

# Evaluation of Candidate Reference Genes for Real-Time Quantitative PCR of Plant Samples Using Purified cDNA as Template

Michael A. Phillips · John C. D’Auria · Katrin Luck · Jonathan Gershenzon

Published online: 14 March 2009

© The Author(s) 2009. This article is published with open access at Springerlink.com

**Abstract** Quantitative real-time polymerase chain reaction (qRT-PCR) is a precise method to measure changes in gene transcript level. Accurate quantification requires careful RNA quality assessment, determination of primer efficiency, and selection of an appropriate reference gene. While many experimental procedures for these purposes have been described for mammalian samples, the direct application of these methods to plant samples often introduces unexpected experimental errors due to the complex and variable nature of the ribosomal RNA species present in typical plant extracts. In this paper, we report a simple procedure for the purification and quantification of complementary DNA (cDNA) after reverse transcriptase reactions by microcapillary electrophoresis. The use of purified cDNA allows template concentrations to be more accurately standardized for SYBR Green PCR reactions and increases amplification efficiencies so that these closely resemble those determined by the standard curve method. These advantages facilitate a more precise evaluation of the transcript levels of candidate reference genes under various experimental conditions without bias from differences in reverse transcriptase efficiency, template loading, or the presence of PCR inhibitors following reverse transcription. Using samples from *Arabidopsis thaliana* and *Picea abies*

(Norway spruce), we demonstrate the value of this approach for selecting reference genes.

**Keywords** Quantitative real-time PCR analysis · Gene expression · Microcapillary electrophoresis · *Arabidopsis thaliana* · *Picea abies*

## Introduction

Quantitative real-time polymerase chain reaction (qRT-PCR) is a sensitive analytical technique for measuring changes in gene transcript level. It is composed of two main steps: the reverse transcription of RNA into complementary DNA (cDNA) and the PCR amplification of a portion of the target cDNA monitored in real-time through the fluorescence of double-stranded DNA where fluorescence is assumed to be directly proportional to amplicon formation. Due to the exponential nature of PCR, the number of cycles required for the amplicon-associated fluorescence to rise above an arbitrary threshold level within the linear phase of the reaction (a quantity known as crossing time, threshold cycle, or simply  $C_t$ ) can be used to calculate the starting concentration of the corresponding transcript in a biological sample. qRT-PCR has become the most widely used method for precise measurements of differential gene expression (review in Wong and Medrano 2005), microarray validation (Rajeevan et al. 2001; Ramakrishnan et al. 2002), and single nucleotide polymorphism typing (Johnson et al. 2004) due to its broad linear range (over six orders of magnitude) and unparalleled sensitivity. Its increasing importance in the characterization of transcripts from laser dissected cells (Kerk et al. 2003), genes with limited expression (Czechowski et al. 2004), and other applications has spurred rapid methodological development (Liu and Saint 2002a, b; Bar et al. 2003; Sanchez

---

M. A. Phillips (✉) · J. C. D’Auria · K. Luck · J. Gershenzon  
Department of Biochemistry,  
Max Planck Institute for Chemical Ecology,  
Hans Knöll Strasse 8,  
07745 Jena, Germany  
e-mail: mapgmy@cid.csic.es

### Present address:

M. A. Phillips  
Center for Research in Agricultural Genomics (CRAG-CSIC),  
C/Jordi Girona 18-26,  
08034 Barcelona, Spain

et al. 2004; Mouritzen et al. 2004). Statistical approaches for the analysis of real-time PCR data (Pfaffl et al. 2002; Rutledge and Cote 2003; Nordgard et al. 2006) and design of highly efficient primers (Marshall 2004; Pattyn et al. 2003; Tichopad et al. 2002) have recently extended the power of this technique.

While measurement of the concentration of a messenger RNA (mRNA) species with qRT-PCR can be performed in a relative (Pfaffl 2001) or absolute (Bustin 2000) manner, relative quantification is usually biologically more relevant and easier to perform in practical terms. However, relative quantification requires an additional reference gene transcript to be measured to normalize template loading differences between samples. Two of the best described techniques for relative quantification are the  $2^{-\Delta\Delta C_t}$  method (Livak and Schmittgen 2001) and the efficiency-corrected method (Pfaffl 2001). In the  $2^{-\Delta\Delta C_t}$  method, the efficiency of primer amplification for both the reference gene and the gene-of-interest are assumed to be the same, where efficiency ( $E$ ) is defined as the fold increase in the number of amplicon molecules per cycle, which is maximally 2. However, since this assumption leads to high error rates in some cases, the efficiency-corrected model is generally preferred. Using this latter model, the fold change in the number of amplicon molecules per cycle is calculated as the ratio of efficiencies of the target gene to the reference gene, each raised to the power of the difference in  $C_t$  measured between treatment and control samples for that gene (Pfaffl 2001):

$$FC = E_{\text{target}}^{\Delta C_{t-c}} / E_{\text{ref}}^{\Delta C_{t-c}}$$

where FC is the fold change.

In this way, for reference genes whose expression varies little between treatment and control conditions, the denominator reflects only differences in template loading.

The use of the efficiency-corrected model for relative quantification in qRT-PCR requires accurate knowledge of amplification efficiency, since small imprecisions in efficiency measurement may lead to large errors in fold change estimation. Amplification efficiencies are influenced by a number of factors including the presence of PCR inhibitors carried over from template preparation and the innate chemical properties of primer and template molecules. Efficiency can be determined by use of a standard curve (Pfaffl 2001) or by analysis of the slope of individual amplification curves (Liu and Saint 2002a, b; Ramakers et al. 2003; Rutledge 2004; Tichopad et al. 2003).

Besides establishing efficiency, two other important steps in the validation of qRT-PCR are selection of an appropriate reference gene(s) and assessment of the quality of biological samples. Use of a reference gene whose expression varies under treatment conditions could result in

a large experimental bias. Numerous reports have described methods for selecting optimal reference genes based on analysis of microarray data (Andersen et al. 2004), absolute quantification using synthetic amplicons as an external reference (Tricarico et al. 2002; Dheda et al. 2004), averaging multiple reference genes (Vandesompele et al. 2002), and statistical analysis of reference gene  $C_t$  variance (Pfaffl et al. 2004). However, determining reference gene expression stability under specific treatment conditions is rarely included as an integral part of the qRT-PCR validation process, although Brunner et al. (2004) compared different tissue and organ types in the analysis of variance of reference gene expression in poplar (*Populus trichocarpa*).

