The lack of a systematic validation of reference genes: a serious pitfall undervalued in reverse transcription-polymerase chain reaction (RT-PCR) analysis in plants

Laurent Gutierrez[1,2,*], Mélanie Mauriat[1], Stéphanie Guénin[3], Jérôme Pelloux[2], Jean-François Lefebvre[3], Romain Louvet[2], Christine Rusterucci[2], Thomas Moritz[1], François Guerineau[2], Catherine Bellini[1,4] and Olivier Van Wuytswinkel[2]

[1]Umeå Plant Science Centre, Department of Forest Genetics and Plant Physiology, Swedish University of Agricultural Sciences, 901 83 Umeå, Sweden
[2]Groupe de Génomique Fonctionnelle des Plantes EA3900, Université de Picardie Jules Verne, Faculté des Sciences, 33 Rue St Leu, 80039 Amiens, France
[3]Centre de Ressources Régionales en Biologie Moléculaire, Université de Picardie Jules Verne, Faculté des Sciences, 33 Rue St Leu, 80039 Amiens, France
[4]Laboratoire de Biologie Cellulaire, Institut National de la Recherche Agronomique, 78026 Versailles Cedex, France

Summary

Reverse transcription-polymerase chain reaction (RT-PCR) approaches have been used in a large proportion of transcriptome analyses published to date. The accuracy of the results obtained by this method strongly depends on accurate transcript normalization using stably expressed genes, known as references. Statistical algorithms have been developed recently to help validate reference genes, and most studies of gene expression in mammals, yeast and bacteria now include such validation. Surprisingly, this important approach is under-utilized in plant studies, where putative housekeeping genes tend to be used as references without any appropriate validation. Using quantitative RT-PCR, the expression stability of several genes commonly used as references was tested in various tissues of *Arabidopsis thaliana* and hybrid aspen (*Populus tremula* × *Populus tremuloides*). It was found that the expression of most of these genes was unstable, indicating that their use as references is inappropriate. The major impact of the use of such inappropriate references on the results obtained by RT-PCR is demonstrated in this study. Using aspen as a model, evidence is presented indicating that no gene can act as a universal reference, implying the need for a systematic validation of reference genes. For the first time, the extent to which the lack of a systematic validation of reference genes is a stumbling block to the reliability of results obtained by RT-PCR in plants is clearly shown.

Introduction

Transcript quantification, by Northern blotting or reverse transcription-polymerase chain reaction (RT-PCR) analysis, requires the use of stably expressed genes, known as references, for accurate normalization of mRNA fractions, as several variables (including the amount of starting material, reverse transcriptase efficiency and overall differences in transcriptional activity between tissues) can affect the strength of the detected signals. In Northern blotting, the quantity of total RNA loaded is commonly estimated by ethidium bromide staining or hybridization of rRNAs with a radiolabelled probe. These controls are not very accurate, and thus do not provide reliable indications of the consistency of the mRNA loadings in each lane. However, as Northern blotting is used to assess large variations in expression levels, errors caused by approximate normalization appear to be acceptable as they will not affect dramatically the expression trends of target genes. In contrast, in microarray analyses, which are also based on hybridization techniques, a wide range of probes and targets are used in a single experiment. This makes the approach a very good tool for monitoring global changes in gene expression, and allows the normalization of each experiment using a range of statistical analyses of the global hybridization signal (Quackenbush, 2002) without the need for reference genes. The approach is very different from

RT-PCR analysis, in which the expression level of each target gene is normalized to the expression of a reference gene that is presumed to be representative of the cDNA concentration in each sample, and subject to the same errors during cDNA preparation as the target gene(s) (Bustin *et al.*, 2005).

A major advantage of RT-PCR is that it can detect very low quantities of the target transcript and can be used to assess discrete variations in expression. The development of real-time technologies, allowing quantitative RT-PCR, has provided complementary tools to microarrays, with consequent improvements in both sensitivity and specificity (Nolan *et al.*, 2006). However, quantitative RT-PCR has only recently been used to monitor gene expression in plants (Gachon *et al.*, 2004), and remains underused, considering its ability to discriminate between the expression of closely related genes and to quantify transcript levels of very weakly expressed genes (Czechowski *et al.*, 2004). This increase in accuracy emphasizes the importance of the use of highly reliable reference genes. For this purpose, statistical algorithms, including the geNorm software used in this study, have been developed recently to identify the best reference genes for use under given experimental conditions (Pfaffl *et al.*, 2002, 2004; Vandesompele *et al.*, 2002). Most quantitative RT-PCR analyses of gene expression in mammals, yeast and bacteria include such evaluations using geNorm software (http://medgen.ugent.be/~jvdesomp/genorm/citations.php).

Surprisingly, the validation of reference genes in plants has received very little attention to date. However, in 2005, clear evidence was presented that several genes are more stably expressed than the commonly used references in *Arabidopsis* (Czechowski *et al.*, 2005). In the study by Czechowski *et al.* (2005), a list of new candidates for use as reference genes was suggested. Surprisingly, citation reports indicate that the importance of these results does not appear to have been recognized. Indeed, over the past 2 years, putative house-keeping genes have continued to be widely used as reference genes for RT-PCR analysis in plants, although they are certainly not the best options. During the past 6 months, more than 95% of RT-PCR analyses published in *The Plant Cell*, *The Plant Journal* and *Plant Physiology* have used improperly validated references (M. Mauriat and L. Gutierrez, unpubl. data). The choice of such genes as references is arbitrary and inappropriate for several reasons. First, their status as 'housekeeping' genes is generally based on unpublished Northern blot or histochemical analysis, methods known to be mainly qualitative. Consequently, transcript normalization using such genes is not consistent with the high accuracy associated with RT-PCR. Second, these genes are often used as references in experimental conditions that differ from those in which their expression stability was tested. As it has been shown that the expression of such housekeeping genes, although constant in some experimental conditions, can vary considerably in other cases (Volkov *et al.*, 2003; Czechowski *et al.*, 2005; Nicot *et al.*, 2005), their systematic use as reference genes may lead to the misinterpretation of results. The arbitrary choice of reference gene can be further illustrated by the frequent use of rRNAs for normalization, although they have been shown to be inappropriate references (Spanakis, 1993; Johnson *et al.*, 1995; Warner, 1999; Hansen *et al.*, 2001; Solanas *et al.*, 2001; Tyler *et al.*, 2004).

