TECHNICAL ADVANCE

Identification and Validation of Reference Genes for Normalization of Transcripts from Virus-Infected *Arabidopsis thaliana*

S. T. Lilly,^{1,2} R. S. M. Drummond,¹ M. N. Pearson,² and R. M. MacDiarmid¹

¹The New Zealand Institute for Plant and Food Research Limited, Auckland, New Zealand; ²School of Biological Sciences, University of Auckland, Auckland, New Zealand

Submitted 11 October 2010. Accepted 12 November 2010.

Real-time quantitative polymerase chain reaction (qPCR) of complementary DNA is now a standard method for studies of gene expression. However, qPCR can identify genuine variation only when transcript quantities are accurately normalized to an appropriate reference. To identify the most reliable reference genes for transcript quantification by qPCR, we describe a systematic evaluation of candidate reference genes of *Arabidopsis thaliana* ecotype Columbia-0 (Col-0). Twelve genes were selected for transcript stability studies by qPCR of complementary DNA prepared from *Arabidopsis* leaf tissue infected with one of five plant viruses (*Cauliflower mosaic virus*, *Tobacco mosaic virus*, *Tomato spotted wilt virus*, *Turnip mosaic virus*, and *Turnip yellow mosaic virus*). The F-box family protein, elongation factor 1- α , sand family protein, and protodermal factor 2 gene transcripts showed the most stable accumulation, whereas a traditionally used reference gene, *Actin8*, showed the least stable accumulation as measured by the geNorm algorithm. The data furnish plant virologists with reference genes for normalization of qPCR-derived gene expression in virus-infected *Arabidopsis* and will be beneficial to the selection and design of primers targeting orthologous genes in other plant species.

Real-time quantitative polymerase chain reaction (qPCR) using complementary DNA (cDNA) as template has enabled the rapid detection of very low quantities of target transcripts and the ability to assess variations in accumulation. However, with improved sensitivity and throughput come greater demands on experimental design. In particular, the accurate quantification of transcript abundance obtained by qPCR analysis is vitally dependent on the use of stably expressed reference genes to normalize for sample-to-sample variation, variation in RNA integrity, reverse-transcriptase efficiency, and variation in cDNA sample loading (Bustin et al. 2009).

A number of reference genes were identified for Northern analysis of *Arabidopsis* gene expression: *Actin2* and -8 (*ACT2* and *ACT8*), β -actin, β -tubulin 2 and 6 (*TUB2* and -6), glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*), elongation factor 1 α (*EF1 α*), and polyubiquitin 10 (*UBQ10*). These genes, frequently referred to as “housekeeping” genes, are required for cell survival and were assumed to undergo little or no vari-

ation in expression (Thellin et al. 1999, 2009). Unfortunately, the transcript levels of many traditionally used housekeeping genes are not always stable, with transcript accumulation being constant under some conditions and highly variable under others (Czechowski et al. 2005; Glare et al. 2002; Lee et al. 2002; Selvey et al. 2001). The ribosomal RNAs 18S and 28S (18S or 28S rRNA) have also been used (Xue et al. 2009). However, ribosomal subunit transcription is affected by both endogenous and exogenous factors, including biotic stress. They are not polyadenylated and, therefore, cannot be compared with cDNA that is derived from reverse transcribed RNA using oligo-dT primers (Huggett et al. 2005).

Normalization of data against unstable references can result in incorrect conclusions and, ultimately, compromised studies of transcript accumulation (Bustin 2010; Czechowski et al. 2005; Guénin et al. 2009; Gutierrez et al. 2008a and b; Huggett et al. 2005). The use of appropriate reference genes is of particular importance when using qPCR of cDNA to measure gene transcript abundance in the context of viral infection. Viruses typically commandeer and, thus, interfere with host cell pathways, the components of which are often encoded by traditional reference genes (Watson et al. 2007). Moreover, different viruses are likely to manipulate different transcription pathways, and the extent to which these pathways are perturbed is likely dependent on the timing of the virus infection and cell type infected. It is unlikely that any cellular gene maintains a constant steady state of transcript accumulation across all conditions and, as such, it is considered nearly impossible to identify a single reference gene for use in qPCR studies (Andersen et al. 2004; Faccioli et al. 2007; Klatte and Bauer 2009; Lee et al. 2002) let alone across a range of virus infections. Accordingly, careful consideration must be given when choosing genes as references for normalization, and the reliability of a panel of potential genes should be determined for each novel set of experimental conditions. Ideally, at least the best three of these genes, from independent pathways, would be used for normalization (Guénin et al. 2009; Klatte and Bauer 2009; Vandesompele et al. 2002). To date, five criteria have been outlined that define an appropriate reference gene as one that is i) stably accumulating across all analyzed samples and unaffected by any experimental treatment; ii) not associated with any pseudogene so as to avoid amplification of genomic DNA; iii) able, on amplification, to reflect variations in RNA quality, quantity, or cDNA synthesis; iv) equivalent in accumulation to that of the target gene transcript; and v) accumulating to moderate amounts (i.e., a threshold cycle [C_q] of 15 to 30) (Wan et al. 2009).

Corresponding author: R. S. M. MacDiarmid;
E-mail: robin.macdiarmid@plantandfood.co.nz

Validating candidate reference genes presents a circular problem because the stability of a reference gene's transcript abundance must be derived by comparison with other reference genes. Statistical algorithms for processing raw qPCR data have been developed to identify the best reference genes to use under given experimental conditions. Vandesompele and associates (2002) developed an algorithm, geNorm, which uses pairwise comparisons and geometric averaging across a series of reference genes and biological samples. Briefly, this algorithm is based on the principle that the expression ratio of two ideal reference genes should be identical in all samples regardless of changes in experimental conditions or cell or tissue type. Any variation suggests instability of either one or both candidate genes. The calculated gene stability value (*M*) is expressed as the average pairwise variation of the log-transformed standard deviation of a given gene transcript accumulation ratio compared with the ratio of all other genes investigated. Genes with the lowest *M* value have the most stable expression, and stepwise exclusion of genes with the highest *M* value enables ranking of genes in order of increasing stability (Czechowski et al. 2005; Guénin et al. 2009; Gutierrez et al. 2008a and b; Vandesompele et al. 2002).

Validating reference genes has received little attention in plant studies despite evidence of the unreliability of conventionally used *Arabidopsis* housekeeping genes for normalization. Recently, two groups demonstrated that several *Arabidopsis* genes are more stably expressed than the traditionally used housekeeping genes (Czechowski et al. 2005; Gutierrez et al. 2008a and b). Another study has highlighted a lack of rigor in plant science publications, where 96.8% of reference genes used in 188 studies were merely putatively stably expressed (Guénin et al. 2009). A few studies have been carried out to determine reliable reference genes in plant cells across various plant species, developmental stages, and root and shoot abiotic and biotic stresses, including inoculation with *Pseudomonas syringae* and *Phytophthora infestans* (Czechowski et al. 2005; Expósito-Rodríguez et al. 2008;

Guénin et al. 2009; Gutierrez et al. 2008a; Reid et al. 2006; Wan et al. 2009). Additionally, reliable reference genes have been determined from mammalian and insect studies (De Ketelaere et al. 2006; Watson et al. 2007; Xue et al. 2009) and plant genes homologous to these might be considered for validation as reference genes.

However, to the best of our knowledge, no detailed analysis of reference genes for qPCR of cDNA has been carried out in *Arabidopsis* plants subjected to diverse viral infections despite *Arabidopsis* becoming a more commonly used model to study plant-virus interactions (Carr and Whitham 2007). We aimed to identify reference genes for qPCR gene expression experiments of cDNA from *Arabidopsis* leaf tissue either uninfected or infected with one of five viruses across an infection time course.

RESULTS

Initial selection of candidate reference genes.