As methods for the purification of RNA have improved in recent years, more emphasis has been placed on assessing the quality and quantity of RNA to be used in qRT-PCR. It has been shown that meaningful gene expression measurements can still be obtained from partially degraded samples (Schoor et al. 2003), especially when normalized to an appropriate reference gene (Fleige et al. 2006). UV spectrophotometry (Manchester 1996) and fluorescent dyes such as RiboGreen (Hashimoto et al. 2004) have been frequently used to quantify total RNA samples, while quality assessment has shifted from reliance on denaturing agarose gel electrophoresis (Sambrook et al. 1989; Mannhalter et al. 2000) to microcapillary electrophoresis (Imbeaud et al. 2005; Mueller et al. 2000). However, in both cases, the integrity of mRNA transcripts is inferred from the ratio of the large to small cytosolic ribosomal RNA (rRNA) fragments based on the assumption that an undegraded sample has a ratio of 2.0. Nevertheless, rRNA ratios alone are not a sufficiently universal indicator of mRNA integrity (Auer et al. 2003). Therefore, other sample characteristics that can be observed by microcapillary electrophoresis are also employed to establish RNA sample quality using post-analysis software that computes parameters such as the RNA integrity number (RIN; Imbeaud et al. 2005) and the degradation factor (DegFact; Auer et al. 2003).

While the prolific use of real-time PCR in medical diagnostics and animal research has exploited the 28S/18S ratios of mammalian rRNA to determine RNA amount and purity, plants have different types of rRNA pools (Dyer 1982; Loening and Ingle 1967). Total RNA extracts from plants typically contain rRNAs from cytosolic, plastidic, and mitochondrial pools, and the relative contributions from these organelles may vary considerably depending on the tissue, developmental stage, and metabolic state. In photosynthetic tissue, for example, plastidic rRNAs (23S, 16S, 5S) may be even more abundant than their cytosolic counterparts (25S, 18S, 5.8S, 5S). Due to the RNA population differences between plants and other organisms, it is not certain whether the computation of RIN or DegFact are reliable indicators of plant mRNA quality and concentration.

In this paper, we describe a method for improving qRT-PCR of plant samples by purification of the cDNA formed by reverse transcription prior to carrying out PCR. This additional step has two principal advantages. The use of purified cDNA improves PCR efficiency in many cases, while quantification of the purified cDNA allows template concentration to be standardized on a more accurate basis than quantification of the total RNA prior to reverse transcription. These benefits in turn allow researchers to more precisely evaluate the variable expression of reference genes among organs, developmental stages, or growing conditions and so improve the selection of the most appropriate reference gene for a given qRT-PCR investigation.

## Materials and Methods

**Plant Growth Conditions and Total RNA Extraction** *A. thaliana* ecotype Columbia 0 seeds were sterilized and plated on 0.8% Phytagar media containing 4.3 g/L MS salts (Duchefa, Haarlem, The Netherlands). All plates were cold-treated at 4°C for 3 days and either grown under long day conditions (16-h light, 8-h dark) under artificial light (light grown treatment) or wrapped in two layers of aluminum foil and incubated in complete darkness at 24°C (etiolated treatment). Plants were grown for 2 weeks and then harvested for RNA extraction as whole seedlings or transferred to soil and grown another 4 weeks until the onset of flowering. Plants grown in this way were dissected into leaves, roots, stems, and flowers prior to RNA extraction. Some plants were grown further until the time of seed set and siliques were harvested for RNA extraction.

*P. abies* (Norway spruce) liquid culture was started from an established embryogenic callus culture grown on solid EGM6 media (Bishop-Hurley et al. 2001) before being transferred to EGM6 liquid media and subcultured at 10-day intervals in an incubating shaker (24°C, 100 rpm) in darkness. Cultured cells were harvested for RNA extraction as described previously (Phillips et al. 2007).

For RNA extraction, cultured cells, fresh leaves, bark or whole seedlings was weighed out in 100 mg aliquots in a tared, chilled 2-mL ground glass Tenbroek cell homogenizer and a 450  $\mu$ L aliquot of RLT buffer (containing guanidine thiocyanate; QIAGEN, Hilden, Germany) with 10  $\mu$ L  $\beta$ -mercaptoethanol per milliliter was added prior to homogenization on ice. Total RNA was then purified as described in the Rneasy plant mini kit (QIAGEN) and digested on column for 20 min with Dnase I (QIAGEN). Each purified total RNA extract was immediately frozen and stored at -20°C except for a 2.5- $\mu$ L aliquot. RNA concentration and 260:280 and 260:230 nm ratios were determined spectrophotometrically using 1  $\mu$ L of this aliquot diluted with 49  $\mu$ L pure water. Based on these

measurements, the remainder of the aliquot was diluted to 100 ng/ $\mu$ L and further analyzed on an Agilent 2100 Bioanalyzer (Palo Alto, California, USA) using a RNA 6000 Nano LabChip<sup>®</sup>. When the quantification data from the spectrophotometric analysis and the results from the Bioanalyzer differed by more than 10%, the analysis was repeated; otherwise, the RNA concentration used for subsequent steps was based on the concentration estimated from the Bioanalyzer analysis.

**Reverse Transcription and cDNA Purification** Superscript III was used according to manufacturer's instructions (Invitrogen, Carlsbad, California, USA) with the following exceptions. A 2–3 $\times$  scaled-up reaction was typically used in which a 40–60  $\mu$ L reaction with 10–20  $\mu$ g total RNA and 100 pmol oligo dT anchor primer [d(T<sub>18</sub>V)] was carried out for 2 h at 50°C in a Thermomixer (Eppendorf, Hamburg, Germany) with shaking at 500 rpm after a 2-min denaturation of total RNA at 65°C. Following cDNA synthesis, half the reaction was removed and diluted to 2.5 ng/ $\mu$ L total RNA with pure water, and 5  $\mu$ L Rnase A (10  $\mu$ g/mL) was added to the remainder. The RNA was digested for 30 min at 37°C, followed by removal of another 1- $\mu$ L aliquot. The remainder was purified on a QIAquick PCR column (QIAGEN) as described after dilution of the sample in 10 vol PB buffer (containing guanidine hydrochloride and isopropanol; QIAGEN). Samples were eluted in 30  $\mu$ L, and 1  $\mu$ L of each sample was subsequently analyzed in triplicate on the Bioanalyzer using an RNA 6000 Pico LabChip<sup>®</sup> without further concentration.

