

RESEARCH ARTICLE

Reliable reference genes for normalization of gene expression data in tea plants (*Camellia sinensis*) exposed to metal stresses

Ming-Le Wang, Qing-Hui Li, Hua-Hong Xin, Xuan Chen, Xu-Jun Zhu, Xing-Hui Li*

Tea Research Institute, College of Horticulture, Nanjing Agricultural University, Nanjing, Jiangsu, People's Republic of China

* lxh@njau.edu.cn



OPEN ACCESS

Citation: Wang M-L, Li Q-H, Xin H-H, Chen X, Zhu X-J, Li X-H (2017) Reliable reference genes for normalization of gene expression data in tea plants (*Camellia sinensis*) exposed to metal stresses. PLoS ONE 12(4): e0175863. <https://doi.org/10.1371/journal.pone.0175863>

Editor: Christian Schönbach, Kumamoto University, JAPAN

Received: December 10, 2016

Accepted: March 31, 2017

Published: April 28, 2017

Copyright: © 2017 Wang et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), 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 financially by the National Natural Science Foundation of China (31570689, 31470690), <http://www.nsf.gov.cn/>; the China Earmarked Fund for Modern Agro-industry Technology Research System (CARS-23), <http://www.moa.gov.cn/>; the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD), <http://www.ec.js>.

Abstract

Tea plants [*Camellia sinensis* (L.) O. Kuntze] are an important leaf-type crop that are widely used for the production of non-alcoholic beverages in the world. Exposure to excessive amounts of heavy metals adversely affects the quality and yield of tea leaves. To analyze the molecular responses of tea plants to heavy metals, a reliable quantification of gene expression is important and of major importance herein is the normalization of the measured expression levels for the target genes. Ideally, stably expressed reference genes should be evaluated in all experimental systems. In this study, 12 candidate reference genes (i.e., *18S rRNA*, *Actin*, *CYP*, *EF-1α*, *eIF-4α*, *GAPDH*, *MON1*, *PP2AA3*, *TBP*, *TIP41*, *TUA*, and *UBC*) were cloned from tea plants, and the stability of their expression was examined systematically in 60 samples exposed to diverse heavy metals (i.e., manganese, aluminum, copper, iron, and zinc). Three Excel-based algorithms (geNorm, NormFinder, and BestKeeper) were used to evaluate the expression stability of these genes. *PP2AA3* and *18S rRNA* were the most stably expressed genes, even though their expression profiles exhibited some variability. Moreover, commonly used reference genes (i.e., *GAPDH* and *TBP*) were the least appropriate reference genes for most samples. To further validate the suitability of the analyzed reference genes, the expression level of a phytochelatin synthase gene (i.e., *CsPCS1*) was determined using the putative reference genes for data normalizations. Our results may be beneficial for future studies involving the quantification of relative gene expression levels in tea plants.

Introduction

Quantification of gene expression levels is an important part of the systematic characterization of gene transcriptional mechanisms and regulatory networks. The quantitative real-time polymerase chain reaction (qRT-PCR) is one of the most commonly used methods to quantify target genes expression levels due to its practical simplicity, specificity, reproducibility, and highly sensitivity in detecting transcripts with low copy numbers [1, 2]. An accepted standard procedure to conduct and interpret qRT-PCR experiments was lacking prior to 2009, which is

edu.cn/. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

when Bustin et al. [3] proposed the MIQE guidelines. Selecting appropriate reference genes is crucial for the reliable quantification of gene expression data [4]. To evaluate the stability of candidate reference genes in various experimental conditions, statistical algorithms have been developed, including geNorm [5], NormFinder [6], and BestKeeper [7]. An ideal reference gene should be stably transcribed under diverse experimental conditions [8]. However, it is unreasonable to expect the expression of any gene to be completely stable in a living cell. Thus, it is necessary to identify suitable reference genes to normalize expression data prior to investigating target genes expression levels.

Tea plants [*Camellia sinensis* (L.) O. Kuntze] originating from the Yunnan–Guizhou Plateau in southwestern China are an important perennial evergreen woody crop of the family *Theaceae* [9, 10]. Being rich in biologically active metabolites, such as tea polyphenols, theanine, and polysaccharides, tea leaves have long been used as the raw materials for dietary supplements, health foods, cosmeceuticals, and especially the production of non-alcoholic caffeine-containing beverages in the world [11, 12]. As sessile organisms, tea plants are continuously exposed to various adverse environmental conditions, such as drought stress [13], heat stress [14], salinity stress [15], and especially heavy metal stresses, which considerably affect tea growth, production, and quality [16–21]. For example, high concentration of Mn decreased tea production [16, 17]. Zn-stress decreased net photosynthetic rate, transpiration rate, stomatal conductance, growth and relative water content of *Camellia sinensis* considerably [18, 19]. Moreover, Yadav and Mohanpuria [20] demonstrated that Cu and Al exposure induces oxidative stress in *C. sinensis*. Additionally, excessive iron can adversely affect the quality of tea [21]. Identifying reliable reference genes under different environmental conditions is imperative for analyzing immediate molecular responses in *C. sinensis* cells.

Several suitable *C. sinensis* reference genes have been investigated, with some inconsistency observed in their expression levels under different experimental conditions [22–24]. *CsPTBI*, *CsEF1*, *CsSAND1*, *CsCLATHRIN1*, and *CsUBC1* are the top five most stably expressed reference genes under six experimental conditions (i.e., diurnal expression in leaves, expression in different organs, expression in leaves/shoots exposed to different cold and short day treatment, expression in shoots treated with an auxin antagonist, and expression in shoots treated with lanolin) [22]. *TUA1* is the most suitable reference gene for analyses of damaged tissues [23]. Additionally, *CsTBP* and *CsTIP41* displayed the maximum stability in tea leaf development, while *CsTBP* is also the most stably expressed gene in response to hormones [24]. However, a systematic approach for selecting reference genes useful for analyzing gene expression levels in *C. sinensis* plants in response to metal stresses has not been developed. Hence, identifying suitable reference genes in *C. sinensis* plants exposed to increasing metal concentrations is necessary, which will provide new information relevant to future research on molecular mechanism studies in tea plants.

In this study, we selected 12 candidate reference genes (i.e., *18S rRNA*, *Actin*, *CYP*, *EF-1 α* , *eIF-4 α* , *GAPDH*, *MON1*, *PP2AA3*, *TBP*, *TIP41*, *TUA*, and *UBC*) that were confirmed to be stably expressed in earlier studies [25–28]. The sequences of the 12 candidate genes were obtained based on our previously generated *C. sinensis* transcriptome sequencing data [29]. Specific details regarding these reference genes are listed in Table 1. We used qRT-PCR to determine gene expression levels in tea leaves exposed to increasing concentrations of metals (i.e., Mn, Al, Cu, Fe, and Zn). Three different algorithms (i.e., geNorm, NormFinder, and BestKeeper) were used to evaluate the expression stability of the candidate reference genes. Additionally, *C. sinensis* phytochelatin synthase (*CsPCS1*), which is important for detoxifying the effects of heavy metals and was found to be up-regulated at its transcript level in response to Cu and Al stresses [20], was used to validate the reliability of the selected reference genes in tea leaves. This study is the first to analyze potential reference genes for investigating tea plants under

Table 1. The characteristics of primers used for quantitative real-time PCR in *C. sinensis*.

