Mol Genet Genomics (2013) 288:671–681
DOI 10.1007/s00438-013-0780-6

ORIGINAL PAPER

Validation of reference transcripts in strawberry (*Fragaria* spp.)

Maureen A. Clancy · Hernan G. Rosli · Srikar Chamala · W. Brad Barbazuk · P. Marcos Civello · Kevin M. Folta

Received: 19 April 2013 / Accepted: 29 August 2013 / Published online: 24 September 2013

Abstract Contemporary methods to assay gene expression depend on a stable set of reference transcripts for accurate quantitation. A lack of well-tested reference genes slows progress in characterizing gene expression in high-value specialty crops. In this study, a set of strawberry (*Fragaria* spp.) constitutively expressed reference genes has been identified by merging digital gene expression data with expression profiling. Constitutive reference candidates were validated using quantitative PCR and hybridization. Several transcripts have been identified that show improved stability across tissues relative to traditional reference transcripts. Results are similar between commercial octoploid strawberry and the diploid model. Our findings also show that while some never-before-used references are appropriate for most applications, even the most stable reference transcripts require careful assessment across the diverse tissues and fruit developmental states before being adopted as controls.

Keywords *Fragaria* · Gene expression · Normalization · Reference genes · Strawberry

Communicated by S. Hohmann.

M. A. Clancy · K. M. Folta (✉)
Horticultural Sciences Department, University of Florida, 1301 Fifield Hall, Gainesville, FL 32611, USA
e-mail: kfolta@ufl.edu

H. G. Rosli
Instituto de Investigaciones Biotecnológicas, Instituto Tecnológico de Chascomús (IIB-INTECH), UNSAM-CONICET, Camino Circunvalación Laguna Km 8.2 (B7130IWA), Chascomús, Buenos Aires, Argentina

S. Chamala · W. B. Barbazuk
Department of Biology, Genetics Institute, University of Florida, Gainesville, FL 32611, USA

W. B. Barbazuk · K. M. Folta
Graduate Program in Plant Cellular and Molecular Biology, University of Florida, Gainesville, FL 32611, USA

P. M. Civello
Facultad de Ciencias Exactas, Universidad Nacional de La Plata (UNLP), 47 y 115, (1900), La Plata, Buenos Aires, Argentina

P. M. Civello
INFIVE (CONICET-UNLP), Instituto de Fisiología Vegetal, Diag. 113 y Calle 61, n° 495, (1900), La Plata, Buenos Aires, Argentina

Introduction

Evaluation of gene expression has progressed from visualization of a hybridization signal to the quantification of transcript accumulation standardized to unvarying reference genes. In the analysis of reverse transcriptase quantitative PCR (RT-qPCR) data, relative steady-state transcript accumulation across experimental variables has typically been normalized to a baseline defined by a well-described constitutive reference gene. It has often been assumed that reference gene transcripts will exhibit little or no variability in their steady-state accumulation between different species, tissues, treatments or developmental states. Traditional references include ribosomal RNA, ubiquitin, actin, and elongation factor transcripts.

However, the reference genes historically employed for normalization in plant biology are not necessarily applicable across species. There is a need to identify vetted reference transcripts for PCR-based expression analyses (Udvardi et al. 2008), especially in organisms with specialized tissues or developmental states. Even in *Arabidopsis thaliana*, new transcripts were identified that expanded and improved the precision of RT-qPCR over

traditional reference transcripts (Czechowski et al. 2005). One emerging theme is that there are typically no universally applicable truly constitutive reference/housekeeping genes, but some have been defined that work well across broad contexts (Klie and Debener 2011).

The goal of the present work is to test the expression stability of empirically selected reference transcripts in strawberry (*Fragaria* spp.) across multiple tissue types, including different ploidy levels, both in wild and cultivated materials. Promising candidates in strawberry were originally defined through analysis of next-generation transcript sequencing data (Folta et al. 2010). A series of constitutively expressed gene candidates were defined based on their approximately equal read representation across strawberry tissues.

This analysis tests the expression of transcripts identified in the earlier study for consistent steady-state levels across strawberry tissue types. The list of candidates includes glyceraldehyde 3-phosphate dehydrogenase (*FaGAPDH1*), a transcript often used in normalization. Two common reference genes, ubiquitin and actin, were assessed for stability via hybridization, and showed substantial variation between tissues, indicating that they would be unsuitable for normalization in strawberry studies. The best transcripts for normalization proved to be those identified from RNA sequencing efforts, and represent mRNAs from some uncharacterized genes. The findings indicate that traditional reference transcripts chosen are generally less reliable and that new candidates derived from high-throughput sequencing provide an improved reference set for RNA analysis in strawberry. Our results show that while expression of these new candidates proved remarkably stable, strawberry's diverse tissues preclude the identification of any single transcript to function as a universal control.

Materials and methods

Plant materials

Plants and dissected tissues of the octoploid cultivated strawberry *F.* × *ananassa* (cv. Strawberry Festival) and the diploid strawberry *F. vesca semperflorens* (Hawaii-4; GRIN accession PI551572, FRA197) were utilized in these trials. The octoploid line was chosen because it was employed in analysis of the strawberry transcriptome (Folta et al. 2010). The diploid line Hawaii-4 is the accession with published genome sequence (Shulaev et al. 2011) and has been widely used in transgenic studies (Haymes and Davis 1998; Oosumi et al. 2006). Strawberry plants were grown in greenhouse conditions, with approximately a 12-h photoperiod and a temperature between 18 and 26 °C. RNA for dot blots and qRT-PCR was isolated using the CTAB-based method as described (Folta et al. 2010). Three independent replicates were obtained from each genotype for each treatment or tissue type.

Candidate selection

The strawberry transcript candidates selected for further evaluation were obtained from digital expression profiling (Folta et al. 2010) and are presented in Table 1. Candidates were identified as transcripts with a statistically equal frequency of reads across tissues tested.

RNA blotting and hybridization

Initial evaluation of each candidate reference transcript employed hybridization probing of RNA dot blots. A 'master mix' was prepared for each RNA sample to ensure reproducible loading of equal RNA quantities into replicate dot blot wells. Total RNA (300 ng RNA per target spot) was diluted in DEPC-treated water and sample master mixes assembled on ice. An equal amount of RNA denaturing solution containing formamide, formaldehyde and MOPS (Sambrook and Russell 2001) was added, the mixture was heated to 65 °C for 10 min, then flash-cooled on ice. Two volumes of ice-cold 10× SSC were added and denatured samples maintained on ice until loading into the wells of the dot blot apparatus. The positively charged nylon membrane (GeneScreen Plus, Perkin Elmer; Waltham, MA, USA) was marked for orientation and placed into the Microsample Filtration Manifold following manufacturer's instructions (Schleicher and Schuell, product No. SRC096/0). Denatured samples were loaded into assigned positions of the manifold and drawn through with a gentle vacuum. After disassembly of the apparatus, applied RNA was fixed to the nylon membrane by exposure to UV light (UV Crosslinker, Spectroline, Westbury, NY, USA).