The aim of this study was to demonstrate clearly the extent to which commonly used reference genes in plants can be variably expressed and can introduce a major bias in results obtained by RT-PCR analyses. In addition, our aim was to demonstrate the inability to validate reference genes accurately by wide-scale gene expression studies, because no gene can act as a universal reference, and therefore to highlight the importance of a systematic validation for each set of experimental conditions. For this purpose, 14 genes were chosen from those most often used as references, and their stability of expression was tested. Their expression levels were analysed in 16 different *Arabidopsis* tissues using quantitative RT-PCR. It was shown that some are clearly not stably expressed, and should be avoided as references in future quantifications based on similar experiments. The impact of the use of such unstable references for gene expression studies was further demonstrated by quantifying the expression of a given target transcript using appropriate and inappropriate reference genes for normalization. The non-universal status of reference genes was highlighted by studying the expression stability of hybrid aspen genes that are putative orthologues of stably expressed genes in *Arabidopsis*. Focusing on the aspen model at the cellular level, it was shown that the systematic validation of reference genes using the geNorm algorithm is an easy and robust solution, allowing great improvement in the accuracy of RT-PCR analysis.

Results and discussion

High variability in the distribution of 'reference' mRNA populations during *Arabidopsis* development

In this study, the expression levels of 14 genes were assessed in 16 types of *Arabidopsis* sample representing various organs in various developmental stages (see 'Experimental procedures'). These genes, selected because of their common use as references in RT-PCR analyses, were *ACT2*, *ACT7*, *ACT8*, *APT1*, *eF1α*, *eIF4A*, *TUB2*, *TUB6*, *TUB9*, *UBQ4*, *UBQ5*, *UBQ10*,

Table 1 Primer sequences and amplicon characteristics for selected *Arabidopsis* genes

Symbol	Gene name	Locus	Forward primer	Reverse primer	Amplicon	
					L (bp)	T_m (°C)
ACT2	Actin 2	At3g18780	AATTACCCGATGGGC A	TCATACTCGGCCTTGGA	367	85.10
ACT7	Actin 7	At5g09810	ACCACTACCGCAGAAC	GCTCATACGGTCAGCA	343	84.75
ACT8	Actin 8	At1 g49240	GAATTACCCGACGGAC A	ACGGTCTGCAATACCT	216	84.15
ACT1	Adenine phosphoribosyltransferase 1	At1 g27450	GTTGAATGTGCTTGCG	CTTTAGCCCCTGTTGG	348	81.10
EF1α	Elongation factor 1-α	At5g60390	TCCAGCTAAGGGTGCC	GGTGGGTACTCGGAGA	264	85.35
eIF4A	Eukaryotic translation initiation factor 4A-1	At3g13920	ACTGACCTCTTAGCTCG	CAGATCGGCCACGTTC	231	84.20
NDUFA8	NADH-ubiquinone oxidoreductase 19-kDa subunit	At5g18800	ACACAAACGAGTTTGATCTG	GTGAGTTCCAAAGACAACTA	219	79.55
TUB2	Tubulin β-2/β-3 chain	At5g62690	CATGGTGGGTTTTGCT	ACTGTCGATTTCACGTTG	261	87.50
TUB6	Tubulin β-6 chain	At5g12250	GAAGTCMGCGTCTGTG	ACTCGCCTTCGTCATC	260	84.55
TUB9	Tubulin β-9 chain	At4g20890	AGTGTCCCTGAGCTAAC	AGT GGGAGCTATATCGC	225	83.40
UBQ4	Polyubiquitin	At5g20620	AGGAAGGGATTCCACC	AGTCCGACCATCCTCA	296	83.90
UBQ5	Ubiquitin extension protein	At3g62250	GACGCTTCATCTCGTCC	CCACAGGTT GCGTTAG	186	83.55
UBQ10	Polyubiquitin	At4g05320	ATCACCCTTGAAGTGGA	GAAACCACCACGAAGAC	194	85.90
UBQ11	Polyubiquitin	At4g05050	GAACCAAGTTCATGTATCGT	ACACTCATCAAACTAAGCAC	219	78.60

