Planta
DOI 10.1007/s00425-012-1682-2

ORIGINAL ARTICLE

A critique of widely used normalization software tools and an alternative method to identify reliable reference genes in red clover (*Trifolium pratense* L.)

Khosro Mehdi Khanlou · Erik Van Bockstaele

Received: 5 April 2012 / Accepted: 25 May 2012

Abstract Determination of appropriate reference genes is crucial to normalization of gene expression data and prevention of biased results in qRT-PCR studies. This study is the first attempt to systematically compare potential reference genes to detect the most constitutively expressed reference genes for accurate normalization in red clover tissues including leaves, stems and roots. To identify the best-suited reference gene(s) for normalization, several statistical algorithms such as geNorm, BestKeeper and NormFinder have been developed. All these algorithms are based on the key assumption that none of the investigated candidate reference genes show systematic variation in their expression profile across the samples being considered. However, this assumption is likely to be violated in practice. The authors therefore suggest a simple and novel stability index based on the analysis of variance model which is free from the assumption made by the algorithms. We assessed the expression stability of eight candidate reference genes including actin (*ACT*), glyceraldehyde-3-phosphate-dehydrogenase (*GADPH*), elongation factor-1alpha (*EF-1α*), translation initiation factor (*EIF-4a*), ubiquitin-conjugating enzyme E2 (*UBC2*), polyubiquitin (*UBQ10*), sand family protein (*SAND*) and yellow-leaf-specific protein 8 (*YLS8*). Our results indicated that *UBC2* and *UBQ10* ranked as the two most stably expressed genes in leaf tissue. *UBC2* and *YLS8* were defined as optimal control genes for stem tissue. *EIF-4a* and *UBC2* were found to be the most stable reference gene for root tissue. *GAPDH* and *SAND* showed relatively low stability in expression study of red clover. When all tested tissues were considered, we observed that *YLS8* and *UBC2* showed remarkable stability in their expression level across tissues.

Keywords qRT-PCR · Red clover · Reference genes · Systematic validation

Abbreviations

ACT	Actin
ANOVA	Analysis of variance
CRG	Candidate reference gene
EF-1α	Elongation factor-1alpha
EIF-4a	Translation initiation factor
GADPH	Glyceraldehyde-3-phosphate-dehydrogenase
qRT-PCR	Quantitative reverse transcription real-time polymerase chain reaction
RG	Reference gene
SAND	Sand family protein
UBC2	Ubiquitin-conjugating enzyme E2
UBQ10	Polyubiquitin
YLS8	Yellow-leaf-specific protein 8

Introduction

Quantitative reverse transcription real-time polymerase chain reaction (qRT-PCR) has become the most popular method for detection and quantification of mRNA transcripts. In comparison with conventional methods of gene expression analysis such as Northern blot, ribonuclease

K. Mehdi Khanlou · E. Van Bockstaele
Department of Plant Production, Ghent University, Coupure Links 653, 9000 Ghent, Belgium

K. Mehdi Khanlou (✉) · E. Van Bockstaele
Plant Sciences Unit, Institute for Agricultural and Fisheries Research (ILVO), Caritasstraat 21, 9090 Melle, Belgium
e-mail: mehdikhanlou.khosro@ugent.be; mehdikhanloo@yahoo.com

Published online: 21 June 2012

protection assay and reverse transcriptase PCR assay, qRT-PCR has the advantages of high sensitivity, specificity, dynamic range and throughput capacity (Vandesompele et al. 2002). However, like conventional methods, the measured gene expression variation in qRT-PCR analysis consists of true biological variation generated by the phenotype or under-investigated phenomena (Vandesompele et al. 2009) as well as non-biological variations caused by differences in the quantity and integrity of the RNA template, RNA recovery and efficiency of cDNA synthesis (Andersen et al. 2004; Exposito-Rodriguez et al. 2008). These experimental errors (non-biological variations) should be eliminated as much as possible to quantify gene expression at a trustworthy level. Numerous computational methods (so-called normalization methods) have been introduced to remove or minimize experimentally caused errors, leaving only the true variants. Among the suggested normalization methods, the use of the internal control gene is most acknowledged. During the pre-genomic era, the genes that have housekeeping roles in basal cell activities were chosen as internal control for expression profile studies. Later on, however, further investigations proved that the expression level of some of the best-known and most commonly used control genes, which were supposed to be uniformly expressed, were actually found to be unstable in expression level across different treatments, biological processes and even across different tissues (Schmittgen and Zakrajsek 2000; Radonic et al. 2004; Hu et al. 2009). In other words, having a housekeeping function does not always mean that the gene has stable expression level in the cell. For normalization purposes, the use of an unstably expressed housekeeping gene can lead to severe misinterpretation of the results (Remans et al. 2008). This has convinced researchers to carefully validate the stability of the expression level of housekeeping genes prior to using them for normalization in qRT-PCR data analysis. The identification of experimentally validated control genes (so-called reference genes) that guarantee the accurate quantification of expression level of the target gene(s) (Lovdal and Lillo 2009) has been a challenge. Several statistical algorithms including geNorm (Vandesompele et al. 2002), BestKeeper (Pfaffl et al. 2004) and NormFinder (Andersen et al. 2004) have been developed to address this conundrum. geNorm is one of the most frequently used algorithms. It is based on the criterion that two ideal reference genes (RGs) should have a minimal expression ratio across the investigated sample set, regardless of cell type or condition. The geNorm software defines the gene expression stability value (M), which is calculated as the mean standard deviation of the logarithmic transformed expression ratios across samples for the particular gene relative to other genes under investigation. Then geNorm eliminates the gene with the highest M value (i.e., the least stable gene) from the panel and recalculates new M values for remaining genes until reaching the last two genes with the smallest M value (i.e., the most stable genes). The BestKeeper program is also an Excel-based tool to select best-suited RGs by performing a pair-wise correlation analysis of candidate reference genes (CRGs) (Pfaffl et al. 2004). First, descriptive statistics of the raw cycle threshold (C_t) values including standard deviation (SD) and coefficient of variance (CV, %) are computed for all the RGs. In the next step, Pearson's coefficients of correlation are calculated for each CRG pair. The highly correlated CRGs, which have a low standard deviation (SD < 1), are then combined into index value (i.e., normalization factor) using the geometric mean of their C_t values. Finally, BestKeeper determines the correlation coefficient of each CRG with the index value (called the "BestKeeper index"), along with the probability (P) value, which indicates the significance of the obtained correlation coefficient. In this way, it can present the contribution of each CRG over BestKeeper index. The CRG that is best correlated to the BestKeeper index is considered to be the most stable. The NormFinder program is another algorithm that allows ranking CRGs according to their expression stability. NormFinder calculates the stability index using analysis of variance (ANOVA) on log-transformed expression values. Using NormFinder, the proportion of both components of overall variance (intra- and intergroup variation) can be taken into account to calculate the normalization factor (NF). To do so, the NormFinder package considers intra- and intergroup variations in determination of the expression stability in a given set of CRGs. However, one should bear in mind that the accuracy of these algorithm outputs rests on the assumption that none of the tested CRGs shows any systematic variation in expression. The existence of CRGs with systematic variations among investigated CRGs could lead to false selection of RGs, which would result in erroneous normalization. We believe that it is unlikely that all investigated CRGs would have unbiased expression, because experimental conditions could influence CRGs expressions. The stability of expression of individual CRGs cannot be presumed prior to being subjected to the condition under investigation. To overcome this problem, we believe that it is necessary to use a statistical method which could evaluate each CRG independent of the assumption of whether the expressions of studied CRGs have systematic variation or not. For this purpose, we have created a new stability index based on the ANOVA model.