Each blot of RNA target samples included 500–600 ng denatured salmon sperm DNA to serve as a non-specific hybridization control. Also, 0.1–1.0 ng of the PCR product corresponding to the hybridization probe was applied as a positive hybridization control to confirm conditions of probe excess. The positive control also served to confirm that eventual signal detection or/and quantitation was not saturated. Either of these events could produce false indications of a constitutive signal.

DNA fragments for use as hybridization probes were synthesized by PCR using the primers presented in Table 1. Radioactive probes were generated by random priming of purified PCR products with the Prime-a-Gene Labeling System (Promega Corp; Fitchburg, WI USA). Denatured radiolabeled probes were incubated overnight at 65 °C with RNA blots in hybridization solution (Church

Table 1 Primer sequences used to synthesize DNA amplicons for radiolabeling

Gene name/Contig number[a]	Transcript expression level[a]	Putative gene function[b]	Amplicon size (bp)	Primer pairs for probe synthesis
Candidate transcripts				
FaCHP1 (contig-21335)	Medium	Conserved hypothetical protein	593	5′-CATCTGCGAGGGCCACACCA-37′/
				5′-TCCTGTGAGTGCCTGGGATGAGA-3′
FaAATP1 (contig-18391)	Medium	Plastid ATP/ADP transport protein	1,134	5′-GCTGGTGAAAGTGGGCCACCA-3′/
				5′-GAGCTGAGGCTGCTGCTGCG-3′
FaHDP1 (contig-26403)	Medium	Homeobox domain protein	625	5′-GCAACGGATCAAACAGACCGCCT-3′/
				5′-TCCAAAGCACATGCACTCACCCT-3′
FaCHP2 (contig-l8191)	Low	Conserved hypothetical protein	690	5′-CCTCCTGCTCTGTCTACCTGCCT-3′/
				5′-AGCTGCTGTCTTGGAACCCAATCA-3′
				5′-GGCGCATGTCCCTCACAGCAA-3′
FaRPS1 (contig-00394)	Low	26S proteasome subunit 7	720	5′-GCAGCCTCTGCAGGTTGCGA-3/
				5′-GCTGGGTCTAGTGTGTCCGGC-3′
FaENP1 (contig-22543)	High	Endoplasmin-like protein	947	5′-GCCAAACGGTTTCTGTTGGCCG-3′/
				5′-TGGGAGTCCAAGGCTGATGGCT-3′
FaAP1 (contig-25571)	High	Aspartic proteinase	937	5′-ACCGCGACGCTGTTCTCTGT-3′/
				5′-TCACGATGCCAGTGGCTCCAA-3′
FaGAPDH1 (contig-16067)	High	Glyceraldehyde 3-phosphate dehydrogenase	457	5′-ACAGTGGGTACACGGAAGGCCA-3′/
				5′-CATGGGGTTCAGTTGGTGCCG-3′
FaCHP3 (contig-19988)	High	Conserved hypothetical protein	774	5′-CCAAGGCGGAGCACGGTCTG-3′/
				5′-AGGGCACAGCAGAACGCCAC-3′
Other transcripts monitored in this study				
UBQ	High	Ubiquitin	703	5′-TTTTCGTGAAAACCCTCACC-3′/
				5′-TGAGCTCTCCACTCCAAGGT-3′
EF1A	Medium	Elongation factor 1 alpha	804	5′-CTGTAACAAGATGGATGGATGCCACC-3′/
				5′-GGCGCATGTCCCTCACAGCAA-3′
HIS	Medium	Histone H1	583	5′-CTTTCCAGTCTCCCCCTTTC-3′/
				5′-CCTTGACCGACTTGGTTTTG-3′
ACT	Medium	Actin	716	5′-AATGGTGAAGGCTGGTTTTG-3′/
				5′-TCTCAGCTCCAATCGTGATG-3′
LHCB/CAB	Variable	Light-harvesting, chlorophyll binding protein	667	5′-AGGCCGAGTCACAATGCGCAA-3′/
				5′-AACAGGGTCGGCTAAGTGGTCG-3′
rbcL	Variable	Ribulose-bisphosphate carboxylase/ oxygenase large subunit	775	5′-AGGCCTTGCGCGCTCTACGT-3′/
				5′-TGCCAAACGTGAATACCGCCGG-3′
psaA	Variable	Photosystem I subunit A	775	5′-GGTTGGGCGGGGCATCAAGT-3′/
				5′-CTTGGGGACG CCCTAACGCG-3′

[a] As defined in Folta et al. (2010)

and Gilbert 1984) according to standard protocols (Sambrook and Russell 2001).

Quantitation of RNA dot blots

Hybridized dot blots were washed extensively under stringent conditions and exposed to Biomax XAR film (Kodak; Rochester, NY, USA) for empirically determined periods to yield optimal images for quantitation. Images were captured from the developed films using a flatbed scanner and subsequent quantitation of individual dot blot signals was performed with ImageJ software (Abràmoff et al. 2004; Rasband 1997).

RNA isolation

Fragaria vesca tissue types processed for RT-qPCR consisted of: young leaf, mature leaf, petiole, crown, flower,

root, fruit, and stolon. For the octoploid strawberry *F.* × *ananassa*, the tissue types were: young leaf, mature leaf, petiole, crown, flower, and root, as well as five individual stages of fruit development. Total RNA was isolated for RT-qPCR using a CTAB-based method described by Chang et al. (Chang et al. 1993), only following LiCl precipitation the pellet was resuspended in 500 µl SSTE (1 M NaCl, 0.5 % sodium dodecyl sulfate, 10 mM Tris–HCl pH 8.0, 1 mM EDTA), transferred to a microcentrifuge tube and vortexed in an equal volume of chloroform:isoamyl alcohol (24:1). The mixture was then centrifuged at 12,000×*g* for 5 min and the supernatant was transferred to a new microcentrifuge tube. Two volumes of 100 % ethanol were added and RNA was precipitated at −80 °C for >30 min, followed by centrifugation at 12,000×*g*. The pellet was washed once in 76 % ethanol containing 0.3 M sodium acetate pH 5.2, and then was resuspended in 20–50 µl dH_2O.