Gene symbol	<i>Arabidopsis</i> locus description	<i>Arabidopsis</i> homolog locus	Primer sequences (5'-3') forward/reverse	Amplicon length (bp)	Melting T _m (°C)	Efficiency (%)	Correlation coefficient (R ²)
18S rRNA	18S ribosomal RNA	ATMG01390	ACACCCCTGGGAATTGGTTT/ GTATGCGCCAATAAGACCAC	106	83.5	102.7	0.998
Actin	Actin 7	AT5G09810	CAGACCGTATGAGCAAGGAA/ GCTTAGGGATGCGAGGATAG	122	82.0	101.0	0.999
CYP	Cyclophilin	AT3G56070	TTTCGCGATGAGAACTTCAA/ CCATCTCCTTCACCACACTG	181	82.5	100.3	0.996
EF-1α	Elongation factor-1α	AT1G07940	CAAGCGTGTTCATCGAGAGAT/ ATACCACGTTTCACGTTTCAGC	108	81.5	105.8	0.997
eIF-4α	Eukaryotic translation initiation factor 4α-1	AT3G13920	TGAGAAGGTTATGCGAGCAC/ GCAACATGTCAAACACACGA	149	83.0	104.3	1.000
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase	AT1G42970	GACTGGAGAGGTGGAAGAGC/ AGCCATTCCAGTCAATTTCC	114	82.5	99.8	0.996
MON1	MONENSIN SENSITIVITY1	AT2G28390	ATTTCTTCGTGGAGAATGG/ GCCATAAACAAGCTCCAAT	160	82.0	98.9	0.999
PP2AA3	Protein phosphatase 2A subunit A3	AT1G13320	CAACATGTTGCTCTGCTTT/ GGGAAAGGAATATGCGCAGA	100	81.0	101.6	0.995
TBP	TATA-box binding protein	AT1G55520	AAGGGATCCAAAGACGACAG/ TGAAATCCTTGAATTTGGCA	149	81.5	98.2	0.997
TIP41	TIP41-like family protein	AT4G34270	CGAAAGAGCCCATTTCTCTTC/ ACGTGTGTCCTCAATCTCA	173	80.5	104.5	0.997
TUA	Tubulin alpha-3	AT5G19770	TTTGGAGCGCTTGTCTGTAG/ TGTGTTCAAGGAGGGAATGA	134	82.0	102.3	0.996
UBC	Ubiquitin-protein ligase	AT4G27960	GACATGTTTCATTGGCAAGC/ ACCTTAGGTGGCTTGAATGG	116	81.0	101.7	0.995
CsPCS1	Phytochelatin synthase	AT5G44070	AATGCCCTTGCTATTGATCC/ CTCCAGAACAGTGAGCCAAA	151	81.0	103.4	0.992

<https://doi.org/10.1371/journal.pone.0175863.t001>

heavy metal stress conditions. Our data may be useful for developing more accurate and reliable protocols to analyze the expression of other tea plant genes in response to metal stresses.

Materials and methods

Plant materials and treatments

Two-year-old tea plants (*C. sinensis* cv. ‘Longjing-changyecha’) were collected from the fields of Gaochun District, Jiangsu Province, China (longitude: 118.57E, latitude: 31.19N). The plants were pre-incubated in a control nutrient solution [30] for 28 days (from September 8, 2015 to October 6, 2015) in a climate-controlled chamber under a 14-h light (24°C)/10-h dark (20°C) photoperiod (light intensity: 220 μmol m⁻² s⁻¹) and relative humidity of 75%. We then added MnSO₄, Al₂(SO₄)₃, CuSO₄, FeSO₄, or ZnSO₄ for final concentrations of 50 μM, 12.5 mM, 13 μM, 210 μM, and 51 μM, respectively. After incubating the treated and control plants for 0, 1, 4, and 7 days, the fully expanded third leaves from the top bud of tea plants were harvested (Figure A in S1 File), immediately frozen in liquid nitrogen, and stored at -80°C.

Total RNA extraction and cDNA synthesis

Total RNA was extracted using the EasyPure[®] Plant RNA Kit (TransGen, Beijing, China). The concentration and purity of RNA samples were measured using the ONE Drop OD-1000 + spectrophotometer (ONE Drop, Shanghai, China). Only samples with an A₂₆₀/A₂₈₀ ratio of

1.8–2.0 and an A_{260}/A_{230} ratio > 2.0 were used for the subsequent cDNA synthesis. The integrity of the purified RNA was further confirmed by 1.2% agarose gel electrophoresis. We generated cDNA from 1 μg total RNA using the *TransScript*[®] All-in-One First-Strand cDNA Synthesis SuperMix for qPCR (One-Step gDNA Removal) (TransGen). The resulting cDNA was diluted 20-fold in distilled deionized water and analyzed in a qRT-PCR assay.

Selection of candidate reference genes and primer design

We selected 12 candidate genes (i.e., *18S rRNA*, *Actin*, *CYP*, *EF-1 α* , *eIF-4 α* , *GAPDH*, *MON1*, *PP2AA3*, *TBP*, *TIP41*, *TUA*, and *UBC*) that were previously determined to be appropriate reference genes for qRT-PCR assays from the TAIR database (<http://www.arabidopsis.org>). Potential homologues of these genes were identified using our *C. sinensis* transcriptome sequencing data [29]. The qRT-PCR primers were designed with Primer PREMIER 5.0 software (*PREMIER* Biosoft International, Palo Alto, CA, USA) according to the Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) guidelines [3]. The target sequences of the 12 *C. sinensis* candidate reference genes were cloned using Taq DNA polymerase (TaKaRa, Dalian, China). The PCR amplification was conducted in a 20 μl sample consisting of 10.9 μl double-distilled H₂O, 2 μl 10 \times PCR Buffer, 1.6 μl dNTPs (2.5 mM each), 1.2 μl MgCl₂ (25 mM), 2 μl template cDNA, 1 μl each primer (10 μM), and 0.3 μl of Taq DNA polymerase. The PCR conditions were as follows: 4 min at 95°C for denaturation; 35 cycles of 30 s at 95°C (denaturation), 30 s at 55°C (annealing), and 30 s at 72°C (extension); and a final step of 5 min at 72°C for extension. The resulting amplicons were purified and cloned into the pEASY-T1 Simple Cloning Vector (TransGen) for sequencing (Genscript, Nanjing, China). Primer sequences and amplicon characteristics are listed in Table 1.