Red clover (*Trifolium pretense* L.) has drawn considerable attention as an important source of secondary metabolites (Saviranta et al. 2008). qRT-PCR is the most robust tool available for in-depth studies of the functions of genes involved in the metabolic pathway of the secondary metabolites in red clover. This tool also enables careful quantification of the expression level of the genes during bioengineering manipulation research, aiming to increase

exploitation of the metabolites' advantages. To date, expression studies performed on red clover employed only one control gene for normalization. No systematic study has been carried out to verify the expression stability of the gene used or to identify more RGs.

The aims of the present study were (1) to discuss the potential pitfalls arising from uncritical use of the above-mentioned statistical algorithms when the existence variation in the expression of CRGs violates the underlying assumption, (2) to propose a novel stability index for ranking CRGs, and (3) to systematically validate a panel of putative RGs and identify the useful genes for data normalization in future gene expression studies among different tissues in red clover.

Materials and methods

Plant materials

Regeneration ability is a prerequisite for any genetic manipulation study. The capacity for plant regeneration in red clover is under genetic control, and not all red clover cultivars have capability of regeneration. We therefore used NEWRC red clover germplasm, which is a highly regenerable population. The seeds were sown in pots containing a mixture of sand, expanded clay, and soil (1:1:1, by weight) and were kept under greenhouse conditions at 25 °C with 16 h per day illumination of 400 μmol m^{-2} s^{-1}. The leaves, stems and roots of ten plants were separately sampled at the late vegetative stage and then immediately frozen at −80 °C.

Total RNA isolation and cDNA preparation

Frozen tissue samples were ground with a Retsch Tissuelyser (Qiagen). Total RNA was isolated using the RNeasy Plant Mini Kit (Qiagen) using the manufacturer's instructions. Residual DNA was removed with TURBO™ DNase (Ambion) following the manufacturer's instructions. The quantity and quality of total RNA determined using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies). Hereafter, equal amounts of the different tissue samples of total RNA (50 ng) were reverse transcribed into cDNA in a total reaction volume of 20 μl using the Superscript VILO cDNA synthesis kit (Invitrogen) according to the manufacturer's protocol and then the cDNA was diluted 1:10 for qRT-PCR.

Primer design and verification of amplified products

Sequences of RGs of *Arabidopsis* were BLASTed against the red clover EST database from the National Center for Biotechnology Information (NCBI) to identify putative orthologues. Primers were designed according to the red clover EST sequences which had presented high similarity to the *Arabidopsis* RGs using Vector NTI Advance™ version 11 (Invitrogen). The GC content of the primers ranges from 45 to 55 % and the melting temperatures (T_m value) are between 59 and 60 °C. Hairpin structures and primer-dimer formation of all the primers were evaluated using Vector NTI Advance™ version 11. All primer pairs were custom-ordered from a commercial supplier (Invitrogen). High Fidelity PCR was performed in 50-μl volumes that contained 100 ng of cDNA, 1× High Fidelity PCR buffer, 0.2 mM dNTP mix, 0.25 μM each primer, 2 mM MgSO4 and 1 U Platinum Taq High Fidelity (Invitrogen). Amplification was carried out in a GeneAmp 9600 thermocycler (Applied Biosystems) with the program 94 °C for 2 min, 30 cycles of 94 °C for 30 s, 60 °C for 30 s, 68 °C for 2 min and 68 °C for 10 min. Running an aliquot of the PCR products on a 1.5 % agarose gel revealed the presence of single products of appropriate size, which confirmed the specificity of all primers (data not shown). To verify the sequences of amplification products, the remaining PCR products were cloned with the TOPO TA Cloning Kit for sequencing (Invitrogen) according to the manufacturer's instructions. Plasmids isolated from the possible positive colonies were purified using a plasmid mini kit (Qiagen). The purified plasmids were directly sequenced using an ABI Prism BigDye Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems) with the M13 forward and reverse universal primers following the manufacturer's instructions. The sequencing PCR was conducted as follows: 1 min at 94 °C followed by 25 cycles of 94 °C for 10 s, 58 °C for 5 s, 60 °C for 1 min with a final extension at 60 °C for 4 min in the above-mentioned thermocycler. The sequencing reactions were subjected to ethanol precipitation to remove excess dye-terminators. The sequencing products were dissolved in formamide and then heat-denatured for 8 min before loading onto a capillary ABI 3130 Genetic Analyzer system. Bases were determined, and sequences assembled using Sequencing Analysis software v.5.2 (Applied Biosystems). The nucleotide sequences obtained were aligned against the reference sequences of red clover EST database using Vector NTI Advance™ version 11.

Determination of primer efficiency

The TOPO-cloned amplicons with identical sequences linearized with the restriction enzyme *Eco*RV. The digested plasmids were separated on 1 % agarose gel, isolated and then purified by Qiagen Gel extraction kit in accordance with the instructions of the manufacturer. The primer efficiencies (E) of each primer pair were calculated from

10-fold dilution series of linearized plasmid containing the respective sequence-verified amplicons over 5 dilution points that were measured in duplicates. The dilution series was used to generate a standard curve by plotting the quantification cycle for each dilution point against log transformed of dilution series of input template (linearized plasmid). The slope produced by the standard curve used to calculate the primer efficiency was according to the formula $E = 10^{(-1/\text{slope})}$, as previously described (Pfaffl 2001).

qRT-PCR conditions

Quantitative real-time PCRs were carried out in 384-well plates using the LightCycler 480 (Roche). The reactions were performed in duplicate and consisted of 5 µl of LightCycler 480 SYBR Green I Master, 1 µl deionized water, 1 µl of each of the 10 µM forward and reverse gene-specific primers and 2 µl 10-fold diluted cDNA as template in a final volume of 10 µl. To ensure the absence of DNA contamination in the reaction mixture, no-template control (NTC) was included in each run. Cycling parameters were 95 °C for 10 min to denature the DNA templates and activate polymerase enzyme, followed by 45 cycles of 95 °C for 10 s, 60 °C for 12 s and a final extension of 72 °C for 10 s. Detection of fluorescence was performed at the end of the annealing period of each cycle. Immediately after the amplification, the specificity of each amplicon was verified via generation of a melting curve by heating to 95 °C for 5 s, cooling to 65 °C for 1 m and slowly heating to 97 °C with a continuous fluorescence data collection of 10 acquisitions per °C. Also, qRT-PCR assays were performed on equivalent amounts of total RNA without reverse transcription (as template) for each sample in accordance to the abovementioned reaction mixture and the thermal profile to verify the absence of genomic DNA contamination in the extracted total RNA.