**Bioanalyzer Data Acquisition and Analysis** Total RNA samples were analyzed on an Agilent Bioanalyzer 2100 and RNA 6000 Nano Labchip<sup>®</sup> using the Expert Software (Agilent, version B.02.02.SI258) to determine the RIN quality number, concentration, and rRNA ratios. Total RNA integrity was further judged with the DegFact program (version 1.41; Auer et al. 2003). Purified cDNA was analyzed again on the Bioanalyzer using a Pico 6000 LabChip<sup>®</sup>. The smear analysis function of the Expert software was used to integrate single stranded cDNA fragments in the range of 250 bp–4 kb (as judged by the RNA6000 ladder; Ambion, Austin, USA). Based on concentration estimates in this range, cDNA templates were diluted to 50 pg/ $\mu$ L with pure water for qRT-PCR analysis. Purified cDNA samples with significantly different size distributions as judged by their electropherogram peak shapes were excluded from further analysis.

**Quantitative Real-Time PCR and Primer Design** All experiments were performed on a Stratagene Mx3000P (La Jolla, California, USA) using SYBR<sup>®</sup> green I with

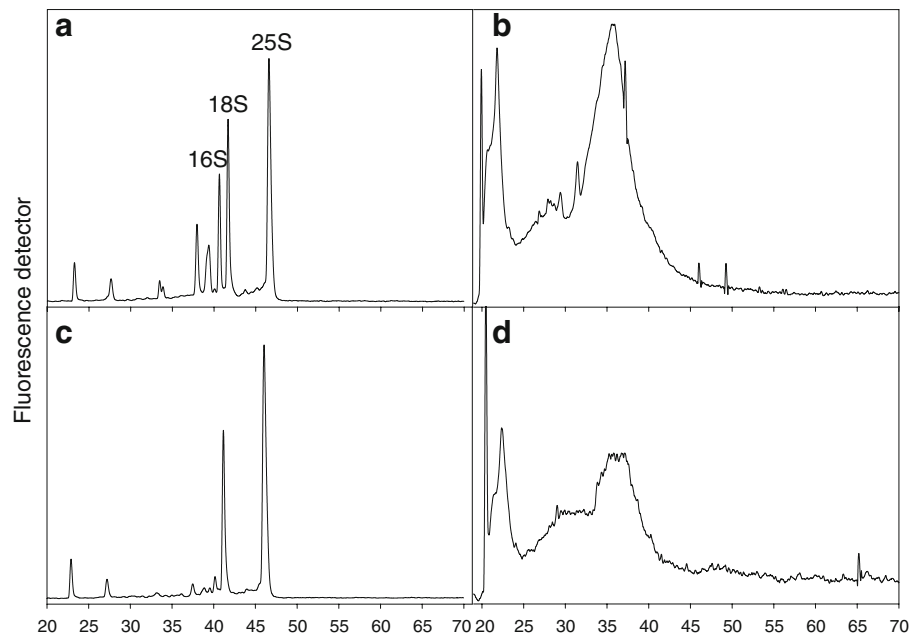
ROX as an internal loading standard. Each 25- $\mu$ L reaction contained either 50 pg purified cDNA or cDNA corresponding to 2.5 ng total RNA. Controls included non-RT controls (where 2.5 ng total RNA without reverse transcription was used to monitor genomic DNA contamination) and non-template controls (water template). PCR thermocycles were run as follows: 96°C denaturation for 6 min followed by 40 cycles of 30 s at 96°C, 30 s at 55°C, and 30 s at 72°C. Fluorescence was read following each annealing and extension phase. All runs were followed by a melting curve analysis. The products of each primer pair were cloned and sequenced at least six times to verify the specificity of the primers. The linear range of template concentration to  $C_t$  value was determined by performing a series of fourfold dilutions (1- to 1,024-fold) using purified cDNA from a minimum of three independent RNA extractions analyzed in three technical replicates. Primer efficiencies for all primer pairs were calculated using the standard curve method (Pfaffl 2001) and individually for each reaction using LinRegPCR (Ramakers et al. 2003). The stability of reference gene expression under different conditions was tested by comparing  $C_t$  values between treatment and control reactions performed with an identical amount of purified cDNA. Descriptive statistics for this comparison were obtained by entering  $C_t$  values from these analyses into the BestKeeper Excel software application (Pfaffl et al. 2004). Treatments for *Arabidopsis* consisted of light-grown seedlings versus etiolated seedlings and comparison of various organs. Reference genes for *A. thaliana* (including sequence numbers; sequences of forward and reverse primers) were actin (At1g49240; Actin28F: GGTAACATTGTGCTCAGTGGTGG; Actin28R: AACGACCTTAATCTTCATGCTGC), desiccation-responsive protein 29 (At5g52310; RD29AF: ATCACTTGGCTCCACTGTTGTTC; RD29AR: ACAAACACACATAAACATC CAAAGTG), adenine phosphoribosyltransferase 1 (At1g27450; APT1F: GTTGCAGGTGTTGAAGCTA GAGGT; APT1R: TGGCACCAATAGCCAACGCAA TAG), RNA polymerase II large subunit (At4g35800; RP21sF: GAAGGCAAAGGAAGGCAGAATCAG; RP21sR: GCAATACTCCACGGAACACCAAG), and ubiquitin (At4g05320; UBQ10F: CACACTTCACTTG GTCTTGCGT; UBQ10R: GTCTTTCCGGTGAGAGT CTTAC). *P. abies* reference genes designed from sequences obtained from a spruce EST project (Ralph et al. 2006) (including sequences of forward and reverse primers) were: ubiquitin (PaUbiF: CGGCAAGCAGTTGGAGGATGG; PaUbiR: CGGAGGACGAGGTGAAGAGTGG), 18S rRNA (Pa18SF: CGGCGGATGTTGCTCTAAG; Pa18SR: TCTGTCAATCCTTACTATGTCTGG), and tubulin (PaTubF: CGTTACCTGCTGCCTGAG, PaTubR: GCTCT GTATTGCTGTGAACC). All primers were designed using BeaconDesigner (version 5.0; PremierBiosoft, Palo Alto,

California, USA) and HPLC-purified (Invitrogen). All other molecule design, sequence analysis, and vector operations were carried out using VectorNTI (version 10, Invitrogen).

