Malays. Appl. Biol. (2019) 48(5): 67–75

SELECTION AND VALIDATION OF REFERENCE GENES FOR qRT-PCR EXPRESSION ANALYSES OF LEAF, FLOWER AND CAPSULES OF *Impatiens balsamina* (BALSAMINACEAE)

LIAN CHEE FOONG[1] and SHEH MAY TAM[1*]

[1]*School of Biosciences, Faculty of Health and Medical Sciences, Taylor's University, 1, Jalan Taylors, 47500 Subang Jaya, Selangor, Malaysia*
[*]*E-mail: shehmay.tam@taylors.edu.my*

Accepted 19 November 2019, Published online 31 December 2019

ABSTRACT

Impatiens balsamina (Balsaminaceae) is both an ornamental and medicinal plant featured in many Asian traditional remedies due to an abundance of selected secondary metabolites with beneficial pharmacological properties. Attempts to improve production of biocompounds have led to characterization of genes, gene expression and pathways involved in secondary metabolites biosynthesis in *I. balsamina*. Gene expression analysis by quantitative reverse transcription PCR (qRT-PCR) technique has been well-established but its reliability is heavily dependent on the selection of suitable reference genes, which information is currently lacking for *I. balsamina,* and other *Impatiens* species. In this study, six typical housekeeping genes including *actin-like protein (ACT), polyubiquitin-A* (*UBQ*), *ubiquitin-conjugated enzyme E2* (*UBC*), *fructose-bisphosphate aldolase* (*FBA*), *elongation factor 1-alpha* (*EF1-α*), and *clathrin light chain 1* (*CLC*) were evaluated as candidate reference genes. Expression stability was tested in four tissue types (leaf, flower, early- and mature stage capsules) of *I. balsamina*. *FBA*, *CLC*, and *EF1-α* were observed to be the most stable reference genes across the tested tissue types. When expression data were normalised by the most stable *FBA* gene, the target secondary metabolite biosynthesis pathway genes showed significant different expression pattern compared to the least stable *UBQ* gene. The use of multiple reference genes in the combination of either *FBA* + *EF1-α* or *FBA* + *CLC* + *EF1-α* was observed to give more reliable results compared with a single gene. This study provides a resource for selection of suitable reference gene(s) for future gene expression normalisation experiments in *I. balsamina.*

Key words: Gene expression, tissues, *Impatiens balsamina*, reference genes, qRT-PCR

INTRODUCTION

The garden- or rose balsam, *Impatiens balsamina* L. (Balsaminaceae) is an annual herbaceous plant indigenous to South Asia that is valued for both its ornamental and traditional medicinal properties (Chua, 2016; Nurul *et al.*, 2010). Phytochemical and pharmacological studies of *I. balsamina* have led to the isolation of various secondary metabolites such as alkaloids, anthocyanins, glycosides, flavonoids/ flavanols, saponins, and terpenoids, with quinones being a dominant class, to which are attributed potent pharmacological activities, accumulating at various levels in different plant tissues (Meenu *et al.*, 2015; Shah *et al.*, 2017; Singh *et al.*, 2017). The potential downstream applications exhibited by selected bioactive secondary metabolites have naturally led to further investigation of related biosynthesis pathway genes (and their expression levels), biosynthesis pathways and molecular mechanisms underlying their production in *I. balsamina,* such as using the quantitative real-time PCR (qRT-PCR) method (Bustin, 2002).

Quantitative real-time PCR (qRT-PCR) provides a rapid, efficient, accurate, and reproducible method to analyse the mRNA transcription level in different samples or tissues (Bustin, 2002; Gachon *et al.*, 2004). However, the sensitivity and reliability of qRT-PCR are dependent on an appropriate normalisation method using one or more experimentally validated reference genes which ideally show consistent expression profile in different experimental conditions i.e. relative quantification (Nolan *et al.*, 2006; VanGuilder *et al.*, 2008). Most reference genes selected are housekeeping genes involved in basal cellular

* To whom correspondence should be addressed.

maintenance due to the assumption that they are uniformly expressed across different tissues, developmental stages, and treatments (Caldana *et al.*, 2007), such as *actin (ACT), elongation factor 1 alpha (EF1-α), polyubiquitin (UBQ), tubulin (TUB), and ubiquitin-conjugated enzyme (UBC*) (Czechowski *et al.*, 2005; Radonić *et al.*, 2004; Stürzenbaum & Kille, 2001; Zhou *et al.*, 2017). However, transcriptional levels of some housekeeping genes may not be consistent, especially in samples across different developmental stages, tissues or experimental conditions, hence potentially misinterpreting expression results (Guénin *et al.*, 2009; Huggett *et al.*, 2005; Radonić *et al.*, 2004; Suzuki *et al.*, 2000; Thellin *et al.*, 1999).