Six µg of total RNA were DNAse I (RQ1 RNase-Free DNAse, Promega; Fitchburg, WI USA) treated according to manufacturer recommendations. DNAse-treated RNA was extracted with an equal volume of a mixture of chloroform:octonol (24:1) and precipitated 1 h at −80 °C with two volumes of absolute ethanol plus 1/10 volume of 3 M sodium acetate (pH 5.2). After 30 min centrifugation at 10,000×*g*, the pellet was washed with 70 % ethanol, dried, and resuspended in 15 µl of water. Final concentration of purified RNA was determined using a NanoVue spectrophotometer (GE Healthcare; Uppsala, Sweden). cDNA synthesis was performed with 2.4 µg of total RNA in a final volume of 70 µl, using oligo-dT primers and reverse transcriptase ImProm-II RT (Promega; Fitchburg WI). Prior to RT-qPCR use, cDNA products were diluted five-fold with RNase-free water.

Sequences used for SYBR green RT-qPCR primer design were derived from candidates with relatively invariable gene expression across different tissue types as previously reported (Folta et al. 2010) and verified by RNA dot blots (Figs. 1, 2). For all candidates, primer pairs were

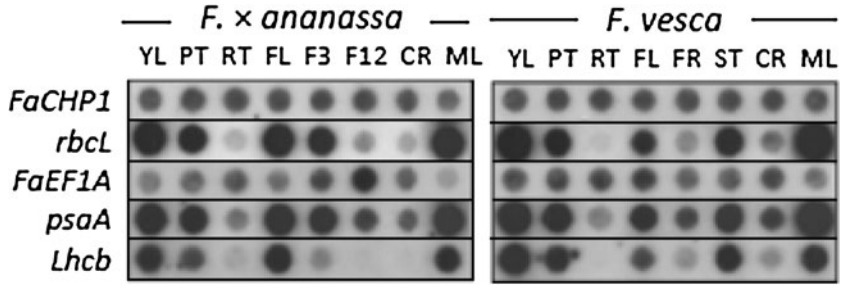

Fig. 1 Demonstration of the dynamic range of the assay. A range of labeled DNA probes were hybridized to RNA derived from multiple tissues from *F* × *ananassa* and *F. vesca*. The tissues tested are expanding leaves (*YL*), petioles (*PT*), roots (*RT*), flowers (*FL*), white unexpanded fruits (*F3*), red over-ripe fruits (*F12*), crowns (*CR*), mature leaves (*ML*), stolons (*ST*), fruit from several developmental stages (*FR*). *FaCHP1* is a candidate constitutively expressed transcript in this study and is shown for comparison

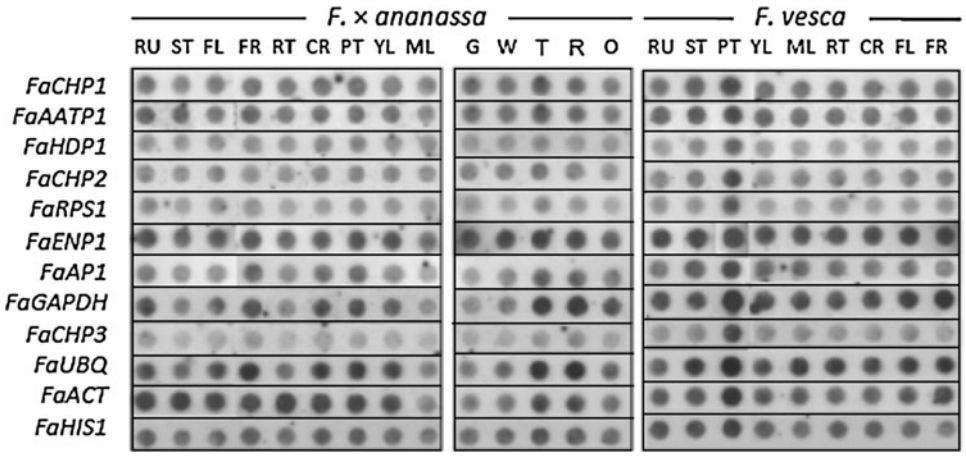

Fig. 2 Hybridization of radiolabeled constitutive reference gene candidate probes against filter-bound targets representing tissues from cultivated strawberry (*F.* × *ananassa*) and diploid (*F. vesca*) strawberry. *RU* runner tip, *ST* stolon, *FL* flower, *FR* "turning fruit", *RT* root, *CR* crown, *PT* petiole, *YL* young expanding leaf, *ML* mature leaf. The center panel represents a fruit ripening series *G* green, *W* white, *T* turning, *R* red, *O* over-ripe. *F. vesca* petiole RNA was consistently overloaded as determined by normalization to a full-length rDNA probe

Table 2 Sequences and amplification properties of RT-qPCR primers for the nine strawberry constitutive candidate transcripts and *FaEF1A*

Transcript designation	Primer sequence forward/reverse	Annealing temperature (°C)	Amplicon length (bp)	Amplification efficiency
FaCHP1	5′-TGCATATATCAAGCAACTTTACACTGA-3′/	60	91	1.935
	5′-ATAGCTGAGATGGATCTTCCTGTGA-3′			
FaAATP1	5′-TGGTACAATCTTCTTGAGGGTTGA-3′/	60	74	1.792
	5′-GCTGATGGTTTGGTGAAGTTGA-3′			
FaHDP1	5′-TGAACCACTGTGATGTAGCACTACTG-3′/	60	85	1.921
	5′-GCTATACCTGAAGATTGGAATAGATTTG-3′			
FaCHP2	5′-TCAAGAAAGGGAAGAGCAGTTTG-3′/	60	85	1.944
	5′-ACCCTCTTCTTCCTCTGCAATTC-3′			
FaRPS1	5′-GCTTGAGCAGAAAGACAGTCTAATTC-3′/	60	81	1.915 (*F.* × *ananassa*)
	5′-CTAGG GTTCTTCTCGTCCTTGATATC-3′			
FaENP1	5′-GCCACGTCTCTTTGACATTGACT-3′/	60	71	1.971
	5′-TTCCGAATGGGCTTTCCA-3′			
FaAP1	5′-GGCCCAACGACAATTATTACTGA-3/	60	81	1.790
	5′-AACCACAGTCTTGCATTCTTGACT-3′			
FaGAPDH1	5′-CATTGAGAGCAGGCAGAACCT-3′/	60	72	1.946
	5′-CCTCATTCAACATCATTCCTAGCA-3′			
FaCHP3	5′-TGTGGTCCAATGCCCATACTATT-3′/	60	76	1.894
	5′-AACGGCTCCTCAGGAAGAGAA-3′			
FaEF1A	5′-GCCCATGGTTGTTGAAAACTTT-3′/	60	69	1.968
	5′-GGCGCATGTCCCTCACA-3′			

Amplification efficiency was determined using LinReg PCR software

designed using Primer Express software (Applied Biosystems; Grand Island NY USA) with default settings, and are detailed in Table 2. Amplicon lengths ranged between 69 and 91 bp.