## Results

Accurate assessment of the quality and quantity of RNA present in each sample is an essential early step in qRT-PCR analysis. Thus, we first determined the validity of standard methods for estimating RNA quality in plant samples that are based on the ratios of RNA species. Analysis of total RNA by microcapillary electrophoresis revealed striking differences between the rRNA populations in our plant samples and those of typical mammalian cells. Instead of the 28S and 18S rRNA bands common to most mammals, most flowering plants, including the model organism *Arabidopsis thaliana*, have ribosomal bands at 25S and 18S (Dyer 1982) as seen in Fig. 1a, c. Hence, we calculated 25S/18S ratios as a crude measure of RNA quality. But plants also have other rRNA species that can be present at high levels. For example, in plastid-rich tissues (including photosynthetic tissues), plastid-specific rRNAs not present in mammalian cells make up a significant fraction of total RNA (Chang et al. 2005; Loening and Ingle 1967; Dyer 1982). These extra RNA species can be easily distinguished in electropherograms of total RNA extracted from *A. thaliana* seedlings grown in continuous light (Fig. 1a) and darkness (Fig. 1c). Seedlings grown in darkness are etiolated, and the plastids in these cells are metabolically less active and hence contain fewer ribosomes. Consequently, the bulk of total RNA in these samples is derived from the cytosol.

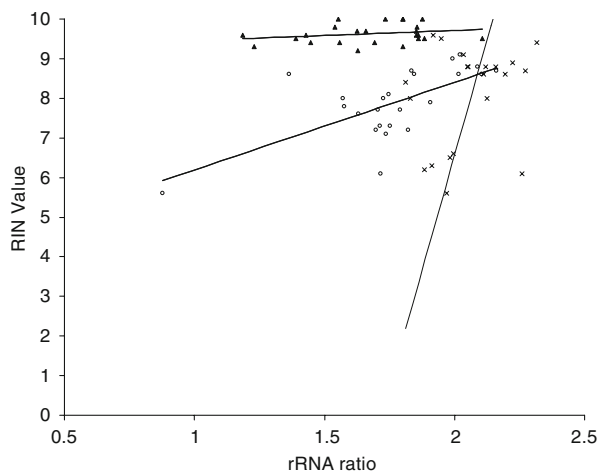
To assess RNA quality in these variable samples, we applied previously reported standards to 24 total RNA samples from light-grown and etiolated *A. thaliana* seedlings, measuring their RIN values (Imbeaud et al. 2005) using the Agilent Expert software and their degradation factors (Auer et al. 2003) with the program DegFact 1.41. Due to the presence of chloroplast rRNA, DegFact was unable to arrive at a degradation factor for the majority of these samples (data not shown) and so rated them as extremely degraded. On the other hand, the calculation of RIN values for each sample was more successful, and these are plotted versus the corresponding 25S/18S ratios in Fig. 2. An additional set of 24 total RNA samples from Norway spruce (*Picea abies*) bark and liquid culture was similarly analyzed. The only statistically significant correlation between RIN value and 25/18S ratio was seen for light-grown *A. thaliana* seedlings ( $r^2 = 0.67$ ,  $p < 0.001$ ). Etiolated seedlings had considerable variation in their calculated RIN values despite 25S/18S ratios which were



**Fig. 1** Total RNA profile of plant tissues as determined by microcapillary electrophoresis can vary significantly depending on growth conditions, but purified cDNA profiles are similar. **a** Total RNA electropherogram from whole, light-grown *Arabidopsis* seedlings (the two large cytosolic rRNA fragments and the 16S plastidic rRNA fragments are indicated). **b** Purified cDNA following reverse

transcription of **a**. **c** Total RNA from whole, dark-grown (etiolated) *Arabidopsis* seedlings. **d** Purified cDNA following reverse transcription of **c**. Samples in **a** and **c** were analyzed on an Agilent Bioanalyzer using an RNA Nano LabChip<sup>®</sup>. Samples in **b** and **d** were analyzed on an RNA Pico LabChip<sup>®</sup>

generally between 1.8 and 2.2, resulting in a low correlation between these parameters ( $r^2 = 0.26$ ,  $p = 0.254$ ). Norway spruce total RNA samples, which were of variable quality based on a range of 25S/18S ratios, consistently scored high RIN values regardless of rRNA ratio ( $r^2 = 0.24$ ,  $p = 0.27$ ).



**Fig. 2** Scatterplot comparing two methods of assessing RNA quality for various plant samples. The 25S/18S rRNA ratio is plotted on the horizontal axis and the RNA integrity (*RIN*) value on the vertical axis. Trend lines were calculated for 24 independent samples each derived from light-grown *Arabidopsis* seedlings (*circle*), etiolated *Arabidopsis* seedlings (*ex*), or Norway spruce cultured cells (*triangle*). Correlation coefficients ( $r^2$ ) are given in the text

Thus, RIN ratings did not prove to be a consistent indicator of plant RNA quality and varied considerably according to species and tissue type.

To circumvent issues of sample assessment based on total RNA characteristics, we next explored the possibility of quantifying the amount of cDNA produced in the reverse transcriptase reactions for standardizing qRT-PCR template concentration. Purified cDNA from various plant species and organs was generated by RNase treatment, and reverse transcriptase reaction components were removed through silica column purification. The resulting purified cDNA was analyzed on a Bioanalyzer Pico RNA LabChip<sup>®</sup>. The electrophoretic trace (Fig. 1b, d) provided information about the concentration and size distribution of cDNA fragments. Degraded RNA carried over from the purification procedure emerged as a distinct low-molecular-weight peak on the left of the electropherogram trace. Using the packaged smear analysis function of the Bioanalyzer Expert software, cDNA fragments in the range of 250bp–4kb were quantified and used as a basis for preparing qRT-PCR templates of identical concentration. In this way, qRT-PCR reactions with a uniform concentration of template could be prepared from photosynthetic (Fig. 1a) or non-photosynthetic (Fig. 1c) tissues even though the original total RNA samples were dominated by highly different rRNA species.