Though it is recognized that systematic screening of reference genes should be performed to avoid biased results during normalisation of gene expression study, no information is available for *I. balsamina* nor any other *Impatiens* species to date (no qRT-PCR study has been reported so far). Therefore, in this study, six typical housekeeping genes were selected by mining of *I. balsamina* transcriptome datasets, namely *fructose-bisphosphate aldolase 3 (FBA), clathrin light chain 1 (CLC), EF1-α*, *UBC*, *UBQ*, and *ACT*. This is followed by systematic evaluation of expression stability of these candidate reference genes in leaf, flower and two developmental stages of capsules (early- and mature stage) of *I. balsamina* using five statistical methods. To validate the stability of the proposed reference genes, the expression levels of eight target genes randomly selected from four major secondary metabolite biosynthesis pathways were examined.

MATERIALS AND METHODS

Preparation of plant material

Impatiens balsamina (variety of pink, multi-petals) plants were grown in a home garden setting in Seri Kembangan, Selangor, Malaysia in the period from 1 July to 8 September 2017, in 1-litre growth bags using a mix of black garden soil and clay (2:1). At 10-weeks old, samples of mature leaves (50-100 mm in length), flowers (fully blossomed), and capsules at two developmental stages, early- (E; length ≤ 12 mm, white seeds) and mature- (M; length 15-18mm, brown seeds) were collected in triplicates and immediately submerged in RNA*later*® solution (Ambion, USA).

Total RNA extraction and cDNA synthesis

Total RNA of each sample was extracted following the optimised protocol reported by Foong *et al.* (2017). Total RNA concentration and purity were measured using a Nanodrop ND1000 spectrophotometer (NanoDrop Technologies, USA). Total RNA samples with $A_{260/280}$ and $A_{260/230}$ ratios ≥ 1.8 were subjected to reverse transcription into cDNA using the Tetro cDNA synthesis kit (Bioline, UK) with oligo $(dT)_{18}$ primer following the manufacturer's instructions.

Selection of candidate reference genes based on transcriptome data

Expression data generated in previous RNA-sequencing (RNA-seq) study (SRA Accession PRJNA526137) was used to select and obtain full sequences of *I. balsamina* candidate reference genes with a coefficient of variation (CV) value lower than 0.3 (Pombo *et al.*, 2017). Primers of six selected candidate reference genes were designed using Primer 3 software (http://bioinfo.ut.ee/primer3-0.4.0/) with the following criteria: melting temperature 55-60°C, GC content 45-55%, primer length 18-24 bp, amplicon length 100-200 bp, and spanning exon-exon junction. The primer sequences are listed in Table 2.

qRT-PCR analysis

qRT-PCR was performed in Eppendorf™ RealTime PCR Cap Strips using a MasterCycler EP Gradient Thermal Cycler (Eppendorf, Germany). Each reaction contained 20 µL total volume, with 300 ng template cDNA, 1× SensiFAST™ SYBR® No-Rox mix (Bioline, UK), and 400 nM forward and reverse primers. The PCR cycling conditions were programmed as 95°C for 2 min, 40 cycles of 95°C for 5 s, 60°C for 10 s and 72°C for 10 s. Melting curves were determined by heating the products from 60°C to 95°C. For each primer pair, a standard curve was generated from a five-fold serial diluted cDNA template to calculate PCR amplification efficiency and regression coefficient (R^2). Each qPCR reaction was carried out in two technical and three biological replicates for all four tissue types. A template-free negative control was also included in all of the batches. For evaluating gene expression stabilities, ΔCt (Livak & Schmittgen, 2001), geNorm (Vandesompele *et al.*, 2002), NormFinder (Andersen *et al.*, 2004), BestKeeper (Pfaffl *et al.*, 2004) were employed using Microsoft Excel and RefFinder (Xie *et al.*, 2012) was conducted in the website https://www.heartcure.com.au/reffinder/.

Normalization of selected target genes

For validation, the expression pattern of eight target genes randomly selected from four main secondary metabolite biosynthesis pathways, namely the mevalonate (MVA-), 2-C-methyl-D-erythritol 4-phosphate (MEP-), Shikimate-, and 1,4-dihydroxy-2-naphthoate (DHNA) pathways were examined by qRT-PCR analysis using the same

methodologies described above in four *I. balsamina* tissue types (leaf, flower, early- and mature stage capsules). The tested target genes included *diphosphomevalonate decarboxylase* (*MVD*; EC: 4.1.1.33), *isopentenyl-diphosphate delta-isomerase* (*IDI*; EC:5.3.3.2), *1-deoxy-D-xylulose-5-phosphate synthase* (*DXS*; EC:2.2.1.7), *1-deoxy-D-xylulose-5-phosphate reductoisomerase* (*DXR*; EC:1.1.1.267), *3-deoxy-7-phosphoheptulonate synthase* (*DAHPS*; EC:2.5.1.54), *shikimate kinase* (*SK*; EC:2.7.1.71), *Isochorismate synthase/2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate synthase/2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase/Osuccinylbenzoate synthase* (*PHYLLO*; EC:5.4.4.2/2.2.1.9/4.2.99.20/4.2.1.113) and *acyl-activating enzyme 14* (*AAE14*; EC: 6.2.1.26). Expression was normalised using (1) the most stably expressed reference gene, (2) the two most stably expressed reference genes combination (according to NormFinder analysis), (3) three most stably expressed reference genes combination and (4) the least stably expressed gene. Relative expression analysis based on $2^{-\Delta\Delta CT}$ method was conducted and the results were presented as log-transformed fold change (Livak & Schmittgen, 2001).