L, length (in base pairs); T_m, melting temperature.
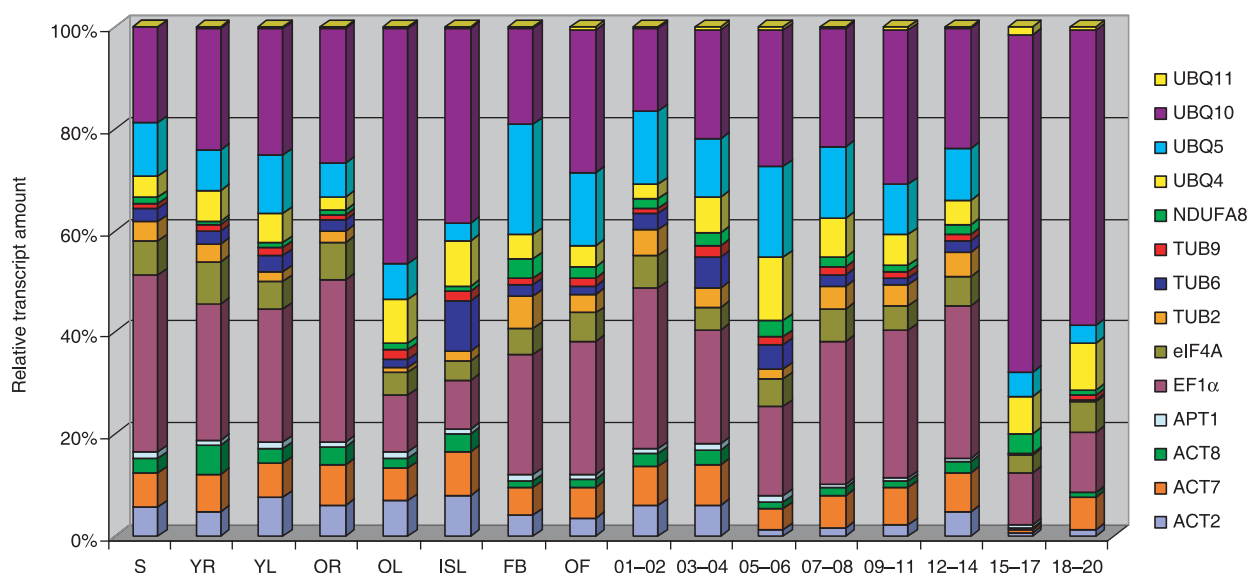

Figure 1 Distribution of transcript populations of 14 selected genes during *Arabidopsis* development. In each developmental stage, the transcript amount of each gene is shown as a ratio relative to the sum of the 14 transcript populations. This representation clearly shows that most of the ratios fluctuate independently of the others, indicating that these 14 genes do not all exhibit stable expression. FB, floral bud; ISL, inflorescence; OF, open flower; OL, old leaf; OR, old root; S, seedling; YL, young leaf; YR, young root; 01–02 to 18–20, siliques (days after flowering). Table 1 gives full names of genes.

UBQ11 and *NDUFA8* (see Table 1 for their full names). A flax (*Linum usitatissimum*) orthologue of the last gene has recently been shown to be a valuable reference gene during seed development (Gutierrez *et al.*, 2006). Figure 1 shows the variations in the relative quantities of the 14 mRNAs during *Arabidopsis* development. Transcript quantities are shown, for each developmental stage, as ratios relative to

the sum of the 14 transcript populations. High variability in the distribution of mRNA populations was found during *Arabidopsis* development. For example, *UBQ10* mRNA represented less than 20% of the total mRNA population in seedlings, but more than 60% in siliques, 18–20 days after flowering (DAF), whereas the relative amount of *eF1α* mRNA declined from ~40% to ~15% in these organs. In addition,
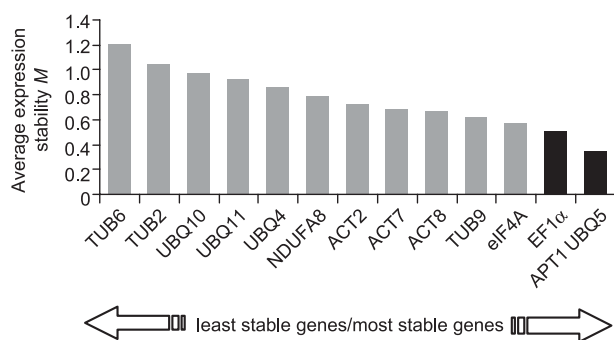
Figure 2 geNorm ranking of expression stability of 14 genes during *Arabidopsis* development. Only three genes (black bars) of the 14 commonly used as references are valid (i.e. *M* < 0.5) under our developmental conditions. Table 1 gives full names of genes.

although it can be seen that the proportion of *ACT7* mRNA remained relatively constant in all organs, except siliques at 15–17 DAF, the proportions of *TUB6* mRNA varied widely. These results clearly suggest that the expression level of each mRNA population is not constant, but, instead, varies in unique temporal and/or spatial patterns. However, the detected variation of each transcript is not absolute, but relative to the sum of the 14 transcript populations, which is not necessarily constant. Statistical analysis of these data is therefore required to evaluate the stability of the expression of each gene under the tested conditions.

Assessment of gene expression stability using statistical analysis by geNorm

Non-normalized expression levels (see 'Experimental procedures') were analysed using geNorm software (Vandesompele *et al.*, 2002), which determines, for each gene, the pair-wise variation with all other genes as the standard deviation of their logarithmically transformed expression ratios. *M*, the measure of the gene expression stability, is the average pair-wise variation of a particular gene compared with that of all other genes. Genes with the lowest *M* values have the most stable expression. Figure 2 shows the ranking of the genes tested, according to their expression stability in our *Arabidopsis* samples. *APT1* and *UBQ5* were the most stably expressed genes, whereas the expression of *TUB6* was quite variable. In addition, *M* values > 0.5 were obtained for 11 of the 14 tested genes (in grey in Figure 2), implying that most of these 'classic' references are far from stably expressed, and cannot be considered as good reference genes under the conditions examined here (Vandesompele *et al.*, 2002). These results are unsurprising because, as mentioned above, it has already been reported that several genes have higher

expression stability (i.e. lower *M* values) than many 'classic' references (Czechowski *et al.*, 2005). More surprisingly, over the past 2 years, little research has taken account of these results, and the so-called 'classic' reference genes are still commonly used for the normalization of RT-PCR data in plants. Therefore, the effect of their use on the accuracy of transcript quantification needs to be tested.

