Identification and evaluation of suitable reference genes for gene expression analysis in rubber tree leaf

xiangyu long (✉ yuxianglong006@163.com)

rubber research institute

Jilai Lu

rubber research institute

Nat N. V. Kav

university of alberta

Yunxia Qin

rubber research institute

Yongjun Fang

rubber research institute

Zhiyong Wang

Hainan university

Research article

Keywords: Reference gene, Expression profiling, Stability, Normalization, Leaf, Rubber tree leaf

Posted Date: June 13th, 2019

DOI: <https://doi.org/10.21203/rs.2.10285/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Version of Record: A version of this preprint was published at Molecular Biology Reports on February 4th, 2020. See the published version at <https://doi.org/10.1007/s11033-020-05288-8>.

Abstract

Background Gene expression profiling is increasingly applied to investigate molecular mechanisms for which, normalization with suitable reference genes is critical. Previously we have reported several suitable reference genes for laticifer samples from rubber, however, little is known about reference genes in leaf. Results The main objective of this current study was to identify some reference genes with stable expression patterns in leaf at various developmental stages, as well as during abiotic (temperature extremes) and biotic stresses. Gene expression profiling experiments in rubber tree leaf identified the ubiquitin-proteasome system as having excellent potential as reference genes. Among a total of 30 tested genes investigated, 24 new (including 11 genes involved in the ubiquitin-proteasome system), 4 previously identified and 2 specific genes, were further evaluated using quantitative real-time PCR. Our results indicated that the new genes had better stability of expression when compared with others. For instance, an ubiquitin conjugating enzyme (RG0099) and three ubiquitin-protein ligases (RG0928, RG2190 and RG0118) expressed stably in all samples, and were confirmed to be suitable reference genes in rubber tree leaf in four different conditions. Finally, we suggest that using more than one reference gene may be appropriate in gene expression studies when employing different software to normalize gene expression data. Conclusion Our findings have significant implications for the reliability of data obtained from genomics studies in rubber tree and perhaps in other species.

Background

Gene expression profiling studies are widely used to investigate plant processes, including metabolic pathways, including regulatory networks that are crucial for plant growth & development as well as in stress responses. These types of genomics-based investigation necessitate the use of appropriate reference genes to normalize gene expression data. Incorrect choice of reference genes in techniques such as real-time quantitative PCR (qRT-PCR), RT-PCR, Northern blotting, ribonuclease protection assay and gene chip can result in the misinterpretation of gene expression data [1-3]. An optimal reference gene should be stably expressed in various tissues and under different experimental conditions without being influenced by the experimental conditions used [4, 5]. Common reference genes that are used to normalize gene expression data include 18S ribosomal RNA (18S rRNA), actin (ACT), glyceraldehyde-3-phosphate dehydrogenase (GAPDH), tubulin (TUB). However, various studies have indicated that they may be unsuitable as reference genes for normalization due to the effects of certain experimental conditions on their expression [6-8]. Many studies have been conducted to screen and identify new reference genes that show stable expression patterns and to determine the appropriateness of genes traditionally used as reference genes in specific species and experimental conditions [9-18].

In rubber trees (*Hevea brasiliensis*), natural rubber synthesis is the main metabolic activity in laticifers which are highly specialized cells that directly influence the yield of rubber [19]. Most studies in molecular biology, especially those using gene expression profiling, have paid close attention to rubber biosynthesis in laticifers [20-25]. Previously, we have evaluated the expression stability of 22 reference genes across abundant laticifer samples, including latex regenerations, hormone treatments, tappings, individual trees and genotypes of *Hevea* [26, 27]. As mentioned earlier, the availability of reliable reference genes is critical for gene expression studies involving glycometabolism and hormone regulation of rubber biosynthesis in laticifers [22, 25, 28, 29].

Leaves as photosynthetic tissues (sources) provide the nutrition and energy for rubber biosynthesis [20]. In recent years, more and more attention has been paid to leaf research, including leaf development and disease-resistance. For instance, due to the fact that sucrose is the precursor for rubber biosynthesis, enzymological characterization and gene expression of sucrose metabolizing enzymes were systematically investigated to understand the regulation of sucrose metabolism during leaf development in the rubber tree [30]. In addition, leaf development has been widely studied within the context of leaf diseases, especially South American leaf blight (SALB), which directly influences rubber yield, [31, 32]. However, as mentioned previously, little is known about reliable reference genes in leaves in order to normalize gene expression data in leaves.

In this study, two transcriptome sequencing profiles of leaves were screened for appropriate reference genes [32, 33]. The expression of selected genes was verified by qRT-PCR under four different experimental conditions. Our results demonstrated that the ubiquitin-proteasome system had significant potential as reference gene in rubber leaf tissues. Another important finding from this study was that the choice of algorithm used to normalize gene expression data had a significant impact on the results. Our findings are discussed within the context of using the appropriate reference gene to normalize gene expression data, which has the potential to significantly affect results from gene expression profiling studies.