RT-qPCR assays and analysis

The RT-qPCR experiments were performed using a StepOnePlus Real Time PCR System (Applied Biosystems; Grand Island NY USA) according to manufacturer's recommendations. Each well contained 7.5 µl 2× EvaGreen qPCR Mastermix-ROX (Applied Biological Materials Inc.; Richmond, BC Canada), 0.45 µl of each of the 5 µM forward and reverse primers, 4 µl of a five-fold dilution of the cDNA synthesis product and 2.6 µl of water. Reactions were run in 96-well plates in triplicate and the default temperature program consisted of 10 min at 95 °C for one cycle and 15 s at 95 °C, 60 s at 60 °C for 40 cycles.

After each run, a melting curve analysis was performed to confirm that the fluorescence signal corresponded to the specific product amplification. To determine the mean PCR efficiency of each pair of primers, LinReg PCR software (Ramakers et al. 2003) was used.

Statistical analysis

Two computational methods were used to assess RT-qPCR results for the strawberry normalization candidates by calculating values that represent relative gene expression stability. In both approaches, average *C*q values obtained for primer/target pairs were converted into relative quantities by the ΔΔ*C*t method (Livak and Schmittgen 2001) using the sample with highest *C*q value as calibrator and imported into the software for analysis. The statistical analyses were performed with results from *F.* × *ananassa* and *F. vesca* tissue samples as a combined set of data, and also separately for the octoploid and the diploid findings.

The first statistical method made use of geNorm^{PLUS} in qbase^{PLUS} (v2.4 package) the next-generation successor to the original geNorm Visual Basic Application for Excel (Hellemans et al. 2007; Vandesompele et al. 2002). The geNorm^{PLUS} method estimates an expression stability measure (*M*) for each reference gene candidate, such that a lower *M* value indicates greater expression stability relative to other candidates tested. The second tool used was the NormFinder algorithm (Andersen et al. 2004) which ranks tested genes according to their calculated stability values and identifies the single optimal normalization gene from a

set of candidates. In addition, NormFinder will identify the best pair of normalization genes when used in tandem in the same RT-qPCR experiment.

Results

RNA blot hybridization

An experiment was performed to verify the dynamic range of the assay and to ensure conditions of probe excess and specific probe binding. DNA dot blots were prepared with increasing amounts of denatured PCR products representing each candidate cDNA in amounts from 1.0 pg to 1.0 ng. An expected increase in signal was detected across the range of target amounts, demonstrating linear hybridization and detection (not shown). Non-specific DNA controls showed negligible signal after long exposures (not shown), never exceeding approximately 1 % of the weakest target hybridizations.

To demonstrate that experiments were performed in probe excess and able to detect expression variation between tissues, an RNA blot hybridization experiment using probes for transcripts known to exhibit tissue-specific variation was performed (Fig. 1). Hybridization signals from probes corresponding to the chloroplast transcripts *psaA* and *rbcL*, and the nuclear gene family *Lhcb* were present at high levels in photosynthetic tissues and were less abundant in tissues like roots or crowns.

Parallel sets of RNA dot blots were hybridized to probes for the strawberry constitutive candidate *FaCHP1* (contig 21335) and *FaEF1A* (Fig. 1). The results demonstrate that the *FaEF1A* transcript level also varies between tissue types while those from *FaCHP1* are more stable. Hybridization results for the octoploid and diploid strawberry genotypes were comparable.

Hybridization across strawberry tissues

Strawberry normalization candidates were hybridized to a comprehensive set of tissue samples. Results of one representative hybridization experiment are presented in Fig. 2 and the quantitative data are shown in Table 3. The results in Fig. 2 show hybridization signals across octoploid (*F.* × *ananassa*) and diploid (*F. vesca*) tissues, as well as in a developmental series of octoploid fruits. A probe against ubiquitin (UBQ) showed considerable variation across tissues, with relatively low steady-state accumulation in stolons, roots and mature leaves. Strawberry fruits also showed substantial variation with the UBQ probe, peaking at the breaker stage. Similar variation was observed in the diploid species but at lower magnitude.

A probe against actin (ACT) also showed variation in transcript abundance between *Fragaria* tissue types, especially in the developmental series for the commercially relevant octoploid fruit. Again the young unexpanded fruits and the overripe fruits showed less transcript accumulation than mature leaves in the cultivated strawberry. A probe directed against histone *H1* (HIS) showed relative stability across conditions. *FaGAPDH1*, a common normalization transcript in other species, showed considerable variation in dot blot analyses.

Table 3 Stability of candidate transcript expression as assessed by hybridization

Candidate	*F.* × *ananas*sa nine tissue types			*F.* × *ananas*sa five stages of fruit development			F. × *ananassa* overall			*F. vesca* nine tissue types		
	Mean	SD	CV %	Mean	SD	CV %	Mean	SD	CV %	Mean	SD	CV %
FaCHP1	1.135	0.084	7.4	1.277	0.168	13.2	1.272	0.119	9.4 %	1.253	0.241	19.3
FaAATP1	1.153	0.103	8.9	1.202	0.227	18.9	1.217	0.153	12.6 %	1.088	0.074	6.8
FaHDPl	1.181	0.104	8.8	1.142	0.161	14.1	1.249	0.143	11.5 %	1.157	0.126	10.9
FaCHP2	1.158	0.129	11.2	1.144	0.094	8.3	1.179	0.119	10.1 %	1.328	0.199	15.0
FaRPS1	1.231	0.170	13.8	1.201	0.131	10.9	1.274	0.165	13.0 %	1.151	0.153	13.3
FaENP1	1.651	0.372	22.5	1.819	0.805	44.2	1.822	0.644	35.4 %	1.215	0.265	21.8
FaAP1	1.295	0.188	14.5	1.166	0.198	17.0	1.368	0.227	16.6 %	1.429	0.476	33.3
FaGAPDH1	1.506	0.308	20.5	1.588	0.498	31.4	1.566	0.388	24.8 %	2.290	0.928	40.5
FaCHP3	1.270	0.221	17.4	1.358	0.301	22.1	1.313	0.250	19.0 %	1.171	0.141	12.0
FaACT	1.225	0.196	16.0	1.215	0.189	15.5	1.268	0.201	15.9 %	1.396	0.401	28.8
FaHIS	1.119	0.094	8.4	1.070	0.091	8.5	1.139	0.096	8.5 %	1.266	0.193	15.3
FaUBQ	1.392	0.270	19.4	1.616	0.616	38.1	1.537	0.415	27.0 %	1.336	0.170	12.7