Quantitative real-time PCR assay

We conducted the qRT-PCR assay using the *TransStart*[®] Tip Green qPCR SuperMix (TransGen) and a CFX96 Real-Time System (Bio-Rad, Hercules, CA, USA). The PCR solution (20 μl) contained 10 μl 2 \times *TransStart*[®] Tip Green qPCR SuperMix, 0.2 μM each primer, 1 μl diluted cDNA, and nuclease-free water. The PCR reaction conditions were as follows: 95°C for 2 min; 40 cycles of 95°C for 10 s, 60°C for 15 s, and 72°C for 20 s; 72°C for 3 min. The final dissociation curve was obtained between 65°C and 95°C to verify primer specificity. Each assay included three technical and biological replicates, and involved a standard curve with six serial dilution points. Amplification efficiencies (E) were calculated using standard curves with satisfactory linear relationships ($R^2 > 0.99$). The following equation was used to calculate the E-value (%): $E = (10^{[-1/\text{slope}]} - 1) \times 100$.

Data analysis

The expression levels of the 12 genes were determined according to the quantification cycle (Cq) value. The raw data are listed in Table A in S1 File. Three different Microsoft Excel-based programs (i.e., geNorm v3.5 [5], NormFinder v0.953 [6], and BestKeeper v1.0 [7]) were used to determine the expression stability of the candidate genes. The raw data were analyzed using BestKeeper. However, for geNorm and NormFinder, the raw Cq values were converted into relative quantity values using the formula $2^{-\Delta\text{Cq}}$ ($\Delta\text{Cq} = \text{each corresponding Cq value} - \text{minimum Cq value}$; Table B in S1 File). Additionally, the fold change in expression of the target gene (*CsPCSI*; GenBank: KY264048) relative to the different reference genes at various time points was analyzed using the $2^{-\Delta\Delta\text{Cq}}$ method, in which $\Delta\Delta\text{Cq} = (\text{Cq}_{\text{Target gene}} - \text{Cq}_{\text{Reference gene}})_{\text{Day } x} - (\text{Cq}_{\text{Target gene}} - \text{Cq}_{\text{Reference gene}})_{\text{Day } 0}$ [31].

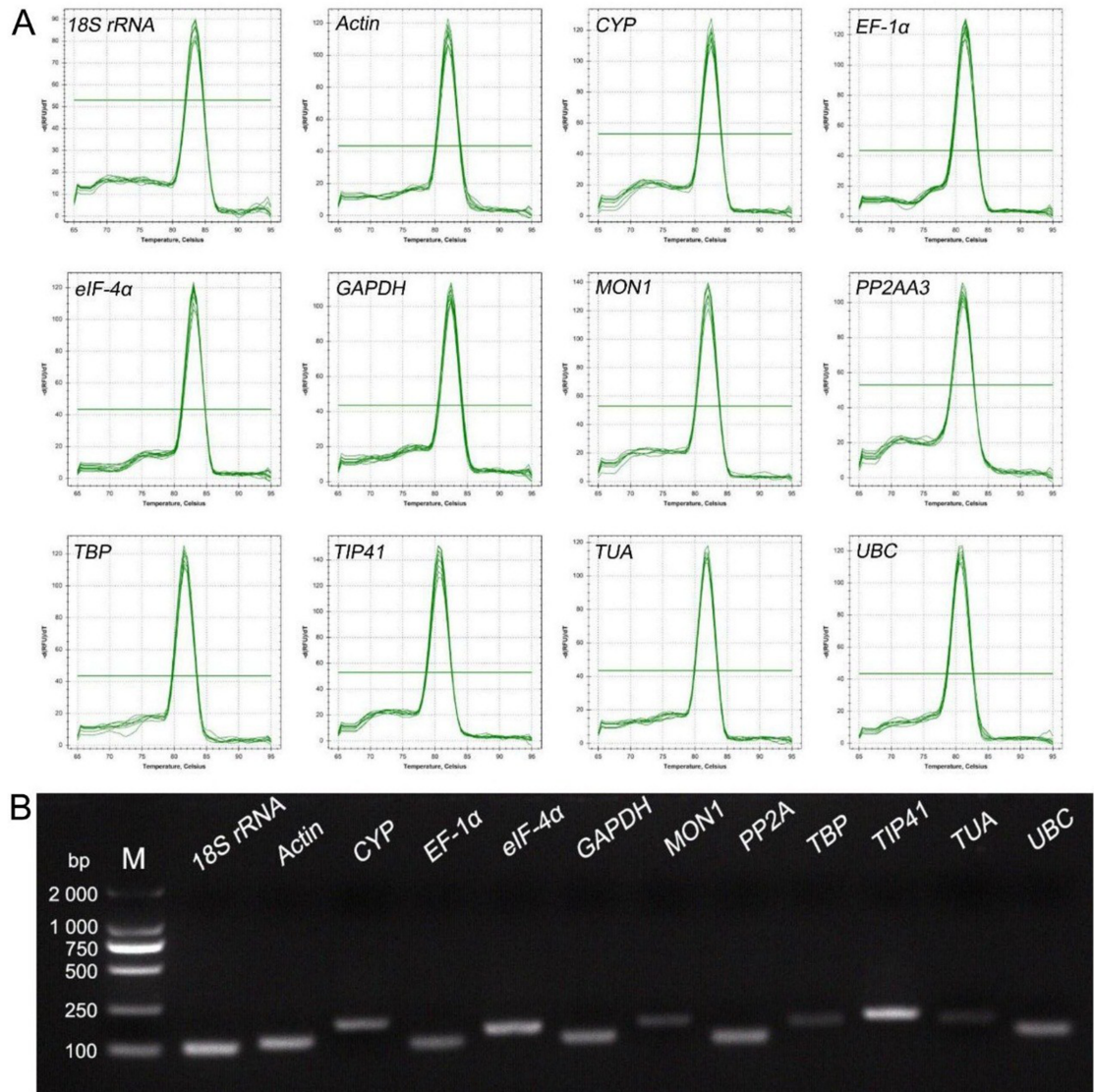


Fig 1. Confirmation of primer specificity and amplicon size. (a) Melting curve analysis of 12 candidate reference genes. (b) Amplification results for 12 candidate genes using a *C. sinensis* cDNA template. M: DL2000 DNA Marker.

<https://doi.org/10.1371/journal.pone.0175863.g001>

Results

Assessment of primer specificity and amplification efficiency

We designed gene-specific primer pairs based on the sequences of the 12 candidate reference genes cloned from *C. sinensis* transcriptome sequences (Figures B–M and Table C in S1 File). Details regarding the 12 genes, primer sequences, amplicon lengths, and melting temperatures are provided in Table 1. The melting curve analysis revealed that all reactions produced a single distinct peak (Fig 1A). All primers amplified with a single PCR product of the expected size according to 1.5% agarose gel electrophoresis (Fig 1B). The estimated PCR amplification efficiencies for the candidate reference genes varied from 98.2% for *TBP* to 105.8% for *EF-1α*, and the correlation coefficients (R^2) ranged from 0.995 for *UBC* to 1.000 for *eIF-4α* (Table 1).