Methods

Plant materials and treatments

The *Hevea* cultivars in Hainan Province, Reyan7-33-97 (synonym: CATAS7-33-97 or RY7-33-97), were provided by the experimental plantation of CATAS (Chinese Academy of Tropical Agricultural Sciences, Danzhou, Hainan, China). One-year-old tissue cultured Reyan7-33-97 plants were selected for the collection of six developmental stages of leaf, including bud, bronze, color-change, pale-green, bright green (mature) and aging. In addition, 6-month-old tissue cultured Reyan7-33-97 plants were used for our biotic and abiotic stress experiments. For the biotic stress experiments, leaves were collected at 0 h, 12 h, 24 h, 48 h, 3d and 5d after inoculating with the pathogenic fungus *Corynespora cassiicola* HccYN57. In the case of abiotic stress experiments, plants were treated with low (4°C) and high (45 °C) temperatures in a growth cabinet with relative humidity of 75-90% and lighting of 16 hours per day, and their leaves were collected at 0 h, 3 h, 6 h, 9 h, 12 h, and 24 h after the imposition of stress. All collected leaves were frozen and stored at -70°C until they were used for RNA isolation.

Total RNA isolation and cDNA synthesis

Total RNA was extracted from leaf tissue using PureLink™ Plant RNA Reagent (Thermo, USA) following the manufacturer's instructions. Genomic DNA was removed after digestion with DNase 1 (Thermo, USA) also according to manufacturer's protocols. RNA concentration and its integrity were determined using a spectrophotometer NanoDrop 2000 and using agarose gel electrophoresis. cDNA was synthesized by reverse transcription using PrimeScript™ RT reagent Kit with gDNA Eraser (Perfect Real Time, Takara, Japan), as per the manufacturer's protocols. Newly synthesized cDNA was diluted 10 fold for use in qRT-PCR experiments.

Collecting candidate reference genes from transcriptome sequencing data

In this study we focused on two reported transcript profiles for screening new candidate reference genes from rubber tree leaf at developmental stages (bronze, color change, pale-green and bright green) as well as under temperature stress (high and low temperature) [32, 33]. Genes were selected as candidate reference genes if they showed reads per kilobase per million mapped reads (RPKM) fold change <1.5. Functional annotations and classifications of these genes were carried out using BLAST NCBI (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) and KEGG (<https://www.genome.jp/kegg/pathway.html>). Finally, 24 genes were chosen as new candidate reference genes to further validate expression stability using qRT-PCR.

Primer design and qRT-PCR assays

A total of 30 tested primer pairs were designed using the OligoAnalyzer web service (<https://sg.idtdna.com/calc/analyzer>), for the 24 new candidate reference genes, 4 previously used reference genes (Eukaryotic translation initiation factor 1A, eIF1Aa; DEAD/DEAH box helicase, RH8; Ubiquitin-protein ligase, UBC2a and UBC2b) and 2 specific genes (ATP citrate lyase, RG1588154 and RG65560). PCR reactions were performed on a CFX96 Touch Real-Time PCR (Bio-Rad, USA), using the SYBR® Premix Ex Taq™ II (Perfect Real Time, TaKaRa, China). Each sample had three technical replicates, and the PCR reaction conditions were as previously described [27]. Data analysis was carried out using CFX Manager (Bio-Rad, USA), including the melting curve, quantification cycle values (Ct), PCR efficiency (E) and correlation coefficients (R²).

Data analysis

The arithmetic mean, geometric mean, standard deviation and variation coefficient of quantification cycle values (Ct) were calculated and statistically analyzed using SPSS 13 (<http://www.spss.com/>). Four software with different algorithm, NormFinder, GeNorm, Bestkeeper and comparative delta Ct method, were used to analyze expression stability of reference genes, according to their instructions (RefFinder, <https://reffinder.net/>) [34-37]. Minitab 15 software (<http://www.minitab.com/>) was used to calculate the Pearson correlation coefficients for ranking results from four different evaluation algorithms.

Results

Selection of candidate reference genes

In this study, we used two previously reported [32, 33] transcript profiles of rubber tree in order to select new candidate reference genes from leaf, including developmental stages (bronze, color change, pale-green and bright green) and abiotic stress (high and low temperature). When the gene expression data was analyzed for those genes where the change in expression less than 1.5 fold, we observed that there were 3,673 and 1,987 transcripts that could be selected from the two experiments, respectively. Among those, 82 genes were observed to exhibit little or no change of expression in both experiments (Fig. S1). In order to identify additional suitable candidate genes, BLASTx was used for gene functional annotation against NCBI and KEGG databases, which resulted in 62 annotated genes involved in five pathways (Table S1). Within those pathways, the genetic information processing pathway was the most represented (39 genes, 47.56%), and included transcription machinery (16 genes, 19.51%), ubiquitin system (13 genes, 15.85%), membrane trafficking (6 genes, 7.32%), chromosome (3 genes, 3.66%) and protein processing (1 gene, 1.22%). Others that were represented included metabolism (11 genes, 13.41%), environmental information (6 genes, 7.32%), organismal system (3 genes, 3.66%) and cellular processes (3 genes, 3.66%) (Fig. 1). These results provided us with the identities of several genes that could be further analyzed using qRT-PCR to narrow down those that may serve as reliable reference genes.