Total RNA representing various tissues from *F.* × *ananassa* cultivar Strawberry Festival and diploid *F. vesca* accession Hawaii four were arrayed on nylon filters and probed with radiolabeled candidate sequences. Hybridization signals were normalized against subsequent re-probing of the same blots with a full-length rDNA probe. SD indicates standard deviation and CV % is the coefficient of variation

The results of Table 3 indicate that the traditional reference transcripts *FaACT* and *FaUBQ* were more variable overall than *FaHIS* which showed comparatively stable expression. Among the strawberry candidates, *FaENP1* and *FaGAPDH1* exhibited the most variability in expression in the octoploid while *FaAP1* and *FaGAPDH1* were the most variable in *F. vesca* tissues.

RT-qPCR profiles of candidate expression

Normalization candidates were next evaluated by RT-qPCR using the primer pairs shown in Table 2. All candidates were amplified using samples from *F. vesca* and *F.* × *ananassa* with the exception of candidate *FaRPS1* which failed to amplify with *F. vesca* template. *C*q values determined from the RT-qPCR using geNorm^{PLUS} are presented in Fig. 3. The relative transcript abundance of the various candidates as indicated by *C*q values is consistent with the results from next-generation transcript sequencing data (Folta et al. 2010). The RT-qPCR results identify *FaGAPDH1* as the most abundant candidate transcript in both *F. vesca* and *F.* × *ananassa*.

Validation of candidate expression stability by RT-qPCR

The performance of the strawberry candidates as RT-qPCR reference standards was assessed using two methods: NormFinder and geNorm^{PLUS} in qbase^{PLUS}. NormFinder analysis ranked the candidates according to relative gene expression stability across tissues (Table 4). *FaGAPDH1* was the top ranked NormFinder candidate for *F. vesca*, with a remarkably high standard error for the estimated stability value, followed by *FaCHP1* and *FaAP1*. In contrast, *FaGAPDH1* placed in the bottom half of the reference candidates in *F.* × *ananassa* expression stability. For the *F.* × *ananassa* tissue set including one intermediate stage of fruit development, NormFinder identified *FaCHP1* and *FaAP1* as the top two most stably expressed candidates, followed by *FaRPS1*, *FaENP1*, and *FaEF1A*. For both the octoploid and diploid strawberry tissue groups NormFinder ranked *FaCHP1*, *FaCHP3*, *FaAP1* and *FaEF1A* in the top five most stable transcripts. Analysis of the *F.* × *ananassa* fruit developmental series ranked *FaENP1* and *FaEF1A* as the most stably expressed candidates, followed by *FaCHP2*, *FaAP1* and *FaCHP1*.

The geNorm^{PLUS} ranking of the strawberrry reference candidates is presented in Fig. 4 with the most stable candidates shown at the right side of each graph. In each case, geNorm^{PLUS} has identified *FaENP1* as the most stably expressed transcript. Top ranked genes include *FaEF1A* and *FaENP1* for *F. vesca*, *FaCHP3* and *FaENP1* for *F.* × *ananassa* and *FaCHP1* and *FaENP1* for the

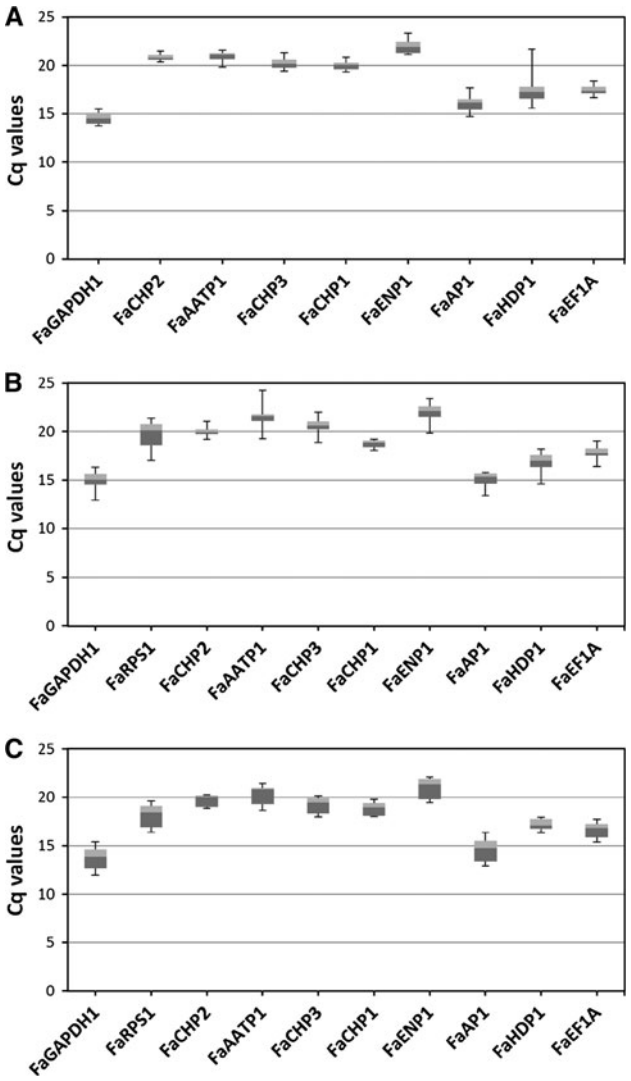

Fig. 3 Expression levels of strawberry reference gene candidates as depicted by box plot graphs of quantification cycle (*C*q) values. **a** *F. vesca* variation across tissue types. **b** *F.* × *ananassa* variation across eight tissue types. In both **a** and **b** unripe fruit was measured. **c** *F* × *ananassa* fruit series. Boxes indicate average *C*q values for each candidate from the 25th to the 75th percentile. *Vertical lines* (*whiskers*) indicate maximum and minimum *C*q values

octoploid fruit series. *M* values below geNorm's arbitrary cutoff of 0.5 denote stable gene expression (Mafra et al. 2012; Rocha-Martins et al. 2012). The average expression stability value is below this threshold for the complete group of candidates across the fruit development series. The *M* value is below this cutoff after elimination of *FaHDP1* from *F.* × *ananassa* and *F. vesca* tissue results.