To evaluate whether purifying reverse-transcribed cDNA for qRT-PCR template not only helped standardize template

concentration but also helped make reaction efficiency more uniform, we synthesized cDNA from a number of plant RNA sources and then purified cDNA from half of each reverse transcriptase reaction mixture. The other unpurified half was diluted with water to a similar cDNA concentration, assuming a 2% conversion of total RNA into cDNA. This assumption was based on typical conversion rates estimated from quantifying cDNA (data not shown) which were consistent with previous literature reports (Jacobson 1987; Sambrook et al. 1989; Groppa and Morse 1993). We then compared the two sets of samples as templates in qRT-PCR assays for various genes (Table 1). Individual reaction efficiencies were calculated for both groups by the standard curve method (Pfaffl 2001). For *P. abies*, individual efficiency estimates were consistently higher when purified templates were used, more closely matching efficiency values obtained by the standard curve method. However, for *A. thaliana* samples, there was no statistically significant difference between purified and unpurified cDNA in PCR reaction efficiency. The efficiencies judged by the standard curve method were higher and approached two.

To determine what effect template purification could have on the variance of reference gene  $C_t$  values, qRT-PCR assays were performed with the above samples for several genes that might be considered good reference gene candidates and the results analyzed with the BestKeeper software (Pfaffl et al. 2004). A good reference gene must be stably expressed between control and treatment conditions to serve as a useful internal standard for target gene quantification. This application compares the stability of multiple candidates under control and experimental conditions using externally provided  $C_t$  data. In this case, we have utilized this software to assess the effect of cDNA purification on candidate reference gene evaluation. The results summarized in Table 2 show that of the six reference genes examined in

this way (four in *A. thaliana*, two in *P. abies*), all but one showed lower standard deviations when purified cDNA was used as a template instead of the diluted RT reaction. Only APT1F had a lower standard deviation when the diluted RT template was used, but the difference was small (0.68 versus 0.84  $C_t$  units). In some cases, such as for *P. abies* tubulin primers, the effect of using a purified template was considerable (standard deviation of 2.54  $C_t$  units for diluted RT template versus 0.82  $C_t$  units for purified cDNA template). These differences can be accounted for by the elimination of PCR inhibitors and the greater precision in template standardization afforded from direct quantification of the cDNA template for PCR instead of measurements of total RNA. Thus, the assessment of reference gene candidates can be greatly facilitated by cDNA purification.

The selection and use of reference genes in RT-PCR has been the subject of many recent, innovative approaches. Andersen et al. (2004) described a mathematical model for identifying reference gene candidates from microarray data. However, in order for this method to be broadly useful, microarray data must be available for exactly the same experimental conditions to be examined by the qRT-PCR study, since a reference gene whose expression does not change during one treatment may change under different conditions. Normalization to sample total RNA in combination with an external standard curve constructed from a synthetic amplicon has also been proposed as an alternative (Tricarico et al. 2002). However, because of the variable nature of plant rRNAs from one tissue to another, this approach is generally unsuitable for quantitative PCR in plant systems. Normalization against a panel of housekeeping genes has also been described (Vandesompele et al. 2002). However, the extensive validation requirements and additional cost and preparation time necessitated by the use multiple reference genes make this approach less than desirable in some contexts. In the absence of perfect

**Table 1** Effect of template purification on the efficiency of qRT-PCR amplification

Organism	Gene	Single amplification curves		Standard curve method	
		E (purified cDNA)	E (unpurified cDNA)	<i>p</i>	<i>E</i>
<i>Picea abies</i>	Ubiquitin	1.95 (0.09)	1.88 (0.04)	<0.001	1.99 (0.01)
	Tubulin	1.85 (0.07)	1.80 (0.08)	<0.001	1.99 (0.01)
<i>Arabidopsis thaliana</i>	Ubiquitin	1.90 (0.05)	1.91 (0.05)	>0.2	1.98 (0.01)
	Actin	1.91 (0.03)	1.86 (0.14)	0.117	2.04 (0.01)
	RP2ls	1.77 (0.04)	1.74 (0.15)	>0.2	1.94 (0.02)
	APT1f	1.87 (0.04)	1.85 (0.06)	>0.2	1.96 (0.07)

Individual efficiencies were compared from single amplification curves for nine independent biological samples prepared in triplicate ( $n=27$ ) using the LinRegPCR software (Ramakers et al. 2003). Template used was purified cDNA (see “Materials and Methods”) or a diluted aliquot of the reverse transcription reactions that was not further purified. Mean efficiencies and standard deviations (in parentheses) are shown for candidate reference genes from *A. thaliana* and *P. abies*. Efficiencies were also determined for the same primers using the standard curve method (Pfaffl 2001). Differences in efficiencies estimated from single curves for purified vs. unpurified templates were determined by the two-tailed *t* test. For  $p>0.05$ , there is no statistically significant difference between the  $C_t$  values for these two groups

**Table 2** Effect of template purification on reference gene evaluation

cDNA template	<i>Arabidopsis thaliana</i>								Norway spruce			
	Purified cDNA				Unpurified cDNA				Purified cDNA		Unpurified cDNA	
	Ubiquitin	Actin	rp2ls	apt1F	Ubiquitin	Actin	rp2ls	apt1F	Ubiquitin	Tubulin	Ubiquitin	Tubulin
Geometric mean	18.1	17.7	24.9	22.0	24.7	24.5	31.1	29.4	17.5	17.0	23.3	23.5
Arithmetic mean	18.1	17.7	24.9	22.0	24.7	24.5	31.1	29.4	17.6	17.1	23.5	23.8
Min	17.5	16.9	23.3	20.1	23.2	23.5	29.1	28.4	15.6	15.1	21.9	21.6
Max	19.1	18.3	27.7	23.6	25.6	25.2	39.9	31.8	19.6	18.8	31.2	34.1
SD	0.46	0.32	0.90	0.84	0.55	0.48	1.65	0.68	0.93	0.82	1.65	2.54
CV	2.55	1.83	3.63	3.81	2.22	1.97	5.29	2.31	5.31	4.82	7.05	10.7
Min (fold)	-1.53	-1.67	-2.40	-3.31	-2.52	-1.84	-2.94	-1.85	-3.58	-3.21	-2.50	-3.07
Max (x-fold)	1.86	1.54	5.17	2.70	1.78	1.58	137	4.39	3.99	2.96	138	500
SD ( $\pm$ x-fold)	1.35	1.23	1.79	1.71	1.42	1.36	2.88	1.55	1.87	1.73	3.02	5.46

Variance and descriptive statistics were measured with the BestKeeper Excel-based application (Pfaffl et al. 2004) using reference gene  $C_t$  values. All values shown are  $C_t$  units, unless otherwise indicated. The threshold values used to determine  $C_t$ s in these experiments were on the order of 1,200–1,400 dR (baseline subtracted fluorescence, arbitrary units), corresponding to dRn (dR normalized to reference dye) of 0.11–0.13

sample template standardization, the search for a suitable reference gene becomes a “circular problem” due to the difficulties in obtaining absolute measurements of reference gene expression.