Expression profiles of candidate reference genes

Based on the expression and annotation results, 24 genes were selected as candidate reference genes which were involved in basic metabolic pathways or had the most stable expression in transcript profiles. These genes represented the transcription machinery (9 genes), ubiquitin system (11 genes), specific protein (1 gene), amino acid metabolism (1 gene), nucleotide metabolism (1 gene) and uncharacterized protein (1 gene). The identities of the genes selected for detailed characterization of expression profiles are provided in Table 1. In addition, 4 previously used reference genes (Eukaryotic translation initiation factor 1A, eIF1Aa; DEAD/DEAH box helicase, RH8; Ubiquitin-protein ligase, UBC2a and UBC2b) and 2 specific genes (ATP citrate lyase, RG1588154 and RG65560) were analyzed in order to ascertain their ability to serve as reliable reference genes (Table 1). Thus, a total of 30 genes were evaluated for their expression stability in 60 leaf samples including various developmental stages, high and low temperature stress, and pathogen stress.

qRT-PCR results showed that the amplification efficiencies (E) for these genes ranged from 0.85 to 1.05, and the correlation coefficients (R²) ranged from 0.997 to 1.000 (Table 1). After removing missing and inconsistent data, the mean Ct value for each gene was calculated for all samples to determine expression levels (Table 2, Fig. 2). Wide ranges of Ct values from 19.56 to 36.97 across all samples were observed for the genes investigated (Table 2). The RG0140 gene had the most abundant transcripts with Ct values 23.71±2.05 and, in contrast, the RG0544 gene had the lowest expression levels with Ct values

31.73±1.89. It was clear that there was approximately 200-fold difference in expression levels between these two genes at either end of the expression spectrum. With respect to variation in levels of expression between different samples, the gene RG1039 exhibited the lowest variability (CV, 5.95) and the gene RG65560 was observed to be the most variable (CV, 11.15).

Assessment of Delta Ct, BestKeeper, Normfinder and Genorm

Four different software packages with different algorithms, Delta Ct [37], BestKeeper [34], Normfinder [35] and Genorm [36], were used to evaluate the expression data for the 30 genes in all samples. Our results suggested that Delta Ct, Normfinder and Genorm provided very similar results for gene expression, specifically the same ranking from 19 to 30. Furthermore, these three software packages exhibited high correlation coefficients (R^2), ranging from 0.9259 to 0.9829 further illustrating the fact that they generally yield similar results (Fig. 3). BestKeeper had the lowest R^2 ranging from 0.066 to 0.080 when compared with other software packages, which indicates an inconsistent ability to accurately calculate gene expression data (Fig. 3). In conclusion, among the software packages evaluated in this study, Delta Ct, Normfinder and Genorm provided more reliable quantification of gene expression when compared with BestKeeper.

Expression stability of candidate reference genes

Our results suggest that there were obvious differences in gene expression during different developmental stages, high and low temperature challenges as well as pathogen stresses for all the genes tested, when Delta Ct, Normfinder and Genorm were used for analysis of the data (Table 3). Based on ranking of expression stability, high and low temperature stress experiments were clustered together, indicating they have similar expression patterns (Fig. S2). Furthermore, in temperature stress experiments, the five most stable genes were RG0623, RG0099, RG1206, RG1293 and RG0470 in high temperature stress treatment, and rUBC2A, RG0427, RG0099, RG1200 and RG1121 in low temperature stress treatment (Table 3). In contrast, the expression of both RG0239 and RG0544 was unstable under temperature stress (Table 3). In disease stress, RG0677, RG1039, RG2190, RG0623 and RG0118 were found to be the most stably expressed genes, whereas RG1376, RG0544, RG1478, RG0470 and RG1206 were the least stable genes. Finally, during leaf development, RG0782 showed the most stability in expression, followed by rEIF1Aa, RG0427, RH8 and RG1121 and the genes RG1478, RG0544, RG1293, RG65560 and RG1206 exhibited lower stability of expression (Table 3).

Taken together, the gene expression data from our four experiments indicate that the genes RG0099, RG0427, RG0928, RG2190, RG0118, RG0677, rEIF1Aa, RG1121, RG0623 and RG0354 were the top ten most stable candidate reference genes whereas the genes RG1478, RG1588154, RG0544, RG1206 and RG65560 were the least stable (Table 3). In top ten stable genes, four experiments only shared one gene, RG0099, that was ranked 2nd in high temperature stress, 3rd in low temperature stress, and 6th in developmental stages & disease stress (Fig. 4, Table 3). In contrast, RG0554 was considered to be unstable in expression, ranking 30th in high temperature stress, 29th in low temperature stress, 27th in developmental stages and disease stress (Fig. 4, Table 3). Among the previously used four reference genes, rEIF1Aa showed a better stability of gene expression in all samples with an overall ranking of 7th. Two ATP citrate lyase genes, RG1588154 and RG65560, were identified to be the least stable, ranking as 27th and 30th respectively when all four experiments are considered.