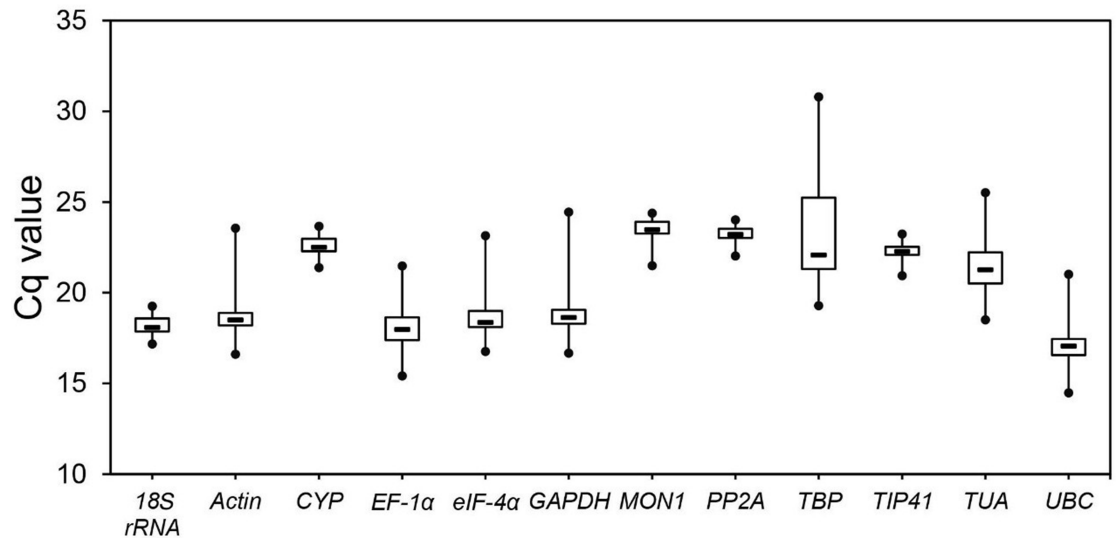


Fig 2. Quantification cycle (Cq) values of the 12 candidate reference genes in *C. sinensis* leaves under metal stresses. The lines across boxes represent the mean Cq values. The boxes indicate the 25th and 75th percentiles, while the whiskers correspond to the maximum and minimum values.

<https://doi.org/10.1371/journal.pone.0175863.g002>

Quantification cycle values of candidate reference genes

The Cq values in qPCR provided an overview of the gene expression levels of 12 candidate reference genes in 60 samples. There were apparent differences in the transcript abundance among genes. The raw data are listed in Table A in [S1 File](#). The mean Cq values of all reference genes ranged from 17.08 to 23.51 ([Fig 2](#)). Low Cq value indicates the high expression levels, conversely mean the low expression levels. Among the 12 analyzed genes, *UBC* exhibited the highest expression levels with a mean Cq of 17.08, while *TBP* had the lowest expression levels with a mean Cq of 23.51 ([Fig 2](#)).

Expression stability of candidate reference genes

To identify the most suitable reference gene, three Microsoft Excel-based algorithms (geNorm, NormFinder, and BestKeeper) were used to analyze the stability of each reference gene. The geNorm analysis indicated that all genes performed well under individual stress conditions, with M values lower than the default limit of 1.5. However, the most suitable reference gene differed among treatments ([Table 2](#)). *PP2AA3* and *TBP* with same M values were the two best reference genes for Mn-treated leaves (Figure N in [S1 File](#)). Additionally, the most appropriate reference genes were *MON1* and *TIP41* in Al-treated leaves, *MON1* and *PP2AA3* in Cu-treated leaves, *EF-1α* and *eIF-4α* in Fe-treated leaves, and *CYP* and *PP2AA3* in Zn-treated leaves. *TUA* exhibited unstable expression in all samples. The pairwise variation [i.e., $V_n/n+1$ ($n \geq 2$)] of the geNorm program was also used to determine the optimal number of reference genes required for data normalizations. Values ($V_2/3$) lower than the recommended threshold of 0.15 indicated that two reference genes were sufficient for normalizing the gene expression data resulting from the exposure to the five metal stress conditions ([Fig 3](#)).

The results of the NormFinder analysis revealed that the genes with the lowest values were the most stably expressed ([Table 2](#)). The three most stably expressed reference genes for all samples were *Actin* (0.129), *EF-1α* (0.134), and *TUA* (0.146). *Actin* was the most stably expressed gene in Mn- and Fe-treated tea leaves ([Table D](#) in [S1 File](#)). In contrast, *18S rRNA*

Table 2. Gene expression stability ranked by geNorm, NormFinder, and BestKeeper software programs. SD: standard deviation; CV: coefficient of variation.

Rank	geNorm		NormFinder		BestKeeper		
	Gene	Stability	Gene	Stability	Gene	SD	CV
1	<i>TIP41</i>	0.22	<i>Actin</i>	0.129	<i>PP2AA3</i>	0.37	1.58
2	<i>MON1</i>	0.22	<i>EF-1α</i>	0.134	<i>CYP</i>	0.39	1.71
3	<i>PP2AA3</i>	0.24	<i>TUA</i>	0.146	<i>MON1</i>	0.42	1.78
4	<i>18S rRNA</i>	0.25	<i>UBC</i>	0.164	<i>TIP41</i>	0.42	1.91
5	<i>CYP</i>	0.27	<i>18S rRNA</i>	0.169	<i>18S rRNA</i>	0.41	2.28
6	<i>UBC</i>	0.59	<i>TIP41</i>	0.172	<i>TUA</i>	1.08	5.05
7	<i>EF-1α</i>	0.75	<i>MON1</i>	0.173	<i>EF-1α</i>	0.91	5.07
8	<i>TUA</i>	0.86	<i>PP2AA3</i>	0.174	<i>UBC</i>	0.87	5.09
9	<i>Actin</i>	0.99	<i>GAPDH</i>	0.181	<i>Actin</i>	1.03	5.47
10	<i>eIF-4α</i>	1.08	<i>eIF-4α</i>	0.189	<i>GAPDH</i>	1.25	6.51
11	<i>GAPDH</i>	1.16	<i>CYP</i>	0.193	<i>eIF-4α</i>	1.37	6.92
12	<i>TBP</i>	1.44	<i>TBP</i>	0.24	<i>TBP</i>	1.72	7.64

<https://doi.org/10.1371/journal.pone.0175863.t002>

was the most stably expressed gene in response to Al and Cu treatments. Moreover, *UBC* (0.138) was likely the most suitable reference gene in Zn-treated leaves.