Impact of the use of inappropriate reference genes on the expression analysis of a target gene

Ranking genes with respect to their stability of expression and claiming that some commonly used reference genes should not be employed for transcript normalization may appear to be an extreme response. However, if a given reference gene has not been properly validated, depending on its use, data on the expression of the target gene(s) may be interpreted incorrectly. The expression pattern of the target gene used here, *eIF4A*, over the course of silique development was analysed using the most stably expressed (*APT1*, *UBQ5*, *eF1α*) and least stably expressed (*TUB6*) genes (Figure 3a). When *APT1*, *UBQ5* and *eF1α* were used for normalization, *eIF4A* appeared to be stably expressed, with only a slight increase in its transcript levels towards the last stage of silique development (18–20 DAF). As *APT1*, *UBQ5*, *eF1α* and *eIF4A* are involved in different, independently regulated, cellular functions, the possibility of the co-regulation of these genes can be excluded. Therefore, the fact that there are only slight differences between the three *eIF4A* expression patterns obtained using *APT1*, *UBQ5* and *eF1α* suggests that they constitute reliable, stably expressed references under our experimental conditions. These findings validate the previous geNorm analysis (Figure 2). In contrast, *eIF4A* expression appears to be quite variable when *TUB6* is used for normalization, highlighting the significant impact of normalization based on this inappropriate reference (Figure 3a).

To confirm this result, the expression of an independent target gene, *At5g02840* (MYB factor), was analysed, which is thought to be stably expressed during silique development (Gutierrez *et al.*, 2006). As expected, normalization of *At5g02840* using *APT1*, *UBQ5* and *eF1α* resulted in a relatively similar constant expression pattern (Figure 3b). In contrast, when *TUB6* was used as the reference gene, the difference between the lowest and highest recorded *At5g02840* expression levels was approximately 100-fold.

With regard to *eIF4A* (Figure 3a), the slight differences between the three *At5g02840* expression patterns obtained using *APT1*, *UBQ5* and *eF1α* as the reference genes
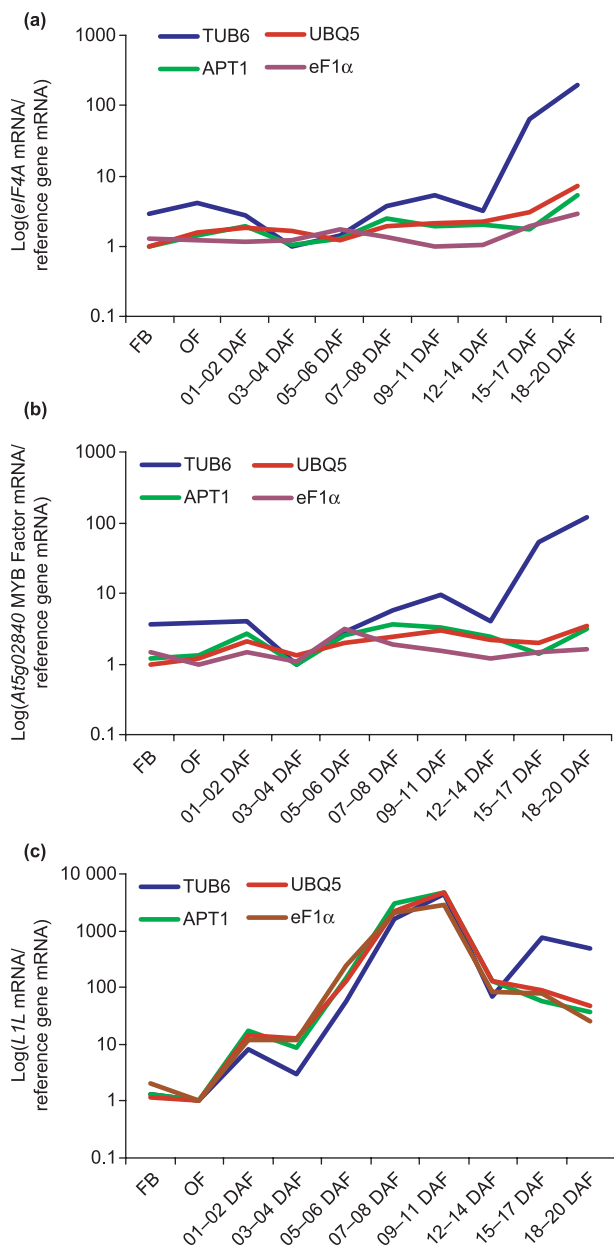
Figure 3 Transcript quantifications during *Arabidopsis* silique development using four different reference genes for normalization. Transcript quantification of *eIF4A* (a), *At5g02840* MYB factor (b) and *LEAFY COTYLEDON1-Like* (*L1L*) (c) using *TUB6*, *APT1*, *UBQ5* or *eF1α* for normalization. Although normalization using *APT1*, *UBQ5* or *eF1α* produced similar expression patterns for each gene, normalization using *TUB6* resulted in a strong bias, dramatically increasing the expression rates during the last stages of silique development. Standard error values are between 4% and 11%; small, uninformative error bars are not shown. FB, floral bud; OF, open flower; 01–02 DAF to 18–20 DAF, siliques days after flowering. Table 1 gives full names of genes.

demonstrate that they constitute reliable, stably expressed references under these experimental conditions (Figure 3b). As the three expression patterns followed the same trend, our results confirm the fairly constant expression of the

At5g02840 MYB factor during silique development. However, the slight differences between the three expression patterns raise questions about the possibility of accurately quantifying weak transcript level variations. Indeed, a difference of approximately twofold was recorded between expression patterns obtained using *APT1* and *UBQ5* (Figure 3a,b). In addition, a difference of approximately 3.5-fold was observed between the expression pattern obtained using *eF1α* and those obtained using *APT1* and *UBQ5* (Figure 3a,b), confirming that *eF1α* is less stably expressed than *APT1* and *UBQ5* (Figure 2). In the literature, differences of less than threefold have frequently been considered to represent significant changes in the expression of a target gene. On the basis of our results, such weak variations could be related simply to the choice of the reference gene used for normalization. It is therefore important to be cautious when interpreting such weak expression changes.