Reference

1. Udvardi MK, Czechowski T, Scheible WR: Eleven golden rules of quantitative RT-PCR. *The Plant cell* 2008, 20(7):1736-1737.
2. Huggett J, Dheda K, Bustin S, Zumla A: Real-time RT-PCR normalisation; strategies and considerations. *Genes and immunity* 2005, 6(4):279-284.
3. Dheda K, Huggett JF, Chang JS, Kim LU, Bustin SA, Johnson MA, Rook GA, Zumla A: The implications of using an inappropriate reference gene for real-time reverse transcription PCR data normalization. *Analytical biochemistry* 2005, 344(1):141-143.
4. Bustin SA: Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. *Journal of molecular endocrinology* 2002, 29(1):23-39.
5. 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. *Journal of experimental botany* 2009, 60(2):487-493.
6. Xu Y, Zhu X, Gong Y, Xu L, Wang Y, Liu L: Evaluation of reference genes for gene expression studies in radish (*Raphanus sativus* L.) using quantitative real-time PCR. *Biochemical and biophysical research communications* 2012, 424(3):398-403.
7. Niu X, Chen M, Huang X, Chen H, Tao A, Xu J, Qi J: Reference Gene Selection for qRT-PCR Normalization Analysis in kenaf (*Hibiscus cannabinus* L.) under Abiotic Stress and Hormonal Stimuli. *Frontiers in plant science* 2017, 8:771.
8. 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.
9. Joseph JT, Poolakkalody NJ, Shah JM: Plant reference genes for development and stress response studies. *Journal of biosciences* 2018, 43(1):173-187.
10. Wan Q, Chen S, Shan Z, Yang Z, Chen L, Zhang C, Yuan S, Hao Q, Zhang X, Qiu D *et al.*: Stability evaluation of reference genes for gene expression analysis by RT-qPCR in soybean under different conditions. 2017, 12(12):e0189405.