The candidate reference genes investigated were selected from two sources; first, the most stable gene transcripts published in studies of transcript accumulation in plants undergoing a range of abiotic and biotic stresses (Table 1), and second, those genes that had been investigated from any host-virus transcriptome microarray data or had previously been used as reference genes for qPCR of cDNA from virus-infected plants (Table 2). Data from microarray experiments were particularly useful because they compare a large number of gene transcripts within a single experiment, thus approaching the ideal of analyzing gene expression data for all expressed genes of an organism from as many different organs and experimental conditions as possible (Czechowski et al. 2005). As such, the publicly available microarray datasets of *Arabidopsis* (Affymetrix ATH1), *Glycine max* (Affymetrix GeneChip), and *Medicago truncatula* (Affymetrix mtr) were a rich source of information for previous authors to select potential qPCR reference genes (Czechowski et al. 2005; Kakar et al. 2008; Libault et al. 2008). Authors of studies investigating the species with no microarray data chose genes from

Table 1. geNorm ranking by *M* value of candidate reference genes from quantitative polymerase chain reaction (qPCR)-validated microarray data from literature^a

Plant species	Validation ^c	Ranking of candidate references from total number of genes investigated in each published study (no.) ^b														No.
		Actin8	APT1	CAC	CYTc2	EF-1α	F-Box	HEL	PDF2	PP2A1	PPR	SAND	TATA	TIP41	YLS8	
<i>Arabidopsis thaliana</i>	Affymetrix ATH1, qPCR ^d	20	...	5	...	21	9	17	6	8	14	1	...	2	7	21
<i>Lycopersicon esculentum</i>	qPCR only ^e	...	7	1=	...	10	...	...	...	...	...	4	3	1=	...	11
<i>Vitis vinifera</i>	qPCR only ^f	8	...	1=	...	7	...	...	...	10	...	5	...	1=	...	14
<i>Gossypium hirsutum</i>	qPCR only ^g	3	...	4	...	6	7	...	...	1	...	...	...	...	...	9
<i>Petunia hybrida</i>	qPCR only ^h	8	...	...	...	4	...	...	...	...	...	2	...	...	...	9
<i>Glycine max</i>	Affymetrix GeneChip, qPCR ⁱ	13	...	...	...	...	2	...	...	...	...	...	...	...	...	15
<i>Lolium perenne</i>	SAGE tags and qPCR ^j	...	...	...	...	1	...	...	...	...	...	...	2	...	...	6
<i>Solanum tuberosum</i>	qPCR only ^k	8	...	...	...	1	...	...	...	...	...	...	...	...	...	8
<i>Medicago truncatula</i>	Affymetrix MTR and qPCR ^l	...	...	...	...	...	...	7	7	...	...	...	...	...	...	8

^a *RDR1* and *RDR6* are not considered candidate reference genes and are omitted from this table.

^b Number of genes validated by qPCR in the referenced publication.

^c Microarray or qPCR validation. Array incorporated complementary DNA (cDNA) from development and shoot, root, hormone, and nutrient stress samples; and qPCR assays incorporated cDNA from development and steady state leaf, shoot, root, and floral tissues.

^d Czechowski and associates (2005).

^e Expósito-Rodríguez and associates (2008).

^f Reid and associates (2006).

^g Artico and associates (2010).

^h Mallona and associates (2010).

ⁱ Libault and associates (2008).

^j Lee and associates (2008). SAGE = serial analysis of gene expression.

^k Nicot and associates (2005).

^l Kakar and associates (2008).

those reported by Czechowski and associates (2005) to validate by qPCR and assess stability of transcript accumulation under their specific conditions (Artico et al. 2010; Expósito-Rodríguez et al. 2008; Lee et al. 2008; Mallona et al. 2010; Nicot et al. 2005; Reid et al. 2006). From this published information, a shortlist of 14 candidate reference genes was selected, including transcripts that were commonly reported as stable by geNorm analysis (Table 1).

Exclusion of genes affected by viral infection.

To determine whether each candidate reference gene had been reported to show transcript changes in the presence of virus infection, published virus transcriptome microarray data were reviewed for changes in their transcript abundance (Alfenas-Zerbini et al. 2009; Ascencio-Ibáñez et al. 2008; Babu et al. 2008; Baebler et al. 2009; Watson et al. 2007; Whitham et al. 2003; Yang et al. 2007) (Table 2). Genes reported for human studies of virus infection, including human immunodeficiency virus and herpes viruses, were also considered where orthologous genes were present in *Arabidopsis* (Watson et al. 2007) (Table 2).

Actin8 and *EF1α* are traditional housekeeping genes used as reference transcripts for qPCR analyses and were included as candidate reference genes despite evidence of increased transcript accumulation under *Plum pox virus* (PPV) infection compared with mock infection of *Arabidopsis* (threefold and fivefold, respectively) (Babu et al. 2008). Cytochrome C2 (*CYTc-2*) was excluded from this study because it showed a very large variation (threefold decrease to a fourfold increase) in response to seven different plant viruses as assessed by *Arabidopsis* Affymetrix microarrays (Ascencio-Ibáñez et al. 2008; Babu et al. 2008; Whitham et al. 2003) and in response to one virus in tomato (Alfenas-Zerbini et al. 2009). Yellow stripe like 8 (*YLS8*) was excluded, albeit due to a small flux (0.7-fold increase) in transcript accumulation in response to *Cabbage leaf curl virus* (CaLCuV) as assessed by microarray analysis of virus-infected *Arabidopsis* plants (Ascencio-Ibáñez et al. 2008). Omitting *CYTc-2* and *YLS8* resulted in a shortlist of 12 candidate reference genes.

Establishing cDNA and primer pairs for qPCR.

Only high-quality RNA that met the following criteria was used in further experiments: i) free of genomic DNA (gDNA) contamination as assessed by PCR, ii) pure as assessed by light absorbance (ratio of 1.8 to 2.0 at 260 or 280 nm) (Nano-drop), and iii) intact as determined by distinct ribosomal banding visualized on a Nano LabChip 6000.

Each RNA sample was tested for integrity and absence of gDNA in triplicate and was subsequently pooled in equimolar proportions. Thus, three biological replicates for each of the 28 treatments were generated.

A 10-fold dilution series (250 ng to 25 pg per qPCR reaction) of each cDNA was prepared to determine the optimum cDNA concentration for subsequent qPCR analyses. Using cDNA at 250 and 25 ng/qPCR resulted in poor amplification of each PCR product, whereas 2.5 ng/qPCR resulted in consistent C_q values in the range of 18 to 35. Therefore, 2.5 ng was selected as the optimal amount of cDNA per reaction.

Primer efficiency values for all candidate genes were calculated with LinRegPCR software and compared with efficiency values calculated using linear regression parameters from five-point standard curves generated via LightCycler 480 analysis (Table 3). Amplification efficiencies were similar from both methods but LinRegPCR offers the benefit of enabling primer efficiencies to be determined within each run; therefore, LinRegPCR primer efficiencies were used for this study. Amplification efficiencies between 1.7 (70% efficiency) and 2.0 (100% efficiency), which is considered acceptable (Pfaffl et al. 2004; Vandesompele et al. 2002), were achieved using 10 of the primer pairs under our PCR conditions.

The transcript accumulation, as determined by the number of cycles needed for amplification related fluorescence to reach a maximal point of a second derivative curve (C_q), was assessed for the candidate reference genes using cDNA template prepared from each of the treatments of *Arabidopsis* plants. Mean and standard error C_q values for each gene amplified from each of the three biological replicates per leaf sample are given in Table 3. *EF1α* was the most abundant reference gene transcript, with

Table 2. Fold change of candidate reference genes following virus infection^a

Host species, virus	Actin8	APT1	CYTc-2	EF1α	TATA	YLS8
<i>Arabidopsis thaliana</i>						
Turnip vein-clearing virus ^b	...	...	2.5–4.0	...	...	...
Oilseed rape mosaic virus	...	...	2.5–4.0	...	...	...
Potato virus X	...	...	2.5–4.0	...	...	...
Cucumber mosaic virus	...	...	2.5–4.0	...	...	...
Turnip mosaic virus (TuMV)	...	...	2.5–4.0	...	...	...
TuMV ^c	Ref	...	...	...	...	...
Plum pox virus ^d	2.97	Ref	–3.3	4.92	...	...
Cabbage leaf curl virus ^e	...	...	0.27	...	...	0.72
<i>Lycopersicon esculentum</i>						
PepYMV ^f	...	Ref	2.00E-26	Ref	...	...
<i>Solanum tuberosum</i>						
Potato virus Y ^{(NTN)g}	...	...	...	Ref	...	...
<i>Homo sapiens</i>						
Human immunodeficiency virus ^h	...	...	...	...	4	...
Herpes simplex 1	...	...	...	...	4	...
Varicella zoster	...	...	...	...	4	...
Cytomegalovirus	...	...	...	...	4	...