Data analysis

Expression levels of the eight housekeeping genes were determined by the number of cycles required for the amplification-related fluorescence to reach a particular threshold level of detection. The raw C_t values were exported into qBASE software version 1.3.5 (Hellemans et al. 2007), corrected by PCR efficiency, and then transformed into relative expression. The relative expressions were then imported into geNorm version 3.5 (Vandesompele et al. 2002) and NormFinder version 0.953 (Andersen et al. 2004) as described in their manual. Unlike geNorm and NormFinder, raw C_t values were imported directly into BestKeeper version 1 (Pfaffl et al. 2004) as inputs. To examine the stability of each CRG independently, the relative expression ratios (R) were calculated for each CRG using the formula $R = E^{-C_{t\ \text{sample}}} / E^{-C_{t\ \text{calibrator}}}$ where E is the efficiency of the gene amplification of corresponding CRG and the calibrator is the sample with the highest expression (or lowest C_t value) within samples of the corresponding CRG. The relative expression ratios were natural log transformed and then subjected to a one-way ANOVA. The mean square of groups (MS_B) which represents intergroup variance caused by difference among tissue samples and mean square errors (MS_W), which represents intragroup variance, were computed. Then the MS_B and MS_W of each CRG were divided by $(-1/\bar{X})$ to calculate intergroup variation index (V_B) and intragroup variation index (V_W) of each gene, respectively, where $\bar{X}$ is average of natural logarithmic transformed relative expression ratios for the corresponding CRG. A stability index was subsequently generated by multiplying V_B by V_W for each gene. The lower the stability value, the more stable the gene is over the investigated tissues. The statistical analysis including correlation analysis and ANOVA were performed using SAS 9.2 (SAS Institute Inc., Cary, NC, USA).

Results

Identification of red clover candidate reference genes

In the present study, eight frequently-used CRGs [i.e., actin (*ACT*), glyceraldehyde-3-phosphate-dehydrogenase (*GAPDH*), elongation factor-1alpha (*EF-1α*), translation initiation factor (*EIF-4a*), ubiquitin-conjugating enzyme E2 (*UBC2*), polyubiquitin (*UBQ10*), sand family protein (*SAND*) and yellow-leaf-specific protein 8 (*YLS8*)] were chosen to identify the most constitutively expressed RGs. Except for *ACT* (Kim et al. 2003; Sullivan 2009a) and *EIF-4a* (Sullivan 2009b), which have already been reported as being used to normalize relative expression of genes of interest in red clover, BLAST was used to determine the sequences of other CRGs. The genes tested were the most commonly used RGs which found to be stably expressed in a genome-wide study in *Arabidopsis* (Czechowski et al. 2005) and a collection of ESTs obtained from 3-week-old red clover plants. The genes with a similarity higher than 1e−92 (E value) were considered as being putative orthologous to the *Arabidopsis* genes (Table 1).

PCR efficiency and amplification specificity

High-quality total RNA from 30 samples was isolated from different tissues including leaves, stem and roots. The obtained RNA was reverse transcribed and then used to assess the expression stability of the CRGs. To confirm the

Table 1 Red clover candidate reference genes and comparison with *Arabidopsis* orthologues

Symbol	GenBank accession number	Primer sequences	*Arabidopsis* orthologous locus	Amplicon length (bp)	Amplification efficiency	Similarity (*E* value)
ACT	AY372368	ATGAGCTTCCTGATGGACAG CCAGCAGCTTCCATTCCAAT	–	100	1.93	–
EF-1α	BB903271	GCGTGTGATTGAGAGGTTTG ACGTTCAGCCTTGAGCTTGT	AT5G60390	94	1.98	2e−166
EIF-4a	BB919542	GACCTGTTGGCTCGTGGTAT TCAGGTTGGGTAGGCAAATC	–	74	1.95	–
GAPDH	BB911051	CGGAATCGTTGAGGGTCTTA TTCCACCTCTCCAGTCCTTG	AT1G13440	95	1.90	2e−157
SAND	BB902537	TTATGCAACAAGGCAAGCTG TCTGAGCGCCAACAAGACTA	AT2G28390	113	1.86	3e−92
UBC2	BB916552	TCCAAACCCAAACTCTCCAG CTGCTCAACAATCTCGCGTA	AT2G02760	97	1.85	2e−129
UBQ10	BB903576	ACCTTGTCTTGCGTCTTCGT TCTTGGATCTTGGCCTTGAC	AT4G05320	121	1.99	1e−169
YLS8	BB917276	CCCTTCCTTGCACCTCTGTA CACCGGAAACAACAACAAGA	AT5G08290	93	1.88	2e−121

specificity of transcript amplification, the PCR-generated products were analyzed using agarose gel electrophoresis after ethidium bromide staining. All primer pairs generated a single band of the desired size. The identities of all PCR products were also verified by TOPO T/A cloning and subsequent sequencing (data not shown). Furthermore, the presence of a single peak in the melting curve, without formation of primer-dimers, also confirmed the good specificity of amplicons generated from the primer sets used in qRT-PCR (figures not shown). To determine the amplification efficiency for each primer pair, a standard curve was generated by 5 serial 10-fold dilutions of the cloned amplicons in a qRT-PCR assay. The threshold cycles (C_t) obtained for duplicates of each dilution were plotted against the log of the arbitrary numbers representing the 10-fold dilutions difference, as the concentration is unknown and not relevant. The slope of the line linking all points was then used to calculate the efficiency. The efficiency values of selected primers ranged from 1.85 for *UBC2* to 1.99 for *UBQ10*, which meets standard guidelines of acceptability (Bookout et al. 2006) (Table 1).

Descriptive statistics and relations of the raw C_t values

In our study, the obtained C_t values varied for each CRG (Fig. 1). Among the selected genes, the gene encoding *UBQ10* had the most abundant transcript level (C_t ranged between 21.57 and 25.45 across all tissues studied). In comparison, the gene encoding *SAND* had the least abundant transcript level (C_t between 29.49 and 35.09 across all tissues considered). Of the achieved C_t values of all of the CRGs studied, none had an invariant level of gene expression. This highlighted the importance of seeking appropriate RGs via statistical approaches to normalize the relative expression of genes of interest for red clover. To describe the degree of relationship, the raw C_t values of investigated CRGs were also used to perform Pearson pairwise correlation tests among genes in each tissue. A highly significant correlation ($P < 0.001$, $r = 0.91$) between *YLS8* and *UBQ10* in leaf tissue was observed. A highly significant correlation ($P < 0.001$, $r = 0.92$) was found between *EF-1α* and *ACT* in stem tissue. Also, *UBC2* and *YLS8* had a highly significant correlation ($P < 0.001$, $r = 0.90$) in root tissue (Table 2).