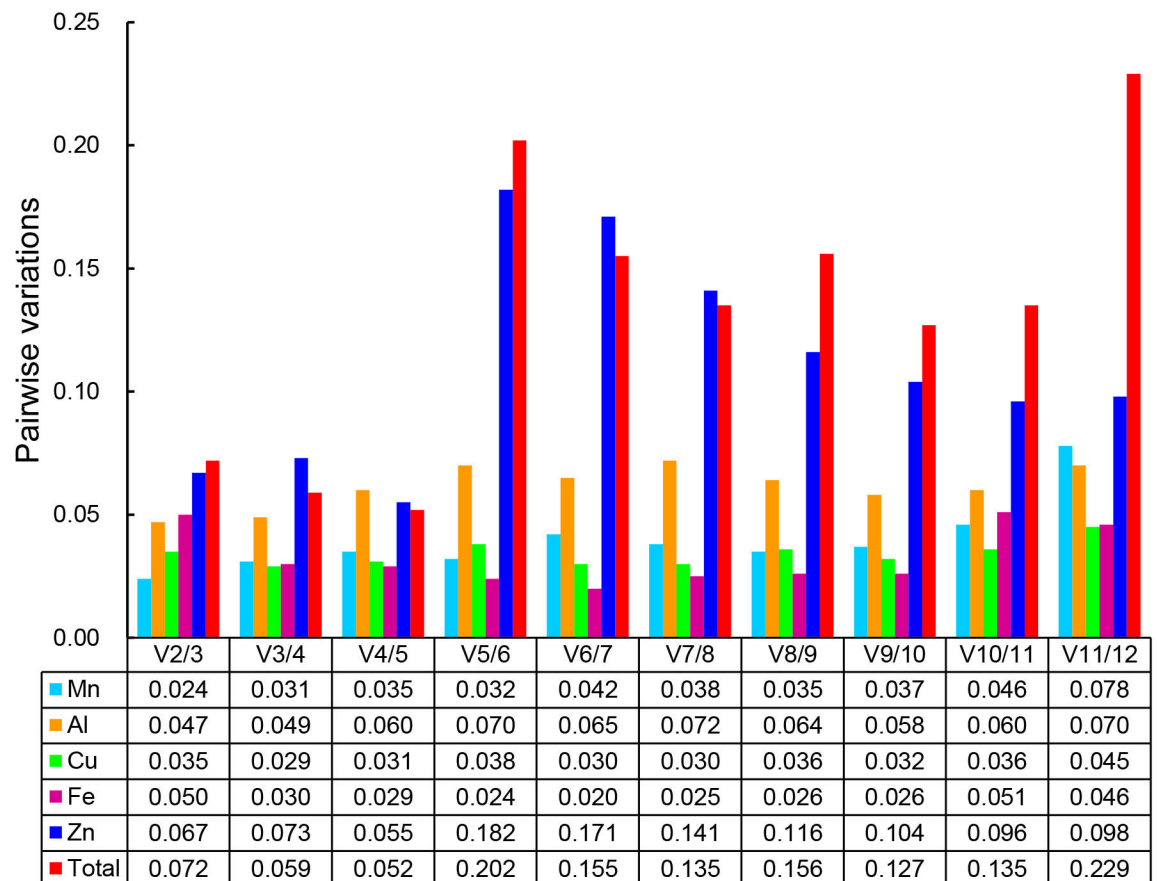


Fig 3. Determination of the optimal number of reference genes required for effective data normalization.

<https://doi.org/10.1371/journal.pone.0175863.g003>

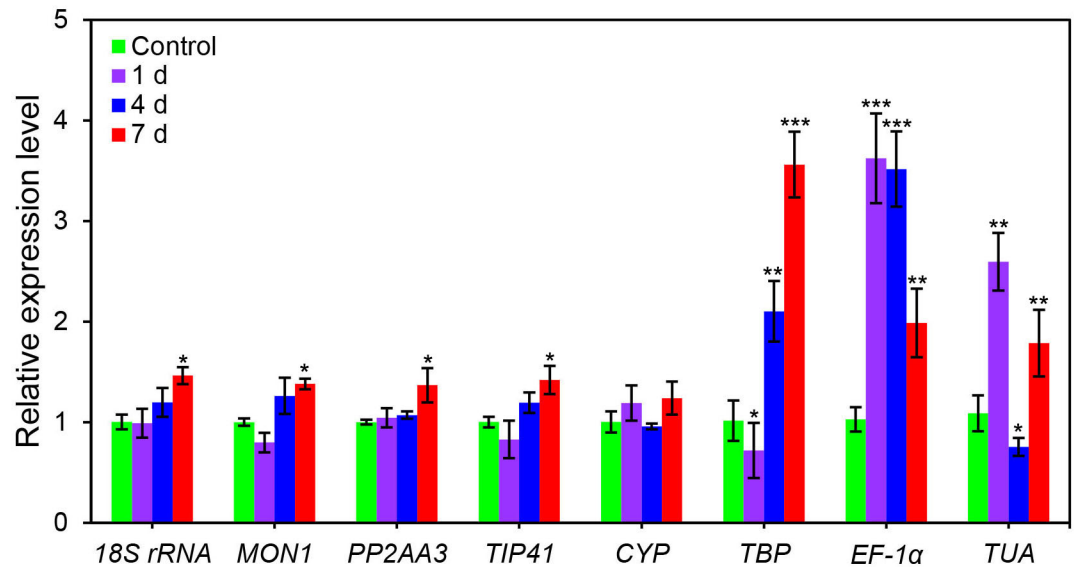


Fig 4. Relative quantification of *CsPCS1* gene expression using candidate reference genes in Al-stressed *C. sinensis* leaves. Data are presented as the means \pm standard deviation of four replicates. Significant differences were determined by Duncan's multiple range test (* $P < 0.05$, ** $P < 0.01$, and *** $P < 0.001$).

<https://doi.org/10.1371/journal.pone.0175863.g004>

Lower standard deviations and coefficients of variation during the BestKeeper analysis corresponded to more stable gene expression. According to the BestKeeper rankings, *PP2AA3* was the most stably expressed gene in Al- and Fe-treated leaves (Table E in S1 File). Additionally, *TIP41* and *MON1* were the most stably expressed genes in Mn- and Cu-treated leaves, respectively. *UBC* expression levels were the most unstable in Zn-treated leaves, which contradicted the NormFinder results that suggested *UBC* is a good reference gene.

Validation of selected reference genes

CsPCS1 gene which is related to metal stresses was selected to further evaluate the reliability of the potential *C. sinensis* reference genes in qRT-PCR assays [20, 32]. The relative *CsPCS1* expression level following an Al treatments was determined using the expression levels of *18S rRNA*, *MON1*, *PP2AA3*, *TIP41*, *CYP*, *TBP*, *EF-1α*, and *TUA* to normalize data (Fig 4). Data normalizations using the more stably expressed reference genes (i.e., *18S rRNA* and *MON1*) resulted in consistent *CsPCS1* expression patterns, with the highest expression level observed following a 7-day exposure to Al. Similar expression patterns were observed when the less stably expressed reference genes were used to normalize data (i.e., *TIP41*, *CYP*, and *TBP*). However, when the least stably expressed genes were used for data normalization (i.e., *EF-1α* and *TUA*), the expression level of *CsPCS1* was considerably biased. This results indicated that the least stable genes *EF-1α* and *TUA* failed to standardize the expression data effectively.

Discussion

In recent years, qRT-PCR has become an outstanding technique for studying gene expression profiles because of its accuracy, sensitivity, and reproducibility. To reach its maximum analytical potential, it is necessary to introduce appropriate internal reference genes or housekeeping genes for normalizing data. Sun et al. [33] were the first researchers to identify suitable *C. sinensis* reference genes for qRT-PCR analyses. They determined that *Csβ-actin* (GenBank: HQ420251) and *CsGAPDH* (GenBank: GE651107) could be used as reference gene during

analyses of different tissues and leaf developmental stages, respectively. Afterwards, *Csβ-actin* and *CsGAPDH* have been widely applied for gene expression analyses in tea plants [14, 34, 35]. However, it is misleading to use previously identified reference genes for normalization without first investigating the stability of their expression levels under specific experimental conditions. Therefore, *18S rRNA*, *Actin*, *CYP*, *EF-1α*, *eIF-4α*, *GAPDH*, *MON1*, *PP2AA3*, *TBP*, *TIP41*, *TUA*, and *UBC* were selected to be validated under the experimental conditions of the present study, because they were observed to be stably expressed under various abiotic stresses, in particular in response to heavy metals [36–39].