With regard to *eIF4A* (Figure 3a), there was a strong apparent increase in the expression of the *At5g02840* MYB factor during the last stage of silique development when *TUB6* was used for normalization (Figure 3b). The maximum expression level occurred at 18–20 DAF, and was 100 times greater than the minimum expression level recorded at 3–4 DAF. This result clearly demonstrates how easy it is to detect an erroneously strong specific expression pattern for a given target gene simply by using an inappropriate reference gene. The normalization of either *eIF4A* or the *At5g02840* MYB factor with *TUB6* resulted in the same specific expression pattern during silique development (Figure 3a,b), whereas normalization with the other reference genes indicated that it could be considered to be constitutive. The *TUB6*-based pattern is clearly an artefact, caused by the great variation in *TUB6* expression, which is high at 3–4 DAF and low at 18–20 DAF (Figure 1).

To test whether the choice of reference gene could influence the interpretation of the expression pattern of a target gene with a well-described, specific expression profile, the same approach was used to quantify the expression of *LEAFY COTYLEDON1-Like* (*L1L*). This gene is strongly expressed during seed development from heart to bent-cotyledon stages (Kwong *et al.*, 2003). As expected, *L1L* transcripts peaked from 5 to 11 DAF, and very similar expression patterns were identified when using *APT1*, *UBQ5* and *eF1α* (Figure 3c). In contrast, normalization using *TUB6* introduced a clear bias in the quantification of *L1L* expression at 5–6 DAF and from 15 DAF onwards. These results show definitively that the use of a non-statistically validated reference gene can dramatically affect the normalization, and thus introduce errors when determining the expression pattern of a target gene.

Table 2 Primer sequences and amplicon characteristics for selected hybrid aspen genes

Locus	Arabidopsis orthologue	Forward primer	Reverse primer	Amplicon	
				L (bp)	T_{m} (°C)
estExt_fgenesh4_pg.C_LG_ll1155	At4g33380	GCGGAAAGAAAAACTGCAAG	TGACAGCACAGCCCAATAAG	126	83.5
estExt_fgenesh4_pm.C_LG_IX0344	At4g34270	GCTGCACTTGCATCAAAAGA	GCAACTTGGCATGACTCTCA	119	81.5
estExt_fgenesh1_pg_v1.C_LG_ll1615	*ANT*	TCTGTCTGTTATGCCCCTCA	CCACCTAGGAAGTCCTCCAGT	119	84.5

L, length (in base pairs); T_{m}, melting temperature.

The validity of a reference gene is non-universal and is highly dependent on the experimental conditions

Although the validation of reference genes has received very little attention in plant studies, it is easy, convenient and feasible for every species for which even limited genomic resources are available. We wished to determine whether a reference validated in one plant could also be valid for a different species. A good starting point was the data included in Czechowski *et al.* (2005). This report was the first to present a list of stably expressed *Arabidopsis* genes (never used as references) for a large range of experimental conditions. In order to assess the validity of the use of such appropriate reference genes in a different plant species, the expression stability of four genes in hybrid aspen (*Populus tremula × Populus tremuloides*) in eight different developmental states (see 'Experimental procedures') was tested. The genes tested were two aspen putative orthologues (see Table 2) of *At4g34270* and *At4g33380* (the most stably expressed genes in *Arabidopsis* according to Czechowski *et al.*, 2005), one aspen putative orthologue of *UBQ* (stably expressed in poplar according to Brunner *et al.*, 2004) and the aspen rRNA *18S*, which is commonly used as a reference in plants. *At4g34270* and *At4g33380* aspen orthologues exhibited higher expression stability than *UBQ* and much more so than *18S* (Figure 4a). The last two genes could not be considered as appropriate reference genes under the conditions applied here, as they had *M* values > 0.5 (in grey in Figure 4a). These results indicate the possible conservation of gene expression stability across different plant species, and seem to validate the database presented in Czechowski *et al.* (2005) as a good tool for finding stably expressed genes in other plant species. Our interest was subsequently focused on finding an appropriate reference gene that could be used at the cellular level, particularly within the aspen cambial region (see 'Experimental procedures'). Surprisingly, under these new experimental conditions, the initial gene ranking (Figure 4a) had to be modified, as *At4g34270* and *UBQ* were the most stably expressed genes throughout the phloem, xylem and
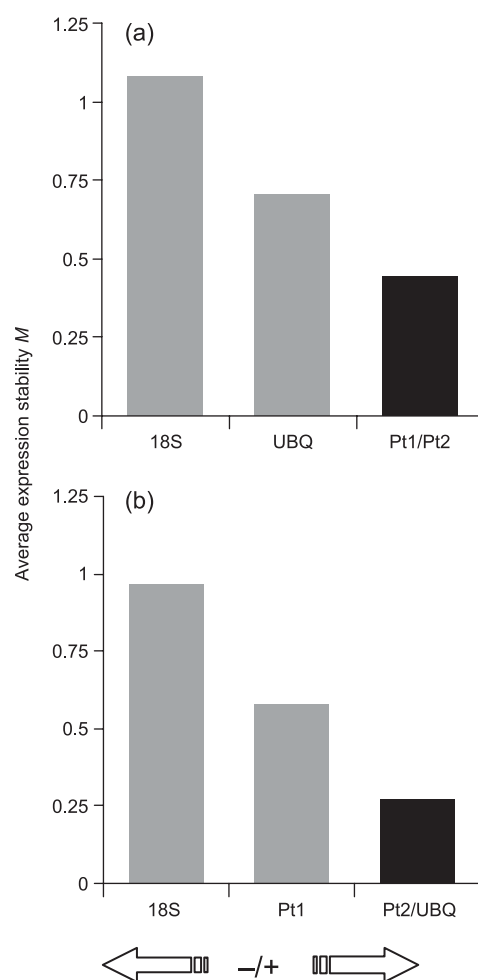