Statistical analysis

All statistical analysis was conducted using R-based packages. One-way analysis of variance (ANOVA) was performed by R function aov using genes and CT value/relative expression value as factors. Post-hoc pairwise comparisons were performed with the R function TukeyHSD (Tukey's Honestly Significant Differences). Significant value was set at $p < 0.05$.

RESULTS AND DISCUSSION

Selection of candidate reference genes, amplification specificity and efficiency of primers

This study sought to evaluate the suitability of candidate reference genes for normalisation of relative gene expression data in different tissues of *I. balsamina* (leaf, flower, early-, and mature stage capsules). Selection of suitable reference genes is based on the criteria of their being stably expressed in tested samples, high amplification efficiency to target genes, and a moderate level of expression (Chi *et al.*, 2016; Kozera & Rapacz, 2013). Based on the FPKM value obtained by transcriptome data (SRA Accession PRJNA526137), the coefficient of variation (CV) was calculated. A total of six unigenes which showed CV $\geq$ 0.3 were selected (Table 1) which were typical housekeeping genes with various cellular functions to avoid co-regulation induced bias. The total RNA extracted from the four tissue samples met the required $A_{260/280}$ and $A_{260/230}$ ratios for qRT-PCR analysis.

Expression levels of the candidate reference genes obtained from qRT-PCR are presented as threshold cycle (Ct) values. High Ct value indicates low expression level of transcript abundance. A range of mean Ct values was observed, from 17.47 to 26.41, suggesting different levels of transcript abundance respectively for each of the six candidate genes across four tissue types. The most highly expressed gene was *UBC*, with an average Ct value of 19.33 ± 1.33 (mean ± SD), followed by *FBA* (20.63 ± 1.29), *UBQ* (21.04 ± 1.90), *EF1-α* (21.59 ± 1.21) and *CLC* (22.23 ± 1.44). The least expressed gene was *ACT*, with an average Ct value of 23.26 ± 2.06. Based on observed standard deviation (SD) values, *ACT* and *EF1-α* showed the highest- and lowest variation in gene expression across the tissues, respectively. Results are presented in a box-and-whisker plot (Figure 1).

Melting curve analysis showed that primer pairs for five genes (*FBA, UBQ, UBC*, *EF1-α* and *CLC*) displayed a single amplification peak indicating a single product formation and no primer dimer contamination in all four tissue types. For *ACT,* two peaks were observed in the melting curve analysis of nine of 120 samples assayed (7.5%) in three tissue types (leaf, flower and mature capsule) and thus excluded from the following expression stability analyses. The presence of a double peak (with consistent, close T_m peaks) is not always indicative of nonspecific amplification, in this case, the appearance of double peaks only in certain samples in the same set of the experiment suggests existence of amplicon sequence heterogeneity (presence of multiple alleles for *ACT*). Subject to further investigation, *ACT* was excluded from our subsequent stability analysis. No amplification was detected in the negative controls. As shown in Table 2, PCR efficiency of the six candidate reference genes ranged from 80.20% (*ACT*) to 98.95% (*EF1-α*) and R^2 values from 0.9940 (*UBA*) to 0.9997 (*EF1-α*), which are comparable to values reported from other similar studies evaluating suitability of candidate reference genes such as tobacco black root-rot suppressive and conducive soils (Almario *et al.*, 2013), *Ganoderma lucidum* (Liu *et al.*, 2017), lentil (*Lens culinaris*) (Sinha *et al.*, 2019), Chinese tallow (*Sapium sebiferum*) (Chen *et al.*, 2017), and tomato-Pseudomonas pathosystem (Pombo *et al.*, 2017).

Expression stability analyses for candidate reference genes

Expression analysis was performed for the five candidate reference genes (excluding *ACT*) across four different tissue types. BestKeeper ranks reference genes according to values of coefficient

Table 1. CV values of six candidate reference genes obtained from transcriptome data of *Impatiens balsamina*

Gene symbol	Unigene ID	FPKM value								Average of FPKM (Mean expression value, MV)	Standard deviation_ FPKM (SD)	CV (SD/MV)	SwissProt annotation description
		Flower 1	Flower 2	Leaf 1	Leaf 2	Early stage capsule 1	Early stage capsule 2	Mature stage capsule 1	Mature stage capsule 2				
UBC	Unigene 8554_All	126.96	129.01	119.76	122.99	144.29	142.32	104.77	106.65	124.59	14.48	0.12	Q9SVD7 Ubiquitin-conjugating enzyme E2 variant 1D
ACT	CL10259. Contig3_All	136.37	131.19	85.99	84.78	144.26	142.04	83.40	84.27	111.54	29.05	0.26	P53492 Actin-7
FBA	Unigene 31064_All	91.23	93.35	70.46	71.38	138.77	134.34	82.26	75.95	94.72	27.16	0.29	Q9ZU52 Probable fructose-bisphosphate aldolase 3, chloroplastic
CLC	CL8233. Contig2_All	88.93	84.69	80.20	81.06	108.90	108.41	71.97	67.54	86.46	15.26	0.18	Q9SKU1 Clathrin light chain 1
EF1-α	Unigene 21294_All	45.86	44.45	49.79	49.72	68.59	67.14	49.90	46.30	52.72	9.58	0.18	P50256 Elongation factor 1-alpha C
UBQ	CL586. Contig5_All	573.87	611.28	339.74	318.23	408.91	384.23	315.67	325.34	409.66	117.99	0.29	P0CG85 Polyubiquitin