A second facet of the geNorm^{PLUS} analysis is the determination of the number of reference genes required for accurate normalization. Figure 5 displays the results of pairwise comparisons of the variability between groupings of the strawberry candidates, The geNorm^{PLUS} V2/3 value represents the variability of the two most stable genes

Table 4 NormFinder ranking of reference gene candidates according to expression stability in *F. vesca* and *F.* × *ananassa* plant tissues including a single fruit developmental stage. A fruit developmental series is shown for *F.* × *ananassa*

F. vesca[a]				*F.* × *ananassa*[a]				*F.* × *ananassa* developing fruit series			
Order	Gene name	Stability value	Standard error	Order	Gene name	Stability value	Standard error	Order	Gene name	Stability value	Standard error
1	*FaGAPDH1*	0.012	0.648	1	*FaCHP1*	0.160	0.075	1	*FaENP1*	0.109	0.057
2	*FaCHP1*	0.112	0.100	2	*FaAP1*	0.162	0.075	2	*FaEF1*	0.113	0.057
3	*FaAP1*	0.195	0.093	3	*FaRPS1*	0.174	0.077	3	*FaCHP2*	0.116	0.058
4	*FaEF1A*	0.198	0.094	4	*FaENP1*	0.178	0.077	4	*FaAP1*	0.119	0.058
5	*FaCHP3*	0.238	0.098	5	*FaEF1A*	0.185	0.078	5	*FaCHP1*	0.155	0.067
6	*FaENP1*	0.252	0.100	6	*FaCHP3*	0.204	0.081	6	*FaRPS1*	0.158	0.068
7	*FaCHP2*	0.522	0.154	7	*FaGAPDH1*	0.254	0.091	7	*FaCHP3*	0.194	0.078
8	*FaAATP1*	0.522	0.154	8	*FaCHP2*	0.391	0.124	8	*FaGAPDH1*	0.236	0.091
9	*FaHDP1*	1.165	0.314	9	*FaAATPl*	0.456	0.141	9	*FaAATP1*	0.288	0.108
				10	*FaHDP1*	0.870	0.254	10	*FaHDP1*	0.431	0.155

[a] Includes fruit only from the 'turning' stage

compared to the three most stable genes. V3/4 represents the variability comparison of the three most stable to the four most stable and so on. The geNorm[PLUS] analysis establishes that simultaneous use of the two most stable as references will result in accurate normalization (Fig. 5). Table 4 shows the results for Normfinder analyses. A value below the cutoff at 0.5 is considered stable expression. The results from Normfinder suggest that a number of transcripts are expressed stably within *F. vesca*, *F.* × *ananassa*, *and F.* × *ananassa* fruits. These include *FaGAPDH1*, *FaCHP1*, *FaAP1*, *FaEF1A*, *FaCHP3* and *FaENP1*.

Discussion

As the methods to analyze gene expression simultaneously track more transcripts with greater sensitivity, the hazards of inadequate normalization become increasingly apparent. Assumptions carried between species undermine analysis, as traditional "housekeeping" genes are not always experimentally validated as appropriate reference genes. Constitutive expression of some typical baseline transcripts is not always as reliable as has often been assumed. For instance, the actin transcript has been shown to be inconsistent in potato (Nicot et al. 2005), poplar (Brunner et al. 2004) and rice (Jain et al. 2006). In a species with diverse tissues like strawberry, identifying a universally applicable constitutive reference transcript is necessary but challenging.

The current study sets aside the preconception that housekeeping gene choices like actin or ubiquitin are evenly expressed. This report expands on high-throughput sequencing data to empirically derive and test a set of new standards for *Fragaria* species. A transcriptome survey in cultivated strawberry identified candidates from contigs composed of even numbers of reads across tissues (Folta et al. 2010). These potential constitutive candidates needed to be confirmed through further study, as high-throughput sequencing could include bias. This report takes these candidates and subjects them to additional analyses. Such genome-based approaches have proven useful in Arabidopsis (Czechowski et al. 2005) where *ACTIN*, *EF1α* and *GAPDH1* showed evidence of inconsistency as reference transcripts. Fig. 2 and Table 3 in this report show that several traditional "constitutive" transcripts are not evenly expressed in strawberry and are inappropriate for normalization, at least between tissues. Similar findings were observed in strawberry relatives like peach (Tong et al. 2009), whereas in apple *ACTIN* and *GAPDH* performed well, at least in the limited tissues tested in a ripening fruit (Schaffer et al. 2007). A study across the diverse tissues of rose floral development found *GAPDH* and *EF1α* to act reliably (Dubois et al. 2011). In postharvest rose, the reference of choice is also tissue or treatment specific, yet only traditional references were tested (Meng et al. 2013). However, an examination of multiple diverse tissues and stressors agrees with the results here from strawberry, as no universal reference transcript was identified (Klie and Debener 2011). The results from these other rosaceous crops mirror those from plants in general, as few transcripts are truly constitutive over broad ranges of tissues, treatments or developmental states.

In this study, independent methods were used to validate earlier RNAseq-based projections. Hybridization controls ensured dynamic range of the assay and conditions of probe excess (Fig. 1). Candidate genes were tested by hybridization for concordance of results with RT-qPCR. When comparing hybridization and RT-qPCR results, we

