





Citation: Kong Q, Yuan J, Gao L, Zhao L, Cheng F, Huang Y, et al. (2015) Evaluation of Appropriate Reference Genes for Gene Expression Normalization during Watermelon Fruit Development. PLoS ONE 10 (6): e0130865. doi:10.1371/journal.pone.0130865

Editor: Robert W Dettman, Northwestern University, UNITED STATES

Received: January 4, 2015
Accepted: May 26, 2015
Published: June 25, 2015

Copyright: © 2015 Kong et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by the National Natural Science Foundation of China (31471894) (QK), the earmarked fund for Modern Agro-industry Technology Research System (CARS-26-16) (ZB), and the Special Fund for Agro-scientific Research in the Public Interest (201203080-2) (ZB). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

RESEARCH ARTICLE

Evaluation of Appropriate Reference Genes for Gene Expression Normalization during Watermelon Fruit Development

Qiusheng Kong, Jingxian Yuan, Lingyun Gao, Liqiang Zhao, Fei Cheng, Yuan Huang, Zhilong Bie*

Key Laboratory of Horticultural Plant Biology, Ministry of Education, College of Horticulture and Forestry Sciences, Huazhong Agricultural University, Wuhan, China

* biezhilong@hotmail.com

Abstract

Gene expression analysis in watermelon (Citrullus lanatus) fruit has drawn considerable attention with the availability of genome sequences to understand the regulatory mechanism of fruit development and to improve its quality. Real-time quantitative reverse-transcription PCR (gRT-PCR) is a routine technique for gene expression analysis. However, appropriate reference genes for transcript normalization in watermelon fruits have not been well characterized. The aim of this study was to evaluate the appropriateness of 12 genes for their potential use as reference genes in watermelon fruits. Expression variations of these genes were measured in 48 samples obtained from 12 successive developmental stages of parthenocarpic and fertilized fruits of two watermelon genotypes by using qRT-PCR analysis. Considering the effects of genotype, fruit setting method, and developmental stage, geNorm determined clathrin adaptor complex subunit (ClCAC), β -actin (ClACT), and alpha tubulin 5 (CITUA5) as the multiple reference genes in watermelon fruit. Furthermore, CICAC alone or together with SAND family protein (CISAND) was ranked as the single or two best reference genes by NormFinder. By using the top-ranked reference genes to normalize the transcript abundance of phytoene synthase (CIPSY1), a good correlation between lycopene accumulation and CIPSY1 expression pattern was observed in ripening watermelon fruit. These validated reference genes will facilitate the accurate measurement of gene expression in the studies on watermelon fruit biology.

Introduction

Watermelon (*Citrullus lanatus*) is a popular and economically important horticultural crop in terms of production and consumption. During development and ripening, watermelon fruits undergo various physiological and biochemical changes, resulting in diverse sizes, colors, shapes, sweetness, textures, and aromas [1]. Watermelon is gaining popularity as a model plant for studying non-climacteric fruits. Studies have focused on the unique metabolic and regulatory networks of watermelon fruit to improve the crucial nutritional attributes of this fruit [2].



Competing Interests: The authors have declared that no competing interests exist.

Moreover, the vast amounts of genomic resources are beneficial for elucidating the biological processes involved in fruit development. Transcriptome sequencing has been conducted to study the transcriptional regulatory networks in watermelon fruit, revealing many key genes involved in sugar, citrulline, and carotenoid metabolisms [1, 3-5]. Target gene expression patterns during fruit development can provide clues to understand its biological functions. Realtime quantitative reverse-transcription PCR (qRT-PCR) is an efficient method to measure gene transcript abundance and to validate gene expression changes detected in RNA-Seq [1, 6] and microarray [2].

In qRT-PCR analysis, gene expression is quantified by normalizing to one or more internal reference genes presumed to be stable throughout a given experiment. Normalization removes considerable non-biological variations associated with the different steps of sample preparations and expression measurements [7]. Accordingly, qRT-PCR accuracy is strongly influenced by internal reference gene stability. The use of unstable reference genes in relative quantification of gene expression could lead to significant biases and misinterpretations of data [8–10]. Moreover, reference genes are not constantly expressed under a wide range of experimental conditions [11, 12]. Thus, systematic validation of reference genes must be conducted on each experiment prior to their use [13, 14]. Statistical algorithms such as geNorm [15] and Norm-Finder [16] have been developed to determine gene expression stability, which greatly simplified the choice of appropriate reference genes for normalization in qRT-PCR analysis.

Suitable reference genes have been identified in watermelon under normal growth conditions as well as biotic and abiotic stresses, providing a good starting point for reference gene selection in qRT-PCR analysis [17]. However, the reference genes in watermelon fruit have not been well characterized in this study. Considering all the organs and tissues under normal growth conditions, at least nine reference genes are required for accurate normalization in watermelon [17]. The use of numerous reference genes is not feasible in practice, particularly when the target genes are few. Meanwhile, an increasing number of evidence shows that reference genes are not universally applicable [12, 18]. Thus, the suitability of candidate reference genes must be carefully evaluated for each experimental system [19]. Stable reference genes have been specifically identified for the fruits of several crops, such as papaya [20] and blueberry [21]. To date, the single, non-validated 18SrRNA has been used as reference gene in qRT-PCR analyses of watermelon fruit [1, 2, 22]. Considering that fruit development is a very complex process, the selection of appropriate reference genes that are not affected by developmental processes is required to study gene expression during fruit development. The identification of the fruit-specific reference genes for watermelon is critical because of its significance.

Fertilization and parthenocarpy are two different approaches for fruit setting that involves distinct gene expressions and regulatory systems during fruit development [23]; thus, they provide good opportunities for selecting reference genes with intrinsic expression stability during fruit development and ripening. In this study, the expression stability of 12 candidate reference genes was evaluated and compared across 48 cDNA samples obtained from 12 developmental stages of parthenocarpic and fertilized fruits of two watermelon genotypes. The appropriate reference genes, which were specific for normalizing the target genes in watermelon fruit by qRT-PCR analysis, were selected using geNorm and NormFinder algorithms.

The carotenoid content is one of the key characteristics to determine the fruit quality of watermelon. Lycopene is the dominant carotenoid in red-fleshed watermelon cultivars, contributing to the red color of watermelon flesh [1]. Phytoene synthase (PSY) catalyzes the first step in carotenoid biosynthetic pathway and is one of the key enzymes that regulate carotenoid contents in watermelon fruits. Two members of PSY gene family, namely, *ClPSY1* (Cla009122) and *ClPSY2* (Cla005425), were identified by transcriptome analysis; however, only *ClPSY1* was associated with the lycopene accumulation in watermelon fruits [1, 5]. Thus, the reliability of



the identified reference genes was further validated by normalizing *ClPSY1* expression and testing its correlation with lycopene accumulation during fruit ripening. The validated reference genes will improve the accuracy of gene expression studies in watermelon fruits.