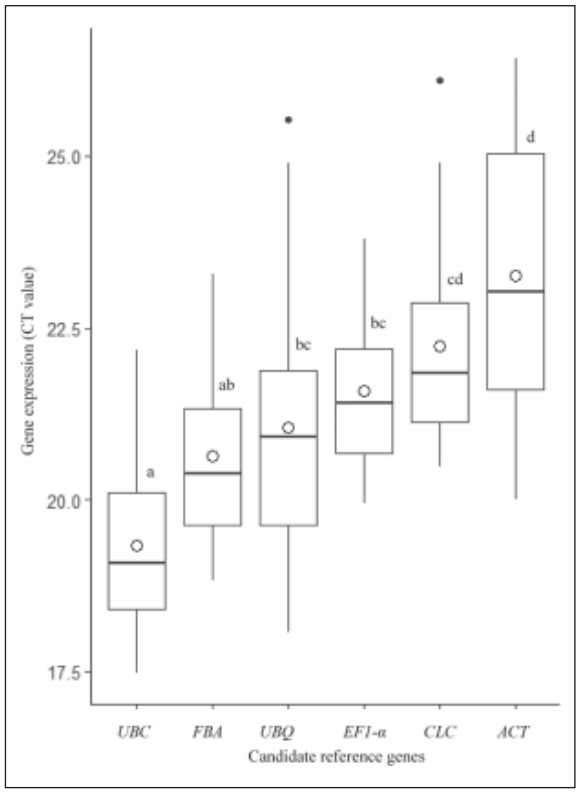

Fig. 1. qRT-PCR Ct values for candidate reference genes. Expression data displayed as Ct values for each reference gene in all samples. See Table 2 for abbreviations of gene names. The boxes represent 25 and 75 percentile range. The vertical lines indicate the value ranges, while the black dots refer to outliers. The horizontal line and black circle (with empty fill) marker in the box indicates the median value and mean value, respectively. One-way ANOVA (Tukey's test) was used to analyse the difference among the CT values of the candidate reference genes. Different letters show the statistical difference, $p < 0.05$.

of variation (CV) and SD calculated from respective Ct values with threshold SD-value < 1.0. Only *EF1-α* showed SD < 1.0 and hence was the most stable gene, although the SD values of both *FBA* (1.04) and *UBC* (1.04) were close (Table 3). NormFinder analysis also detected the best reference gene to be *FBA* (0.283), while the least stable gene was *UBC* (0.497) (Table 3). *FBA* and *EF1-α* were also selected by NormFinder as the best combination of genes with the stability value of 0.221. geNorm results showed that all five candidate reference genes have expression stability (*M)* values below 1.5 (Figure 2), with *CLC* and *FBA* (*M* values = 0.66) ranked as the most stably expressed genes across different tissues, in contrast to *UBQ* (*M* value = 1.13).

Due to their distinct statistical algorithms, it is expected that different ranking results could be obtained from geNorm, NormFinder and BestKeeper as shown in some previous studies (Dai *et al.*, 2017; Fan *et al.*, 2013; Guo *et al.*, 2014), similar to the ranking results as observed in this study. For example, geNorm, NormFinder and ΔCt showed *FBA* as the most stable gene across different tissues in *I. balsamina*, but it was ranked third in BestKeeper analysis. Therefore, RefFinder was used to combine the results generated by ΔCt, BestKeeper, geNorm, and NormFinder, thus providing an integrated evaluation of the overall suitability of the candidate reference genes (Table 4). RefFinder results ranked the candidates in the following order (from the most stable to the least stable): *FBA* > *CLC* > *EF1-α* > *UBC* > *UBQ*.