Analysis of gene expression stability using geNorm

The relative expression values calculated by qBASE for the eight CRGs were subjected to the geNorm algorithm to evaluate their reliability as references (Fig. 2). The *M* values for all CRGs across tissues were <1; this fell below the default limit of $M = 1.5$ (Vandesompele et al. 2002). The *UBQ10* and *YLS8* genes ranked highest in the samples gathered from leaves, with *M* values of 0.29 and 0.33, respectively (Fig. 2a). In stems, the *ACT* and *EF-1α* genes proved to be the best candidates for normalization, with *M* values of 0.14 and 0.18, respectively (Fig. 2b). We determined that *UBC2* and *YLS8* genes were most stably expressed in the red clover root samples, with an *M* value of 0.28 and 0.31, respectively (Fig. 2c). Notably,

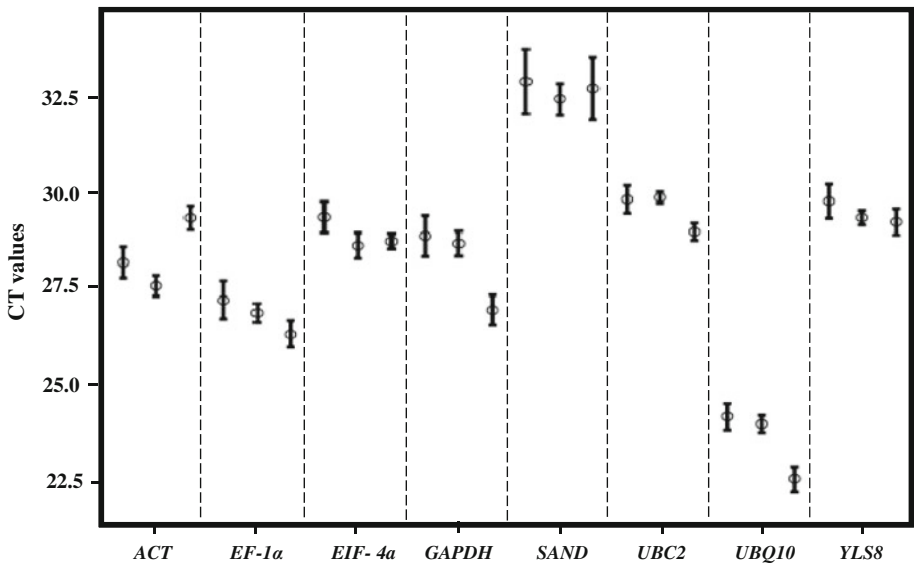

Fig. 1 Distribution of cycle threshold (C_t) values for the candidate reference genes in red clover. *Error bars* show the range of C_t values of leaf (*left*), stem (*center*) and root (*right*) for each gene

Table 2 Pairwise correlation among candidate reference genes in leaf (a), stem (b) and root (c)

	ACT	*EF-1α*	*EIF-4a*	*GAPDH*	*SAND*	*UBC2*	*UBQ10*	*YLS8*
a								
ACT	1							
EF-1α	0.79**	1						
EIF-4a	0.78*	0.71**	1					
GAPDH	0.62**	0.64**	0.49*	1				
SAND	0.44	0.32	0.04	0.05	1			
UBC2	0.39	0.37	0.06	0.15	0.84**	1		
UBQ10	0.78**	0.87**	0.66**	0.69**	0.51*	0.66**	1	
YLS8	0.82**	0.79**	0.66**	0.47*	0.70**	0.68**	0.91**	1
b								
ACT	1							
EF-1α	0.92**	1						
EIF-4a	0.77**	0.74**	1					
GAPDH	0.82**	0.82**	0.77**	1				
SAND	−0.13	−0.08	−0.48*	0.03	1			
UBC2	0.66**	0.41	0.57**	0.32	−0.31	1		
UBQ10	0.64**	0.49*	0.30	0.39	0.17	0.56*	1	
YLS8	0.75**	0.66**	0.51*	0.46*	0.15	0.66**	0.59**	1
c								
ACT	1							
EF-1α	0.71**	1						
EIF-4a	0.54*	0.67**	1					
GAPDH	0.68**	0.83**	0.50*	1				
SAND	0.34	0.40	0.11	0.52*	1			
UBC2	0.62**	0.82**	0.48*	0.59**	0.67**	1		
UBQ10	0.40	0.69**	0.41	0.55*	0.30	0.68**	1	
YLS8	0.59**	0.72**	0.32	0.56**	0.55*	0.90**	0.83**	1

Correlation analyses were carried out based on the C_t values of the eight candidate reference genes

* Correlations below the significance threshold of $P > 0.05$. ** Correlations below the significance threshold of $P > 0.01$

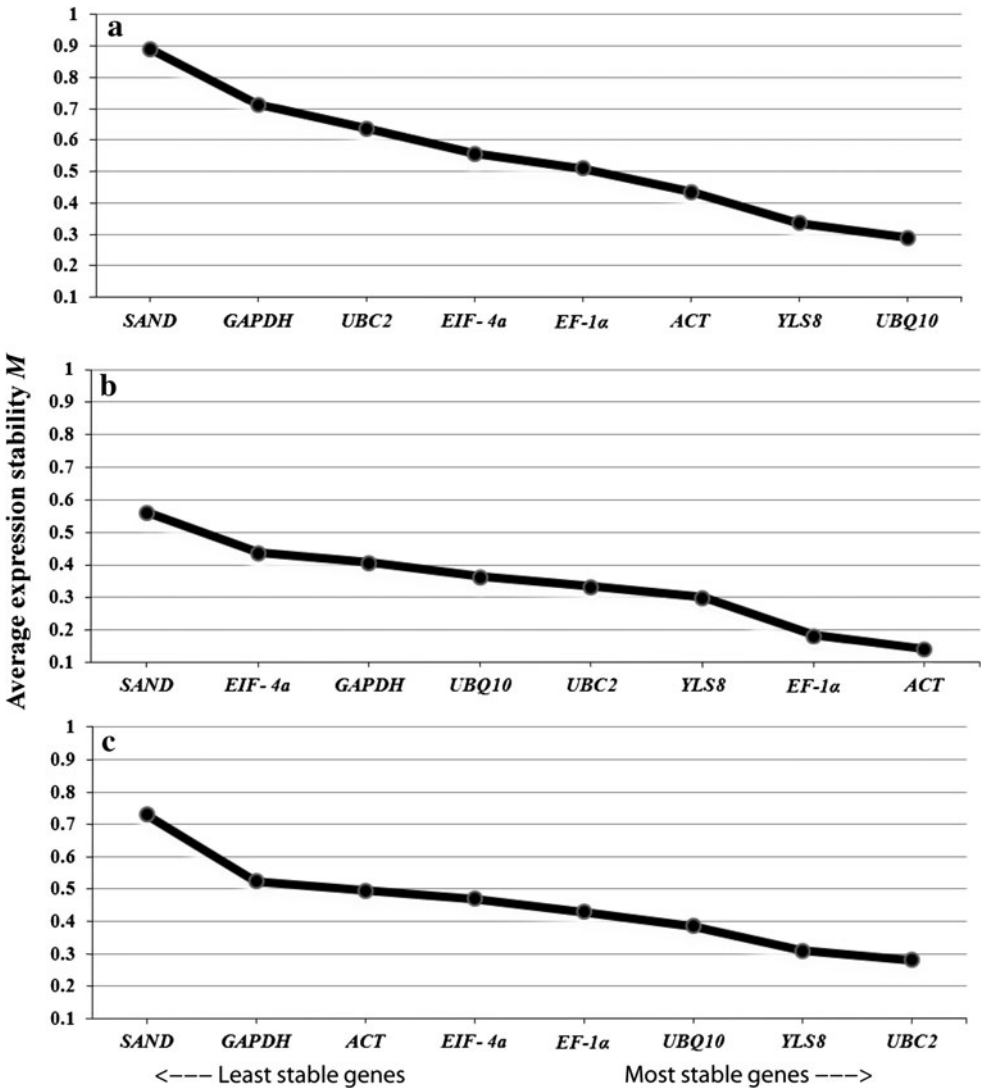

Fig. 2 Average expression stability values (*M*) of red clover candidate reference genes. *M* values of the candidate reference genes were calculated using the geNorm algorithm. Ranking of the stability was carried out on leaf (**a**), stem (**b**) and root (**c**). Lower *M* values indicate the more stably expressed genes

using geNorm, *SAND* was found to be the least stable among the genes examined in any of the red clover tissues.