Materials and Methods

Plant materials and treatments

Watermelon (C. lanatus) inbred line 97103 and commercial F₁ hybrid 8424 were used in the present study. Both were mature at 30 days after pollination and had red flesh. Seedlings at the third true-leaf stage were transplanted in a plastic greenhouse at Huazhong Agricultural University, Wuhan, China (east longitude 113°41′-115°05′, north latitude 29°58′-31°22). Fertilizer application and pest control were implemented following the standard commercial practices. Parthenocarpic fruits were induced by the application of *N*-(2-chloro-4-pyridyl)-*N*'-phenylurea (CPPU). One day before anthesis, the female flowers were isolated with paper bags to prevent natural pollination. On the day of anthesis, the paper bags were removed, and CPPU (Shiteyou, China) was applied by spraying unpollinated ovaries at a concentration of 10 μM in the morning following the manufacturer's instruction. Afterward, the female flowers were covered again until the success of fruit set. Simultaneously, artificial pollination was performed on other female flowers by hand. The pollinated and CPPU-treated female flowers were tagged, and only one fruit per plant was allowed to develop. Fruits were harvested throughout the process of fruit development at 0, 1, 3, 5, 7, 9, 12, 18, 23, 27, 30, and 35 days after anthesis (DAA). Three fruits were randomly harvested at each sampling point, and each fruit represented a biological replication. The ovaries were sampled at 0, 1, and 3 DAA. Beyond 3 DAA, the harvested fruits were cut longitudinally, and flesh tissue was obtained from the center of the fruit. All the samples were immediately frozen in liquid nitrogen and then stored at -80°C for subsequent RNA extraction and lycopene measurements.

Total RNA isolation, cDNA synthesis, DNA extraction, and PCR amplification

The 11 golden rules of qRT-PCR were observed in the experiments [24]. Total RNA was isolated from the frozen samples by following the TransZol (TransGen, China) protocol. The integrity of the extracted RNA was verified using 2% agarose gel electrophoresis. The concentration and purity of RNA were measured using NanoDrop 2000 spectrophotometer (Thermo Scientific, China). RNA samples with $A_{260}/A_{280} > 1.8$ and $A_{260}/A_{230} > 2.0$ were used for subsequent cDNA synthesis. Before cDNA synthesis, genomic DNA (gDNA) was eliminated using PrimeScript RT Reagent Kit with gDNA Eraser (Perfect Real Time) (TaKaRa, China) according to the manufacturer's instructions. An equal amount of total RNA (1 μ g) from each sample was used for each 20 μ L reverse transcription reaction system. gDNA was extracted from the samples using Plant Genomic DNA Kit (Tiangen, China) and amplified using 2 × PCR Reagent (Tiangen, China). The amplified products were resolved by 2% agarose gel electrophoresis.

Candidate reference gene selection and primer design

Reference genes that had been validated on other crops, particularly in fruits, were selected as candidates. The descriptions of these genes are provided in Table 1. For each candidate reference gene, BLASTN was conducted in the Cucurbit Genomics Database (http://www.icugi.org) against watermelon CDS (v1) by using *Arabidopsis* homolog as a query. The sequence of the best-matched gene with its structural information was downloaded. Primer3Plus (http://primer3plus.com/cgi-bin/dev/primer3plus.cgi) was employed for primer design. To control



Table 1. Description of the candidate reference genes, primer sequences, and PCR amplification characteristics.

Gene Name	Gene description	Gene ID ^a	Forward primer sequence (5'-3')	Reverse primer sequence (5'-3')	Product size (bp)	E (%)	R ²
CICAC	Clathrin adaptor complex subunit	Cla016178	GAACTTGGCACCTGTCCTGT	GAACAGTGCAACAGCCTCAA	147	104.3	0.999
CIPP2A	Protein phosphatase 2A regulatory subunit A	Cla021905	GTGATTATGTGGACCGTGGA	TCCAGACATTTGCATTTCCA	180	95.9	1.000
CIRAN	A member of RAN GTPase gene family	Cla012277	TTGAAAAGAAATACGAACCAACC	ATCCCGTAAACCACCGAACT	128	103.0	0.999
CIRPS15	Cytosolic ribosomal protein S15	Cla021565	AAGCTGCGAAAAGCGAAAC	TAATGGCCAATCATCTCAGG	170	92.8	0.992
CISAND	SAND family protein	Cla001870	TGCAAACATAAGGTTATCAGTCTTG	GCATACAAAAACGCCATAGGA	168	96.7	0.998
CITBP2	TATA binding protein 2	Cla011119	CCAGAGTTATTCCCTGGATTG	CCTGGACATGCGCCTTAG	109	93.4	0.998
CITIP41	TIP41-like family protein	Cla016074	CAAGCTCTCGCTGAAAAAGG	GAGACTCTGAGCTTTTGGGTTT	112	96.4	1.000
CITUA5	Alpha tubulin 5	Cla003129	GATGGTATGATGCCCAGTGA	CCGGTAGGCTCCAGTTCTAA	156	100.2	1.000
CITUB	β-tubulin	Cla022418	CAGCACTCCTAGCTTTGGTGA	CGGGGAAATGGGATTAGATT	136	96.9	0.999
CIUPL7	Ubiquitin-protein ligase 7	Cla017746	TGGCAAACGACATGTTATTGA	TTGAAAAAGGCAATATCTGTCG	128	105.7	1.000
CIACT b	β-actin		CCATGTATGTTGCCATCCAG	GGATAGCATGGGGTAGA	GCA	104.0	0.999
CI18SrRNA	18SrRNA		AGCCTGAGAAACGGCTACCACATC	ACCAGACTCGAAGAGCCCG	GTAT	94.8	0.998

^a Watermelon gene ID in the Cucurbit Genomics Database (http://www.icugi.org).

doi:10.1371/journal.pone.0130865.t001

the potential gDNA contaminations in cDNA samples, one primer sequence was designed to cover an exon–exon junction. Then, the generated primer pairs were aligned against all water-melon CDS by Primer-BLAST (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) to confirm its specificity on genome scale. PCR amplification specificity for each gene was further verified by 2% agarose gel electrophoresis, with gDNA and cDNA as templates. Finally, the watermelon species Cl was added as a prefix to the gene name to designate the watermelon orthologous gene. For more comparable results, the known primer sequences of ClACT and 18SrRNA, which had been used as reference genes in watermelon, were used in this study [1, 17].