Figure 4 Expression stability of four aspen genes fluctuates according to developmental stage. (a) Ranking by geNorm of gene expression stability at eight different aspen developmental states. (b) Ranking by geNorm of gene expression stability within the aspen cambial region. See 'Experimental procedures' for description of aspen tissues harvested in each experiment. The geNorm software ranks genes from least (–) to most (+) stably expressed. In order to simplify the figure, the *At4g33380* aspen orthologue is identified as *Pt1* and *At4g34270* as *Pt2*.

cambial regions, whereas *At4g33380* and *18S* aspen orthologues exhibited higher expression variability (Figure 4b). Under this second set of experimental conditions, the last two genes could not be considered as appropriate reference

genes, as they had *M* values > 0.5 (in grey in Figure 4b). These results clearly show that, even if correlations can be identified between the expression stability of orthologous genes, reference genes are not universal, either among plant species or tissues. This demonstrates the need to validate reference genes for each set of experimental conditions before using them to quantify changes in the expression of a particular target gene. Data obtained by wide-scale gene expression analyses can be used as a starting point to choose candidates, but they should not replace a systematic validation of reference genes.

A systematic validation of reference genes by geNorm restores RT-PCR accuracy

In order to test the robustness of the previous assessment of gene expression stability by geNorm (Figure 4), the expression pattern of the *AINTEGUMENTA* (*ANT*) aspen orthologue was analysed within the aspen cambial region (Figure 5). This gene is very highly expressed in the cambium (Schrader *et al.*, 2004; R. Bhalerao and A. Karlberg, Umea Plant Science Centre, unpubl. data), and more precisely peaks in the cambial stem cells. An accurate quantification of *ANT* aspen orthologue transcripts would allow the identification of the section containing these cell layers from a series of sections, and therefore the use of gene expression as a spatial marker for RT-PCR analysis throughout the aspen cambial region. When the *18S* and *At4g33380* aspen orthologues were used as references, the *ANT* aspen orthologue expression level increased from section 4 to section 7, corresponding to the cambium, but it was not possible to clearly determine which one of the sections contained the cambial stem cells (Figure 5a). However, when the *UBQ* and *At4g34270* aspen orthologues were used for normalization, they gave similar but more precise patterns and allowed the restriction of the peak of expression of the *ANT* aspen orthologue to section 5, identifying, by this means, the exact position of the cambial stem cells (Figure 5b). This example demonstrates the robustness of a systematic validation of reference genes using the geNorm algorithm, and its capacity to restore the accuracy of RT-PCR analysis.

Conclusions

The use of putative housekeeping genes is acceptable for non-quantitative techniques when qualitative changes are being measured. These RNAs, expressed at relatively high levels in all cells, make ideal positive controls to determine whether the gene of interest is switched off or on. However, the advent of RT-PCR and real-time technology has placed
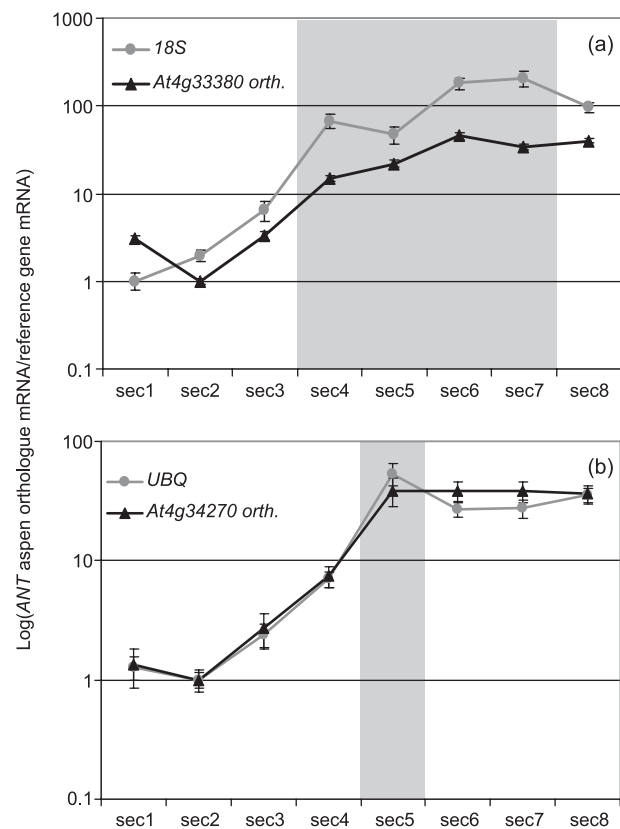

Figure 5 Transcript quantification of *AINTEGUMENTA* (*ANT*) aspen orthologue within the aspen cambial region using four different reference genes for normalization. Expression of the *ANT* aspen orthologue increases in the cambium and specifically peaks in the cambial stem cells. The peak of expression is not specific to one section, but is located within sections 4–7 (in grey) when *ANT* aspen orthologue expression is normalized using *18S* or *At4g33380* aspen orthologues (a). When appropriate reference genes are used, *UBQ* or *At4g34270* aspen orthologues, the peak is exactly localized in section 5 (in grey), allowing this section to be identified as containing cambial stem cells (b). Error bars: ± standard error.

the emphasis on quantitative changes, and should have prompted a re-evaluation of the use of these reference genes (Bustin *et al.*, 2005). This has not been the case in plant analyses, in which arbitrarily chosen 'classic' reference genes have continued to be used for this purpose.