^a Ref = genes that were used as a reference gene in the given study. Candidate reference genes (*CAC*, *F-BOX*, *HEL*, *PDF2*, *PP2A1*, *PPR*, *SAND*, and *TIP41*) were omitted from this table as no data were available.

^b Whitham and associates (2003).

^c Yang and associates (2007).

^d Babu and associates (2008).

^e Ascencio-Ibáñez and associates (2008).

^f Alfenas-Zerbini and associates (2009).

^g Baebler and associates (2009).

^h Watson and associates (2007).

amplification-related fluorescence reaching a maximum at cycle 20.3 ± 0.14 . In contrast, *TIP41*-like family protein (*TIP41*) and Clathrin adaptor complex (*CAC*) were the least abundant transcripts, reaching maximum amplification fluorescence at cycles 32 to 34. Given that the qPCR protocol of this study used 40 cycles of amplification, C_q values of 35 were considered late to amplify (late call) and as such, unreliable. Helicase (*HEL*), *YSL8*, and pentatricopeptide repeat-containing protein (*PPR*) consistently resulted in C_q values greater than 35 (Table 3) and were not included for further validation. *TIP41* was initially included for further validation because of its ranking as a stable gene transcript under developmental, shoot, root, hormone, and

nutrient stress conditions (Table 1) (Czechowski et al. 2005; Expósito-Rodríguez et al. 2008; Reid et al. 2006) but it had a low average primer efficiency (1.4 ± 0.3) and a relatively high C_q (33.76 ± 0.26). Furthermore, calculation of quantity (Q) values in preparation for geNorm analysis demonstrated that *TIP41* consistently gave very low Q values (0.002). Therefore, *TIP41* was removed from further analysis, resulting in a final list of nine candidate reference genes. The nine candidate reference genes and two genes used for comparison, RNA-dependent RNA polymerase 1 (*RDR1*) and *RDR6*, are listed by their *Arabidopsis* Gene Initiative (AGI) locus identifier number, gene descriptors, and function in Table 4.

Table 3. Primers and amplicon size for candidate reference genes and *RDR1* and *RDR6*

AGI ^a	Annotation	5'→3' Primer sequences	LinReg (mean ± SE) ^b	Efficiency LC480 ^b	Mean C_q (± SE) ^c	Tm (°C) ^d	Amplicon size (bp)
At1g49240	<i>Actin8</i>	F-(1,078–1,097) CCCAAAAGCCAACAGAGAGA R-(1,313–1,292) CATCACCAGAGTCCAACACAAT	1.9 ± 0.04	1.987	24.3 ± 0.26	68.00 69.00	...
At1g27450	<i>APT1</i>	F-(1,777–1,796) GATTGTTGAATGTGCTTGCG R-(2,012–1,993) AGCATCAACACTCGATCCAA	1.7 ± 0.02	1.929	28.3 ± 0.32	66.00 66.00	...
At5g46630	<i>CAC</i>	F-(201–220) ATCGCACCTACCGAGATGAC R-(514–493) CGATGTAGACATTGCTGATTCG	1.7 ± 0.04	1.927	32.1 ± 0.34	70.00 69.00	...
At5g60390	<i>EF1α</i>	F-(922–941) CACCACTGGAGGTTTTGAGG R-(1,158–1,140) TGGAGTATTGGGGGTGGT	2.0 ± 0.01	1.945	20.3 ± 0.14	70.00 68.00	...
At5g15710	<i>F-BOX</i>	F-(1,463–1,482) GGCTGAGAGGTTTCGAGTGTT R-(1,602–1,583) GGCTGTTGCATGACTGAAGA	2.0 ± 0.01	1.858	29.0 ± 0.32	70.00 68.00	...
At1g58050	<i>HEL</i>	F-(8,121–8,142) GGGAGAAGGAAGGAGAATTTTG R-(8,316–8,295) GACATAGTCCCGCGCATAATA	N/A	N/A	35+	69.00 69.00	...
At1g13320	<i>PDF2</i>	F-(3,791–3,811) TCATTCCGATAGTCGACCAAG R-(4,020–4,000) TTGATTTGCGAAATACCGAAC	1.8 ± 0.02	1.932	27.4 ± 0.17	69.00 65.00	...
At1g59830	<i>PP2A1</i>	F-(1,464–1,483) CTGATCTCAAGAGCGCATCA R-(1,686–1,678) TCTAGGTGCTGGATCGAATTG	2.0 ± 0.00	2.226	25.6 ± 0.22	68.00 69.00	...
At1g62930	<i>PPR</i>	F-(1,846–1,866) CATGATGGGAGATTGGAAAAA R-(1,967–1,947) AAAGGCGATTACCTCACATA	N/A	2.537	35+	65.00 67.00	...
At2g28390	<i>SAND</i>	F-(3,376–3,396) GTTGGGTCACACCAGATTTTG R-(3,611–3,591) GCTCCTTGCAAGAACACTTCA	2.0 ± 0.01	1.993	25.3 ± 0.11	69.00 69.00	...
At1g55520	<i>TATA-BP</i>	F-(1,922–1,942) TTCAGGAAGGTCCAGCAATAA R-(2,136–2,116) CCAAGGAATATAACCCACCA	2.0 ± 0.01	1.997	26.4 ± 0.17	67.00 69.00	...
At4g34270	<i>TIP41</i>	F-(1,746–1,767) TGAAAACTGTTGGAGAGAAGCA R-(2,066–2,045) TGCTTAATCACTGGGAAGCCTCT	1.4 ± 0.3	N/A	33.76 ± 0.26	68.00 69.00	...
At1g14790 ^e	<i>RDR1</i> ^e	F-(2,750–2,769) CGCCAAGAACGTCTGAACCA R-(2,962–2,943) TTCGCGATGATCCCTAAACT	2.0 ± 0.01	1.863	28.2 ± 0.28	66.00 66.00	...
At3g49500 ^e	<i>RDR6</i> ^e	F-(2,139–2,158) TGCTGAACAGGGAAACACTG R-(3,246–3,227) CTTCTGTCATGGCTCCATAC	1.9 ± 0.01	1.912	27.2 ± 0.48	68.00 70.00	...

^a *Arabidopsis* Gene Initiative (AGI) locus identifier number.

^b Primer efficiency, mean ± standard error (SE). N/A = no amplification.

^c Threshold cycle.

^d Melting temperature.

^e Potential target genes of interest.