11. Kanakachari M, Solanke AU, Prabhakaran N, Ahmad I, Dhandapani G, Jayabalan N, Kumar PA: Evaluation of Suitable Reference Genes for Normalization of qPCR Gene Expression Studies in Brinjal (*Solanum melongena* L.) During Fruit Developmental Stages. *Applied biochemistry and biotechnology* 2016, 178(3):433-450.
12. Wu D, Dong J, Yao YJ, Zhao WC, Gao X: Identification and evaluation of endogenous control genes for use in quantitative RT-PCR during wheat (*Triticum aestivum* L.) grain filling. *Genetics and molecular research : GMR* 2015, 14(3):10530-10542.
13. Lin Y, Zhang C, Lan H, Gao S, Liu H, Liu J, Cao M, Pan G, Rong T, Zhang S: Validation of potential reference genes for qPCR in maize across abiotic stresses, hormone treatments, and tissue types. *PloS one* 2014, 9(5):e95445.
14. Lin F, Jiang L, Liu Y, Lv Y, Dai H, Zhao H: Genome-wide identification of housekeeping genes in maize. *Plant molecular biology* 2014, 86(4-5):543-554.
15. Ji Y, Tu P, Wang K, Gao F, Yang W, Zhu Y, Li S: Defining reference genes for quantitative real-time PCR analysis of anther development in rice. *Acta biochimica et biophysica Sinica* 2014, 46(4):305-312.
16. Bevitore R, Oliveira MB, Grossi-de-Sa MF, Lanna AC, da Silveira RD, Petrofeza S: Selection of optimized candidate reference genes for qRT-PCR normalization in rice (*Oryza sativa* L.) during Magnaporthe oryzae infection and drought. *Genetics and molecular research : GMR* 2014, 13(4):9795-9805.
17. Dekkers BJ, Willems L, Bassel GW, van Bolderen-Veldkamp RP, Ligterink W, Hilhorst HW, Bentsink L: Identification of reference genes for RT-qPCR expression analysis in Arabidopsis and tomato seeds. *Plant & cell physiology* 2012, 53(1):28-37.
18. Kudo T, Sasaki Y, Terashima S, Matsuda-Imai N, Takano T, Saito M, Kanno M, Ozaki S, Suwabe K, Suzuki G *et al*: Identification of reference genes for quantitative expression analysis using large-scale RNA-seq data of Arabidopsis thaliana and model crop plants. *Genes & genetic systems* 2016, 91(2):111-125.
19. Tang C, Yang M, Fang Y, Luo Y, Gao S, Xiao X, An Z, Zhou B, Zhang B, Tan X *et al*: The rubber tree genome reveals new insights into rubber production and species adaptation. 2016, 2(6):16073.
20. Tang C, Huang D, Yang J, Liu S, Sakr S, Li H, Zhou Y, Qin Y: The sucrose transporter HbSUT3 plays an active role in sucrose loading to laticifer and rubber productivity in exploited trees of Hevea brasiliensis (para rubber tree). *Plant, cell & environment* 2010, 33(10):1708-1720.
21. Tang C, Xiao X, Li H, Fan Y, Yang J, Qi J, Li H: Comparative analysis of latex transcriptome reveals putative molecular mechanisms underlying super productivity of Hevea brasiliensis. *PloS one* 2013, 8(9):e75307.
22. Liu S, Lan J, Zhou B, Qin Y, Zhou Y, Xiao X, Yang J, Gou J, Qi J, Huang Y *et al*: HbNIN2, a cytosolic alkaline/neutral-invertase, is responsible for sucrose catabolism in rubber-producing laticifers of Hevea brasiliensis (para rubber tree). *The New phytologist* 2015, 206(2):709-725.
23. Chao J, Yang S, Chen Y, Tian WM: Transcript Profiling of Hevea brasiliensis during Latex Flow. *Frontiers in plant science* 2017, 8:1904.
24. Liu JP, Hu J, Liu YH, Yang CP, Zhuang YF, Guo XL, Li YJ, Zhang L: Transcriptome analysis of Hevea brasiliensis in response to exogenous methyl jasmonate provides novel insights into regulation of jasmonate-elicited rubber biosynthesis. *Physiology and molecular biology of plants : an international journal of functional plant biology* 2018, 24(3):349-358.
25. Long X, He B, Fang Y, Tang C: Identification and Characterization of the Glucose-6-Phosphate Dehydrogenase Gene Family in the Para Rubber Tree, Hevea brasiliensis. *Frontiers in plant science* 2016, 7:215.
26. Li H, Qin Y, Xiao X, Tang C: 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 science : an international journal of experimental plant biology* 2011, 181(2):132-139.
27. Long X, He B, Gao X, Qin Y, Yang J, Fang Y, Qi J, Tang C: Validation of reference genes for quantitative real-time PCR during latex regeneration in rubber tree. *Gene* 2015, 563(2):190-195.
28. Qin Y, Huang Y, Fang Y, Qi J, Tang C: Molecular characterization and expression analysis of the small GTPase ROP members expressed in laticifers of the rubber tree (Hevea brasiliensis). *Plant physiology and biochemistry : PPB* 2014, 74:193-204.
29. Chao J, Yang S, Chen Y, Tian WM: Evaluation of Reference Genes for Quantitative Real-Time PCR Analysis of the Gene Expression in Laticifers on the Basis of Latex Flow in Rubber Tree (Hevea brasiliensis Muell. Arg.). *Frontiers in plant science* 2016, 7:1149.
30. Zhu J, Qi J, Fang Y, Xiao X, Li J, Lan J, Tang C: Characterization of Sugar Contents and Sucrose Metabolizing Enzymes in Developing Leaves of Hevea brasiliensis. *Frontiers in plant science* 2018, 9:58.
31. Lieberei R: South American leaf blight of the rubber tree (Hevea spp.): new steps in plant domestication using physiological features and molecular markers. *Annals of botany* 2007, 100(6):1125-1142.
32. Fang Y, Mei H, Zhou B, Xiao X, Yang M, Huang Y, Long X, Hu S, Tang C: De novo Transcriptome Analysis Reveals Distinct Defense Mechanisms by Young and Mature Leaves of Hevea brasiliensis (Para Rubber Tree). *Scientific reports* 2016, 6:33151.