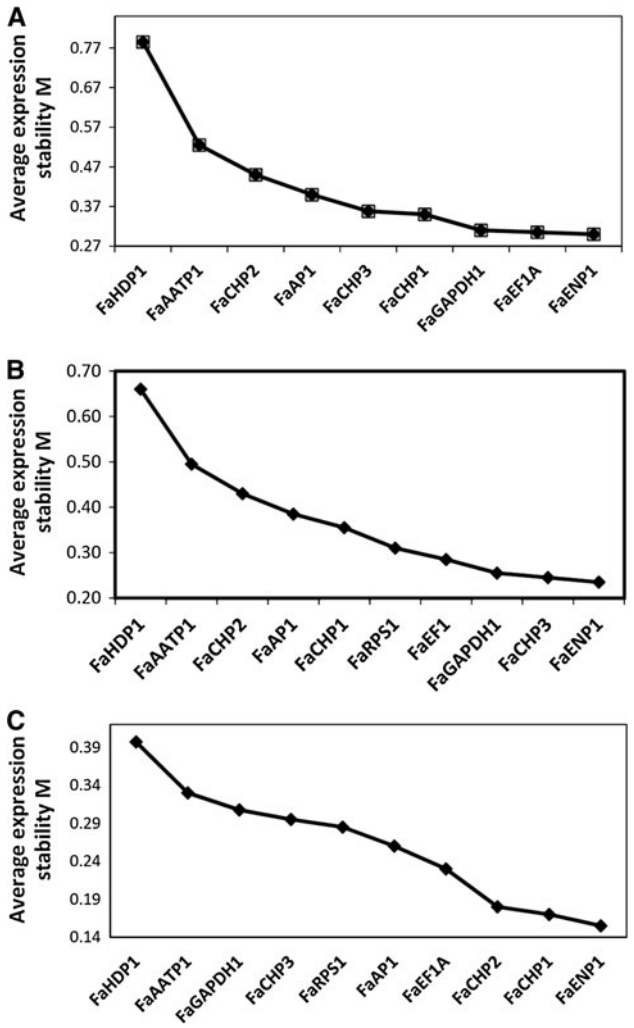

Fig. 4 Expression stability of reference gene candidates as ranked by geNorm^{PLUS} from least stable (high *M* value) to most stable (low *M* value). **a** *F. vesca* variation across tissue types including 'turning' fruit. **b** *F.* × *ananassa* variation across eight tissue types, containing 'turning' fruit. **c** *F.* × *ananassa* fruit series. Data points represent the average expression stability of the remaining candidates after progressive elimination of less stable members. Candidates are displayed by increasing stability from left to right, with the most stably expressed reference genes at the right

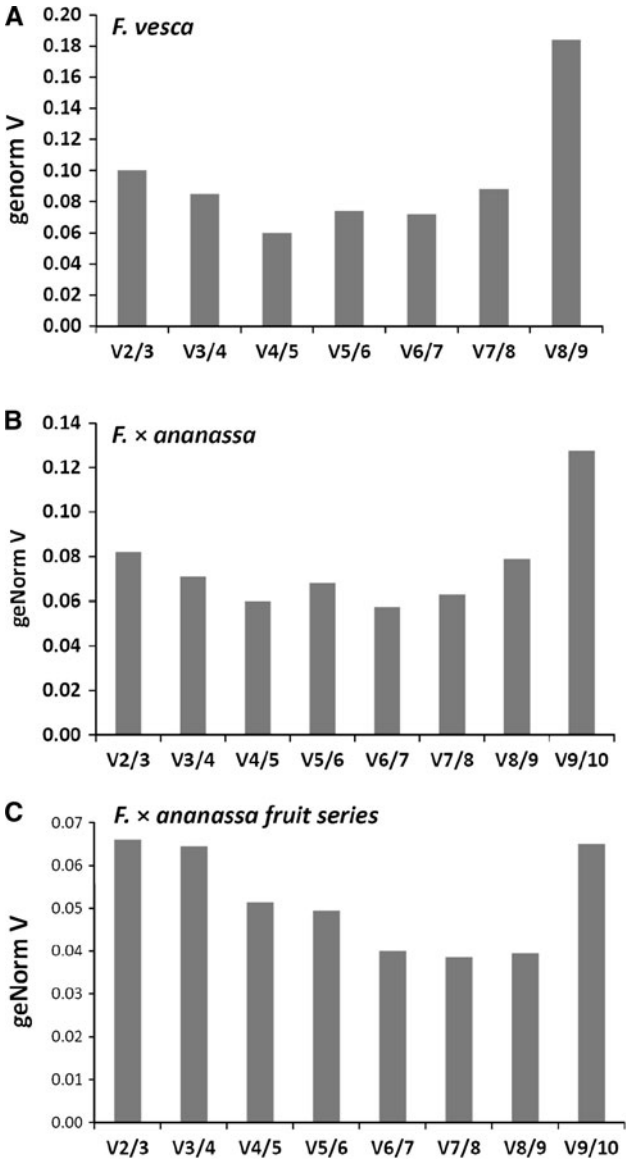

Fig. 5 Determination of the optimal number of reference genes by geNorm^{PLUS} analysis. Pairwise variation (*V*) values of <0.15 signify that no additional candidates are needed for accurate standardization. In these data sets, the use of the two most stable candidates is sufficient for accurate standardization

found some degree of discrepancy regarding the stability of the candidates. In hybridization experiments, *FaAATP1* and *FaHDP1* exhibited low variability of expression in both *Fragaria* species (Table 3). However, these two were ranked least stable of the candidates by both Normfinder and geNorm^{PLUS}, although based on the arbitrary stability cutoff value of 0.5, *FaAATP* did meet geNorm^{PLUS} criteria for stable expression. Together all the candidates, with the exception of *FaHDP1*, were judged stable by geNorm^{PLUS}. Another candidate, the *FaGAPDH1* transcript, was the most variable in *F. vesca* hybridization, but judged most stable by Normfinder although with a large SD. *FaGAPDH1* also was ranked much lower by geNorm^{PLUS}. It should be noted that *FaGAPDH1* was the most abundant of the strawberry candidates and relative expression levels are important considerations when selecting a reference gene.

Differences between hybridization and RT-qPCR results may be due to homoeologs in the octoploid plants that are not detected with the same efficiency using the RT-qPCR primers. On the other hand, hybridization tolerates a lower global similarity so paralogs and homoeologs are likely detected. It is also a formal possibility that RT-qPCR could detect a partial transcript or be affected by a heterogenous pool of differentially spliced variants. The use of two contrasting detection methods is strength of this work.

Ranking the constitutive candidates with two RT-qPCR statistical methods agreed in some cases, yet not in others. The central reason for the difference is that all of the candidates, by definition, are relatively stable in expression. The variation observed is represented in different ways by the various algorithms used to describe it. According to geNorm^{PLUS}, *FaENP1* is the most stable candidate in both the diploid and octoploid strawberry tissues as well as the most stable for the *F.* × *ananassa* fruit development series (Fig. 5). In contrast, Normfinder also places *FaENP1* first in stability for developing octoploid fruit but fourth for other octoploid tissues and sixth overall in *F. vesca* tissues (Table 4). The two statistical methods are in agreement in ranking *FaAATP1* and *FaHDP1* as the least stable candidates in both *Fragaria* species as well as developing fruit. Viewed as groups, there are many similarities among the highly ranked candidates identified by either method. For *F.* × *ananassa* tissues, five of the top six candidates are the same by both methods although the relative ranking varies. Similarly, all of the six most stable candidates are the same by both methods for the octoploid fruit series although in different orders. Results for *F. vesca* also find the same top six and bottom three candidates using either Normfinder or geNorm^{PLUS}. The evaluation of *FaGAPDH1* expression stability reveals some significant differences between tissues tested. Although *FaGAPDH1* was the top candidate in *F. vesca* tissues according to Normfinder and ranked third by geNorm^{PLUS}, both Normfinder and geNorm^{PLUS} ranked this abundant transcript as relatively unstable throughout octoploid fruit development. Differences between stability estimation rankings determined by geNorm^{PLUS} and Normfinder have been reported previously (Pérez et al. 2008; Steinau et al. 2006) and are to be expected based on the different designs of the algorithms in the alternative approaches (Andersen et al. 2004).