Table 4. *Arabidopsis* Gene Initiative (AGI) locus identifier number, descriptor, and function of genes used in this study from information publically available from The *Arabidopsis* Information resource (TAIR)

AGI	Descriptor	Function
At1g49240	<i>Actin8</i>	Member of a subclass of actins composed of <i>Actin2</i> and <i>Actin8</i> . Component of the cytoskeleton. Involved in cell motility and signaling.
At1g27450	<i>APT1</i> : adenine phosphoribosyl transferase 1	Involved in the one-step salvage of adenine to AMP.
At5g46630	<i>CAC</i> : clathrin adaptor complex	Medium subunit family protein. Functions in protein binding and intracellular protein transport.
At5g60390	<i>EF1α</i> : elongation factor 1-α	Involved in calmodulin binding. Translation elongation factor activity.
At5g15710	<i>F-BOX</i> : F-box family protein	Molecular functions and biological processes unknown. Possible involvement in protein degradation via the proteasome.
At1g13320	<i>PDF2</i> : protodermal factor 2	Encodes a 65-kDa regulatory subunit of serine/threonine protein phosphatase 2A (PP2A). Also known as protein phosphatase type 2A regulator subunit 3 (PP2AA3).
At1g59830	<i>PP2A1</i> : protein phosphatase 2A subunit 3	Encodes one of three 65-kDa isoforms of the catalytic subunit of protein phosphatase 2A.
At2g28390	<i>SAND</i> : sand family protein	Membrane protein with putative role in vesicle traffic and late endocytosis (<i>Saccharomyces cerevisiae</i> ; <i>Caenorhabditis elegans</i>).
At1g55520	<i>TATA</i> : TATA-box binding protein	Required for basal transcription. Facilitates recruitment of TFIID to the promoter. Forms preinitiation complex with RNA polymerase.
At1g14790	<i>RDR1</i> : RNA-dependent RNA polymerase 1	Promotes turnover of viral RNAs in infected plants. Amplifies RNAi. Involved in the production of <i>Cucumber mosaic virus</i> small interfering RNAs.
At3g49500	<i>RDR6</i> : RNA-dependent RNA polymerase 6	Involved in trans-acting siRNA and other siRNA biogenesis. Required for post-transcriptional gene silencing and natural virus resistance.

Average transcript accumulation stability (M) as calculated by geNorm.

Q values of the remaining nine candidate reference genes and *RDR1* and *RDR6*, for all three biological replicates, were analyzed using geNorm to generate a stability measure (M). Genes with the lowest stability of transcript accumulation, highest M , were excluded in a stepwise manner until the two most stably expressed genes remained. This process enabled ranking of genes in order of increasing transcript stability across the mock inoculation and virus infection time course (Fig. 1A). Many of the candidate reference gene transcripts (*Actin8*, *APT1*, protein phosphatase 2A subunit 1 [*PP2A1*], *CAC*, and TATA-box binding protein [*TATA*]) had low stability, with M values of 4.30 to 2.54, above the default limit of 1.5 suggested by geNorm. *Actin8* and *APT1* were the first excluded as the least stable genes (M values of 4.30 and 3.20, respectively). Protodermal factor 2 (*PDF2*), sand family protein (*SAND*), *EF1 α* , and an *Arabidopsis* F-box family protein (*F-BOX*) were identified as the most stable genes, with M values of 1.46, 1.32, 1.27, and 1.10, respectively. The *RDR1* gene transcript was less stable ($M = 3.36$) than all candidate reference genes, with the exception of *Actin8*, while *RDR6* was less stable than all other genes ($M = 4.82$).

Average transcript stability and the range of M values of each gene across the time course (2, 3, 7, and 14 days postinoculation [dpi]) of inoculations are shown in Figure 1B. Among the candidate reference genes, *SAND* has the most stable transcript accumulation across the time course, followed by *PDF2*, *CAC*, and *F-BOX*. *PP2A1* has the least stable transcript accumulation whereas *EF1 α* , which is very stable across all treatments and time points, has only moderate stability across the time course of treatments.

Average transcript stability, M , and the range of M values of each gene across all treatments (mock, no Carborundum, and *Cauliflower mosaic virus* [CaMV], *Tomato spotted wilt virus* [TSWV], *Turnip mosaic virus* [TuMV], *Turnip yellow mosaic virus* [TYMV], and *Tobacco mosaic virus* [TMV] inoculation) are shown in Figure 1C. *PP2A1* had the most stable transcript accumulation in response to each treatment, followed by *SAND* and *PDF2*. *Actin8* and *APT1* had the least stable transcript accumulation in response to each treatment ($\Delta M = 5.04$ and 3.10, respectively).

Quantity (Q) of transcript accumulation.

The raw value of rate of transcript accumulation (Q) changes for each gene, depending on the stress the plant incurred, such that no gene transcript accumulated equally across all times and treatments. In order to determine how all the genes accumulate over the time course of all treatments; average Q values were plotted. Stable accumulation of a gene transcript across a time course would theoretically result in a straight and level line with minimal error. Average Q and the range of Q values and range of standard error for each gene are presented in Figure 2A. The results demonstrate that *CAC*, *PDF2*, *TATA*, and *SAND* have lines that are straight compared with *Actin8*. *Actin8* and *APT1* are two of the least stably accumulating transcripts across time, with a range of ΔQ of 0.04 and 0.10, respectively, and both have a large range of error (0.04 and 0.03, respectively). Of the four most stable gene transcripts, as determined by M , *PDF2* has the most stable accumulation ($\Delta Q = 0.02$) and the smallest error (0.02) while *EF1 α* is the least stable in transcript accumulation ($\Delta Q = 0.05$) and has the largest range of error (0.04).

Fig. 1. Ranking of stability and transcript accumulation of candidate reference genes. **A**, geNorm ranking of nine candidate reference genes based on quantitative polymerase chain reaction data. Four technical replicates of three biological replicates of *Arabidopsis* leaf tissue across treatment inoculations: mock inoculation, mock inoculated without Carborundum (no carb), *Cauliflower mosaic virus* (CaMV), *Tomato spotted wilt virus* (TSWV), *Turnip mosaic virus* (TuMV), *Turnip yellow mosaic virus* (TYMV), and *Tobacco mosaic virus* (TMV) across 2, 3, 7, and 14 days postinoculation (dpi) were included in the analysis. **B**, Average transcript accumulation stability (M) across the time course of all treatments. **C**, Average transcript accumulation stability across each treatment. The four candidate reference genes that were identified as most stable in A are indicated by asterisks in remaining panels.

A range of stable accumulation of a gene transcript across treatments would theoretically result in a plot with similar data points and error for each treatment. Average Q and the range of Q values and range of standard error for each gene are presented in Figure 2B. The results demonstrate that *RDR1* is the most stable in accumulation across all treatments, with the smallest range of standard error, while *EF1 α* , *APT1*, and *Actin8* have the most erratic accumulation.

Choice of reference genes affects normalization.

The accumulation of *RDR1* and *RDR6* transcripts were used as examples to show the effect of using different normalization genes to estimate relative accumulation. The transcript quantities of *RDR1* and *RDR6* were determined by qPCR in the same samples used for the selection of reference genes. Average Q

for *RDR1* and *RDR6*, ranging from 0.03 to 1.0 and 0.005 to 1.0, respectively, across time and treatment prior to normalization, are shown in Figure 3A and D. When transcript quantities of these two genes are normalized with the normalization factor generated by geNorm for the three most stable genes (*F-BOX*, *EF1 α* , and *SAND*), the range for *RDR1* was between 0.16 and 1.0 and the standard error increased (Fig. 3B) compared with raw data while the range for *RDR6* increased to between 0.01 and 1.0 (Fig. 3E) and the error remained similar to raw data (Fig. 3D). Interestingly, *RDR6* reacted very differently to inoculation with CaMV than any other virus inoculation. When transcript quantities were normalized instead with the normalization factor generated for the two least stable candidate reference genes (*Actin8* and *APT1*), *RDR1* transcript accumulation was overestimated further to between 0.2 and

Fig. 2. Relative accumulation stability of candidate reference genes. Transcript accumulation of each candidate reference gene, *RDR1* and *RDR6*, across **A**, the time course or **B**, all treatments. Error bars are standard error of mean quantities (Q) of three biological replicates (A, $n = 12$; B, $n = 18$). For each gene, the range of Q values is given at the top of the graph and the range of standard error is shown in brackets. The four most stably accumulating genes are identified by asterisks.

1.0 and the reported error was greater in most cases (Fig. 3C). For *RDR6*, transcript quantity remains in a similar range (0.01 to 1.0) but the error was underestimated, most notably for all treatments at 14 dpi and TSWV at 2 dpi (Fig. 3F).