Table 2. Primer sequences and candidate reference gene parameters of *Impatiens balsamina* obtained from qRT-PCR analysis

Gene abbreviation	Description	Primer sequences (Forward/Reverse)	Amplicon length (bp)	Tm (°C)	Amplification efficiency (%)	Regression Coefficient (R^2)
ACT	Actin-like conserved site-containing protein	5'-ATGGAACTGGGATGGTCAAG/ 5'-CCGTGCTCAATTGGGTATTT	187	76.65	80.20	0.9988
CLC	Clathrin light chain 1	5'-ACTCTCAGCGATCTCCGTTC/ 5'-CTGAAGCCGCATTTCTTTCT	227	78.39	81.33	0.9955
UBQ	Polyubiquitin-A	5'-CCGATACCATCGATAACGTC/ 5'-TATTGTAATCGGCGAGTGTG	123	78.03	87.59	0.9940
UBC	Ubiquitin-conjugating enzyme E2	5'-CGAGAAGCCTCCAACTGTTC/ 5'-ACACCATTTCTTTCCGCAAC	161	75.53	92.94	0.9992
FBA	Fructose-bisphosphate aldolase 3	5'-CGTTATGCTGCCATTTCTCA/ 5'-AGGGGATGCTTTTTCCTTGT	222	75.48	97.09	0.9996
EF1-α	Elongation factor 1-alpha	5'-ATCCGGTAGCATACGTGAGG/ 5'-CGCTGAGATGTCCTCCTCTT	163	75.84	98.95	0.9997

Table 3. Ranking order of the candidate reference genes stability in *Impatiens balsamina* as obtained by BestKeeper and NormFinder analysis

Rank	BestKeeper			NormFinder	
	Gene name*	CV [% CP]	std dev [± CP]	Gene name*	S
1	*EF1-α*	4.44	0.96	*FBA*	0.283
2	*FBA*	5.03	1.04	*CLC*	0.376
3	*CLC*	5.11	1.14	*EF1-α*	0.299
4	*UBC*	5.36	1.04	*UBQ*	0.465
5	*UBQ*	6.94	1.46	*UBC*	0.497
				Best combination of genes	*FBA* and *EF1-α* (S = 0.221)

* See Table 2 for abbreviations of genes. CV = Coefficient of variation; S = Expression stability value.

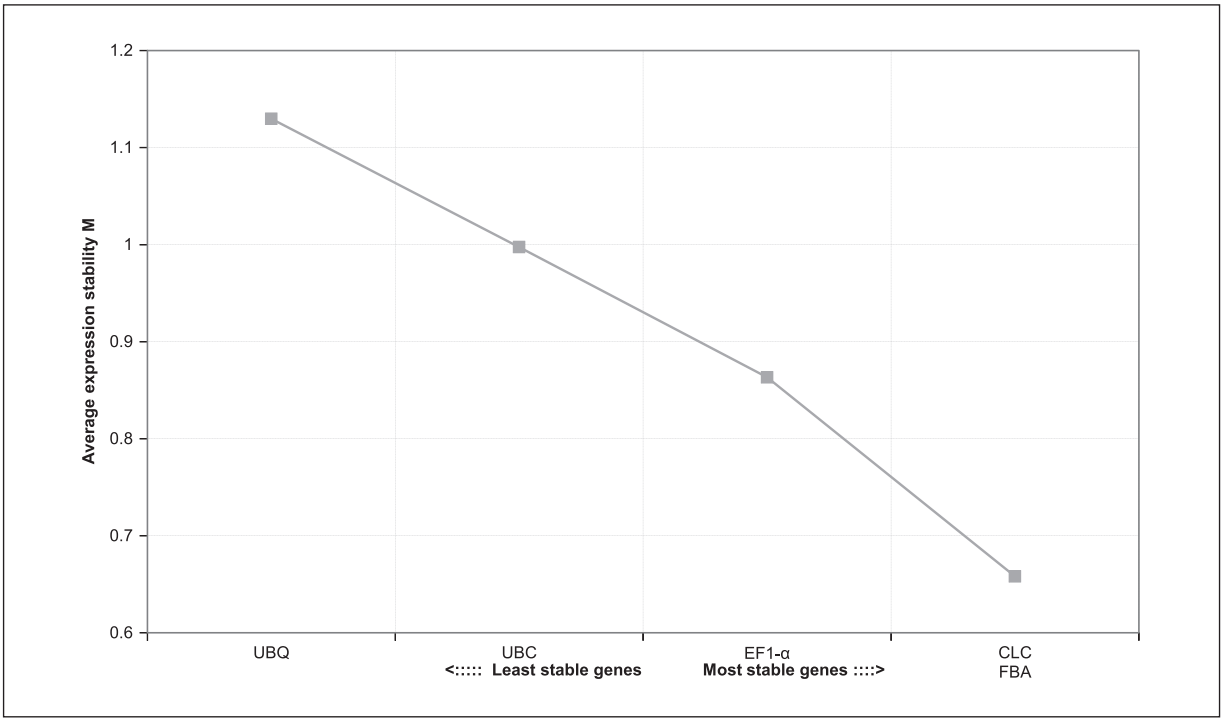

Fig. 2. Average gene expression stability values (M) and ranking of five reference genes by geNorm analysis. Gene expression values of five candidate reference genes across four different tissues, i.e. the leaf, flower, early- and mature stage capsules of *Impatiens balsamina*, were obtained and analysed using geNorm software package. The least stable to the most stable genes were arranged from left to right. See Table 2 for abbreviations of genes.