qRT-PCR analysis

Approximately 80 ng of cDNA was used as template for qRT-PCR analysis. cDNA samples were amplified by using $1\times Top$ Green qPCR SuperMix (TransGen, China) and 0.2 μM of each primer in a final volume of 10 μL . The reactions were performed on a LightCycler480 System (Roche, Switzerland) with two technical replicates for each of the three biological replicates described previously. The negative controls without templates were also included in each run. PCR cycling conditions consisted of an initial denaturation at 94°C for 30 s, followed by 40 cycles of denaturation at 95°C for 5 s, annealing at 55°C for 15 s, and extension at 72°C for 10 s. Melting cure analysis was performed at the end of amplification to verify its specificity. PCR

^b The primer pairs of *CIACT* previously published by Kong et al. [17] and 18SrRNA previously published by Guo et al. [1] were used here.



amplification efficiency for each gene was determined by amplifying fivefold serial dilution of a mixture of all cDNA samples (from 800 ng μL^{-1} to 1.28 ng μL^{-1}).

Evaluation of expression stability

Expression levels depicted as crossing point (Cp) values were recorded for the tested reference genes after each qRT-PCR reaction. A boxplot was drawn using R package (http://www.rproject.org/) to show the expression variation of each gene. The amplification efficiency (E) was calculated using the following formula: $E = (10^{-1/\text{slope}} - 1)$, where the slope is the standard curve slope generated by amplifying the fivefold serial dilution of pooled cDNA samples. NormFinder [16] and geNorm [15] were used to evaluate the gene expression stability. GeNorm calculated not only the gene expression stability value (M) for each gene, but also the pairwise variation values (V) to determine the minimum number of reference genes required for reliable normalization. No additional genes were required for normalization when the pairwise variation $(V_{n/n+1})$ was below 0.15 [15]. Using a model-based approach, NormFinder considered the variations across groups and calculated the expression stability for each reference gene [16]. Lower stability values in both algorithms implied the higher expression stability of the genes. Given that the input data were supposed to be on a linear scale, the raw Cp values were corrected by PCR efficiency and converted into relative expression quantities Q by using the formula: $Q = (1 + E)^{(\min Cp - \text{sample}Cp)}$ before data entry. To evaluate the effects of genotypes and fruit setting methods on the expression stability of the candidate reference genes, 48 diverse samples were further categorized into different subgroups according to the treatments during data analysis.

Determination of *CIPSY1* expression and lycopene contents during fruit ripening

The samples derived from the pollinated fruit of 8424 at 10, 18, 23, 27, 30, and 35 DAA were used to analyze CIPSY1 (Cla009122) expression patterns, which was associated with lycopene accumulation in watermelon fruit [5]. The primer sequences used for ClPSY1 were F: 5'-CTAG CAGATGGCCGGTGT-3' and R: 5'-GCCCTCTTTGTGAAGTTGTTG-3', which were designed and validated according to the methods mentioned above. The top-ranked reference genes identified in this study, as well as 18SrRNA, were used for the normalization. The relative expression levels of *ClPSY1* were evaluated using $2^{-\triangle\triangle Cp}$ method. Normalization factor was calculated as the geometric mean of two or multiple reference genes and used for normalization. The fruit flesh was still white in color at 10 DAA but turned to pink at 18 DAA. Accordingly, the lycopene contents were measured at 18, 23, 27, 30, and 35 DAA. Extraction of carotenoids and measurement of lycopene contents were conducted as described by Liu et al. [25]. Waters 1525 reversed phase HPLC equipped with 2996 photodiode array detector, 717 plus autosampler, Empower Chromatography Manager software (Waters, USA), and YMC C30 carotenoid column (150 mm × 4.6 mm i.d., 3 µm, YMC, USA) was used for lycopene detection and quantification. HPLC-grade lycopene standard was obtained from Sigma-Aldrich (Shanghai, China). Three biological and two technical replicates were adopted in qRT-PCR analysis and lycopene content measurement for each sample.

Results

Amplification specificity and efficiency of candidate reference genes

The candidate reference genes included 10 reference genes that were previously validated in fruits of other crops, *Cl18SrRNA*, and *ClACT*. To prevent the interferences of potential gDNA



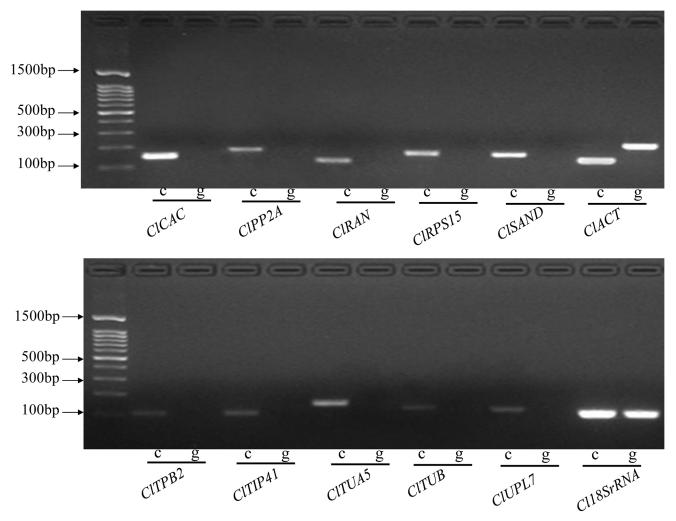


Fig 1. PCR amplification products of the candidate reference genes on cDNA and gDNA templates. "c" represents the cDNA template. "g" represents the gDNA template.

doi:10.1371/journal.pone.0130865.g001

contaminations in cDNA samples on the expression analysis, one primer was designed to cover an exon–exon junction for each gene. For more comparable results, primer pairs of *Cl18SrRNA* and *ClACT* published previously were also used in this study. PCR-amplification specificity of each primer pair was verified by agarose gel electrophoresis using cDNA and gDNA templates. Each reference gene amplified a specific product on cDNA templates, but no products were detected on gDNA templates except for *ClACT* and *Cl18SrRNA*, demonstrating the success of primer design (Fig 1). Moreover, *ClACT* amplified a larger band containing an intron on gDNA, but not on cDNA, indicating the absence of gDNA contamination in cDNA samples. *Cl18SrRNA* amplified the same bands on both gDNA and cDNA templates. A single peak with no visible primer–dimer formation was observed in melting-curve analysis, further confirming the specific amplification of each reference gene (S1 Fig). No signals were detected in the controls without templates in qRT-PCR reactions. PCR-amplification efficiencies of the candidate genes ranged from 92.8% (*ClRPS15*) to 105.7% (*ClUPL7*). The determination coefficients (*R*²) of the standard curve regression equation varied from 0.992 (*ClRPS15*) to 1.000 (*ClPP2A*). These parameters indicated that qRT-PCR systems established for the candidate