DISCUSSION

For qPCR to be accurate, specific conditions and appropriate internal reference genes for normalization must be determined. A reference gene should show minimal change, whereas a gene of interest may change greatly over the course of an experiment. Normalization with multiple reference genes is becoming the minimum standard. Furthermore, a “minimum standard for the provision of information for qPCR experiments” (MIQE) has been proposed (Bustin 2010; Bustin et al. 2009). This minimum standard will encourage detailed auditing of the entire qPCR workflow from experimental design through technical details such as nucleic acid extraction, reverse transcription, qPCR target information, oligonucleotides, PCR cycling protocol, validation of appropriate reference genes, and data analysis. Bustin (2010) and Pfaffl (2010) propose that a current lack of rigorous standards of practice for qPCR has resulted in misinterpretation

of data and publication of erroneous conclusions. By implementing the MIQE guidelines, qPCR can become a robust, accurate, and reliable method for reporting nucleic acid quantification. This is important if confidence in published results and conclusions is to be maintained.

Here, we evaluated the stability of accumulation of nine candidate reference genes, including two traditional housekeeping genes, in *Arabidopsis* subject to five diverse plant viral infections. DNA and RNA viruses with diverse replication, gene expression, and movement strategies were used. This diversity lends itself to a study of gene transcript accumulation in *Arabidopsis* in response to virus infection (Whitham et al. 2003).

Our findings show that *Actin8*, one of the most commonly used reference genes, showed the highest variability in transcript accumulation. These results caution against the use of *Actin8* as a reference gene in virus-infected *Arabidopsis*. Similar results have been observed and reported from validation of reference genes in *Arabidopsis* and other plant species subject to various other stresses. For example, in *Arabidopsis* qPCR-validated data, Czechowski and associates (2005) ranked *Actin8* 20th for stability of 21 genes investigated. Nicot and associates (2005) ranked *Actin8* as least stable of eight genes investigated

Fig. 3. Comparison of raw and normalized accumulation of *RDR1* and *RDR6* gene transcripts. Transcript accumulation across time course of all treatments (days postinoculation [dpi]) of **A** through **C**, *RDR1* and **D** through **F**, *RDR6* following **A** and **D**, no normalization; **B** and **E**, normalization to the three most stable references (*F-BOX*, *EF1α*, and *SAND*); or **C** and **F**, normalized to the two least stable references (*Actin8* and *APT1*), using the normalization factor generated by geNorm. Data were rescaled to 1.0, where 1.0 was the largest quantity (Q) value calculated for each gene transcript across all treatments.

in potato exposed to a biotic (late blight) and two abiotic (cold and salt) stresses. Similarly, Libault and associates (2008) ranked an actin ortholog of the *Arabidopsis* actin family as 13th most stable of 15 genes studied in soya bean subject to 130 different conditions. Mallona and associates (2010) ranked an *Actin* of petunia eighth for stability of nine genes investigated. Our results agree with these earlier findings where *Actin8* shows poor transcript stability. Some studies have reported various actins with high transcript stability, suggesting that *Actin8* might be considered a suitable reference gene given specific experimental conditions (Artico et al. 2010; Gutha et al. 2010; Wan et al. 2009). However, recent studies have revealed that the actins do not satisfy basic requirements for application as reference genes (Gutierrez et al. 2008a; Nicot et al. 2005; Thellin et al. 2009). This may be explained in part by the fact that actins, major components of cytoplasmic microfilaments in eukaryotic cells, participate in other cellular functions in addition to the provision of structural support and cellular shape (Stürzenbaum and Kille 2001).

PDF2, *SAND*, *EF1 α* , and *F-BOX* act in distinct pathways and can be considered appropriate reference genes for normalization given our experimental conditions for virus infection in *Arabidopsis*. These genes were stable across the experimental conditions and provided a range of transcript accumulation for comparison with genes of interest. The primer pairs used in this study to target *F-BOX*, *EF1 α* , *SAND*, and *PDF2* have high identity to genes of other species, especially *Brassica* and *Nicotiana* spp. and *Vitis vinifera*. Therefore, although not demonstrated in this study, they are likely candidates for use in other plant species.

Of our four most stably accumulating transcripts, *PDF2* was the fourth, a finding similar to Czechowski and associates (2005) where *PDF2* was ranked sixth of 21 transcripts of *Arabidopsis* ($M < 0.75$). *PDF2* encodes a 65-kDa regulatory subunit of serine or threonine protein phosphatase and functions to maintain the identity of meristem L1 layer cells, and overexpression can delay flowering (Abe et al. 2003).

Our results show *SAND* to be the third most stable transcript, similar to the findings of Czechowski and associates (2005), where *SAND* ranked first of 21 *Arabidopsis* transcripts ($M < 0.75$), and Remans and associates (2008), where *SAND*, in combination with *F-BOX* and *YLS8*, was identified as one of the best candidates for normalization of data from *Arabidopsis* tissue exposed to cadmium (Cd) and copper (Cu) ($M < 0.4$). Mallona and associates (2010) ranked *SAND* second most stable of nine transcripts in petunia ($M = 0.5$). The encoded protein is thought to have a function in vesicle trafficking and late endocytosis as determined from *Saccharomyces cerevisiae* and *Caenorhabditis elegans* gene orthologs (Poteryaev and Spang 2005). Given that *SAND* plays a role in vesicle traffic, it may not be a suitable reference gene in studies of other pathogens where vesicle traffic may play an important role in defense.

EF1 α is frequently used as a reference gene and was one of the four most stably accumulating transcripts in this study. Our finding that *EF1 α* was stable in transcript accumulation is consistent with results from some other groups. For example, Nicot and associates (2005) ranked *EF1 α* as the most stable of 8 genes ($M < 0.3$) and Gutierrez and associates (2008a) ranked *EF1 α* second most stable of 14 genes studied in *Arabidopsis* ($M < 0.5$). Wan and associates (2009) ranked *EF1 α* second most stable of 10 transcripts studied in cucumber ($M = 0.8$). However, Czechowski and associates (2005) reported that *EF1 α* was the least stable of 21 transcripts of *Arabidopsis* ($M > 1.5$) and Expósito-Rodríguez and associates (2008) ranked *EF1 α* 10th of 11 transcripts of tomato. However, it should be noted that *EF1 α* might not be a suitable reference gene in the context of all known plant virus infections. For example, *EF1 α* has been

shown to have a 4.92-fold change in response to PPV infection (Babu et al. 2008), is one of 226 suppressed genes in response to TuMV (Yang et al. 2007), and shows variability in response to five viruses in tomato (Mascia et al. 2010). The *EF1 α* protein is involved in calmodulin binding and has translation elongation factor activity (The *Arabidopsis* Information Resource [TAIR]).

F-BOX was identified as the most stable transcript in this study. This is consistent again with Remans and associates (2008), where *F-BOX* and *SAND*, along with *YLS8*, were identified as the best candidates for data normalization ($M < 0.3$). Our results are also consistent with the findings of Libault and associates (2008), where *F-BOX* was ranked most stable of 15 transcripts tested from soya bean ($M < 0.3$). Our results are in contrast, however, with those of Artico and associates (2010), who ranked *F-BOX* seventh of nine ($M > 0.8$). The molecular functions and biological processes of this particular *F-BOX* protein in *Arabidopsis* are currently unknown (TAIR).

Recently, Mascia and associates (2010) examined the stability of eight candidate reference genes in tomato subjected to four different plant viruses at 9 dpi and one plant virus at 30 dpi. High variability was demonstrated for two widely used reference genes, 18S rRNA and *EF1 α* . However, three commonly used reference genes—*Actin 3* (*ACT3*), *GAPDH*, and *ubiquitin 3* (*UBI*)—showed high levels of stability throughout the infected tissues. Of the two novel candidate reference genes evaluated, *uridylate kinase* (*UK*) did not respond to virus or viroid infection. Therefore, *UK* might be another appropriate candidate reference gene for trial as a transcript for normalization across diverse plant–virus interactions.