Analysis of gene expression stability using BestKeeper

The raw C_t value generated by a real-time PCR platform was used as input file in the Excel-based spreadsheet of BestKeeper applet to determine the stability of eight CRGs. The result of the correlation of each eight CRG with BestKeeper Index is presented in Table 3. These results showed that *UBQ10* ($r = 0.96$) in leaves, *ACT* ($r = 0.94$) in stems and *YLS8* ($r = 0.91$) in roots presented the highest coefficient of correlation with the BestKeeper index (i.e., the highest stability). *SAND* demonstrated the highest variation in all tissues, especially in leaves and roots with a standard deviation of greater than 1 (SD > 1). *SAND* was therefore excluded from the BestKeeper index calculation in leaf and root tissue.

Analysis of gene expression stability using NormFinder

The NormFinder algorithm was performed on the data where the sample sets are considered as three groups; the sample set was divided into three groups based on the tissue type (leaves, stems and roots). Plotting inter- and intragroup variations for each gene showed that *ACT* had the highest intergroup variation and *SAND* presented the highest intragroup variation (Fig. 3). The stability index for each gene was calculated according to corresponding inter- and intragroup variations. The obtained result suggested that *YLS8* (stability index of 0.115) and *EF-1α* (stability index of 0.172) was the best combination of two RGs (Table 4).

Table 3 The expression stability values for red clover reference genes calculated by the BestKeeper algorithm

Tissue	BestKeeper versus	*ACT*	*EF-1α*	*EIF-4a*	*GAPDH*	*SAND*	*UBC2*	*UBQ10*	*YLS8*
Leaf	Coeff. of corr. (*r*)	0.88	0.86	0.66	0.65	0.66	0.68	0.96	0.95
	P value	0.001	0.001	0.001	0.002	0.002	0.001	0.001	0.001
Stem	Coeff. of corr. (*r*)	0.94	0.89	0.73	0.86	0.13	0.61	0.71	0.81
	P value	0.001	0.001	0.001	0.001	0.603	0.005	0.001	0.001
Root	Coeff. of corr. (*r*)	0.74	0.89	0.56	0.84	0.72	0.91	0.75	0.87
	P value	0.001	0.001	0.011	0.001	0.001	0.001	0.003	0.001

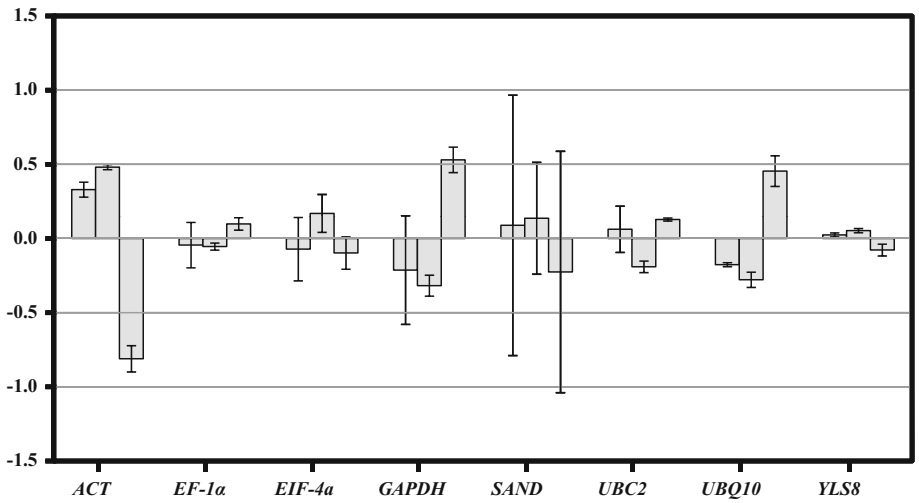

Fig. 3 Plotting the inter- and intragroup variances calculated by NormFinder for candidate reference genes in three tissues of red clover. For each gene, the intergroup variations of leaf (*left*), stem (*center*) and root (*right*) are shown as *histogram bars* and intragroup variation as *error bars*

Table 4 Stability values of candidate reference genes according to their expression profiles across tested tissues (leaf, stem and root) of red clover including as calculated by NormFinder

Candidate reference gene	Stability index
ACT	0.612
EF-1α	0.172
ELF-4a	0.268
GAPDH	0.487
SAND	0.428
UBC2	0.224
UBQ10	0.385
YLS8	0.115

Analysis of gene expression stability using ANOVA

Each investigated gene was evaluated independently for possible differences in expression among tissues using one-way ANOVA. *UBQ10* had the highest *F* value (or lowest *P* value) at 18.33, indicating a significant difference in gene expression among tissues. *SAND* had the lowest *F* value (0.22), indicating no difference among tissues tested (Table 5). Also, both intra and intergroup variances were broken down into their component parts across tissues. The results show that *UBC2* had the lowest variance (0.070 and 0.011 in leaf and stem, respectively), *EIF-4a* had the lowest variance in root (0.022) and *SAND* had the highest variances in all three tissues. The stability indexes were calculated as described in "Materials and methods." A CRG with a low stability index is considered to be most stable across all tested tissues. Among the CRGs under study, *YLS8* (stability index of 0.028) was the most stable; *UBC2* was second most stable (stability index of 0.07) and *EIF-4a* was third most stable (stability index of 0.12) (Table 5).

Discussion

Numerous sources of variation, including the total RNA content of the sample, the number of cells in the starting material, the RNA extraction efficiency, differential enzymatic efficiencies and transcriptional activity resources, could introduce bias in evaluating the expression profile of the investigated genes. Therefore, a strategy is required to minimize the effect of any non-biological variation as much as possible in order to monitor the actual intrinsic variation of desirable genes across a given sample set. For this purpose, several strategies (so-called normalization

Table 5 Determination of the stability of candidate reference genes based on analysis of variance (ANOVA) technique