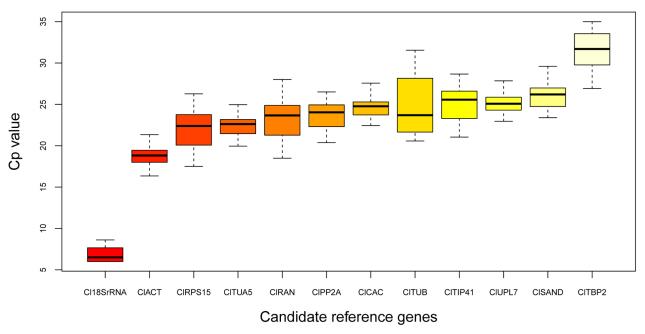


Fig 2. Boxplot analysis of expression variations of the tested reference genes in all 48 samples. The line across the box is the median. The boxes are 25/75 percentiles. Whisker caps are the minimum and maximum values.

doi:10.1371/journal.pone.0130865.g002

reference genes were highly specific and efficient. The primer sequences and amplification characteristics of the candidate reference genes are presented in <u>Table 1</u>.

Expression patterns of candidate reference genes

Similar to artificial pollination, 100% fruit set was obtained with CPPU treatment. Pollinated fruits had normal seeds, whereas CPPU treatments resulted in the development of seedless fruits. Expression profiles of the 12 candidate reference genes in watermelon fruits were obtained by qRT-PCR analysis. Expression variations among the two genotypes and two fruit setting methods in 12 successive development stages were found (Fig 2). The mean Cp values for the tested genes varied from 6.8 (Cl18SrRNA) to 31.9 (ClTBP2) and were mainly between 21.0 and 26.0. Cl18SrRNA, ClUPL7, and ClTUA5 exhibited smaller expression variations among the selected reference genes (< 5 cycles), whereas ClTBP2, ClRPS15, ClRAN, and ClTUB demonstrated relatively higher expression variations (> 8 cycles). Direct comparison of raw Cp values could not accurately estimate the expression stability of each reference gene because of the existence of non-biological variations in experimental processes. Consequently, statistical approaches were required to evaluate the expression stability of the candidate reference genes.

Expression stability of candidate reference genes

The widely used tools of geNorm and NormFinder were used to assess the expression stability of the tested genes. Considering the effects of genotypes, fruit setting methods and developmental phases on the 48 samples, geNorm determined that *ClCAC* and *ClACT* was the best pair of reference genes (<u>Table 2</u>). Pairwise variation analysis showed that at least three reference genes, including *ClCAC*, *ClACT*, and *ClTUA5*, were required for more reliable normalization when the proposed cut-off value of 0.15 was adopted (<u>S2 Fig</u>). To study the effects of genotype and fruit setting method on expression stability of the reference genes and to identify more



Table 2. Expression stability of the candidate reference genes determined by geNorm.

Ranking			97103					84	24		97103 and 8424			
	All samples	M n = 48 ^a	Pollination	M n = 12	CPPU treatment	M n = 12	Pollination	M n = 12	CPPU treatment	M n = 12	Pollination	M n = 24	CPPU treatment	M n = 24
1	CICAC ^b	0.41	CISAND b	0.21	CICAC b	0.26	CICAC ^b	0.26	CICAC ^b	0.21	CICAC b	0.46	CICAC ^b	0.35
1	CIACT b	0.41	CIACT ^b	0.21	CIACT ^b	0.26	CIACT ^b	0.26	CIACT b	0.21	CIACT ^b	0.46	CIACT ^b	0.35
2	CITUA5 b	0.53	CICAC	0.31	CITUA5 ^b	0.41	CITUA5 ^b	0.4	CIPP2A	0.37	CIUPL7 b	0.56	CITUA5 ^b	0.45
3	CIUPL7	0.58	CIPP2A	0.34	CIPP2A	0.48	CIUPL7	0.46	CISAND	0.41	CITUA5 b	0.62	CIPP2A	0.51
4	CISAND	0.63	CITIP41	0.41	CIUPL7	0.56	CISAND	0.6	CITUA5	0.42	CISAND	0.66	CISAND	0.56
5	CIPP2A	0.67	CIRPS15	0.45	CISAND	0.62	CIPP2A	0.68	CIUPL7	0.48	CIPP2A	0.74	CIUPL7	0.6
6	CITIP41	0.78	CIUPL7	0.48	CITIP41	0.7	CITBP2	0.76	CITIP41	0.63	CITIP41	0.85	CITIP41	0.7
7	CIRPS15	0.86	CITUA5	0.56	CIRPS15	0.78	CITIP41	0.91	CITBP2	0.75	CIRPS15	0.92	CIRPS15	8.0
8	CIRAN	0.97	CIRAN	0.62	CIRAN	0.85	CIRPS15	1.03	CIRPS15	0.85	CIRAN	1.01	CIRAN	0.92
9	CI18SrRNA	1.08	CI18SrRNA	0.69	CI18SrRNA	0.95	CIRAN	1.12	CIRAN	0.94	CI18SrRNA	1.12	CI18SrRNA	1.04
10	CITBP2	1.21	CITBP2	1	CITBP2	1.1	CI18SrRNA	1.26	CI18SrRNA	1.08	CITBP2	1.27	CITBP2	1.15
11	CITUB	1.41	CITUB	1.22	CITUB	1.37	CITUB	1.44	CITUB	1.3	CITUB	1.45	CITUB	1.38

^a The sample size.

doi:10.1371/journal.pone.0130865.t002

stable reference genes under specific fruit development conditions, 48 samples were further divided into different subgroups according to the treatments listed in <u>Table 2</u>. In the inbred line 97103, the best pairs of reference genes for pollination and CPPU treatments were *ClSAND* and *ClACT*; and *ClCAC* and *ClACT*. In F₁ hybrid 8424, *ClCAC* and *ClACT* were ranked as the best pair of genes regardless of fruit setting methods; they were also the best pair of genes for different fruit setting treatments regardless of genotypes. However, pairwise variations showed that different minimum numbers of reference genes (from 2 to 4) were required for more accurate normalization in the different subgroups (<u>Table 2</u>, <u>S2 Fig</u>). *ClTUB* was identified as the least stable gene under all conditions.

