACd, Labate CA, Marino CL, Monteiro-Vitorello CdB, Monte DdC, Giglioti E, Kimura ET, Romano E, Kuramae EE, Lemos EGM, Almeida ERP, Jorge EC, Albuquerque EVS, Silva FR, Vinecky F, Sawazaki HE, Dorry HFA, Carrer H, Abreu IN, Batista JAN, Teixeira JB, Kitajima JP, Xavier KG, Lima LM, Camargo LEA, Pereira LFP, Coutinho LL, Lemos MVF, Romano MR, Machado MA, Costa MMC, Grossi-de-Sá MF, Goldman MHS, Ferro MIT, Tinoco MLP, Oliveira MC, Van Sluys M-A, Shimizu MM, Maluf MP, Eira MTS, Guerreiro Filho O, Arruda P, Mazzafera P, Mariani PDSC, Oliveira RLBC, Harakava R, Balbao SF, Tsai SM, Mauro SMZ, Santos SN, Siqueira WJ, Costa GGL, Formighieri EF, Carazzolle MF, Pereira GAG: **Brazilian coffee genome project: an EST-based genomic resource.** *Brazilian Journal of Plant Physiology* 2006, **18**:95-108.
64. **Coffee Genome Project** [<http://www.lge.ibi.unicamp.br/cafe/>]
65. **Solanaceae Genome Project (SOL)** [<http://www.sgn.cornell.edu/content/coffee.pl>]
66. Pfaffl MW: **Quantification strategies in real-time PCR.** *The real-time PCR encyclopaedia A-Z of quantitative PCR* 2004:87-120.
67. Ramakers C, Ruijter JM, Deprez RHL, Moorman AFM: **Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data.** *Neuroscience Letters* 2003, **339**:62-66.
68. Pfaffl MW: **A new mathematical model for relative quantification in real-time RT-PCR.** *Nucleic Acids Res* 2001, **29**:e45.
69. **geNorm Software** [<http://medgen.ugent.be/~jvdesomp/genorm/>]
70. Majerowicz N, Söndahl M: **Induction and differentiation of reproductive buds in Coffea arabica L.** *Brazilian Journal of Plant Physiology* 2005, **17**:247-254.
71. Pezzopane JRM, Pedro Júnior MJ, Thomaziello RA, Camargo MBPd: **Escala para avaliação de estádios fenológicos do cafeeiro arábica.** *Bragantia* 2003, **62**:499-505.
72. Morais H, Caramori PH, Kogushi MS, Ribeiro AMdA: **Escala fenológica detalhada da fase reprodutiva de Coffea arabica.** *Bragantia* 2008, **67**:257-260.
73. De Castro RD, Marraccini P: **Cytology, biochemistry and molecular changes during coffee fruit development.** *Brazilian Journal of Plant Physiology* 2006, **18**:175-199.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
[http://www.biomedcentral.com/info/publishing\\_adv.asp](http://www.biomedcentral.com/info/publishing_adv.asp)