Genes	Source of variation	MS (Variance) components			Total MS (Variance)	F	Sig.	$(-1/\bar{X})$	V_B^1	V_W^2	Stability index
		Leaves	Stems	Roots							
ACT	Between groups	0.066	1.201	1.826	3.093	15.00**	0.000	0.959	3.22	0.21	0.692
	Within groups	0.104	0.043	0.059	0.206						
EF-1α	Between groups	0.337	0.011	0.463	0.811	2.86	0.074	0.704	1.15	0.40	0.461
	Within groups	0.165	0.042	0.076	0.283						
EIF-4a	Between groups	0.411	0.161	0.058	0.630	3.21	0.056	1.009	0.62	0.19	0.120
	Within groups	0.102	0.072	0.022	0.196						
GAPDH	Between groups	0.916	0.477	2.711	4.104	12.70**	0.000	0.574	7.13	0.56	4.011
	Within groups	0.167	0.062	0.092	0.323						
SAND	Between groups	0.079	0.106	0.001	0.188	0.22	0.799	0.629	0.29	1.32	0.394
	Within groups	0.384	0.089	0.356	0.831						
UBC2	Between groups	0.118	0.164	0.559	0.841	7.67**	0.002	1.142	0.73	0.09	0.070
	Within groups	0.070	0.011	0.029	0.110						
UBQ10	Between groups	0.766	0.361	2.183	3.310	18.33**	0.000	0.741	4.46	0.24	1.087
	Within groups	0.079	0.033	0.069	0.181						
YLS8	Between Groups	0.183	0.020	0.081	0.284	1.43	0.255	1.364	0.20	0.14	0.028
	Within Groups	0.113	0.019	0.066	0.198						

Intergroup variance (MS_B) and intragroup variance (MS_W) are broken down into the contributing components (tissues)

** Significant at the 0.01 level

[1] Between group variance divided by mean

[2] Within group variance divided by mean

strategies) were proposed to obtain accurate data using qRT-PCR. It appears that the use of RGs could offer more reliability and accuracy compared to other strategies; for this reason, this is now the most preferred way of normalization (Vandesompele et al. 2009). However, internal RGs should have invariant levels of gene expression across a sample set. The use of unstably expressed RGs expressed constitutively could increase errors introduced by the technical variations, and could also add more bias to the result. Therefore, prior to using internal control genes for normalization, their expression stability should be statically validated across all of the treatments/categories being studied. Several statistical algorithms such as geNorm, BestKeeper and NormFinder have been developed to identify consistently expressed genes. Each of them has pitfalls; the strength of each algorithm lies in the circumstances of its use. These algorithms also do not independently assess CRGs for their expression stability. Therefore, we believe that in addition to commonly used algorithms such as geNorm, NormFinder and BestKeeper, statistical methods which can evaluate CRGs independently are needed to provide credence to the selection of the "best" suited CRGs for superior transcript normalization of gene expression.

geNorm, developed by Vandesompele et al. (2002), is the most commonly used normalization algorithm. geNorm has two main advantages: (1) the raw data are used as an input file and does not need to be normality distributed, and (2) geNorm does not need a large sample size to provide reliable results (Serrano et al. 2011). The analysis relies on the principle that the expression ratio of two 'ideal' RGs should be identical in all samples, regardless of the experimental condition or cell type (Vandesompele et al. 2002). geNorm top ranks those genes with the highest degree of similarity in their expression profile, rather than with minimal variation (Andersen et al. 2004), because geNorm investigates the expression ratio of candidate genes and does not take the variation across sample sets into account. It is therefore crucial that the two 'ideal' RGs top ranked by geNorm should independently be regulated and not co-regulated, because co-regulated genes have a high similarity in their expression profile. To avoid selection of co-regulated genes, the tested CRGs are chosen from distinct biological processes and metabolic pathways. However, it is also possible that two CRGs show a strong bias toward either down- or up-regulation during the experiment, regardless of their origins in independent cellular functions. In other words, the experimental condition could regulate them in such a way that the genes have high similarity of expression. As a result, geNorm erroneously top ranks these genes, regardless of whether they have a high level of variation or not. However, based on the generally accepted criterion, RGs should show invariant levels of gene expression in a given test set and

experimental design. Since there is no way to presume any systematic variation of CRGs which can be caused by experimental conditions before performing the experiment, we therefore believe that independent evaluation of each CRG is necessary to reduce the risk of artificial selection of variant CRGs. Upon comparison of the results of a correlation study (Table 2) with geNorm results (Fig. 2), we observed that the CRGs with a high correlation are chosen as most stable RGs by geNorm. However, when the within-group variations of CRGs for each tissue are taken into account (Table 5), we observed that in leaves, *UBQ10* (variation of 0.079) and *YLS8* (variation of 0.113) ranked as the second and fifth out of eight CRGs. In stem tissue, *ACT* and *EF-1a* (correlation of 0.92) (Table 2) are found to be most stable by geNorm (Fig. 2). However, *EF-1a* (0.042) and *ACT* (0.043) ranked as the fourth and fifth among CRGs. The situation is better in root tissue because the selected CRGs by geNorm, *UBC2* and *YLS8* (correlation of 0.90) (Table 2), ranked as the second and third invariant CRGs. BestKeeper uses C_t values (instead of relative expressions) as input to identify the best normalizers among the investigated candidate reference. For this purpose, BestKeeper employs Pearson correlation analysis, which is a parametric method. Therefore, the Pearson correlation coefficient is only valid for normally distributed data with a homogeneous variance. If the data do not meet these assumptions, using a parametric test can thus lead to false results. Although Pfaffl et al. (2004) stated that the C_t values seem to be best estimators of the expression levels, some reports warn of subjecting C_t values to parametric methods (Schmittgen et al. 2000) due to C_t values being logarithmic in nature (Pelch et al. 2010), if ever normally distributed without a linear scale (Kortner et al. 2011). BestKeeper was developed based on the assumption of normal distribution of input data (C_t values). To calculate the BestKeeper index in the original paper, Pfaffl et al. (2004) used three CRGs, *UBQ*, *GAPD* and *β-ACT*. When we subjected the C_t values of those three CRGs to the Shapiro–Wilk normality test, we realized that *UBQ* (*P* values of 0.016) and *GAPD* (*P* values of 0.006) do not have normal distribution and *β-ACT* (*P* values of 0.052) has only a poor normal distribution. Violating normal distribution assumptions raises questions about the accuracy of the results obtained by parametric statistical methods in BestKeeper. Although the authors promised to develop a new version of BestKeeper software using Spearman and Kendall Tau correlations for non-parametric data, those versions are not yet available. The determination of standard deviation for each investigated CRG is considered as the major advantage of BestKeeper. BestKeeper then eliminates the CRGs which show high standard deviation (SD $\geq$ 1) from further calculation. However, Schmittgen et al. (2000) believe that presentation of statistical data calculated from the raw C_t values falsely represents the error and should be avoided. As earlier mentioned, BestKeeper employs pairwise correlation analysis to determine suitable RGs; BestKeeper top ranks the RGs which display similar expression patterns across samples as well as geNorm (Ponton et al. 2011). In our study, the results indicated that the CRGs were chosen as highly stable RGs in different tissues by geNorm. They were also top ranked by BestKeeper despite the existence of CRGs with less variation in the tested tissues, but they were considered to be less stable CRGs when using geNorm and BestKeeper. This result was in agreement with the previous report that *GAPDH*, which showed the lowest overall CV of all genes, was ranked second to last by geNorm and last by BestKeeper (Kortner et al. 2011).