NormFinder is more powerful in detecting inter- and intra-group variations. Therefore, the samples were divided into different subgroups based on genotypes and fruit setting methods in calculations. The results are shown in Table 3. Considering the four subgroups of two genotypes and their respective fruit setting methods, NormFinder ranked CICAC as the single best reference gene and CICAC and CISAND as the combination of two best genes. The effects of genotype and fruit setting method on the expression stability of the candidate reference genes were also evaluated. CICAC alone or together with CIRAN was ranked as the single best reference gene or the combination of two best genes during fruit development of 97103. Similarly, CISAND alone or together with CICAC was ranked as the single or two best reference genes during fruit development of 8424 regardless of fruit setting methods. During the development of pollinated fruits, which included genotypes 97103 and 8424, CICAC was the single best reference gene, whereas CIPP2A and CIACT were the two best genes. By contrast, CISAND alone or together CICAC were ranked as the single or the two best reference genes in the parthenocarpic fruit induced by CPPU treatment. CITBP2 and CITUB were ranked as the least stable genes under different conditions.

Cl18SrRNA, frequently used as reference gene in watermelon fruit, was ranked both by geNorm and NormFinder as second to fifth from the bottom in all samples and different subsets,

b The multiple reference genes determined by pairwise variation analysis when the cut-off value 0.15 was adopted (S2 Fig).



Table 3. Expression stability of the candidate reference genes determined by NormFinder.

Ranking	All samples (n = 48 ^a)		97103 (n = 24)		8424 (n = 24)		Pollination (n = 24)		CPPU treatment (n = 24)	
	Gene name	М	Gene name	М	Gene name	М	Gene name	М	Gene name	М
1	CICAC ^b	0.16	CICAC ^b	0.11	CISAND ^b	0.09	CICAC	0.14	CISAND ^b	0.11
2	CISAND b	0.18	CIACT	0.17	CICAC b	0.16	CIPP2A b	0.15	CICAC b	0.16
3	CIACT	0.22	CIRAN ^b	0.17	CITUA5	0.17	CISAND	0.2	CIACT	0.19
4	CIPP2A	0.3	CITIP41	0.18	CIACT	0.2	CIACT b	0.27	CITIP41	0.21
5	CITIP41	0.32	CIRPS15	0.2	CIPP2A	0.22	CITUA5	0.34	CIUPL7	0.27
6	CITUA5	0.34	CISAND	0.22	CITIP41	0.26	CIUPL7	0.36	CIPP2A	0.3
7	CIRPS15	0.41	CITUA5	0.24	CIRAN	0.28	CITIP41	0.39	CIRPS15	0.32
8	CIUPL7	0.42	CIUPL7	0.29	CIUPL7	0.29	CI18SrRNA	0.41	CITUA5	0.33
9	CI18SrRNA	0.46	CIPP2A	0.32	CIRPS15	0.29	CIRPS15	0.53	CIRAN	0.42
10	CIRAN	0.49	CI18SrRNA	0.32	CITBP2	0.3	CIRAN	0.57	CITBP2	0.46
11	CITUB	0.62	CITBP2	0.45	CI18SrRNA	0.42	CITUB	0.6	CI18SrRNA	0.46
12	CITBP2	0.73	CITUB	0.55	CITUB	0.49	CITBP2	1.08	CITUB	0.61

^a The sample size.

doi:10.1371/journal.pone.0130865.t003

indicating that this gene was unsuitable for normalization in qRT-PCR analysis in watermelon fruit.

Validation of top-ranked reference genes

To validate the reliability of the top-ranked reference genes, expression pattern of *ClPSY1* and lycopene accumulation were analyzed in the pollinated fruits of 8424. The single (*ClCAC* as determined by NormFinder), pair (*ClCAC* and *ClACT* as determined by geNorm, and *ClPP2A* and *ClACT* as determined by NormFinder), and multiple best reference genes (*ClCAC*, *ClACT*, *ClUPL7*, and *ClTUA5* as determined by geNorm) identified in the pollination subset regardless of genotypes were used to normalize the expressions of *ClPSY1*. *Cl18SrRNA*, the widely used reference gene in watermelon fruit, was used as the control. The results are presented in Fig 3.

CIPSY1 was expressed at 10 DAA with the raw Cp value of 24.4. Thus, the transcript abundance of CIPSY1 at 10 DAA was set as the control. By using the top-ranked reference genes for normalization, the same CIPSY1 expression trends were observed, gradually increasing after 10 DAA, peaking at 27 DAA, and then decreasing (Fig 3). By contrast, different CIPSY1 expression levels were observed using different top-ranked reference genes. The highest expression level was observed using the best single reference gene CICAC, whereas the lowest expression level was obtained using the combination of CIPP2A and CIACT as determined by NormFinder. Moderate expression levels were found using the pair and multiple reference genes as determined by geNorm, with only slight difference on expression levels noted between the two observations. However, using Cl18SrRNA for normalization, distinct CIPSY1 expression pattern was observed, which remained stable from 10 DAA to 18 DAA, sharply increasing to the highest level at 23 DAA and then decreased as the fruit matured.

Lycopene accumulation was measured from 18 DAA, whereas the fruit flesh was white at 10 DAA. The results showed that lycopene content gradually increased as the fruit ripened and reached the highest level of 39.08 $\mu g \cdot g^{-1}$ fresh weight at 27 DAA and then decreased and remained stable from 30 DAA to 35 DAA. Good correlations were found between lycopene

^b The combination of two best genes.



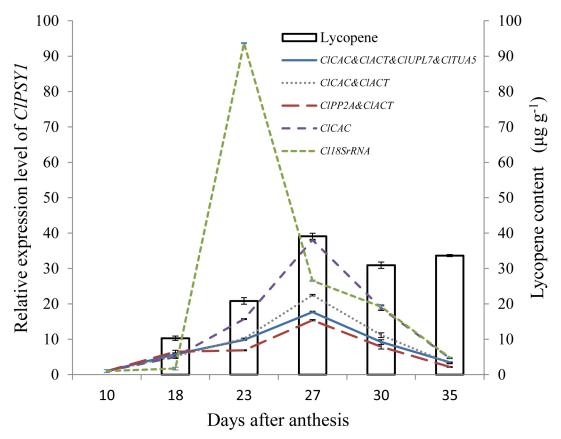


Fig 3. Lycopene accumulation and expression profiles of CIPSY1 during fruit ripening. Geometric mean was calculated for the two or multiple reference genes, and used for normalization. CI18SrRNA was used as control. The results are depicted as mean \pm SE (n = 6).

doi:10.1371/journal.pone.0130865.g003

biosynthesis until 27 DAA and *ClPSY1* expression by using the top-ranked reference genes for normalization. However, when normalized by *18SrRNA*, no correlation was observed between *ClPSY1* expression pattern and lycopene accumulation in ripening fruit.