In this study, it has been demonstrated that the use of inappropriate references can dramatically change the interpretation of the expression pattern of a given target gene, and thus introduce flaws in our understanding of the role of such a gene. Therefore, the lack of validation of reference genes in plants questions the accuracy of some results obtained by RT-PCR using such non-valid references. It has been shown that, instead of a universal validation using wide-range expression data, the systematic validation of reference genes by geNorm is the answer. It is easy and

robust, allowing this major pitfall to be overcome in plants. In addition, as slight differences may still be observed in the apparent expression patterns of a target gene obtained using different appropriate validated reference genes, weak apparent changes in gene expression should be treated with caution. In such cases, it may be wise to use at least two geNorm-validated reference genes involved in distinct cellular functions to confirm the weak expression changes.

Experimental procedures

Plant material

Arabidopsis thaliana plants (Columbia ecotype) were grown under light photoperiods of 16 h with day/night temperatures of 22/18 °C. Entire, 7-day-old *Arabidopsis* seedlings (S), 3-week-old leaves (young leaves), 3-week-old roots (young roots), 6-week-old leaves (old leaves), 6-week-old roots (old roots), 6-week-old stems (inflorescence), floral buds (FB), open flowers (OF) and siliques 1–2, 3–4, 5–6, 7–8, 9–11, 12–14, 15–17 and 18–20 DAF were harvested. To determine silique age, individual flowers were tagged with coloured threads on the day they opened.

Hybrid aspen (*Populus tremula* L. × *Populus tremuloides* Michx. Clone T89) were grown under light photoperiods of 18 h at a constant temperature of 18 °C. Vegetative tissues corresponding to eight different developmental states were harvested: leaves and internodes at three different levels on the plant (i.e. three different developmental stages), the shoot apex and roots of a 3-month-old hybrid aspen. Longitudinal tangential cryosections (3 × 15 mm) were also taken across the wood-forming tissues of a 12-year-old aspen tree (*Populus tremula*) growing in northern Sweden, as described in Uggla *et al.* (2001). A set of eight samples characteristic of the different cell types within the cambial region were collected as described in Israelsson *et al.* (2005). All tissues were immediately frozen in liquid nitrogen and stored at –80 °C.

RNA extraction and cDNA synthesis

For *Arabidopsis*, RNA extraction and cDNA synthesis were performed for two biological replicates, as described in Gutierrez *et al.* (2006). All cDNA samples were tested by PCR using specific primers flanking an intron sequence to confirm the absence of genomic DNA contamination (Louvet *et al.*, 2006). For aspen, RNA extractions from the different organs of two biological replicates were performed using an Aurum Total RNA Kit (Bio-Rad, Hercules, CA, USA), following the manufacturer's instructions. RNA extractions from the wood cryosections were performed as described in Israelsson *et al.* (2005). cDNA synthesis was performed using an iScript cDNA Synthesis Kit (Bio-Rad).

Quantitative RT-PCR analysis in *Arabidopsis*

Specific primers were designed for reference gene coding sequences (see Table 1) using LC Probe Design© software (Roche, Basle, Switzerland). Each primer sequence was chosen because it bore no similarity to any data in *Arabidopsis* databases available via the BLASTn program (Altschul *et al.*, 1997). An annealing temperature of 60 °C was used for all primers. Table 1 lists the primer sequences designed for each gene and amplicon characteristics. To ensure that the fluorescence signal was derived from the single intended amplicon, a melting curve analysis was added to each PCR programme, and each PCR product was run on a 4% agarose gel. Real-time PCR was performed in a Roche LightCycler® using a FastStart DNA MasterPLUS SYBR Green I Kit (Roche), according to the manufacturer's recommended protocol. For each gene, quantitative PCR was performed in triplicate. Crossing threshold (CT) values, the number of PCR cycles for which the accumulated fluorescence signal in each reaction crosses a threshold above the background, were acquired by LightCycler® software 3.5 (Roche) using the second derivative maximum method. CT values are a function of the amplification efficiency of PCR, which was found to be between 1.92 and 1.99 depending on the primer pairs. These data were exported into RelQuant© software (Roche), which provides efficiency-corrected normalized quantification results. The results are expressed as the target/reference ratio of the sample, and are therefore corrected for sample heterogeneities and detection-derived variation. The efficiency-corrected quantification performed by RelQuant© is based on relative standard curves describing the PCR efficiencies of each primer pair (Larionov *et al.*, 2005). The relative standard curves were determined and used for each analysis.

Quantitative RT-PCR analysis in aspen

The BLAST program of the JGI database (DoE Joint Genome Institute and Poplar Genome Consortium; http://genome.jgi-psf.org/Poptr1/Poptr1.home.html) was used to identify estExt_fgenesh4_pm.C_LG_IX0344, estExt_fgenesh4_pg.C_LG_II1155 and estExt_fgenesh1_pg_v1.C_LG_II1615, the putative aspen orthologues of *At4g34270*, *At4g33380* and *ANT Arabidopsis* genes, respectively. *At4g34270* and *At4g33380* are two genes found to be stably expressed in *Arabidopsis* (Czechowski *et al.*, 2005). Using these *Populus trichocarpa* coding sequences, primers were designed using Primer 3 software. Primer sequences and amplicon characteristics are shown in Table 2. In addition, specific primers for *18S* (Israelsson *et al.*, 2005) and *UBQ* (Brunner *et al.*, 2004) were used. Quantitative RT-PCRs, with an annealing temperature of 57 °C, were performed in triplicate in a 96-well plate in a Bio-Rad iCycler iQ™ Real-Time PCR Detection System (Bio-Rad) using iQ™ SYBR Green Supermix (Bio-Rad), according to the manufacturer's recommended protocol. The PCR efficiency was found to be between 1.90 and 2.13, depending on the primer pairs, and the CT values for each sample were acquired by iCycler iQ™ software 3.0 (Bio-Rad).