Table 4. Ranking order of the candidate reference genes stability in *Impatiens balsamina* as determined by RefFinder analysis

Method	Delta CT	BestKeeper	Normfinder	geNorm	Recommended comprehensive ranking
1	*FBA*	*EF1-α*	*FBA*	*CLC* | *FBA*	*FBA*
2	*CLC*	*UBC*	*CLC*		*CLC*
3	*EF1-α*	*FBA*	*EF1-α*	*EF1-α*	*EF1-α*
4	*UBQ*	*CLC*	*UBQ*	*UBC*	*UBC*
5	*UBC*	*UBQ*	*UBC*	*UBQ*	*UBQ*

*See Table 2 for abbreviations of genes.

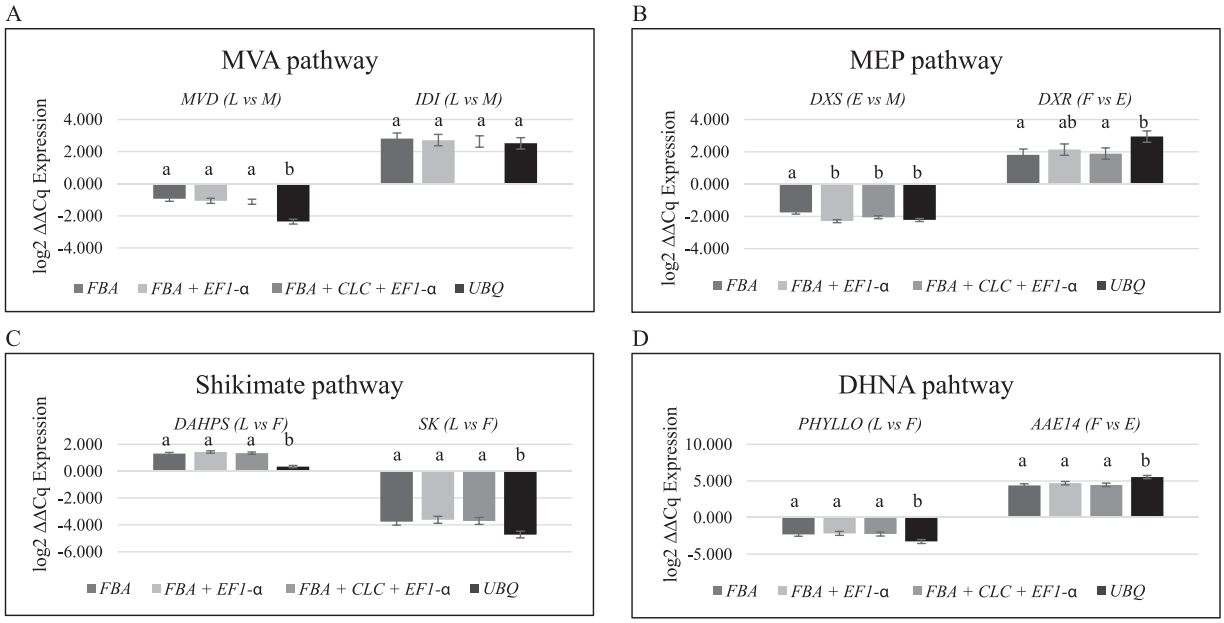

Fig. 3. Relative quantification of eight target genes expression in different tissues using validated reference genes for normalisation. The validated reference gene(s), *FBA* (the most stable reference gene based on four out of five statistical package analysis in this study), *FBA + EF1-á* (the best combination based on NormFinder analysis), *FBA + CLC + EF1-á* (the three most stable reference genes based on RefFinder analysis), and *UBQ* (the least stable reference gene based on most RefFinder analysis), were used as normalisation factors for analysing genes involved in the (A) mevalonate- (MVA-; *MVD, IDI*), (B) 2-C-methyl-D-erythritol 4-phosphate (MEP-; *DXS, DXR*), (C) Shikimate- (*DAHPS, SK*), and (D) 1,4-dihydroxy-2-naphthoate (DHNA-; *PHYLLO, AAE14*) pathways. See Materials and Methods as well as Table 2 for abbreviations of genes. Results are presented as a mean log-transformed fold change (obtained by three biological and two technical replicates) in relative expression of random pairwise tissue comparison between leaf (L), flower (F), early- (E) and mature- (M) stage capsules. One-way ANOVA (Tukey's test) was used to analyse the difference among the CT values of the target gene using different individual/combination of candidate reference genes. Different letters (for the same target gene) show the statistical difference, $p < 0.05$.