Discussion

Accurate interpretation of qRT-PCR results mainly depends on the use of stable reference genes for data normalization to minimize non-biological variations between samples. Consequently, systematic validation of reference genes is essential to produce reliable data in qRT-PCR analysis [13, 19, 24]. At present, systematic validation of reference genes is always conducted across a wide range of organ samples in different experimental conditions [17, 26]. However, increasing evidence shows that the best reference genes varied significantly depending on experimental conditions and organs assayed; compiling a list of suitable genes that could be used as references across a wide range of experimental conditions was impossible [19, 27]. Therefore, validating references under specific experimental condition is essential. To date, expression studies performed on watermelon fruits have used only *18SrRNA* for normalization [1, 2, 22]. No systematic study has been carried out to verify the expression stability of the reference gene used in watermelon fruit or to identify more reference genes. Thus, screening and validation of appropriate reference genes specifically for watermelon fruit are urgently needed to facilitate the accurate expression measurement of target genes.



In this study, systematic validation of optimal reference genes was conducted for watermelon fruit considering the effects of fruit setting methods, genotypes, and fruit developmental stages. Although many algorithms have been developed to aid in the selection of appropriate reference genes, geNorm and NormFinder proved to be sufficient for validation of reference genes [20]. Considering all 48 samples, geNorm determined ClCAC and ClACT as the best pair of reference genes. ClCAC, ClACT, and ClTUA5 were also required to calculate a reliable normalization factor. NormFinder ranked CICAC alone or together with CISAND as the single or two best genes. Different results were expected because of the different algorithms adopted in the two methods. GeNorm is based on the principle that the expression ratio between two ideal internal control genes is identical in all samples regardless of the experimental condition [15]. By contrast, NormFinder is a model-based approach that estimates expression variations and ranks the candidate reference genes with minimal estimated intra- and inter-group variations [16]. In this study, all 48 samples were divided into different subgroups according to genotypes and fruit setting methods to comply with the model-based algorithm of NormFinder. Consequently, NormFinder potentially provided a more robust measurement of gene expression stability [16]. Meanwhile, CAC and SAND were also recommended in tomato fruit [28]. Suitable reference genes were also systemically validated in fruits in other studies. However, different results were obtained in different species. For example, CAC and UBQ2 were the most stable genes in banana pulp [29]. In papaya fruit, EIF, TBP1, and TBP2 exhibited good performance [20]. ACT was recommended in lychee fruit [30]. These results demonstrate that screening for appropriate reference genes is a prerequisite before qRT-PCR analysis. However, the reference genes identified in other studies were good candidates, and verifying the stability of these candidates was an efficient approach to find the most suitable references [31, 32].

After dividing the samples into different groups according to the treatments and calculating the expression stability of the candidate genes, differences on the rankings of the candidate reference genes were observed between different genotypes and fruit setting methods. In genotype 97103, geNorm ranked ClSAND and ClACT as well as ClCAC and ClACT as the best pairs of reference genes for the pollination and CPPU treatments. Although, the best pair of reference genes was ClCAC and ClACT for both pollination and CPPU treatment in 8424, the minimum number of reference genes to calculate the reliable normalization factor was 4 and 3, respectively. When the data from the two genotypes were integrated, the minimum number of reference genes required to calculate the normalization factor was different (Table 2). Based on the results determined by NormFinder, the differences on the single best gene or combination of two best genes were more pronounced between watermelon genotypes or fruit setting methods (Table 3). Thus, both the genotypes and fruit setting methods affected the rankings of the candidate references. The suitable reference genes were also found to be genotype dependent in several crops, such as mustard [12], grapevine [33], strawberry [34], and sweet potato [35]. These findings partially explained why different best reference genes were obtained in different experiments on the same crop. As a result, appropriate reference genes should be validated or selected under specific experimental conditions. However, the validation of reference genes is very tedious and impossible to conduct before each experiment, particularly when the number of target genes to be normalized is limited. Thus, specialized studies on reference gene validation are necessary to test the suitability of candidate reference genes as many as possible. To obtain representative and widely applicable results, the experiments should be designed to include the affecting factors as many as possible [34]. In this study, aside from the fruit developmental stages, the effects of genotypes and fruit setting methods were also considered and evaluated, which nearly included all the potential factors that affect watermelon fruit development under normal growth condition, to provide a comprehensive evaluation of the expression stability of the selected reference genes.



Transcriptional regulation has shown to be a major mechanism that controls the biosynthesis of specific carotenoids [5]. *ClPSY1* is one of the key genes that control lycopene biosynthesis in watermelon fruits [1, 5]. Thus, the correlation of *ClPSY1* expression pattern and lycopene biosynthesis during the ripening of watermelon fruit can offer a reliable measurement to test the reliability of top-ranked reference genes, and good correlations were found between lycopene accumulation and *ClPSY1* expression by using the top-ranked reference genes for normalization. However, when normalized by *Cl18SrRNA*, no correlation was observed between *ClPSY1* expression pattern and lycopene biosynthesis in fruit ripening. These results demonstrate that the top-ranked reference genes in this study are suitable for normalization during watermelon fruit development.

Nevertheless, different expression levels of *ClPSY1* were observed using different top-ranked reference genes. The highest expression level was observed by using *ClCAC*, whereas the lowest expression level was obtained by using *ClPP2A* and *ClACT* as determined by NormFinder. Moderate expression levels were found using the pair and multiple reference genes as determined by geNorm. Similar results were also observed in previous studies [17, 36]. Considering that the variation in the average of multiple genes was smaller than the variation in individual genes, the normalization factor calculation based on the geometric mean of multiple reference genes provided more accurate and reliable normalization of expression data [15, 16]. Accordingly, the multiple reference genes identified in this study were recommended to replace the single and non-validated *18SrRNA* for more reliable normalization results in watermelon fruit.

Considering the significant effect of reference genes on the normalization results, primer specificities of the candidate reference genes have been stressed in many studies [16, 24]. Reference genes with primers spanning the exon–exon junction were suggested to control the potential gDNA contamination [37]. Meanwhile, primer pairs flanking an intron were also designed for candidate reference genes [16, 17, 38]. In this study, the primers spanning the exon–exon junction were developed for 10 selected candidate reference genes, along with *ClACT* with amplicon spanning an intron developed in our previous study [17]. In practice, *ClACT* can be used first to test the existence of gDNA contamination in cDNA samples. In the absence of gDNA contaminations, both the multiple reference genes of *ClCAC*, *ClACT*, and *ClTUA5* and the best combination of *ClCAC* and *ClSAND* can be used for normalization. However, in the presence of gDNA contaminations, the best combination of *ClCAC* and *ClSAND* is preferred for normalization if the primers of the target genes also spanned the exon–exon junction. Thus, these reference genes provided a powerful strategy for normalization in qRT-PCR analysis for watermelon fruits.