In this investigation, we quantified template levels based on cDNA amount rather than total RNA and then evaluated the suitability of various reference gene candidates on a condition-by-condition basis, assuming that certain candidate reference genes would be better suited for some experimental conditions than others. For an assortment of typical candidate reference genes,  $\Delta C_t$  values for various experimental treatments and organ types of interest were

determined (Table 3). These samples were chosen to address two typical applications in quantitative PCR: to measure target gene expression changes within a single organ during an experimental treatment and among two or more organs under similar environmental conditions. These different experimental questions may require different reference genes. For example, in comparing etiolated vs. light-grown *A. thaliana* seedlings, we found that ubiquitin expression changed the least ( $\Delta C_t=0.75$ ), making it an attractive reference gene for these experimental conditions, whereas *RP2ls* was a rather poor candidate under the same circumstances ( $\Delta C_t=4.08$ ). However, the expression of

**Table 3** Comparison of reference gene expression in *A. thaliana* among different organs and growth conditions using purified cDNA templates

		Ubiquitin	Actin	RD29AF	APT1F	RP2ls
Exp 1	Control conditions	Light-grown				
	Treatment	Dark grown				
	$\Delta C_t$	0.75	1.58	1.36	1.29	4.08
	Fold error	1.62	2.78	2.07	2.25	10.24
Exp 2	Control organ	Leaves				
	Organ A	Roots				
	$\Delta C_t$	0.12	1.70	0.68	1.21	0.42
	Fold error	1.08	2.99	1.44	2.13	1.27
	Organ B	Flowers				
	$\Delta C_t$	0.30	2.68	1.40	1.78	0.02
	Fold error	1.21	5.66	2.12	3.04	1.01
	Organ C	Siliques				
	$\Delta C_t$	3.65	4.23	7.28	1.35	2.72
	Fold error	10.41	15.46	49.58	2.33	4.72
	Organ D	Stems				
	$\Delta C_t$	0.46	1.07	1.16	1.34	0.58
Fold error	1.34	1.99	1.86	2.32	1.39	

$\Delta C_t$  reflects the average  $C_t$  difference in cycles between a control and a treatment conditions or between two organs. Fold error was calculated as  $E^{-\Delta C_t}$ , where  $E$  is the standard curve efficiency for those primers and  $\Delta C_t$  was determined from the average difference in  $C_t$  values for that gene between two conditions or organs

*RP21s* was extremely uniform between leaf and flower, making it an ideal candidate to compare organ-specific gene expression between these two organs.

## Discussion

An essential component of optimal reference gene selection for qRT-PCR is the determination of variation in expression levels across treatments (Brunner et al. 2004; Pfaffl et al. 2004). A stably expressed reference gene should produce constant  $C_t$  values under experimental and control conditions over many independent replicates if the PCR conditions are constant. Variations in  $C_t$  values may arise from several sources, including the presence of PCR inhibitors in the reaction, differences in reverse transcriptase efficiency, tissue matrix specific factors, loading error, pipetting error, and of course natural variation in the expression of that gene. Choosing a stably expressed reference gene is therefore of critical importance to obtaining accurate and reproducible qRT-PCR results using the relative quantification method. However, recent evaluations of reference gene performance (Vandesompele et al. 2002; Pfaffl et al. 2004) have indicated that the use of traditional reference genes, such as GAPDH, ACTIN, and 18S rRNA, may introduce unexpected experimental error in many cases. Here, we have described an improved method for evaluating candidate reference genes using PCR reactions with purified cDNA as a template.

Other approaches for selecting and employing reference genes in a qRT-PCR have been reported in recent years. For example, Vandesompele et al. (2002) concluded that the geometric averaging of multiple reference genes could reduce the effects of single reference gene variation, while Pfaffl et al. (2004) described a method for the statistical evaluation of reference gene stability to select an optimal group of reference genes among a list of candidates (Pfaffl et al. 2004). However, the use of multiple reference genes may be impractical for high throughput applications because of the increase in sample number, cost, and setup. The use of external standard curves using purified PCR product has been described to identify kinetic outliers during the evaluation of reference genes (Bar et al. 2003) through absolute quantification and amplicon copy number calculation, yet this method still does not provide any additional information about the stability of a reference gene under experimental conditions. Normalization of qRT-PCR results to total RNA amount would eliminate the circular problem in verifying reference gene expression stability on an absolute scale, permitting normalization to a single reference gene. However, this method frequently generates unacceptable error because it ignores differences in the mRNA content of total RNA (partic-

ularly prominent in plant samples) and in reverse transcriptase efficiency.

Here, we have demonstrated that normalization of qRT-PCR results of plant samples on the basis of the same amount of purified cDNA is a feasible approach to reference gene evaluation on an absolute scale. The purification and analysis of cDNA by microcapillary electrophoresis using a Bioanalyzer (Agilent) has previously been described for the standardization of fluorescently labeled samples prior to microarray hybridizations (Grissom et al. 2005). However, recent advances in this technique, including the introduction of high sensitivity Pico RNA LabChips<sup>®</sup>, and RNase digestion of the total RNA has made possible the complete purification and quantification of cDNA obtained in reverse transcriptase reactions. By standardizing the amount of template in qRT-PCR reactions, we eliminate the circular problem of relative measurements of reference gene expression and enable the direct comparison of reference gene  $C_t$  values between treatment and control samples. Using this technique, we observed that amplification efficiencies of individual reactions using purified cDNA templates were in many cases higher than reactions using unpurified template at similar cDNA concentrations, possibly due to the removal of PCR inhibitors such as DTT and unused oligo dT primer used in the formation of cDNA. However, this effect seemed to be species-specific, possibly reflecting the abundance of secondary metabolites carried over from the initial RNA extractions. Nonetheless, this technique allowed us to more accurately measure the natural variance of candidate genes by established means (Pfaffl et al. 2004) independent of the contributions of inhibitors or template loading errors.