33. Huang Y, Fang Y, Long X, Liu L, Wang J, Zhu J, Ma Y, Qin Y, Qi J, Hu X *et al*: Characterization of the rubber tree metallothionein family reveals a role in mitigating the effects of reactive oxygen species associated with physiological stress. *Tree physiology* 2018, 38(6):911-924.
34. Pfaffl MW, Tichopad A, Prgomet C, Neuvians TP: Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper–Excel-based tool using pair-wise correlations. *Biotechnology letters* 2004, 26(6):509-515.
35. 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 research* 2004, 64(15):5245-5250.
36. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F: Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome biology* 2002, 3(7):Research0034.
37. Silver N, Best S, Jiang J, Thein SL: Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR. *BMC molecular biology* 2006, 7:33.
38. Gutierrez L, Mauriat M, Guenin S, Pelloux J, Lefebvre JF, Louvet R, Rusterucci C, Moritz T, Guerinneau F, Bellini C *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 biotechnology journal* 2008, 6(6):609-618.
39. Galeano E, Vasconcelos TS, Ramiro DA, De Martin Vde F, Carrer H: Identification and validation of quantitative real-time reverse transcription PCR reference genes for gene expression analysis in teak (*Tectona grandis* L.f.). *BMC research notes* 2014, 7:464.
40. Gill G: SUMO and ubiquitin in the nucleus: different functions, similar mechanisms? *Genes & development* 2004, 18(17):2046-2059.
41. Yamao F: Ubiquitin system: selectivity and timing of protein destruction. *Journal of biochemistry* 1999, 125(2):223-229.
42. Sato T, Maekawa S, Yasuda S, Yamaguchi J: Carbon and nitrogen metabolism regulated by the ubiquitin-proteasome system. *Plant signaling & behavior* 2011, 6(10):1465-1468.
43. Ye J, Jin CF, Li N, Liu MH, Fei ZX, Dong LZ, Li L, Li ZQ: Selection of suitable reference genes for qRT-PCR normalisation under different experimental conditions in *Eucommia ulmoides* Oliv. *Scientific reports* 2018, 8(1):15043.
44. Liu J, Huang S, Niu X, Chen D, Chen Q, Tian L, Xiao F, Liu Y: Genome-wide identification and validation of new reference genes for transcript normalization in developmental and post-harvested fruits of *Actinidia chinensis*. *Gene* 2018, 645:1-6.
45. Pabuayon IM, Yamamoto N, Trinidad JL, Longkumer T, Raorane ML, Kohli A: Reference genes for accurate gene expression analyses across different tissues, developmental stages and genotypes in rice for drought tolerance. *Rice (New York, NY)* 2016, 9(1):32.
46. Niu L, Tao YB, Chen MS, Fu Q, Li C, Dong Y, Wang X, He H, Xu ZF: Selection of Reliable Reference Genes for Gene Expression Studies of a Promising Oilseed Crop, *Plukenetia volubilis*, by Real-Time Quantitative PCR. *International journal of molecular sciences* 2015, 16(6):12513-12530.
47. Storch TT, Pegoraro C, Finatto T, Quecini V, Rombaldi CV, Girardi CL: Identification of a novel reference gene for apple transcriptional profiling under postharvest conditions. *PLoS one* 2015, 10(3):e0120599.
48. Wan H, Yuan W, Ruan M, Ye Q, Wang R, Li Z, Zhou G, Yao Z, Zhao J, Liu S *et al*: Identification of reference genes for reverse transcription quantitative real-time PCR normalization in pepper (*Capsicum annuum* L.). *Biochemical and biophysical research communications* 2011, 416(1-2):24-30.
49. Gu CS, Liu LQ, Xu C, Zhao YH, Zhu XD, Huang SZ: Reference gene selection for quantitative real-time RT-PCR normalization in *Iris. lactea* var. *chinensis* roots under cadmium, lead, and salt stress conditions. *TheScientificWorldJournal* 2014, 2014:532713.
50. Hu Y, Fu H, Qiao H, Sun S: Validation and Evaluation of Reference Genes for Quantitative Real-Time PCR in *Macrobrachium nipponense*. 2018, 19(8).
51. Wang M, Wang Q, Zhang B: Evaluation and selection of reliable reference genes for gene expression under abiotic stress in cotton (*Gossypium hirsutum* L.). *Gene* 2013, 530(1):44-50.
52. 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 physiology* 2005, 139(1):5-17.
53. Moraes GP, Benitez LC, do Amaral MN, Vighi IL, Auler PA, da Maia LC, Bianchi VJ, Braga EJ: Evaluation of reference genes for RT-qPCR studies in the leaves of rice seedlings under salt stress. *Genetics and molecular research : GMR* 2015, 14(1):2384-2398.
54. Xia W, Mason AS, Xiao Y, Liu Z, Yang Y, Lei X, Wu X, Ma Z, Peng M: Analysis of multiple transcriptomes of the African oil palm (*Elaeis guineensis*) to identify reference genes for RT-qPCR. *Journal of biotechnology* 2014, 184:63-73.
55. Tang F, Chu L, Shu W, He X, Wang L, Lu M: Selection and validation of reference genes for quantitative expression analysis of miRNAs and mRNAs in Poplar. *Plant methods* 2019, 15:35.

56. Chen J, Huang Z, Huang H, Wei S, Liu Y, Jiang C, Zhang J, Zhang C: Selection of relatively exact reference genes for gene expression studies in goosegrass (*Eleusine indica*) under herbicide stress. *Scientific reports* 2017, 7:46494.

57. Zhao S, Torres A, Henry RA, Trefely S, Wallace M, Lee JV, Carrer A, Sengupta A, Campbell SL, Kuo YM *et al*: ATP-Citrate Lyase Controls a Glucose-to-Acetate Metabolic Switch. *Cell reports* 2016, 17(4):1037-1052.

58. Cheng Y, Pang X, Wan H, Ahammed GJ, Yu J, Yao Z, Ruan M, Ye Q, Li Z, Wang R *et al*: Identification of Optimal Reference Genes for Normalization of qPCR Analysis during Pepper Fruit Development. *Frontiers in plant science* 2017, 8:1128.

59. Long XY, Wang JR, Ouellet T, Rocheleau H, Wei YM, Pu ZE, Jiang QT, Lan XJ, Zheng YL: Genome-wide identification and evaluation of novel internal control genes for Q-PCR based transcript normalization in wheat. *Plant molecular biology* 2010, 74(3):307-311.