Calculation of the non-normalized expression levels used in Figure 1 and statistical analysis of gene expression stability

For each gene in each set of replicate samples, the non-normalized expression level was calculated using $X/E^{(\text{mean CT})}$, where E is the PCR efficiency, 'mean CT' is the average of the three CT values, one from each replicate, and X is the number of amplicon molecules obtained at CT. The fluorescence threshold used for CT evaluation was constant for all quantitative PCR analyses, so that all CT values are related to the same amount of DNA that could be quantified by an arbitrary

constant *C*. *X* was calculated from *C/L*, where *L* is the amplicon length in base pairs. *X* is therefore expressed relative to the number of amplicon molecules at CT. $X/E^{(\text{mean CT})}$ accurately evaluates the amount of target cDNA (i.e. the non-normalized expression level), because this term includes the PCR efficiency (which is different for each amplicon) and the amplicon length. This is not the case in other classical formulae, for example $2^{(40 - \text{mean CT})}$. The ratios of each transcript population shown in Figure 1 were obtained using $X/E^{(\text{mean CT})}$ by calculating the relative amount of cDNA as a function of the sum of the 14 cDNA populations for each developmental stage. The *M* values of gene expression stability were evaluated by analysing each non-normalized expression level using the geNorm algorithm (http://medgen.ugent.be/~jvdesomp/genorm/), according to Vande-sompele *et al*. (2002).

Acknowledgements

References

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. **25**, 3389–3402.

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

Bustin, S.A., Benes, V., Nolan, T. and Pfaffl, M.W. (2005) Quantitative real-time RT-PCR – a perspective. *J. Mol. Endocrinol*. **34**, 597–601.

Czechowski, T., Bari, R.P., Stitt, M., Scheible, W.R. and Udvardi, M.K. (2004) Real-time RT-PCR profiling of over 1400 Arabidopsis transcription factors: unprecedented sensitivity reveals novel root- and shoot-specific genes. *Plant J*. **38**, 366–379.

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

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

Gutierrez, L., Conejero, G., Castelain, M., Guenin, S., Verdeil, J.L., Thomasset, B. and Van Wuytswinkel, O. (2006) Identification of new gene expression regulators specifically expressed during plant seed maturation. *J. Exp. Bot*. **57**, 1919–1932.

Hansen, M.C., Nielsen, A.K., Molin, S., Hammer, K. and Kilstrup, M. (2001) Changes in rRNA levels during stress invalidates results

from mRNA blotting: fluorescence in situ rRNA hybridization permits renormalization for estimation of cellular mRNA levels. *J. Bacteriol*. **183**, 4747–4751.

Israelsson, M., Sundberg, B. and Moritz, T. (2005) Tissue-specific localization of gibberellins and expression of gibberellin-biosynthetic and signaling genes in wood-forming tissues in aspen. *Plant J*. **44**, 494–504.

Johnson, M.L., Redmer, D.A. and Reynolds, L.P. (1995) Quantification of lane-to-lane loading of poly(A) RNA using a biotinylated oligo(dT) probe and chemiluminescent detection. *Biotechniques*, **19**, 712–715.

Kwong, R.W., Bui, A.Q., Lee, H., Kwong, L.W., Fischer, R.L., Goldberg, R.B. and Harada, J.J. (2003) Leafy cotyledon1-like defines a class of regulators essential for embryo development. *Plant Cell*, **15**, 5–18.

Larionov, A., Krause, A. and Miller, W. (2005) A standard curve based method for relative real time PCR data processing. *BMC Bioinformatics*, **6**, 62.

Louvet, R., Cavel, E., Gutierrez, L., Guenin, S., Roger, D., Gillet, F., Guerineau, F. and Pelloux, J. (2006) Comprehensive expression profiling of the pectin methylesterase gene family during silique development in *Arabidopsis thaliana*. *Planta*, **224**, 782–791.

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

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

Pfaffl, M.W., Horgan, G.W. and 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.

Pfaffl, M.W., Tichopad, A., Prgomet, C. and Neuvians, T.P. (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.

Quackenbush, J. (2002) Microarray data normalization and transformation. *Nat. Genet*. **32** (Suppl.), 496–501.

Schrader, J., Nilsson, J., Mellerowicz, E., Berglund, A., Nilsson, P., Hertzberg, M. and Sandberg, G. (2004) A high-resolution transcript profile across the wood-forming meristem of poplar identifies potential regulators of cambial stem cell identity. *Plant Cell*, **16**, 2278–2292.

Solanas, M., Moral, R. and Escrich, E. (2001) Unsuitability of using ribosomal RNA as loading control for Northern blot analyses related to the imbalance between messenger and ribosomal RNA content in rat mammary tumors. *Anal. Biochem*. **288**, 99–102.

Spanakis, E. (1993) Problems related to the interpretation of autoradiographic data on gene expression using common constitutive transcripts as controls. *Nucleic Acids Res*. **21**, 3809–3819.

Tyler, L., Thomas, S.G., Hu, J., Dill, A., Alonso, J.M., Ecker, J.R. and Sun, T.P. (2004) Della proteins and gibberellin-regulated seed germination and floral development in Arabidopsis. *Plant Physiol*. **135**, 1008–1019.

Uggla, C., Magel, E., Moritz, T. and Sundberg, B. (2001) Function and dynamics of auxin and carbohydrates during earlywood/latewood transition in Scots pine. *Plant Physiol*. **125**, 2029–2039.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N.,

De Paepe, A. and Speleman, F. (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol*. **3**, RESEARCH0034.

Volkov, R.A., Panchuk, I.I. and Schoffl, F. (2003) Heat-stress-dependency and developmental modulation of gene expression: the potential of house-keeping genes as internal standards in mRNA expression profiling using real-time RT-PCR. *J. Exp. Bot*. **54**, 2343–2349.

Warner, J.R. (1999) The economics of ribosome biosynthesis in yeast. *Trends Biochem. Sci*. **24**, 437–440.