The evaluation of typical plant reference genes employing this approach revealed innate differences in expression among organs that could compromise their use in qRT-PCR studies. While it was relatively simple to choose a reference gene which was stably expressed under different environmental conditions (such as etiolated versus light-grown *Arabidopsis* seedlings), finding one that was consistently expressed in different organs proved more difficult.

There are some drawbacks to evaluating reference genes for qRT-PCR using purified cDNA as a template following reverse transcriptase reactions. The yield of cDNA purification from silica gel columns is low (approximately 25%). Therefore, a large amount of total RNA is typically required to obtain enough purified cDNA to perform the requisite number of replicate analyses, making analysis of laser-dissected cells or other limited tissues difficult by this method. In addition, the inclusion of additional sample processing steps following the reverse transcriptase reaction and the microcapillary electrophoresis analytical step lead to an increase in sample workup time, which may be impractical when a large number of samples are involved. However, the relatively



similar PCR reaction efficiencies for purified and unpurified templates means that following careful evaluation and selection of reference genes using purified cDNA, relative quantification can probably be performed thereafter without the additional processing steps of cDNA purification. Studies of gene expression in Arabidopsis and Norway spruce liquid cultures are currently being carried out using these methods.

**Acknowledgments** This work was supported by the Max Planck Society and a Humboldt Foundation Fellowship (J.D.). The authors thank the institute greenhouse staff for raising the plants used in this study and Irmgard Seidl-Adams for a critical reading of the manuscript.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

- Andersen CL, Jensen JL, Orntoft TF (2004) Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res* 64:5245–5250 doi:10.1158/0008-5472.CAN-04-0496
- Auer H, Lyianarachchi S, Newsom D, Klisovic MI, Marcucci G, Kornacker K (2003) Chipping away at the chip bias: RNA degradation in microarray analysis. *Nat Genet* 35:292–293 doi:10.1038/ng1203-292
- Bar T, Stahlberg A, Muszta A, Kubista M (2003) Kinetic outlier detection (KOD) in real-time PCR. *Nucleic Acids Res* 31:e105 doi:10.1093/nar/gng106
- Bishop-Hurley SL, Zabkiewicz RJ, Grace L, Gardner RC, Wagner A, Walter C (2001) Conifer genetic engineering: transgenic *Pinus radiata* (D. Don) and *Picea abies* (Karst) plants are resistant to the herbicide Buster. *Plant Cell Rep* 20:235–243 doi:10.1007/s002990100317
- Brunner AM, Yakovlev IA, Strauss SH (2004) Validating internal controls for quantitative plant gene expression studies. *BMC Plant Biol* 4:14 doi:10.1186/1471-2229-4-14
- Bustin SA (2000) Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *J Mol Endocrinol* 25:169–193 doi:10.1677/jme.0.0250169
- Czechowski T, Bari RP, Stitt M, Scheible WR, Udvardi M (2004) Real-time PCR profiling of over 1400 Arabidopsis transcription factors: unprecedented sensitivity reveals novel root- and shoot-specific genes. *Plant J* 38:366–379 doi:10.1111/j.1365-3113X.2004.02051.x
- Chang IF, Szick-Miranda K, Pan SQ, Bailey-Serres J (2005) Proteomic characterization of evolutionarily conserved and variable proteins of Arabidopsis cytosolic ribosomes. *Plant Physiol* 137:848–862 doi:10.1104/pp.104.053637
- Dheda K, Huggett JF, Bustin SA, Johnson MA, Rook G, Zumla A (2004) Validation of housekeeping genes for normalizing RNA expression in real-time PCR. *Biotechniques* 37:112–119
- Dyer TA (1982) Nucleic acids and proteins in plants. In: Partheier P, Boulter D (eds) *Encyclopedia of Plant physiology*, vol 14b. Springer, Berlin, pp 171–191
- Fleige S, Walf V, Huch S, Prgomet C, Sehm J, Pfaffl MW (2006) Comparison of relative mRNA quantification models and the impact of RNA integrity in quantitative real-time RT-PCR. *Biotechnol Lett* 28:1601–1613 doi:10.1007/s10529-006-9127-2
- Grissom SF, Lobenhofer EK, Tucker CJ (2005) A qualitative assessment of direct-labeled cDNA products prior to microarray analysis. *BMC Genomics* 6:36 doi:10.1186/1471-2164-6-36
- Groppe JC, Morse DE (1993) Isolation of full-length RNA templates for reverse transcription from tissues rich in RNase and proteoglycans. *Anal Biochem* 210:337–343
- Hashimoto JG, Beadles-Bohling AS, Wiren KM (2004) Comparison of RiboGreen and 18S rRNA quantitation for normalizing real-time RT-PCR expression analysis. *Biotechniques* 36:54–56 58–60
- Imbeaud S, Gaudens E, Boulanger V, Barlet X, Zaborski P, Eveno E, Mueller O, Schroeder A, Auffray C (2005) Towards standardization of RNA quality assessment using user-independent classifiers of microcapillary electrophoresis traces. *Nucleic Acids Res* 33:e56 doi:10.1093/nar/gni054
- Jacobson A (1987) Purification and fractionation of poly(A)+RNA. *Methods Enzymol* 152:254–261 doi:10.1016/0076-6879(87)52028-6
- Johnson MP, Haupt LM, Griffiths LR (2004) Locked nucleic acid (LNA) single nucleotide polymorphism (SNP) genotype analysis and validation using real-time PCR. *Nucleic Acids Res* 32:E55 doi:10.1093/nar/gnh046
- Kerk NM, Ceserani T, Tausta SL, Sussex IM, Nelson TM (2003) Laser capture microdissection of cells from plant tissues. *Plant Physiol* 132:27–35 doi:10.1104/pp.102.018127
- Liu WH, Saint DA (2002a) A new quantitative method of real time reverse transcription polymerase chain reaction assay based on simulation of polymerase chain reaction kinetics. *Anal Biochem* 302:52–59 doi:10.1006/abio.2001.5530
- Liu WH, Saint DA (2002b) Validation of a quantitative method for real time PCR kinetics. *Biochem Biophys Res Commun* 294:347–353 doi:10.1016/S0006-291X(02)00478-3
- Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>-ΔΔC<sub>T</sub></sup> (T). *Methods* 25:402–408
- Loening UE, Ingle J (1967) Diversity of RNA components in green plant tissues. *Nature* 215:363–367 doi:10.1038/215363a0
- Manchester KL (1996) Use of UV methods for measurement of protein and nucleic acid concentrations. *Biotechniques* 20:968–970
- Mannhalter C, Koizar D, Mitterbauer G (2000) Evaluation of RNA isolation methods and reference genes for RT-PCR analyses of rare target RNA. *Clin Chem Lab Med* 38:171–177 doi:10.1515/CCLM.2000.026
- Marshall OJ (2004) PerlPrimer: cross-platform, graphical primer design for standard, bisulphite and real-time PCR. *Bioinformatics* 20:2471–2472 doi:10.1093/bioinformatics/bth254
- Mouritzen P, Nielsen PS, Jacobsen N, Noerholm M, Lomholt C, Pfundheller HM, Ramsing NB, Kauppinen S, Tolstrup N (2004) The ProbeLibrary (TM)—expression profiling 99% of all human genes using only 90 dual-labeled real-time PCR probes. *Biotechniques* 37:492–495
- Mueller O, Hahnenberger K, Dittmann M, Yee H, Dubrow R, Nagle R, Ilsley D (2000) A microfluidic system for high-speed reproducible DNA sizing and quantitation. *Electrophoresis* 21:128–134 doi:10.1002/(SICI)1522-2683(20000101)21:1<128::AID-ELPS128>3.0.CO;2-M
- Nordgard O, Kvaloy JT, Farmen RK, Heikkila R (2006) Error propagation in relative real-time reverse transcription polymerase chain reaction quantification models: the balance between accuracy and precision. *Anal Biochem* 356:182–193 doi:10.1016/j.ab.2006.06.020
- Pattyn F, Speleman F, De Paepe A, Vandesompele J (2003) RTPrimerDB: the real-time PCR primer and probe database. *Nucleic Acids Res* 31:122–123 doi:10.1093/nar/gkg011