Expression levels were determined as Cq values by qRT-PCR. The mean Cq values of the genes ranged from 17.08 (*UBC*)–23.51 (*TBP*), and the Cq values for all the tested samples were between 15 and 30. Here, the Cq values of *PP2AA3*, *18S rRNA*, and *CYP* ($Cq_{\max} - Cq_{\min} < 2$ cycles; SD = 0.47, 0.50, and 0.46, respectively) were distributed more centrally than those of the other candidate genes, whereas the Cq value of *TBP* showed the highest variation ($Cq_{\max} - Cq_{\min} > 10$ cycles, SD = 3.03). The Cq value comparison can provide a rough estimate on stability of gene expression, but not sufficient for accurate evaluation of expression patterns of reference genes [40, 41].

Three statistical algorithms (i.e., geNorm, NormFinder, and BestKeeper) were further used to determine which reference gene is best suited for transcript normalization in tea leaves exposed to heavy metals. Our geNorm results ($V_2/3$ value < 0.15) revealed that two reference genes were sufficient for qRT-PCR data normalizations for the aforementioned five heavy metal stresses, indicating that the addition of third gene had no significant effect for normalization. In Mn-treated tea leaves, the geNorm analysis indicated that *PP2AA3*, *TBP*, and *UBC* were the most appropriate reference genes. Although *TBP* was identified as the best reference gene according to the geNorm results, it was only the seventh most suitable gene based on the BestKeeper analysis. NormFinder, geNorm, and BestKeeper ranked *Actin* as the best, third-best, and sixth-best reference gene, respectively. Additionally, *TIP41* was identified as the best reference gene by the BestKeeper program, while it was only the eighth-best and seventh-best reference gene according to geNorm and NormFinder, respectively. Based on these results, *PP2AA3* combined with *TBP* or *UBC* were recommended as the best combination of stable reference genes for normalizing gene expression levels in Mn-treated tea leaves. Similarly, *18S rRNA* combined with *PP2AA3* or *TIP41* are sufficient for analyses of Al-treated tea leaves. *MON1* combined with *18S rRNA* or *TIP41* should be used for Cu treatments, *eIF-4α* combined with *Actin* or *CYP* are suitable for Fe treatments. Finally, *PP2AA3* combined with *CYP* or *18S rRNA* are sufficient for analyzing Zn-treated tea leaves.

During previous analyses of gene expression levels in tea plant, *CsGAPDH* (GenBank: GE651107) was often used to normalize qRT-PCR data [13, 42, 43] because it was stably expressed in many other plant species [44–46]. However, according to geNorm and BestKeeper results of our study, *GAPDH* was identified as the least stably expressed reference gene in Zn- and Mn-treated tea leaves. These results highlight the fact that there is no single universal reference gene that is stably expressed under all experimental conditions [47–49] and that reference genes should be reconfirmed according to specific experimental conditions [50, 51].

To validate the suitability of potential reference genes, the expression profiles of *CsPCS1* was assessed in Al-treated tea leaves, with *18S rRNA*, *MON1*, *PP2AA3*, *TIP41*, *CYP*, *TBP*, *EF-1α*, and *TUA* as internal reference genes. The *CsPCS1* expression patterns were similar when the gene expression data were normalized using the most stably expressed reference gene (i.e., *18S rRNA*) and less stably expressed reference genes (i.e., *PP2AA3*, *MON1*, and *TIP41*). In contrast, when *TBP*, *EF-1α*, and *TUA* were used for data normalizations, the expression patterns and transcript levels were obviously different from those obtained following data normalizations with *18S rRNA* and other suitable reference genes. Hence, using a stable

reference gene is a prerequisite for accurate relative quantifications of gene expression levels [52, 53].

Conclusions

To the best of our knowledge, this is the first report describing the identification of suitable reference genes for qRT-PCR analyses in *C. sinensis* leaves exposed to heavy metal stress. The stability analysis of gene expression by geNorm, NormFinder, and BestKeeper indicated that no single gene was stably expressed under different metal stress conditions. We have identified *PP2AA3*, *UBC*, and *TBP* as suitable reference genes under Mn stress. *18S rRNA*, *PP2AA3*, and *TIP41* were the most stably expressed genes under Al stress. *MON1*, *TIP41*, and *18S rRNA* were the most stable ones under Cu stress; *eIF-4 α* , *CYP*, and *Actin* under Fe stress; and *PP2AA3*, *CYP*, and *18S rRNA* under Zn stress. Therefore, reference genes should be selected according to the specific metal stress being investigated. Moreover, the stability of previously used reference genes should be re-assessed to increase the accuracy of expression data and to avoid error propagation under certain experimental conditions. Additionally, the analysis of *CsPCS1* expression levels confirmed the importance of selecting appropriate reference genes for the normalization of qRT-PCR data. The reference genes selected in this study provide more choices for further gene expression analysis and functional studies in *C. sinensis*.

Supporting information

S1 File. Contains Figures. A-N and Tables A-E.
(DOCX)

Acknowledgments

We would like to thank Dr. Hai-Bin Wang (College of Horticulture, Nanjing Agricultural University, Nanjing, Jiangsu, People's Republic of China) for his helpful advice on data analysis.

Author Contributions

Conceptualization: XHL MLW.

Data curation: MLW QHL.

Formal analysis: MLW QHL HHX XJZ.

Funding acquisition: XHL.

Investigation: MLW QHL.

Methodology: MLW.

Project administration: XC XHL.

Resources: XHL.

Software: HHX XJZ.

Supervision: XC XHL.

Validation: MLW XHL.

Visualization: MLW QHL.

Writing – original draft: MLW QHL.

Writing – review & editing: XC XHL.