Tables

Table 1 Description of the 24 candidate reference genes, 4 previous reference gens and 2 specific genes

Gene Name	Gene ID	Accession Number	Gene Function	Primer Sequence F(5'-3')	Primer Sequence R(5'-3')
RG1293	scaffold1293_128736	XM_021799024	50S ribosomal protein L24, chloroplastic	GAATGTAGCAAGTCGGGTGG	TTTCGTGTTCTGCTCTTCCTT
RG1039	scaffold1039_22604	XM_021794249	Splicing factor 3B subunit 4	ATGGGTCACAATTCAATGGCAT	AGGGGTATGATGTATGGTTAC
RG2011	scaffold2011_47037	XM_021804641	Transcription and mRNA export factor SUS1	AAAGACGAAATGAAAGCTCTTTGC	AGGAATAGATGCTCTGCCCTT
RG0347	scaffold0347_1131868	XM_021827178	Transcription elongation factor SPT4 homolog 2	CGCTGGTTGAGAATCGGAAG	TTGGCGGAATGTATTGTACTC
RG1200	scaffold1200_31568	XM_021797521	Transcription factor GTE9	ATGGCTTGGGAAGTTTAAAGTTC	TCAATCAATTTCTCCCTCTTC
RG0140	scaffold0140_1151167	XM_021815322	Eukaryotic translation initiation factor 1A	GGAAGATGCACAAGAAGGTTTGG	TCCTCATCAAGACCTCCAGC
RG0239	scaffold0239_1243650	XM_021821646	Pre-rRNA-processing protein TSR2 homolog	TTAGACAGGCGGCAGATAAC	GTTCACCTGATTTGATTTCGA
RG1121	scaffold1121_251572	XM_021795919	General transcription factor IIE subunit 2	CCACAAACACGGCAAAGAGG	TCTGGAAGATGTGCATTTGTC
RG0623	scaffold0623_52574	XM_021781311	Peptidyl-prolyl cis-trans isomerase E	ACGCATTAAGGGTGGTGAAC	ATCTTCCTTTTCTCCTTCTCG
RG1478	scaffold1478_36769	XM_021801104	Ubiquitin protein ligase DRIP2 E3	CAGTTTCTCCACAGGCAGATAG	GTAAGAAGCTGCTAAGGGATG
RG0520	scaffold0520_199136	XM_021835094	Ubiquitin-protein ligase At3g02290 E3	CTGTCCAACATGCCTTGAAGA	GGTGTCACCTCTTTCCATCCA
RG0928	scaffold0928_351560	XM_021791601	Ubiquitin-protein ligase RING1 E3	ACGATGCTGATTATGAGAGTCGG	AGCTCCACCATCACTGCTAG
RG1690	scaffold1690_30417	XM_021802842	Ubiquitin-protein ligase RNF170 E3	TGTTCAGAAAGTGCGAGAGTC	CAAACAATCTAACATCCCCAC
RG0470	scaffold0470_347639	XM_021833577	Ubiquitin-conjugating enzyme E2 28	GTAAGTAGAAATACACCCCCCG	GCTTCCAAGCGTATTTTGAAG
RG0099	scaffold0099_128107	XM_021812245	Ubiquitin-fold modifier-conjugating enzyme 1	CCGTTACTTATCCATCTACTGCTC	CAAAGTGGCTTGAAATGAACC
RG0354	scaffold0354_501312	XM_021827483	Ubiquitin-like protein ATG12	AGCATAATTGACCAGCAGTTTACC	ACTGCCGATGCTCCTATACTC
RG2190	scaffold2190_4544	XR_002493952	F-box protein SKIP19/Ubiquitin-protein ligase E3	GTCAATATGGACGGACATTTAGGG	TCAGATGATCCCACGTCAGC
RG0782	scaffold0782_240563	XM_021787005	F-box protein SKIP8/Ubiquitin-protein ligase E3	GCCATTGTGAATGCCAATACAG	CCTTGATCCCAATTAAATGCGA
RG1376	scaffold1376_117308	XM_021800037	F-box/LRR-repeat protein 14 /Ubiquitin-protein ligase E3	CTTTGCTCAGCTCCGTATCTG	ACAAGAGGCTTTAACCCATCA
RG0118	scaffold0118_304761	XM_021814014	NEP1-interacting protein-like 2/Ubiquitin-protein ligase E3	ACCATCTGCTGTACAATATGC	ATGATCCATGTCTTAGGAGCC
RG0677	scaffold0677_270674	XR_002491008	Histone-lysine N-methyltransferase, H3 lysine-9 specific SUVH3	AGAGTTGACTTATGACTATGGGACT	ACCAAAACAACCTCGGCATC
RG0544	scaffold0544_697128	XM_021836138	Pentatricopeptide repeat-containing protein (unkown)	TGTCTCCTGATAGAAGATGTTGTG	ATGTGCTCTGACAGTTTGTGG
RG1206	scaffold1206_49642	XM_021797646	Protein CURVATURE THYLAKOID 1A	AACTCAGTACCTTTGCTGCC	ATATCTGTTGCTAATTCCTTTC
RG0427	scaffold0427_670390	XM_021831455	Bis(5'-adenosyl)-	GCACAGAAAATTGGCAGTCG	ATGTGGATATGAACATGGGGC