- Pfaffl MW (2001) A new mathematical model for relative quantification in real-time PCR. *Nucleic Acids Res* 29:2002–2007 doi:10.1093/nar/29.9.e45
- Pfaffl MW, Horgan GW, Dempfle L (2002) Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Res* 30:e36 doi:10.1093/nar/30.9.e36
- Pfaffl MW, Tichopad A, Prgomet C, Neuvians TP (2004) Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper—Excel-based tool using pair-wise correlations. *Biotechnol Lett* 26:509–515 doi:10.1023/B:BILE.0000019559.84305.47
- Phillips MA, Walter MH, Ralph S, Dabrowska P, Luck K, Urós EM, Boland W, Strack D, Rodríguez-Concepción M, Bohlmann J, Gershenzon J (2007) Functional identification and differential expression of 1-deoxy-D-xylulose 5-phosphate synthase in induced terpenoid resin formation of Norway spruce (*Picea abies*). *Plant Mol Biol* 65:243–257 doi:10.1007/s11103-007-9212-5
- Rajeevan MS, Vernon SD, Taysavang N, Unger ER (2001) Validation of array-based gene expression profiles by real-time (kinetic) RT-PCR. *J Mol Diagn* 3:26–31
- Ralph SG, Yueh H, Friedmann M, Aeschliman D, Zeznik JA, Nelson CC, Butterfield YSN, Kirkpatrick R, Liu J, Jones SJM, Marra MA, Douglas CJ, Ritland K, Bohlmann J (2006) Conifer defence against insects: microarray gene expression profiling of Sitka spruce (*Picea sitchensis*) induced by mechanical wounding or feeding by spruce budworms (*Choristoneura occidentalis*) or white pine weevils (*Pissodes strobi*) reveals large-scale changes of the host transcriptome. *Plant Cell Environ* 29:1545–1570 doi:10.1111/j.1365-3040.2006.01532.x
- Ramakers C, Ruijter JM, Deprez RHL, Moorman AFM (2003) Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neurosci Lett* 339:62–66 doi:10.1016/S0304-3940(02)01423-4
- Ramakrishnan R, Dorris D, Lublinsky A, Nguyen A, Domanus M, Prokhorova A, Gieser L, Touma E, Lockner R, Tata M, Zhu XM, Patterson M, Shippy R, Sendera TJ, Mazumder A (2002) An assessment of Motorola CodeLink (TM) microarray performance for gene expression profiling applications. *Nucleic Acids Res* 30:e3
- Rutledge RG (2004) Sigmoidal curve-fitting redefines quantitative real-time PCR with the prospective of developing automated high-throughput applications. *Nucleic Acids Res* 32:e178
- Rutledge RG, Cote C (2003) Mathematics of quantitative kinetic PCR and the application of standard curves. *Nucleic Acids Res* 31:e93 doi:10.1093/nar/gng093
- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular cloning: a laboratory manual*, 2nd edn. Cold Spring Harbor Laboratory Press, NY
- Sanchez JA, Pierce KE, Rice JE, Wangh LJ (2004) Linear-after-the-exponential (LATE)-PCR: an advanced method of asymmetric PCR and its uses in quantitative real-time analysis. *Proc Natl Acad Sci U S A* 101:1933–8 doi:10.1073/pnas.0305476101
- Schoor O, Weinschenk T, Hennenlotter J, Corvin S, Stenzl A, Rammensee HG, Stevanovic S (2003) Moderate degradation does not preclude microarray analysis of small amounts of RNA. *Biotechniques* 35:1192–1201
- Tichopad A, Dilger M, Schwarz G, Pfaffl MW (2003) Standardized determination of real-time PCR efficiency from a single reaction set-up. *Nucleic Acids Res* 31:e122 doi:10.1093/nar/gng122
- Tichopad A, Dzidic A, Pfaffl MW (2002) Improving quantitative real-time RT-PCR reproducibility by boosting primer-linked amplification efficiency. *Biotechnol Lett* 24:2053–2056 doi:10.1023/A:1021319421153
- Tricarico C, Pinzani P, Bianchi S, Paglierani M, Distanti V, Pazzagli M, Bustin SA, Orlando C (2002) Quantitative real-time reverse transcription polymerase chain reaction: normalization to rRNA or single housekeeping genes is inappropriate for human tissue biopsies. *Anal Biochem* 309:293–300 doi:10.1016/S0003-2697(02)00311-1
- Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 3:research0034
- Wong ML, Medrano JF (2005) Real-time PCR for mRNA quantitation. *Biotechniques* 39:75–85 doi:10.2144/05391RV01