References

1. Radonic A, Thulke S, Mackay IM, Landt O, Siegert W, Nitsche A. Guideline to reference gene selection for quantitative real-time PCR. *Biochem Biophys Res Commun*. 2004; 313(4): 856–862. PMID: [14706621](https://pubmed.ncbi.nlm.nih.gov/14706621/)
2. Remans T, Keunen E, Bex GJ, Smeets K, Vangronsveld J, Cuypers A. Reliable gene expression analysis by reverse transcription-quantitative PCR: reporting and minimizing the uncertainty in data accuracy. *Plant Cell*. 2014; 26(10): 3829–3837. <https://doi.org/10.1105/tpc.114.130641> PMID: [25361954](https://pubmed.ncbi.nlm.nih.gov/25361954/)
3. 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. <https://doi.org/10.1373/clinchem.2008.112797> PMID: [19246619](https://pubmed.ncbi.nlm.nih.gov/19246619/)
4. 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. <https://doi.org/10.1105/tpc.108.059774> PMID: [18664615](https://pubmed.ncbi.nlm.nih.gov/18664615/)
5. 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.
6. 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. <https://doi.org/10.1158/0008-5472.CAN-04-0496> PMID: [15289330](https://pubmed.ncbi.nlm.nih.gov/15289330/)
7. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res*. 2001; 29(9): e45. PMID: [11328886](https://pubmed.ncbi.nlm.nih.gov/11328886/)
8. Dheda K, Huggett JF, Chang JS, Kim LU, Bustin SA, Johnson MA, et al. The implications of using an inappropriate reference gene for real-time reverse transcription PCR data normalization. *Anal Biochem*. 2005; 344(1): 141–143. <https://doi.org/10.1016/j.ab.2005.05.022> PMID: [16054107](https://pubmed.ncbi.nlm.nih.gov/16054107/)
9. Kumar A, Chawla V, Sharma E, Mahajan P, Shankar R, Yadav SK. Comparative transcriptome analysis of Chinari, Assamica and Cambod tea (*Camellia sinensis*) types during development and seasonal variation using RNA-Seq technology. *Sci Rep*. 2016; 6: 37244. <https://doi.org/10.1038/srep37244> PMID: [27853259](https://pubmed.ncbi.nlm.nih.gov/27853259/)
10. Liu M, Tian H-I, Wu J-H, Cang R-R, Wang R-X, Qi X-H, et al. Relationship between gene expression and the accumulation of catechin during spring and autumn in tea plants (*Camellia sinensis* L.). *Hortic Res*. 2015; 2: 15023. <https://doi.org/10.1038/hortres.2015.23> PMID: [26506527](https://pubmed.ncbi.nlm.nih.gov/26506527/)
11. Kumar D, Gulati A, Sharma U. Determination of theanine and catechin in *Camellia sinensis* (Kangra tea) leaves by HPTLC and NMR techniques. *Food Anal Method*. 2016; 9(6): 1666–1674.
12. Sharangi AB. Medicinal and therapeutic potentialities of tea (*Camellia sinensis* L.)—A review. *Food Res Int*. 2009; 42(5–6): 529–535.
13. Liu SC, Jin JQ, Ma JQ, Yao MZ, Ma CL, Li CF, et al. Transcriptomic analysis of tea plant responding to drought stress and recovery. *PLoS ONE*. 2016; 11(1): e0147306. <https://doi.org/10.1371/journal.pone.0147306> PMID: [26788738](https://pubmed.ncbi.nlm.nih.gov/26788738/)
14. Liu ZW, Wu ZJ, Li XH, Huang Y, Li H, Wang YX, et al. Identification, classification, and expression profiles of heat shock transcription factors in tea plant (*Camellia sinensis*) under temperature stress. *Gene*. 2016; 576(1): 52–59.
15. Wang YX, Liu ZW, Wu ZJ, Li H, Zhuang J. Transcriptome-wide identification and expression analysis of the NAC gene family in tea plant [*Camellia sinensis* (L.) O. Kuntze]. *PLoS One*. 2016; 11(11): e0166727. <https://doi.org/10.1371/journal.pone.0166727> PMID: [27855193](https://pubmed.ncbi.nlm.nih.gov/27855193/)
16. Ishibashi Y, Matsuo H, Baba Y, Nagafuchi Y, Imato T, Hirata T. Association of manganese effluent with the application of fertilizer and manure on tea field. *Water Res*. 2004; 38(12): 2821–2826. <https://doi.org/10.1016/j.watres.2004.04.006> PMID: [15223275](https://pubmed.ncbi.nlm.nih.gov/15223275/)
17. Li QH, Li Y, Wu XY, Zhou L, Zhu XJ, Fang WP. Metal transport protein 8 in *Camellia sinensis* confers superior manganese tolerance when expressed in yeast and *Arabidopsis thaliana*. *Sci Rep-Uk*. 2017; 7: 39915.
18. Mukhopadhyay M, Mondal TK. Effect of zinc and boron on growth and water relations of *Camellia sinensis* (L.) O. Kuntze cv. T-78. *Natl Acad Sci Lett*. 2015; 38(3): 283–286.
19. Mukhopadhyay M, Das A, Subba P, Bantawa P, Sarkar B, Ghosh P, et al. Structural, physiological, and biochemical profiling of tea plantlets under zinc stress. *Biol Plantarum*. 2013; 57(3): 474–480.