Validation of selected reference genes

To examine the effect of reference genes on the normalisation of target gene expression, the relative expression levels of eight secondary metabolite biosynthesis pathway genes across four different tissues of *I. balsamina* were normalised to the most and least stable reference genes, as well as the best combination set as determined by NormFinder software (*FBA + EF1-α*), and the combination of three most stable reference genes observed in this study (*FBA + CLC + EF1-α*). One of the tested target genes (*IDI*) showed consistent results regardless of which individual or combination of reference genes. Statistical analysis showed the expressions of five target genes (*MVD, DXR, DAHPS, PHYLLO, AAE14*) were significantly different when normalised using either only *FBA* or the combination of *FBA + EF1-á* and *FBA + CLC + EF1-á* compared to the least stable gene *UBQ* alone (Figure 3). Similarly, it was observed that expression of *DXR* gene was significantly different when normalised using the best individual or combination reference genes compared to the least stable *UBQ* gene. However, relative expression level of *DXS* gene showed more consistent results when using multiple reference genes as well as *UBQ*, suggesting that the use of three reference genes in qRT-PCR could increase the accuracy and sensibility of gene expression analysis among different tissues of *I. balsamina*. It was suggested that normalisation of relative gene expression data be calculated based on geometric mean of at least three reference genes, because no single gene that shows stable expression under all sets of experimental conditions has been reported (Die & Rowland, 2013; Schmid *et al.*, 2003; Vandesompele *et al.*, 2002). Chi *et al.* (2016) evaluated nine general housekeeping genes for their expression stability by qRT-PCR in four adzuki bean cultivars, three different tissues, four abiotic stress and one biotic stress and found different groups of best reference genes for different conditions. Our results also indicate that the analysis of target gene expression can be affected by different reference genes, and suggest two or three reference genes should be used for qRT-PCR data normalisation.

CONCLUSIONS

For qRT-PCR data normalisation in *I. balsamina* across various tissues (leaf, flower, capsule) and developmental stages (capsule early- and mature), the most suitable reference genes found in this study were *FBA*, *CLC* and *EF1-α*. Moreover, the combination of *FBA* + *EF1-α* and *FBA* + *CLC* + *EF1-α* were appropriate to be used for normalising qRT-PCR data. Future gene expression studies for other experimental treatments or tissues of *Impatiens* species may benefit from the inclusion of the top-ranked reference genes found in this study.

ACKNOWLEDGEMENTS

This work was supported by a Taylor's Research Grant Scheme (TRGS/MFS/1/2017/SBS/002) from Taylor's University and funded in part by a Fundamental Research Grant Scheme (FRGS/2/2014/SG05/TAYLOR/02/1) from the Ministry of Higher Education, Malaysia.

REFERENCES

Almario, J., Moënne-Loccoz, Y. & Muller, D. 2013. Monitoring of the relation between 2, 4-diacetylphloroglucinol-producing Pseudomonas and *Thielaviopsis basicola* populations by real-time PCR in tobacco black root-rot suppressive and conducive soils. *Soil Biology and Biochemistry*, **57:** 144-155.

Andersen, C.L., Jensen, J.L. & Ørntoft, T.F. 2004. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Research*, **64(15):** 5245-5250.

Bustin, S. 2002. INVITED REVIEW Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. *Journal of Molecular Endocrinology*, **29:** 23-39.

Caldana, C., Scheible, W.-R., Mueller-Roeber, B. & Ruzicic, S. 2007. A quantitative RT-PCR platform for high-throughput expression profiling of 2500 rice transcription factors. *Plant Methods*, **3(1):** 7.

Chen, X., Mao, Y., Huang, S., Ni, J., Lu, W., Hou, J. & Wang, Q. 2017. Selection of suitable reference genes for quantitative real-time PCR in *Sapium sebiferum*. *Frontiers in Plant Science*, **8:** 637.

Chi, C., Shen, Y., Yin, L., Ke, X., Han, D. & Zuo, Y. 2016. Selection and validation of reference genes for gene expression analysis in *Vigna angularis* using quantitative real-time RT-PCR. *PLOS ONE*, **11(12):** e0168479.

Chua, L.S. 2016. Untargeted MS-based small metabolite identification from the plant leaves and stems of *Impatiens balsamina*. *Plant Physiology and Biochemistry*, **106:** 16-22.

Czechowski, T., Stitt, M., Altmann, T., Udvardi, M.K. & Scheible, W.-R. 2005. Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. *Plant Physiology*, **139(1):** 5-17.

Dai, T.-M., Lü, Z.-C., Liu, W.-X. & Wan, F.-H. 2017. Selection and validation of reference genes for qRT-PCR analysis during biological invasions: The thermal adaptability of *Bemisia tabaci* MED. *PLOS ONE*, **12(3):** e0173821.

Die, J.V. & Rowland, L.J. 2013. Superior cross-species reference genes: a blueberry case study. *PLOS ONE*, **8(9):** e73354.

Fan, C., Ma, J., Guo, Q., Li, X., Wang, H. & Lu, M. 2013. Selection of reference genes for quantitative real-time PCR in bamboo (*Phyllostachys edulis*). *PLOS ONE*, **8(2):** e56573.

Foong, L.C., Anthony Ho, S.H., Lim, Y.M. & Tam, S.M. 2017. A modified CTAB-based protocol for total RNA extraction from the medicinal plant *Impatiens balsamina* (Balsaminaceae) for next-generation sequencing studies *Malaysian Applied Biology*, **46(11)**.

Gachon, C., Mingam, A. & Charrier, B. 2004. Real-time PCR: what relevance to plant studies? *Journal of Experimental Botany*, **55(402):** 1445-1454.