The NormFinder algorithm uses a model-based approach which takes variations across subgroups into account and avoids artificial selection of co-regulated genes (Andersen et al. 2004). NormFinder, which is less sensitive to co-regulation of the RGs, enables ranking of all reference gene candidates based on intra- and intergroup variations. Users should understand that the intergroup variation calculated by NormFinder does not show systematic variation across the sample subgroups of a single CRG. In reality, the intergroup variation shows to which degree the data from group 1 of a particular CRG varies compared to the mean variation of data from group 1 for all tested CRGs (Kortner et al. 2011). Therefore, the accuracy of the result achieved by NormFinder is still based on the assumption that the average of the tested candidate genes should show no systematic variation. In the original paper, the stability of expression of the CRGs subjected to NormFinder was already proved by microarray analysis (Andersen et al. 2004). In practice, such microarray expression data are usually not available. If the tested CRGs show a similar systematic variation, the mean expression will be biased toward the variant genes. In this case, the top-ranked genes will not be the most stable; in reality, they are top ranked because of their similar expression to the biased mean expression. In our study, the NormFinder algorithm ranked *YLS8* and *EF-1α* as the best combination of RGs across tissues (Table 4). However, a visual inspection of the descriptive plot of C_t values (Fig. 1) and also the result of the stability index achieved by ANOVA (Table 5) revealed that *EF-1α* occupied the fifth position out of eight CRGs, thus it cannot be considered as stable reference gene. Although *ACT* ranks as the least stable CRG across tissues (Table 4) with the highest intergroup variation by NormFinder (Fig. 3), Fig. 1 and Table 5 show that *ACT* was not the least stable gene. This result is in accordance with the study of Kortner et al. (2011), who reported that *GAPDH* (the lowest overall CV of all genes) was ranked as the worst-scoring CRG by NormFinder.

All of the abovementioned algorithms are limited by the assumption that none of the tested CRGs vary systematically, but this assumption does not always reflect reality. CRGs must therefore be tested with statistical methods which are free from this assumption. In previous studies, the application of standard parametric statistical techniques (i.e., *t* test, regression, ANOVA) for independent assessment of expression of CRGs were reported that used C_t or relative expression ratios as input data. The most notable examples of these previous reports is the study performed by Brunner et al. (2004) who proposed a stability index calculated by multiplication of the coefficient of variation (CV) with the slope of regression (*b*). To determine intragroup variation, the raw C_t values were subjected to ANOVA to calculate CV via division of MSE by the mean, and then were multiplied by 100. However, as mentioned above, Schmittgen et al. (2000) disagreed with calculation of CV from the raw C_t values. To account for intergroup variation, Brunner et al. (2004) calculated the slope of regression for each gene across different groups based on their C_t values. However, the regression slope (*b*) presents only the linear part of intergroup variation. But intergroup variation consists of linear variation and nonlinear variation. Thus, we believe that the application of the linear regression slope (*b*) could cause inaccurate estimation of intergroup variation. Also, in fact, although plotting C_t is informative, the raw C_t values do not have a linear scale and also are not corrected for primer efficiency (Kortner et al. 2011). Therefore, the direct use of raw C_t values for ANOVA and linear regression could lead to poor estimation and erroneous results. A recent study conducted by Kortner et al. (2011) is another example of independent evaluation of CRGs using relative expression ratios as input file. They proposed total CV as the stability index. They argued in cases where two CRGs have the same total CV, the gene with lower intergroup variation is a suitable reference gene for normalization. However, it is unknown to which degree the intra- and intergroup variation contribute to the total variation (CV). They therefore suggest another stability index, the *F* value, which is the ratio between the intergroup and intragroup variation. They claimed that stable RGs should have small total CV and *F* values. However, this cannot always be correct. Imagine two CRGs (*x* and *y*), in which the intergroup variation of gene *x* is two times bigger than its intragroup variation. The *F* value of gene *x* thus equals to 2. The intergroup variation of gene *y* is two times bigger than the intergroup variation of gene *x*, and the intragroup variation of gene *y* is four times bigger than intergroup variation of gene *x*. Therefore, the *F* value of gene *y* equals 1. Although gene *y* has a smaller *F* value, its total CV will be much higher than gene *x*. That is why Kortner et al. (2011) were faced with discrepancies when ranking genes with total CV and *F* values. Most importantly, they performed parametric statistical methods on relative expression ratios which have no symmetrical scale (Gorte et al. 2011). Because subjecting C_t values or relative expression ratios to standard parametric statistical techniques is an egregious error that has also occurred in previous studies. It is now important to ask what kind of data can be subjected to statistical methods. When taking the formula used to calculate the *M* value in geNorm into account, one sees that the relative expression ratios were log 2 transformed. By default, NormFinder uses the logarithmic transformation of the raw relative quantities (Rytkonen et al. 2010). We therefore decided to perform natural log transformation on the relative expression ratios to model fold changes in an additive way (Szabo et al. 2004) and to satisfy the assumption of normality for ANOVA. We then propose a novel stability index based on ANOVA using the natural logarithm of the relative expression ratios as input data. We divided MS_W into its components (different tissues) which enabled us to distinguish the suitable CRGs with the lowest variation in each tissue. Moreover, when observing the components of MS_B, one can identify the tissue which has the lowest or highest contribution in intergroup variation in the corresponding CRG. Such information can help researchers in preliminary identification of candidate genes. If one wants to conduct a new qRT-PCR experiment for expression analysis in some of the tissues tested in the current study, the genes showing a lower contribution in intergroup variation can be appropriate CRGs in corresponding tissues. Also, because the components of stability index (V_B and V_W) are known in our study (Table 5), if one encounters two CRGs with very close or the same stability index, the gene with the lower V_B can be simply chosen as proper RG. Another advantage of the proposed stability index is that a limited number of CRGs can be evaluated in all algorithms, whereas any number of CRGs can be evaluated by our proposed stability index. Also, unlike NormFinder, inter- and intragroup variations do not have the same weight when computing its stability index. In our proposed stability index, these two components are equally weighted.

Selecting reliable RGs is of crucial importance for qRT-PCR studies. Several available algorithms, including geNorm, BestKeeper and NormFinder, have been developed to identify reliable RGs among the CRGs being studied. All of the algorithms are based on the assumption that none of the tested CRGs show variability in their expression across all the samples being considered. Observation in practice disproves this assumption. In our study, we revealed that in cases where the underlying assumption is violated, the application of geNorm, BestKeeper and NormFinder for selecting the best RG(s) could greatly heighten the risk of artificial selection. We propose an alternative, model-based method for computing stability

of CRGs expression which is free from the assumption used in these analysis tools. Simplicity of calculation and accuracy are the major advantages of our model-based stability index as compared with the previously developed model-based methods, because their application is either too intricate (Szabo et al. 2004; Serrano et al. 2011) or they do not take the efficiency of PCR into consideration (Brunner et al. 2004). In this study, we have attempted to alert users to the assumption and limitations of the available algorithms. Of course, if the assumption is met, the outputs of these algorithms can be reliable.