20. Yadav SK, Mohanpuria P. Responses of *Camellia sinensis* cultivars to Cu and Al stress. *Biol Plantarum*. 2009; 53(4): 737–740.
21. Liu XW, Gao XY, He YQ, Gao X, Xiao C, Wu GX, et al. Effect of several trace elements on the tea plant physiological and tea quality. *Guangdong Agricultural Sciences*. 2010; 6: 162–165.
22. Hao XY, Horvath DP, Chao WS, Yang YJ, Wang XC, Xiao B. Identification and evaluation of reliable reference genes for quantitative real-time PCR analysis in tea plant (*Camellia sinensis* (L.) O. Kuntze). *Int J Mol Sci*. 2014; 15(12): 22155–22172. <https://doi.org/10.3390/ijms151222155> PMID: 25474086
23. Ma QP, Hao S, Chen X, Li XH. Validation of reliability for reference genes under various abiotic stresses in tea plant. *Russ J Plant Physiol*. 2016; 63(3): 423–432.
24. Wu ZJ, Tian C, Jiang Q, Li XH, Zhuang J. Selection of suitable reference genes for qRT-PCR normalization during leaf development and hormonal stimuli in tea plant (*Camellia sinensis*). *Sci Rep*. 2016; 6: 19748. <https://doi.org/10.1038/srep19748> PMID: 26813576
25. Nicot N, Hausman JF, Hoffmann L, Evers D. Housekeeping gene selection for real-time RT-PCR normalization in potato during biotic and abiotic stress. *J Exp Bot*. 2005; 56(421): 2907–2914. <https://doi.org/10.1093/jxb/eri285> PMID: 16188960
26. 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. <https://doi.org/10.1007/s00425-011-1410-3> PMID: 21505864
27. Jain M, Nijhawan A, Tyagi AK, Khurana JP. Validation of housekeeping genes as internal control for studying gene expression in rice by quantitative real-time PCR. *Biochem Biophys Res Commun*. 2006; 345(2): 646–651. <https://doi.org/10.1016/j.bbrc.2006.04.140> PMID: 16690022
28. Martins PK, Mafra V, De Souza WR, Ribeiro AP, Vinecky F, Basso MF, et al. Selection of reliable reference genes for RT-qPCR analysis during developmental stages and abiotic stress in *Setaria viridis*. *Sci Rep*. 2016; 6: 28348. <https://doi.org/10.1038/srep28348> PMID: 27321675
29. Wang WD, Xin HH, Wang ML, Ma QP, Wang L, Kaleri NA, et al. Transcriptomic analysis reveals the molecular mechanisms of drought-stress-induced decreases in *Camellia sinensis* leaf quality. *Front Plant Sci*. 2016; 7: 385. <https://doi.org/10.3389/fpls.2016.00385> PMID: 27066035
30. Wan Q, Xu RK, Li XH. Proton release by tea plant (*Camellia sinensis* L.) roots as affected by nutrient solution concentration and pH. *Plant Soil Environ*. 2012; 58(9):429–434.
31. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods*. 2001; 25(4): 402–408. <https://doi.org/10.1006/meth.2001.1262> PMID: 11846609
32. Cobbett CS. Phytochelatin biosynthesis and function in heavy-metal detoxification. *Curr Opin Plant Biol*. 2000; 3(3): 211–216. PMID: 10837262
33. Sun M, Wang Y, Yang D, Wei C, Gao L, Xia T, et al. Reference genes for real-time fluorescence quantitative PCR in *Camellia sinensis*. *Chin Bull Bot*. 2010; 45(5): 579–587.
34. Zhu XJ, Li QH, Hu JY, Wang ML, Li XH. Molecular cloning and characterization of spermine synthesis gene associated with cold tolerance in tea plant (*Camellia sinensis*). *Appl Biochem Biotech*. 2015; 177(5): 1055–1068.
35. Zhu XJ, Zhao Z, Xin HH, Wang ML, Wang WD, Chen X, et al. Isolation and dynamic expression of four genes involving in shikimic acid pathway in *Camellia sinensis* 'Baicha 1' during periodic albinism. *Mol Biol Rep*. 2016; 43(10): 1119–1127. <https://doi.org/10.1007/s11033-016-4045-4> PMID: 27553670
36. Remans T, Smeets K, Opdenakker K, Mathijsen D, Vangronsveld J, Cuypers A. Normalisation of real-time RT-PCR gene expression measurements in *Arabidopsis thaliana* exposed to increased metal concentrations. *Planta*. 2008; 227(6): 1343–1349. <https://doi.org/10.1007/s00425-008-0706-4> PMID: 18273637
37. Basa B, Solti A, Sarvari E, Tamas L. Housekeeping gene selection in poplar plants under Cd-stress: comparative study for real-time PCR normalisation. *Funct Plant Biol*. 2009; 36(12): 1079–1087.
38. Migocka M, Papierniak A. Identification of suitable reference genes for studying gene expression in cucumber plants subjected to abiotic stress and growth regulators. *Mol Breeding*. 2011; 28(3): 343–357.
39. Huang GY, Wang YS. Expression analysis of type 2 metallothionein gene in mangrove species (*Bru-guiera gymnorrhiza*) under heavy metal stress. *Chemosphere*. 2009; 77(7): 1026–1029. <https://doi.org/10.1016/j.chemosphere.2009.07.073> PMID: 19716152
40. Jiang Q, Wang F, Li MY, Ma J, Tan GF, Xiong AS. Selection of suitable reference genes for qPCR normalization under abiotic stresses in *Oenanthе javanica* (Bl.) DC. *PLoS One*. 2014; 9(3): e92262. <https://doi.org/10.1371/journal.pone.0092262> PMID: 24651080

41. Tian C, Jiang Q, Wang F, Wang GL, Xu ZS, Xiong AS. Selection of suitable reference genes for qPCR normalization under abiotic stresses and hormone stimuli in carrot leaves. *PLoS One*. 2015; 10(2): e0117569. <https://doi.org/10.1371/journal.pone.0117569> PMID: 25658122
42. Wang Y, Shu Z, Wang W, Jiang X, Li D, Pan J, et al. *CsWRKY2*, a novel WRKY gene from *Camellia sinensis*, is involved in cold and drought stress responses. *Biol Plantarum*. 2016; 60(3): 443–451.
43. Deng WW, Zhang M, Wu JQ, Jiang ZZ, Tang L, Li YY, et al. Molecular cloning, functional analysis of three cinnamyl alcohol dehydrogenase (CAD) genes in the leaves of tea plant, *Camellia sinensis*. *J Plant Physiol*. 2013; 170(3): 272–282. <https://doi.org/10.1016/j.jplph.2012.10.010> PMID: 23228629
44. Huis R, Hawkins S, Neutelings G. Selection of reference genes for quantitative gene expression normalization in flax (*Linum usitatissimum* L.). *BMC Plant Biol*. 2010; 10: 71. <https://doi.org/10.1186/1471-2229-10-71> PMID: 20403198
45. Li HP, Qin YX, Xiao XH, Tang CR. Screening of valid reference genes for real-time RT-PCR data normalization in *Hevea brasiliensis* and expression validation of a sucrose transporter gene *HbSUT3*. *Plant Sci*. 2011; 181(2): 132–139. <https://doi.org/10.1016/j.plantsci.2011.04.014> PMID: 21683878
46. Janska A, Hodek J, Svoboda P, Zamecnik J, Prasil IT, Vlasakova E, et al. The choice of reference gene set for assessing gene expression in barley (*Hordeum vulgare* L.) under low temperature and drought stress. *Mol Genet Genomics*. 2013; 288(11): 639–649. <https://doi.org/10.1007/s00438-013-0774-4> PMID: 23979536
47. Dekkers BJW, Willems L, Bassel GW, van Bolderen-Veldkamp RP, Ligterink W, Hilhorst HWM, et al. Identification of reference genes for RT-qPCR expression analysis in Arabidopsis and tomato seeds. *Plant Cell Physiol*. 2012; 53(1): 28–37. <https://doi.org/10.1093/pcp/pcr113> PMID: 21852359
48. Gong L, Yang Y, Chen Y, Shi J, Song Y, Zhang H. *LbCML38* and *LbRH52*, two reference genes derived from RNA-Seq data suitable for assessing gene expression in *Lycium barbarum* L. *Sci Rep*. 2016; 6: 37031. <https://doi.org/10.1038/srep37031> PMID: 27841319
49. Lovdal T, Lillo C. Reference gene selection for quantitative real-time PCR normalization in tomato subjected to nitrogen, cold, and light stress. *Anal Biochem*. 2009; 387(2): 238–242. <https://doi.org/10.1016/j.ab.2009.01.024> PMID: 19454243
50. Dos Santos LF, Silva RJS, Do Amaral DOJ, De Paula MFB, Falcao LL, Legavre T, et al. Selection of reference genes for expression study in pulp and seeds of *Theobroma grandiflorum* (Willd. ex Spreng.) Schum. *PLoS One*. 2016; 11(8): e0160646. <https://doi.org/10.1371/journal.pone.0160646> PMID: 27501324
51. Taylor CM, Jost R, Erskine W, Nelson MN. Identifying stable reference genes for qRT-PCR normalisation in gene expression studies of narrow-leafed lupin (*Lupinus angustifolius* L.). *PLoS One*. 2016; 11(2): e0148300. <https://doi.org/10.1371/journal.pone.0148300> PMID: 26872362
52. Goulao LF, Fortunato AS, Ramalho JC. Selection of reference genes for normalizing quantitative real-time PCR gene expression data with multiple variables in *Coffea* spp. *Plant Mol Biol Rep*. 2012; 30(3): 741–759.
53. Guenin 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. <https://doi.org/10.1093/jxb/ern305> PMID: 19264760