RDR1 and *RDR6* were included for comparison to reference genes. *Arabidopsis* requires host *RDR1* and *RDR6* to produce viral secondary small interfering (si)RNAs. Wang and associates (2010) found that both *RDR1* and *RDR6* were involved in the production of *Cucumber mosaic virus* siRNAs and in the amplification of RNAi in infected plants. Given the putative roles of *RDR1* and *RDR6* in the course of plant viral infection, unstable M values were expected in our study. Accordingly, our findings show that the stability of *RDR1* falls between the two least stable candidate reference genes, *Actin8* and *APT1*, whereas *RDR6* is the least stable of all genes considered.

When separating the data across the time course of treatments, our results for candidate reference gene stability, as measured by M , show that *SAND*, *PDF2*, and *F-BOX* are the best combination of three reference genes for normalization. When separating out data across inoculation treatments, our results for candidate reference gene stability, as measured by M , show that *PP2A1*, *SAND*, and *PDF2* are the best combination for normalization. A difference in ranking of appropriate sets of reference genes to use for normalization of different datasets is not unexpected, given that no single gene or set of genes demonstrated constant transcript accumulation for all experimental conditions (Czechowski et al. 2005; Remans et al. 2008; Vandesompele et al. 2002).

Rather than investigating stability, by separating out measures of raw accumulation, or Q , across the time course of treatments, we demonstrated that *Actin8* and *APT1* were two of the least stable in transcript accumulation across time. *PDF2* had the most stable Q value across time. *PDF2* had one of the most stable Q values across the inoculations, while *Actin8* and *APT1* were again among the least stable in Q value across our treatments. *F-BOX* and *SAND* were moderately stable in transcript accumulation. These results serve to emphasize the unreliability of *Actin8* and *APT1* in our study.

Variation in the stability and transcript accumulation of reference genes between sample groups reduces the sensitivity of an assay to detect changes in transcript accumulation of target genes of interest and may produce artificial changes. To dem-

onstrate the effect of using appropriate versus inappropriate reference genes to normalize data, we normalized our potential genes of interest, *RDR1* and *RDR6*, to the normalization factor generated for the three most stable and to the two least stable candidate reference genes. Our results demonstrate a trend toward differences in reported transcript accumulation depending on which set of genes was used to generate the normalization factor. Interestingly, *RDR6* reacts very differently to inoculation with CaMV at all time points. This might be a function of CaMV being a double-stranded DNA virus and warrants further investigation. Our findings emphasize that, if a chosen reference gene has a large transcript fluctuation, normalization will lead to erroneous transcript accumulation profiles of target genes of interest.

Our analyses have shown that different experimental conditions tested require different sets of reference genes. These results are consistent with other groups (Czechowski et al. 2005; Thellin et al. 2009; Vandesompele et al. 2002) and further emphasize the importance of validating reference genes for each experimental condition (for example, different tissues or organs) or for plant viruses that are phloem limited. Recently, a study investigated six putative stably expressed genes in qPCR of cDNA from phloem-limited *Grapevine leafroll associated virus 3*-infected grape leaves (Gutha et al. 2010). A combination of two genes, *Actin* and *NAD5*, was identified as suitable candidates for normalization of data in both virus-free and virus-infected leaves. A further candidate reference gene might include sucrose/proton symporter gene 2 or phloem-specific reporter genes in transgenic plants (Ckurshumova et al. 2009). Changes in transcript accumulation within parenchyma cells adjacent to infected cells could be referenced to the gene transcripts identified in this study.

Selection of appropriate reference genes is necessary when performing real-time qPCR. These reference genes should not vary in the investigated cells or tissues throughout experimental conditions. Reference gene stability varies considerably, and this variability can and does lead to incorrect conclusions being drawn from data. It is recommended that at least three validated reference genes be used to calculate a normalization factor (Guénin et al. 2009; Klatte and Bauer 2009; Vandesompele et al. 2002). Under our experimental conditions of virus-infected *Arabidopsis*, transcript accumulation of *SAND*, *PDF2*, *EF1 α* , and *F-BOX* showed a high level of stability and, although it is imperative that reference genes be validated for every new set of experimental conditions, the genes investigated here provide a starting point for those wishing to evaluate gene transcript accumulation of virus-infected *Arabidopsis* by qPCR. Furthermore, these data will be beneficial to the selection and design of primer pairs targeting orthologous genes in other plant species and will aid in the selection of suitable genes for normalization in studies of diverse plant–pathogen interactions.

MATERIALS AND METHODS

Virus isolates, plant material, and inoculation.

CaMV in *Brassica campestris* subsp. *pekinensis* (Pak Choy) freeze-dried leaf material, TSWV in fresh chrysanthemum, and TuMV in freeze-dried turnip were kindly gifted by J. Fletcher and R. Lister (Plant & Food Research, Lincoln, New Zealand). TYMV in fresh *B. campestris* cv. *chinensis* and TMV in frozen *Nicotiana tabacum* cv. White Burley were obtained from the School of Biological Sciences, University of Auckland, New Zealand. Upon receipt, the virus isolates were mechanically inoculated to *Arabidopsis*, as described below, as a method for maintaining fresh stocks for inoculation of experimental plants. Virus identity was confirmed by sequencing of reverse-transcriptase PCR products.

Arabidopsis thaliana plants (ecotype Columbia-0) were maintained in a shaded greenhouse set for $20 \pm 2^\circ\text{C}$ and a short-day, 8-h photoperiod with 400-W Philips SON-T Agro lights to provide photosynthetically active radiation of 120 to $150 \mu\text{M m}^{-2} \text{s}^{-2}$. At 4 weeks of age, plants were inoculated with each virus in 0.1 M potassium phosphate (K_2HPO_4) buffer (pH 7.4) containing 0.1% (wt/vol) anhydrous sodium sulfite (Na_2SO_3) and Carborundum (as an abrasive). Two forms of mock-inoculated plants (both without infectious homogenate, one without Carborundum) were created as controls. At 2, 3, 7, and 14 dpi, all leaf tissue of three plants per treatment was harvested and immediately frozen in liquid nitrogen and then transferred to -80°C for storage until the RNA was isolated.

RNA isolation, DNase I digest, and cDNA synthesis.

Total RNA was isolated from no greater than 100 mg of frozen leaf tissue using an RNeasy mini kit (Qiagen, Hilden, Germany) according to the manufacturer's protocols. The concentration and purity of the RNA samples was determined using Nanodrop (ND-1000 Spectrophotometer; Nanodrop Technologies Inc., Wilmington, DE, U.S.A.). All samples had an absorbance ratio (absorbance at 260/280 nm) of between 1.8 and 2.2. Following quantification, all RNA samples were normalized to $2.5 \mu\text{g } \mu\text{l}^{-1}$. The integrity and intactness of RNA samples was determined by micro fluidics assessment with a Bioanalyser 2100 Nano LabChip 6000 (Agilent Technologies, Santa Clara, CA, U.S.A.) and most were found to have RNA integrity numbers of 7.0 or greater. Any gDNA contamination was removed using DNase I (amplification grade; Invitrogen, San Diego, CA, U.S.A.) before $2.5 \mu\text{g}$ of RNA was reverse transcribed into cDNA using SuperScript III RNase H[−] reverse transcriptase (Invitrogen) according to the manufacturer's protocols.

Primer design.

To ensure maximum specificity and efficiency during qPCR amplification under a standard set of reaction conditions, primers were designed using the Primer3 software, version 0.1 (S. Rozen and H. J. Skaletsky,). Parameters optimized were melting temperatures of 65°C , with a maximum difference of 3°C ; length of 20 to 24 bases; GC content of 45 to 50%; and product length of 100 to 150 bp. The PCR product of each primer pair spanned an intron and would be amplified from all the known splice variants (as annotated on TAIR). The AGI locus identifier number, descriptor, primer sequences, primer efficiencies, mean C_q , melting temperature, and amplicon size for each candidate are presented in Table 3.

Testing candidate reference gene primers with standard PCR.