To the best of our knowledge, in previous expression studies of red clover, only *ACT* or *EIF-4a* have been used as a normalizer in leaf samples. However, neither *ACT* nor *EIF-4a* has been statistically evaluated for their expression stability. The present study is the first systematic comparison of potential RGs to determine the best-performing genes for accurate normalization in red clover tissues (i.e., leaves, stems and roots). For leaf tissue, *UBC2* and *UBQ10* showed the least variation of the eight CRGs and were thus identified as being the most stable RGs in leaf. For stem tissue, *UBC2* and *YLS8* presented the lowest variations and were thus determined to be the proper RGs for stems. For red clover root tissue, *EIF-4a* and *UBC2* showed the lowest variability; we therefore believe these to be suitable RGs for the root of red clover. It should be bear in mind that the red clover plants used for this study were kept under greenhouse conditions (see "Materials and methods" above). Sampling was performed in the late vegetative stage of the tested plants without any treatment applied. The abovementioned CRGs, which were found to be the best RGs for their corresponding tissues, can therefore only be used as normalizers for expression studies of any genes of interest needing to be evaluated under the same particular conditions. For mRNA expression normalization of the combined plant tissues (leaf, stem and root combined), *YLS8* and *UBC2* showed the highest stability for red clover kept under the abovementioned conditions and sampling protocol. *YLS8* and *UBC2* were therefore the most appropriate RGs for normalization of any genes needed to be assessed across tissues of red clover grown under these conditions. In this study, we sampled three tissue types from 10 non-treated red clover plants to evaluate the expression stability of eight CRGs using the commonly used statistical algorithms such as geNorm, BestKeeper and NormFinder as well as a new method proposed in this study. Our proposed method can be also used to identify reliable RGs among CRGs of a tissue (e.g., leaves) that has been subjected to different treatments. For this purpose, similar to the way that each tissue was considered as a group in this study, each treatment can be considered as a group. The CRGs with low within-group variations in a treatment are considered as suitable RGs in the corresponding treatment. To determine the most stable RGs across treatments, the stability index should be calculated for each CRG, in the same way that we calculated the stability index across tissues in this study. The genes with a low stability index are considered to be the best RGs across the treatments applied.

Acknowledgments This work was funded by the Institute for Agricultural and Fisheries Research (ILVO), Merelbeke, Belgium. The first author was supported by the Ministry of Science, Research and Technology of Iran. The authors gratefully acknowledge Laurence Desmet and Ariane Staelens for excellent technical assistance and Miriam Levenson for English-language correction.

References

Andersen CL, Jensen JL, Orntoft TF (2004) Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. Cancer Res 64:5245–5250

Bookout AL, Cummins CL, Mangelsdorf DJ, Pesola JM, Kramer MF, Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K (2006) High-throughput real-time quantitative reverse-transcription PCR. In: Current protocols in molecular biology. John Wiley, Sons, Inc., New York, pp 15.18.11–15.18.28

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

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

Exposito-Rodriguez M, Borges AA, Borges-Perez A, Perez JA (2008) Selection of internal control genes for quantitative real-time RT-PCR studies during tomato development process. BMC Plant Biol 8:131

Gorte M, Horstman A, Page RB, Heidstra R, Stromberg A, Boutilier K, Yuan L, Perry SE (2011) Microarray-based identification of transcription factor target genes. In: Ling Y (ed) Plant transcription factors: methods and protocols. Methods Mol Biol 754:119–141

Hellemans J, Mortier G, De Paepe A, Speleman F, Vandesompele J (2007) qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. Genome Biol 8:R19

Hu RB, Fan CM, Li HY, Zhang QZ, Fu YF (2009) Evaluation of putative reference genes for gene expression normalization in soybean by quantitative real-time RT-PCR. BMC Mol Biol 10:93

Kim BG, Kim SY, Song HS, Lee C, Hur HG, Kim SI, Ahn JH (2003) Cloning and expression of the isoflavone synthase gene (IFS-Tp) from *Trifolium pratense*. Mol Cells 15:301–306

Kortner TM, Valen EC, Kortner H, Marjara IS, Krogdahl A, Bakke AM (2011) Candidate reference genes for quantitative real-time PCR (qPCR) assays during development of a diet-related enteropathy in Atlantic salmon (*Salmo salar* L.) and the potential pitfalls of uncritical use of normalization software tools. Aquaculture 318:355–363

Lovdal T, Lillo C (2009) Reference gene selection for quantitative real-time PCR normalization in tomato subjected to nitrogen, cold, and light stress. Anal Biochem 387:238–242

Pelch KE, Schroder AL, Kimball PA, Sharpe-Timms KL, Davis JW, Nagel SC (2010) Aberrant gene expression profile in a mouse model of endometriosis mirrors that observed in women. Fertil Steril 93:1615–1627

Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res 29:e45

Pfaffl MW, Tichopad A, Prgomet C, Neuvians TP (2004) Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper—Excel-based tool using pair-wise correlations. Biotechnol Lett 26:509–515

Ponton F, Chapuis MP, Pernice M, Sword GA, Simpson SJ (2011) Evaluation of potential reference genes for reverse transcription-qPCR studies of physiological responses in *Drosophila melanogaster*. J Insect Physiol 57:840–850

Radonic A, Thulke S, Mackay IM, Landt O, Siegert W, Nitsche A (2004) Guideline to reference gene selection for quantitative real-time PCR. Biochem Biophys Res Comm 313:856–862

Remans T, Smeets K, Opdenakker K, Mathijsen D, Vangronsveld J, Cuypers A (2008) Normalisation of real-time RT-PCR gene expression measurements in *Arabidopsis thaliana* exposed to increased metal concentrations. Planta 227:1343–1349

Rytkonen KT, Renshaw GMC, Ashton KJ, Williams-Pritchard G, Leder EH, Nikinmaa M (2010) Elasmobranch qPCR reference genes: a case study of hypoxia preconditioned epaulette sharks. BMC Mol Biol 11:27

Saviranta NM, Anttonen MJ, von Wright A, Karjalainen RO (2008) Red clover (*Trifolium pratense* L.) isoflavones: determination of concentrations by plant stage, flower colour, plant part and cultivar. J Sci Food Agric 88:125–132

Schmittgen TD, Zakrajsek BA (2000) Effect of experimental treatment on housekeeping gene expression: validation by real-time, quantitative RT-PCR. J Biochem Biophys Methods 46:69–81

Schmittgen TD, Zakrajsek BA, Mills AG, Gorn V, Singer MJ, Reed MW (2000) Quantitative reverse transcription-polymerase chain reaction to study mRNA decay: comparison of endpoint and real-time methods. Anal Biochem 285:194–204

Serrano M, Moreno-Sanchez N, Gonzalez C, Marcos-Carcavilla A, Van Poucke M, Calvo JH, Salces J, Cubero J, Carabano MJ (2011) Use of maximum likelihood-mixed models to select stable reference genes: a case of heat stress response in sheep. BMC Mol Biol 12:36

Sullivan ML (2009a) Phenylalanine ammonia lyase genes in red clover: expression in whole plants and in response to yeast fungal elicitor. Biol Plant 53:301–306

Sullivan ML (2009b) A novel red clover hydroxycinnamoyl transferase has enzymatic activities consistent with a role in phaselic acid biosynthesis. Plant Physiol 150:1866–1879

Szabo A, Perou CM, Karaca M, Perreard L, Quackenbush JF, Bernard PS (2004) Statistical modeling for selecting housekeeper genes. Genome Biol 5:R59

Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol 3(0034):1

Vandesompele J, Kubista M, Pfaff MW, Logan J, Edwards K, Saunders N (2009) Reference gene validation software for improved normalization. Chapter 4. In: Logan J, Edwards K, Saunders N (eds) Real-time PCR: Current technology and applications. Caister Academic Press, London