Guénin, S., Mauriat, M., Pelloux, J., Van Wuytswinkel, O., Bellini, C. & Gutierrez, L. 2009. Normalization of qRT-PCR data: the necessity of adopting a systematic, experimental conditions-specific, validation of references. *Journal of Experimental Botany*, **60(2):** 487-493.

Guo, J., Ling, H., Wu, Q., Xu, L. & Que, Y. 2014. The choice of reference genes for assessing gene expression in sugarcane under salinity and drought stresses. *Scientific Reports*, **4:** 7042.

Huggett, J., Dheda, K., Bustin, S. & Zumla, A. 2005. Real-time RT-PCR normalisation; strategies and considerations. *Genes and Immunity*, **6(4):** 279.

Kozera, B. & Rapacz, M. 2013. Reference genes in real-time PCR. *Journal of applied genetics*, **54(4):** 391-406.

Liu, Y., Wang, Y., Guo, F., Zhan, L., Mohr, T., Cheng, P., Huo, N., Gu, R., Pei, D. & Sun, J. 2017. Deep sequencing and transcriptome analyses to identify genes involved in secoiridoid biosynthesis in the Tibetan medicinal plant *Swertia mussotii*. *Scientific Reports*, **7:** 43108.

Livak, K.J. & Schmittgen, T.D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta}$CT method. *Methods*, **25(4):** 402-408.

Meenu, B., Neeraja, E., Greeshma, R. & Alexeyena, V. 2015. *Impatiens balsamina*: An overview. *Journal of Chemical and Pharmaceutical Research*, **7(9):** 16-21.

Nolan, T., Hands, R.E. & Bustin, S.A. 2006. Quantification of mRNA using real-time RT-PCR. *Nature Protocols*, **1(3):** 1559.

Nurul, A., Nur Arina, H., Subhash, J.B. & Farida, H.S. 2010. Total Phenolic Content and RAPD Analysis of Garden Balsam (*Impatiens balsamina* L.) Accessions from Malaysia. *Middle-East Journal of Scientific Research*, **5(6):** 454-463.

Pfaffl, M.W., Tichopad, A., Prgomet, C. & Neuvians, T.P. 2004. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper–Excel-based tool using pair-wise correlations. *Biotechnology Letters*, **26(6):** 509-515.

Pombo, M.A., Zheng, Y., Fei, Z., Martin, G.B. & Rosli, H.G. 2017. Use of RNA-seq data to identify and validate RT-qPCR reference genes for studying the tomato-Pseudomonas pathosystem. *Scientific Reports*, **7:** 44905.

Radonić, A., Thulke, S., Mackay, I.M., Landt, O., Siegert, W. & Nitsche, A. 2004. Guideline to reference gene selection for quantitative real-time PCR. *Biochemical and Biophysical Research Communications*, **313(4):** 856-862.

Schmid, H., Cohen, C.D., Henger, A., Irrgang, S., Schlöndorff, D. & Kretzler, M. 2003. Validation of endogenous controls for gene expression analysis in microdissected human renal biopsies. *Kidney International*, **64(1):** 356-360.

Shah, K.N., Verma, P. & Suhagia, B. 2017. A phyto-pharmacological overview on Jewel Weed. *Journal of Applied Pharmaceutical Science,* **7(08):** 246-252.

Singh, P., Singh, R., Sati, N., Ahluwalia, V. & Sati, O.P. 2017. Phytochemical and Pharmacological Significance of Genus: *Impatiens*. *International Journal of Life Sciences and Scientific Research*, **3(1):** 868-881.

Sinha, R., Sharma, T. & Singh, A.K. 2019. Validation of reference genes for qRT-PCR data normalisation in lentil (*Lens culinaris*) under leaf developmental stages and abiotic stresses. *Physiology and Molecular Biology of Plants*, **25(1):** 123-134.

Stürzenbaum, S.R. & Kille, P. 2001. Control genes in quantitative molecular biological techniques: the variability of invariance. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, **130(3):** 281-289.

Suzuki, T., Higgins, P.J. & Crawford, D.R. 2000. Control selection for RNA quantitation. *BioTechniques*, **29(2):** 332-337.

Thellin, O., Zorzi, W., Lakaye, B., De Borman, B., Coumans, B., Hennen, G., Grisar, T., Igout, A. & Heinen, E. 1999. Housekeeping genes as internal standards: use and limits. *Journal of Biotechnology*, **75(2-3):** 291-295.

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(7):** research0034-0031.

VanGuilder, H.D., Vrana, K.E. & Freeman, W.M. 2008. Twenty-five years of quantitative PCR for gene expression analysis. *BioTechniques*, **44(5):** 619-626.

Xie, F., Xiao, P., Chen, D., Xu, L. & Zhang, B. 2012. miRDeepFinder: a miRNA analysis tool for deep sequencing of plant small RNAs. *Plant Molecular Biology*, **80(1):** 75-84.

Zhou, Z., Cong, P., Tian, Y. & Zhu, Y. 2017. Using RNA-seq data to select reference genes for normalizing gene expression in apple roots. *PLOS ONE*, **12(9):** e0185288.