A 25- μl reaction volume PCR was employed using cDNA at $2.5 \text{ ng } \mu\text{l}^{-1}$, 12.5 μl of GoTaq Green Master Mix (Promega Corp., Madison, WI, U.S.A.), 2 pmol each gene-specific primer (Table 3), and UltraPure DNase/RNase-free distilled water (Invitrogen) to 25 μl . Amplification of cDNA employed the following conditions: initial 5 min of denaturation at 94°C ; followed by 35 cycles of denaturation 94°C for 30 s, annealing at 60°C for 30 s, and extension at 72°C for 30 s; followed by a final extension for 5 min at 72°C . Negative control reactions, one omitting reverse transcriptase and one omitting cDNA template, were performed to confirm the absence of genomic DNA or other contaminants.

The presence of single products of appropriate size was confirmed by electrophoresis through a 1.0% agarose gel that was subsequently stained with ethidium bromide and viewed under ultraviolet light with a Gel Doc imager (Bio-Rad, Hercules, CA, U.S.A.).

qPCR conditions and analysis.

qPCR reactions were preformed in 384-well plates with a LightCycler 480 Real-Time PCR system using LightCycler 480 SYBR Green I Master (Roche Applied Science, Branchburg, NJ, U.S.A.). Reactions were in 10- μ l volumes containing 2 μ l of primer pair (5 μ M each primer), 3 μ l of cDNA (2.5 ng/ μ l), and 5 μ l of SYBR Green I Master mix reagent. A Biomek 3000 Robot (Beckman Coulter, Fullerton, CA, U.S.A.) was used to aliquot all reagents, primers, and samples. Each qPCR reaction was completed in triplicate. Reactions were run using the manufacturer's recommended cycling parameters, consisting of a 95°C preincubation for 5 min and 40 amplification cycles of denaturing at 95°C for 10 s, primer-dependent annealing at 65°C, and elongation at 72°C for 6 s (calculated by dividing largest amplicon length by 25). Fluorescence values were acquired at the end of the elongation phase. Amplification was followed by a single melting-curve-analysis cycle of 95°C for 5 s and 65°C for 1 min with continuous acquisition of fluorescence data at five acquisitions per degree Centigrade at a ramp rate of 0.11°C s⁻¹ to 97°C. Finally, the samples were cooled at 40°C for 10 s.

Primer efficiency and geNorm analysis of gene transcript accumulation stability.

Primer efficiency values were calculated with LinRegPCR (11.0) (Ruijter et al. 2009). The raw fluorescent data were baseline corrected and log transformed, and the window-of-linearity determined. Linear regression analysis was then used to fit a straight line through the data points within the window-of-linearity, and the PCR efficiency was calculated from the slope of the line for each individual sample. The mean PCR efficiency was calculated for each primer pair. LightCycler data output was prepared for LinRegPCR using a custom-written script (Python version 2.6.3; J. McRae, Plant & Food Research, Auckland, New Zealand). The efficiency of amplification from the primers was also calculated for all candidate genes using linear regression parameters from five-point standard curves generated via LightCycler 480 analysis.

To identify the transcript accumulation stability value (M), of each candidate reference gene, C_q values were converted to raw, nonnormalized transcript abundance or Q , using the comparative C_q (ΔC_q) method and the calculated PCR efficiency, and rescaling the data such that the highest relative accumulation value is 1. Stability of gene transcript accumulation was analyzed with geNorm v3.5 analysis software according to instructions in the operator's manual.

In order to determine whether the choice of reference genes affects the normalization of genes of interest, we normalized *RDR1* and *RDR6* to the three most stable candidate reference genes and the two least stable as determined by our geNorm analysis.

ACKNOWLEDGMENTS

We thank J. McRae for his assistance with the Beckman Coulter robot parameters and for providing the custom-written Python V 2.6.3 script, J. Bowen for assistance in setting up initial qPCR plates, N. de Silva for assistance with biometrics and statistics, and D. Cohen and J. Bowen for critical reading of this manuscript and valuable comments. This research was funded by Ministry of Science and Innovation (formerly, The New Zealand Foundation for Research, Science and Technology) grant C06X0710.

LITERATURE CITED

Abe, M., Katsumata, H., Komeda, Y., and Takahashi, T. 2003. Regulation of shoot epidermal cell differentiation by a pair of homeodomain proteins in *Arabidopsis*. *Development* 130:635-643.

- Alfenas-Zerbini, P., Maia, I. G., Fávoro, R. D., Cascardo, J. C. M., Brommonschenkel, S. H., and Zerbini, F. M. 2009. Genome-wide analysis of differentially expressed genes during the early stages of tomato infection by a potyvirus. *Mol. Plant-Microbe Interact.* 22:352-361.
- Andersen, C. L., Jensen, J. L., and Ørntoft, T. F. 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.* 164:5245-5250.
- Artico, S., Nardelli, S. M., Brilhante, O., Grossi-de-Sa, M. F., and Alves-Ferreira, M. 2010. Identification and evaluation of new reference genes in *Gossypium hirsutum* for accurate normalisation or real-time quantitative RT-PCR data. *BMC Plant Biol.* 10:49.
- Ascencio-Ibáñez, J. T., Sozzani, R., Lee, T.-J., Chu, T.-M., Wolfinger, R. D., Cella, R., and Hanley-Bowdoin, L. 2008. Global analysis of *Arabidopsis* gene expression uncovers a complex array of changes impacting pathogen response and cell cycle during geminivirus infection. *Plant Physiol.* 148:436-454.
- Babu, M., Griffiths, J. S., Huang, T.-S., and Wang, A. 2008. Altered gene expression changes in *Arabidopsis* leaf tissues and protoplasts in response to *Plum pox virus* infection. *BMC Genomics* 9:325-346.
- Baebler, S., Krečič-Stres, H., Rotter, A., Kogovšek, P., Cankar, K., Kok, E. J., Gruden, K., Kovač, M., Žel, J., Pompe-Novak, M., and Ravnkar, M. 2009. PVY^{NTN} elicits a diverse gene expression response in different potato genotypes in the first 12 h after inoculation. *Mol. Plant Pathol.* 10:263-275.
- Bustin, S. 2010. Why the need for qPCR publication guidelines?—the case for MIQE. *Methods* 50:217-226.
- Bustin, S. A., Benes, V., Garson, J. A., Hellems, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M. W., Shipley, G. L., Vandesompele, J., and Wittwer, C. T. 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem.* 55:4611-4622.
- Carr, T., and Whitham, S.A. 2007. An emerging model system: *Arabidopsis* as a viral host plant. *Plant Cell Monogr.* 7:159-183.
- Ckurshumova, W., Koizumi, K., Chatfield, S. P., Sanchez-Buelna, S. U., Gangaeva, A. E., McKenzie, R., and Berleth, T. 2009. Tissue-specific GAL4 expression patterns as a resource enabling targeted gene expression, cell type-specific transcript profiling and gene function characterization in the *Arabidopsis* vascular system. *Plant Cell Physiol.* 50:141-150.
- Czechowski, T., Stitt, M., Altmann, T., Udvardi, M. K., and Scheible, W.-F. 2005. Genome-wide identification and testing of superior reference genes for transcript normalization in *Arabidopsis*. *Plant Physiol.* 139:5-17.
- De Ketelaere, A., Goossens, K., Peelman, L., and Burvenich, C. 2006. Technical note: validation of internal control genes for gene expression analysis in bovine polymorphonuclear leukocytes. *J. Dairy Sci.* 89:4066-4069.
- Expósito-Rodríguez, M., Borges, A. A., Borges-Peréz, A., and Pérez, J. A. 2008. Selection of internal control genes for quantitative real-time RT-PCR studies during tomato development process. *BMC Plant Biol.* 8:131-143.
- Faccioli, P., Ciceri, G. P., Provero, P., Stanca, A. M., Morcia, C., and Terzi, V. 2007. A combined strategy of 'in silico' transcriptome analysis and web search engine optimisation allows an agile identification of reference genes suitable for normalisation in gene expression studies. *Plant Mol Biol.* 63:679-688.
- Glare, E. M., Divjak, M., Bailey, M. J., and Walters, E. H. 2002. Beta-actin and GAPDH housekeeping gene expression in asthmatic airways is variable and not suitable for normalising mRNA levels. *Thorax* 57:754-756.
- Guénin, S., Mauriat, M., Pelloux, J., van Wuytswinkel, O., Bellini, C., and Gutierrez, L. 2009. Normalization of qRT-PCR data: the necessity of adopting a systematic, experimental conditions-specific, validation of references. *J. Exp. Bot.* 60:487-493.
- Gutha, L. R., Casassa, L. F., Harbertson, J. F., and Naidu, R. A. 2010. Modulation of flavonoid biosynthetic pathway genes and anthocyanins due to virus infection in grapevine (*Vitis vinifera* L.) leaves. *BMC Plant Biol.* 10:187.
- Gutierrez, L., Mauriat, M., Guénin, S., Pelloux, J., Lefebvre, J. F., Louvet, R., Rusterucci, C., Moritz, T., Guérineau, F., Bellini, C., and van Wuytswinkel, O. 2008a. The lack of a systemic validation of reference genes: a serious pitfall undervalued in reverse transcription-polymerase chain reaction (RT-PCR) analysis in plants. *Plant Biotechnol. J.* 6:609-618.
- Gutierrez, L., Mauriat, M., Pelloux, J., Bellini, C., and van Wuytswinkel, O. 2008b. Letters to the editor: towards a systematic validation of references in real time RT-PCR. *Plant Cell* 20:1734-1735.
- Huggett, J., Dheda, K., Bustin, S., and Zumla, A. 2005. Real-time RT-PCR