The current work tests only variation between tissue types. This is a rigorous test, as strawberry tissues are diverse with fruits featuring extensive transcriptome variation throughout the ripening process. (Aharoni and O'Connell 2002; Bombarely et al. 2010) and ranging expression in between tissues (Folta et al. 2010). Changes occurring due to development or stress, or those observed in minor tissue types (e.g., stamens, achenes, pollen) may lead to significant variation with even the most stable transcripts. Such caveats are consistent with the unifying theme of RT-qPCR reference gene trials—there may be no perfect reference gene. Every experiment should consider at least one additional gene as an internal reference, as suggested by the geNorm^{PLUS} analysis (Fig. 5), and possibly a complementary hybridization experiment to compare quantitation across methods.

This study merges three independent methods to derive a series of RT-qPCR controls. Digital gene expression data, hybridization, and RT-qPCR all were shown to produce similar results for the majority of the empirically selected candidate transcripts, while others (including some traditional RT-qPCR reference genes) are shown to be less reliable. These new normalization genes will be of utility as RT-qPCR continues to grow as the preferred method of expression measurement and strawberry gains traction as a functional system for gene activity in the Rosaceae.

References

Abràmoff MD, Magalhães PJ, Ram SJ (2004) Image processing with ImageJ. Biophotonics Int 11:36–42

Aharoni A, O'Connell AP (2002) Gene expression analysis of strawberry achene and receptacle maturation using DNA microarrays. J Exp Bot 53:2073–2087

Andersen CL, Jensen JL, Ørntoft 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

Bombarely A, Merchante C, Csukasi F, Cruz-Rus E, Caballero JL, Medina-Escobar N, Blanco-Portales R, Botella MA, Muñoz-Blanco J, Sánchez-Sevilla JF et al (2010) Generation and analysis of ESTs from strawberry (Fragaria xananassa) fruits and evaluation of their utility in genetic and molecular studies. BMC Genomics 11:503

Brunner AM, Yakovlev IA, Strauss SH (2004) Validating internal controls for quantitative plant gene expression studies. BMC Plant Biol 4:14

Chang S, Puryear J, Cairney J (1993) A simple and efficient method for isolating RNA from pine trees. Plant Mol Biol Report 11:113–116

Church GM, Gilbert W (1984) Genomic sequencing. Proc Natl Acad Sci 81:1991–1995

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

Dubois A, Remay A, Raymond O, Balzergue S, Chauvet A, Maene M, Pécrix Y, Yang S-H, Jeauffre J, Thouroude T et al (2011) Genomic approach to study floral development genes in *Rosa* sp. PLoS ONE 6:e28455

Folta KM, Clancy MA, Chamala S, Barbazuk WB, Dhingra A, Brunings A, Gomide L, Kulathinal RJ, Peres N, Davis TM (2010) A transcript accounting from diverse tissues of a cultivated strawberry (Fragaria × ananassa). Plant Genome 3:99–105

Haymes KM, Davis TM (1998) Agrobacterium mediated transformation of 'Alpine' Fragaria vesca, and transmission of transgenes to R1 progeny. Plant Cell Rep 17:279–283

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

Jain M, Nijhawan A, Tyagi AK, Khurana JP (2006) Validation of housekeeping genes as internal control for studying gene expression in rice by quantitative real-time PCR. Biochem Biophys Res Commun 345:646–651

Klie M, Debener T (2011) Identification of superior reference genes for data normalisation of expression studies via quantitative PCR in hybrid roses (Rosa hybrida). BMC Res Notes 4:518

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{\Delta\Delta CT}$ Method. Methods 25:402–408

Mafra V, Kubo KS, Alves-Ferreira M, Ribeiro-Alves M, Stuart RM, Boava LP, Rodrigues CM, Machado MA (2012) Reference genes for accurate transcript normalization in citrus genotypes under different experimental conditions. PLoS ONE 7:e31263

Meng Y, Li N, Tian J, Gao J, Zhang C (2013) Identification and validation of reference genes for gene expression studies in postharvest rose flower (Rosa hybrida). Sci Hortic 158:16–21

Nicot N, Hausman J-F, Hoffmann L, 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

Oosumi T, Gruszewski HA, Blischak LA, Baxter AJ, Wadl PA, Shuman JL, Veilleux RE, Shulaev V (2006) High-efficiency transformation of the diploid strawberry (*Fragaria vesca*) for functional genomics. Planta 223:1219–1230

Pérez R, Tupac-Yupanqui I, Dunner S (2008) Evaluation of suitable reference genes for gene expression studies in bovine muscular tissue. BMC Mol Biol 9:79

Ramakers C, Ruijter JM, Deprez RHL, Moorman AFM (2003) Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. Neurosci Lett 339:62–66

Rasband WS (1997) ImageJ. US National Institutes of Health, Bethesda

Rocha-Martins M, Njaine B, Silveira MS (2012) Avoiding Pitfalls of Internal Controls: validation of reference genes for analysis by qRT-PCR and western blot throughout rat retinal development. PLoS ONE 7:e43028

Sambrook J, Russell DW (2001). Molecular cloning: a laboratory manual: CSHL press

Schaffer RJ, Friel EN, Souleyre EJF, Bolitho K, Thodey K, Ledger S, Bowen JH, Ma J-H, Nain B, Cohen D et al (2007) A genomics approach reveals that aroma production in apple is controlled by ethylene predominantly at the final step in each biosynthetic pathway. Plant Physiol 144:1899–1912

Shulaev V, Sargent DJ, Crowhurst RN, Mockler TC, Folkerts O, Delcher AL, Jaiswal P, Mockaitis K, Liston A, Mane SP et al (2011) The genome of woodland strawberry (*Fragaria vesca*). Nat Genet 43:109–116

Steinau M, Rajeevan MS, Unger ER (2006) DNA and RNA references for qRT-PCR assays in exfoliated cervical cells. J Mol Diagn 8:113–118

Tong ZG, Gao ZH, Wang F, Zhou J, Zhang Z (2009) Selection of reliable reference genes for gene expression studies in peach using real-time PCR. BMC Mol Biol 10:13

Udvardi MK, Czechowski T, Scheible W-R (2008) Eleven golden rules of quantitative RT-PCR. The Plant Cell Online 20:1736–1737

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 biology 3:research0034