Conclusions

In this study, the major factors including genotypes, fruit setting methods, and fruit developmental stages that potentially affect the reference gene expression stability in watermelon fruits have been considered to develop reliable and widely applicable results. The normalization factor calculated from the geometric means of *ClCAC*, *ClACT*, and *ClTUA5* as determined by geNorm or *ClCAC* and *ClSAND* as identified by NormFinder can be used to normalize the target genes in watermelon fruit to improve the accuracy and reliability of gene expression studies. Moreover, this study provides a list of potential reference genes with specific primers spanning exon–exon junction that can be easily tested in other reference gene studies for watermelon.

Supporting Information

S1 Fig. Amplification specificity of the primer pairs verified by melting curve analysis. (PDF)



S2 Fig. Pairwise variation analyses of the candidate reference genes by geNorm. (PDF)

Acknowledgments

We are thankful to Dr. Yong Xu (National Engineering Research Center for Vegetables, Beijing, China) for providing seeds of watermelon inbred line 97103.

Author Contributions

Conceived and designed the experiments: QK ZB. Performed the experiments: JY LG LZ. Analyzed the data: QK JY. Contributed reagents/materials/analysis tools: FC YH. Wrote the paper: QK.

References

- Guo S, Liu J, Zheng Y, Huang M, Zhang H, Gong G, et al. Characterization of transcriptome dynamics during watermelon fruit development: sequencing, assembly, annotation and gene expression profiles. BMC Genomics. 2011; 12:454. doi: 10.1186/1471-2164-12-454 PMID: 21936920
- Wechter WP, Levi A, Harris KR, Davis AR, Fei Z, Katzir N, et al. Gene expression in developing watermelon fruit. BMC Genomics. 2008; 9:275. doi: 10.1186/1471-2164-9-275 PMID: 18534026
- Levi A, Davis A, Hernandez A, Wechter P, Thimmapuram J, Trebitsh T, et al. Genes expressed during the development and ripening of watermelon fruit. Plant Cell Rep. 2006; 25(11):1233–1245. PMID: 16802118
- Guo S, Zhang J, Sun H, Salse J, Lucas WJ, Zhang H, et al. The draft genome of watermelon (*Citrullus lanatus*) and resequencing of 20 diverse accessions. Nat Genet. 2013; 45(1):51–58. doi: 10.1038/ng.2470 PMID: 23179023
- Grassi S, Piro G, Lee JM, Zheng Y, Fei Z, Dalessandro G, et al. Comparative genomics reveals candidate carotenoid pathway regulators of ripening watermelon fruit. BMC Genomics. 2013; 14(1):781.
- Fan M, Huang Y, Zhong Y, Kong Q, Xie J, Niu M, et al. Comparative transcriptome profiling of potassium starvation responsiveness in two contrasting watermelon genotypes. Planta. 2014; 239(2):397–410. doi: 10.1007/s00425-013-1976-z PMID: 24185372
- Nolan T, Hands RE, Bustin SA. Quantification of mRNA using real-time RT-PCR. Nat Protoc. 2006; 1 (3):1559–1582. PMID: <u>17406449</u>
- Gutierrez L, Mauriat M, Guénin S, Pelloux J, Lefebvre J-F, Louvet R, et al. The lack of a systematic validation of reference genes: a serious pitfall undervalued in reverse transcription-polymerase chain reaction (RT-PCR) analysis in plants. Plant Biotechnol J. 2008; 6(6):609–618. doi: 10.1111/j.1467-7652. 2008.00346.x PMID: 18433420
- Mafra V, Kubo KS, Alves-Ferreira M, Ribeiro-Alves M, Stuart RM, Boava LP, et al. Reference Genes for Accurate Transcript Normalization in Citrus Genotypes under Different Experimental Conditions. PLoS ONE. 2012; 7(2): e31263. doi: 10.1371/journal.pone.0031263 PMID: 22347455
- Ferguson BS, Nam H, Hopkins RG, Morrison RF. Impact of Reference Gene Selection for Target Gene Normalization on Experimental Outcome Using Real-Time qRT-PCR in Adipocytes. PLoS ONE. 2010; 5(12):e15208. doi: 10.1371/journal.pone.0015208 PMID: 21179435
- Le DT, Aldrich DL, Valliyodan B, Watanabe Y, Ha CV, Nishiyama R, et al. Evaluation of Candidate Reference Genes for Normalization of Quantitative RT-PCR in Soybean Tissues under Various Abiotic Stress Conditions. PLoS ONE. 2012; 7(9):e46487. doi: 10.1371/journal.pone.0046487 PMID: 23029532
- Chandna R, Augustine R, Bisht NC. Evaluation of Candidate Reference Genes for Gene Expression Normalization in Brassica juncea Using Real Time Quantitative RT-PCR. PLoS ONE. 2012; 7(5): e36918. doi: 10.1371/journal.pone.0036918 PMID: 22606308
- Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, et al. The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. Clin Chem. 2009; 55 (4):611–622. doi: 10.1373/clinchem.2008.112797 PMID: 19246619
- 14. Jacob F, Guertler R, Naim S, Nixdorf S, Fedier A, Hacker NF, et al. Careful Selection of Reference Genes Is Required for Reliable Performance of RT-qPCR in Human Normal and Cancer Cell Lines. PLoS ONE. 2013; 8(3):e59180. doi: 10.1371/journal.pone.0059180 PMID: 23554992



- Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol. 2002; 3(7):research0034.1–0034.11. PMID: 12184816
- 16. Andersen CL, Jensen JL, Orntoft TF. 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. 2004; 64(15):5245–5250. PMID: 15289330
- 17. Kong Q, Yuan J, Gao L, Zhao S, Jiang W, Huang Y, et al. Identification of Suitable Reference Genes for Gene Expression Normalization in qRT-PCR Analysis in Watermelon. PLoS ONE. 2014; 9(2):e90612. doi: 10.1371/journal.pone.0090612 PMID: 24587403
- Fan C, Ma J, Guo Q, Li X, Wang H, Lu M. Selection of Reference Genes for Quantitative Real-Time PCR in Bamboo (*Phyllostachys edulis*). PLoS ONE. 2013; 8(2):e56573. doi: 10.1371/journal.pone. 0056573 PMID: 23437174
- Guénin S, Mauriat M, Pelloux J, Van Wuytswinkel O, Bellini C, Gutierrez L. Normalization of qRT-PCR data: the necessity of adopting a systematic, experimental conditions-specific, validation of references. J Exp Bot. 2009; 60(2):487–493. doi: 10.1093/jxb/ern305 PMID: 19264760
- Zhu X, Li X, Chen W, Chen J, Lu W, Chen L, et al. Evaluation of New Reference Genes in Papaya for Accurate Transcript Normalization under Different Experimental Conditions. PLoS ONE. 2012; 7(8): e44405. doi: 10.1371/journal.pone.0044405 PMID: 22952972
- Die JV, Rowland LJ. Superior Cross-Species Reference Genes: A Blueberry Case Study. PLoS ONE. 2013; 8(9):e73354. doi: 10.1371/journal.pone.0073354 PMID: 24058469
- Li Q, Li P, Sun L, Wang Y, Ji K, Sun Y, et al. Expression analysis of beta-glucosidase genes that regulate abscisic acid homeostasis during watermelon (*Citrullus lanatus*) development and under stress conditions. J Plant physiol. 2012; 169(1):78–85. doi: 10.1016/j.jplph.2011.08.005 PMID: 21940067
- Li J, Wu Z, Cui L, Zhang T, Guo Q, Xu J, et al. Transcriptome Comparison of Global Distinctive Features Between Pollination and Parthenocarpic Fruit Set Reveals Transcriptional Phytohormone Cross-Talk in Cucumber (Cucumis sativus L.). Plant Cell Physiol. 2014; 55(7):1325–1342. doi: 10.1093/pcp/ pcu051 PMID: 24733865
- Udvardi MK, Czechowski T, Scheible WR. Eleven golden rules of quantitative RT-PCR. Plant Cell. 2008; 20(7):1736–1737. doi: 10.1105/tpc.108.061143 PMID: 18664613
- 25. Liu C, Zhang H, Dai Z, Liu X, Liu Y, Deng X, et al. Volatile chemical and carotenoid profiles in watermelons Citrullus vulgaris (Thunb.) Schrad (Cucurbitaceae) with different flesh colors. Food Sci Biotechnol. 2012; 21(2):531–541.
- 26. Kong Q, Yuan J, Niu P, Xie J, Jiang W, Huang Y, et al. Screening Suitable Reference Genes for Normalization in Reverse Transcription Quantitative Real-Time PCR Analysis in Melon. PLoS ONE. 2014; 9(1):e87197. doi: 10.1371/journal.pone.0087197 PMID: 24475250
- 27. Warzybok A, Migocka M. Reliable Reference Genes for Normalization of Gene Expression in Cucumber Grown under Different Nitrogen Nutrition. PLoS ONE. 2013; 8(9):e72887. doi: 10.1371/journal.pone.0072887 PMID: 24058446
- 28. Exposito-Rodriguez M, Borges A, Borges-Perez A, Perez J. Selection of internal control genes for quantitative real-time RT-PCR studies during tomato development process. BMC Plant Biol. 2008; 8 (1):131.
- Chen L, Zhong HY, Kuang JF, Li JG, Lu WJ, Chen JY. Validation of reference genes for RT-qPCR studies of gene expression in banana fruit under different experimental conditions. Planta. 2011; 234 (2):377–390. doi: 10.1007/s00425-011-1410-3 PMID: 21505864
- Zhong HY, Chen JW, Li CQ, Chen L, Wu JY, Chen JY, et al. Selection of reliable reference genes for expression studies by reverse transcription quantitative real-time PCR in litchi under different experimental conditions. Plant Cell Rep. 2011; 30(4):641–653. doi: 10.1007/s00299-010-0992-8 PMID: 21301853
- Gutierrez L, Mauriat M, Pelloux J, Bellini C, van Wuytswinkel O. Towards a systematic validation of references in real-time RT-PCR. Plant Cell. 2008; 20(7):1734–1735. doi: 10.1105/tpc.108.059774 PMID: 18664615
- Chao WS, Doğramaci M, Foley ME, Horvath DP, Anderson JV. Selection and Validation of Endogenous Reference Genes for qRT-PCR Analysis in Leafy Spurge (Euphorbia esula). PLoS ONE. 2012; 7 (8):e42839. doi: 10.1371/journal.pone.0042839 PMID: 22916167
- 33. Monteiro F, Sebastiana M, Pais MS, Figueiredo A. Reference Gene Selection and Validation for the Early Responses to Downy Mildew Infection in Susceptible and Resistant Vitis vinifera Cultivars. PLoS ONE. 2013; 8(9):e72998. doi: 10.1371/journal.pone.0072998 PMID: 24023800
- Amil-Ruiz F, Garrido-Gala J, Blanco-Portales R, Folta KM, Muñoz-Blanco J, Caballero JL. Identification and Validation of Reference Genes for Transcript Normalization in Strawberry (Fragaria × ananassa)



- Defense Responses. PLoS ONE. 2013; 8(8):e70603. doi: <u>10.1371/journal.pone.0070603</u> PMID: 23940602
- Park SC, Kim YH, Ji CY, Park S, Jeong JC, Lee HS, et al. Stable Internal Reference Genes for the Normalization of Real-Time PCR in Different Sweetpotato Cultivars Subjected to Abiotic Stress Conditions. PLoS ONE. 2012; 7(12) e51502. doi: 10.1371/journal.pone.0051502 PMID: 23251557
- 36. Gimeno J, Eattock N, Van Deynze A, Blumwald E. Selection and Validation of Reference Genes for Gene Expression Analysis in Switchgrass (*Panicum virgatum*) Using Quantitative Real-Time RT-PCR. PLoS ONE. 2014; 9(3):e91474. doi: 10.1371/journal.pone.0091474 PMID: 24621568
- Czechowski T, Stitt M, Altmann T, Udvardi MK, Scheible WR. Genome-Wide Identification and Testing
 of Superior Reference Genes for Transcript Normalization in Arabidopsis. Plant Physiol. 2005; 139
 (1):5–17. PMID: 16166256
- Reddy DS, Bhatnagar-Mathur P, Cindhuri KS, Sharma KK. Evaluation and Validation of Reference Genes for Normalization of Quantitative Real-Time PCR Based Gene Expression Studies in Peanut. PLoS ONE. 2013; 8(10):e78555. doi: 10.1371/journal.pone.0078555 PMID: 24167633