- normalisation; strategies and considerations. *Genes Immun.* 6:279-284.
- Kakar, K., Wandrey, M., Czechowski, T., Gaertner, T., Scheible, W.R., Stitt, M., Torres-Jerez, I., Xiao, Y., Redman, J. C., Wu, H. C., Cheung, F., Town, C. D., and Udvardi, M. K. 2008. A community resource for high-throughput quantitative RT-PCR analysis of transcription factor gene expression in *Medicago truncatula*. *Plant Methods.* 8:4:18.
- Klatte, M., and Bauer, P. 2009. Accurate real-time reverse transcription quantitative PCR. *Methods Mol. Biol.* 479:61-77.
- Lee, J. M., Roche, J. R., Donaghy, D. J., Thrush, A., and Sathish, P. 2008. Validation of reference genes for quantitative RT-PCR studies of gene expression in perennial ryegrass (*Lolium perenne* L.) *BMC Mol. Biol.* 11:8.
- Lee, P. D., Sladek, R., Greenwood, C. M., and Hudson, T. J. 2002. Control genes and variability: absence of ubiquitous reference transcripts in diverse mammalian expression studies. *Genome Res.* 12:292-297.
- Libault, M., Thibivilliers, S., Bilgin, D. D., Radwan, O., Benitez, M., Clough, S. J., and Stacey, G. 2008. Identification of four soybean reference genes for gene expression normalization. *Plant Genome* 1:44-54.
- Mallona, I., Lischewski, S., Weiss, J., Hause, B., and Egea-Cortines, M. 2010. Validation of reference genes for quantitative real-time PCR during leaf and flower development in *Petunia hybrida*. *BMC Plant Biol.* 10:4.
- Mascia, T., Santovito, E., Gallitelli, D., and Cillo, F. 2010. Evaluation of reference genes for quantitative reverse-transcription polymerase chain reaction normalization in infected tomato plants. *Mol. Plant Pathol.* 11:805-816.
- Nicot, N., Hausman, J.-F., Hoffmann, L., and Evers, D. 2005. Housekeeping gene selection for real-time RT-PCR normalization in potato during biotic and abiotic stress. *J. Exp. Bot.* 56:2907-2914.
- Pfaffl, M. W. 2010. Guest editor's introduction: the ongoing evolution of qPCR. *Methods* 50:215-216.
- 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 pairwise correlations. *Biotechnol. Lett.* 26:509-515.
- Poteryaev, D., and Spang, A. 2005. A role of SAND-family proteins in endocytosis. *Biochem. Soc. Trans.* 33:606-608.
- Reid, K. E., Olsson, N., Schlosser, J., Peng, F., and Lund, S. T. 2006. An optimized grapevine RNA isolation procedure and statistical determination of reference genes for real-time RT-PCR during berry development. *BMC Plant Biol.* 6:27-38.
- Remans, T., Smeets, K., Opdenakker, K., Mathijsen, D., Vangronsveld, J., and Cuyper, A. 2008. Normalisation of real-time RT-PCR gene expression measurements in *Arabidopsis thaliana* exposed to increased metal concentrations. *Planta* 227:1343-1349.
- Ruijter, J. M., Ramakers, C., Hoogaars, W., Bakker, O., van den Hoff, M. J. B., Karlen, Y., and Moorman, A. F. M. 2009. Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Res.* 37:45.
- Selvey, S., Thompson, E. W., Matthaei, K., Lea, R. A., Irving, M. G., and Griffiths, L. R. 2001. Beta-actin—an unsuitable internal control for RT-PCR. *Mol. Cell. Probes* 15:307-311.
- Stürzenbaum, S. R., and Kille, P. 2001. Control genes in quantitative molecular biological techniques: the variability of invariance. *Comp. Biochem. Physiol. Part B Biochem. Mol. Biol.* 130:281-289.
- Thellin, O., Zorzi, W., Lakaye, B., De Borman, B., Coumans, B., Hennen, G., Grisar, T., Igout, A., and Heinen, E. 1999. Housekeeping genes as internal standards: use and limits. *J. Biotechnol.* 75:291-295.
- Thellin, O., ElMoualij, B., Heinen, E., and Zorzi, W. 2009. A decade of improvements in quantification of gene expression and internal standard selection. *Biotechnol. Adv.* 27:323-333.
- 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.* 3:34-45.
- Wan, H., Zhao, Z., Qian, C., Sui, Y., Malik, A. A., and Chen, J. 2009. Selection of appropriate reference genes for gene expression studies by quantitative real-time polymerase chain reaction in cucumber. *Anal. Biochem.* 399:257-261.
- Wang, X. B., Qing W., Ito, T., Cillo, F., Li, W.-X., Chen, X., Yu, J.-L., and Ding, S.-W. 2010. RNAi-mediated viral immunity requires amplification of virus-derived siRNAs in *Arabidopsis thaliana*. *PNAS* 107:484-489.
- Watson, S., Mercier, S., Bye, C., Wilkinson, J., Cunningham, A. L., and Harman, A. N. 2007. Determination of suitable housekeeping genes for normalisation of quantitative real time PCR analysis of cells infected with human immunodeficiency virus and herpes viruses. *Viol. J.* 4:130-135.
- Whitham, S. A., Quan, S., Chang, H.-S., Cooper, B., Estes, B., Zhu, T., Wang, X., and Hou, Y.-M. 2003. Diverse RNA viruses elicit the expression of common sets of genes in susceptible *Arabidopsis thaliana* plants. *Plant J.* 33:271-283.
- Xue, J.-L., Salem, T. Z., Turney, C. M., and Cheng, X.-W. 2009. Strategy of the use of 28 rRNA as a housekeeping gene in real-time quantitative PCR analysis of gene transcription in insect cells infected with viruses. *J. Virol. Methods* 163:210-215.
- Yang, C., Guo, R., Nettleton, D., Peng, J., Carr, T., Yeakley, J. M., Fan, J.-B., and Whitham, S. A. 2007. Spatial analysis of *Arabidopsis thaliana* gene expression in response to *Turnip mosaic virus* infection. *Mol. Plant-Microbe Interact.* 4:358-370.

AUTHOR-RECOMMENDED INTERNET RESOURCES

Broad Institute software archive:
www-genome.wi.mit.edu/genome_software/other/primer3.html
 geNorm website: medgen.ugent.be/~jvdesomp/genorm

There were errors on page 296 in the original published manuscript. In the *Potato virus Y*^{(NTN)g} row in Table 2, Ref was moved to the EF1 α column. On the same page, reference to Baebler et al. 2009 regarding use of CYTc-2 as a reference gene has been removed. Corrections were made on March 29, 2011.