triphosphatase

eIF1Aa	Previous reference genes	HQ268022	Eukaryotic translation initiation factor 1A	GCGTGACTATCAGGACGACAA	CAAGACCTCCAGCAATACCCTC
RH8		HQ323244	DEAD/DEAH box helicase	TCACAGGGTTGGTAGATCAG	CCAAGCTCTTGCTCAATCC
UBC2A		HQ323246	Ubiquitin-protein ligase	CATTTATGCGGATGGAAGCA	CAGGGGAGTTTGGATTGGA
UBC2B		HQ323247	Ubiquitin-protein ligase	CGACCAAGTTTTTCATTTGCGGGTG	AGTCTCTTCTTTGCTGGGGTT
RG1588154	Specific genes	XM_021813008	ATP citrate lyase	CACCAAGTGAGAGAATTGCAGATG	GCCCTCAGCCTATCAACAGTA
RG65560		XM_021836811	ATP citrate lyase	GCCTGAAGCAACAATGACCG	TAGATGGTTATCAGGAGTCAG

Table 2 Expression profile of candidate reference genes

Gene Name	n	Geo Mean [Ct]	AR Mean [Ct]	min [Ct]	max [Ct]	std dev [+/- Ct]	CV [% Ct]
RG1293	60	24.84	24.93	21.01	32.24	2.08	8.34
RG1039	60	27.12	27.16	24.38	32.18	1.62	5.95
RG2011	60	29.33	29.42	25.2	34.26	2.29	7.78
RG0347	60	26.90	26.98	22.41	31.21	2.06	7.65
RG1200	60	26.93	27.01	23.06	32.52	2.11	7.79
RG0140	60	23.62	23.71	19.56	28.69	2.05	8.63
RG0239	60	29.21	29.27	25.11	34.20	1.85	6.32
RG1121	60	26.84	26.92	22.84	31.96	2.04	7.59
RG0623	60	27.33	27.41	22.95	32.09	2.13	7.77
RG1478	60	29.75	29.8	26.76	34.49	1.78	5.99
RG0520	60	28.69	28.78	24.63	34.37	2.22	7.71
RG0928	60	26.36	26.42	22.97	30.91	1.79	6.77
RG1690	60	28.67	28.75	25.51	35.60	2.08	7.23
RG0470	60	24.93	25.02	21.07	31.17	2.15	8.59
RG0099	60	27.03	27.11	23.20	32.53	2.02	7.44
RG0354	60	28.87	28.95	24.68	35.48	2.23	7.69
RG2190	60	28.48	28.55	24.77	33.28	2.06	7.23
RG0782	60	30.93	31.04	26.50	37.88	2.60	8.37
RG1376	60	29.69	29.75	26.10	34.51	1.96	6.58
RG0118	60	27.65	27.71	24.12	32.44	1.90	6.86
RG0677	60	27.53	27.6	23.75	32.39	1.99	7.20
RG0544	60	31.67	31.73	28.29	36.97	1.89	5.96
RG1206	60	23.81	23.92	19.86	33.69	2.39	10.01
RG0427	60	25.95	26.01	22.72	31.02	1.79	6.86
eIF1Aa	60	25.94	26.03	21.54	30.83	2.18	8.39
RH8	60	27.84	27.95	23.55	34.03	2.46	8.81
UBC2A	60	26.32	26.44	21.02	31.66	2.54	9.59
UBC2B	60	26.96	27.01	23.54	30.82	1.73	6.40
RG1588154	60	27.95	28.12	21.79	34.20	3.01	10.72
RG65560	60	28.84	29.03	21.81	35.09	3.24	11.15

Table 3 Comprehensive ranking of candidate reference genes according to Delta Ct, Normfinder and Genorm

Gene	Ranking of candidate reference genes				
	Four experiments	Disease experiment	Low temperature experiment	High temperature experiment	Development experiment
RG0099	1	6	3	2	6
RG0427	2	15	2	9	3
RG0928	3	9	8	7	16
RG2190	4	3	12	21	13
RG0118	5	5	15	8	12
RG0677	6	1	17	23	10
eIF1Aa	7	18	14	20	2
RG1121	8	12	5	25	5
RG0623	9	4	20	1	23
RG0354	10	8	26	17	8
RG0347	11	14	25	11	9
RG0140	12	20	10	24	7
RG1200	13	7	4	26	11
RG0520	14	10	27	12	19
RG0470	15	29	21	5	14
UBC2B	16	24	22	13	22
RG1039	17	2	6	6	25
RG1690	18	22	28	16	17
RG2011	19	11	24	18	21
UBC2A	20	16	1	15	18
RG1376	21	26	7	19	24
RG0782	22	19	23	28	1
RGRH8	23	13	19	27	4
RG0239	24	23	30	29	20
RG1293	25	25	9	4	28
RG1478	26	28	18	22	26
RG1588154	27	17	11	10	15
RG0544	28	27	29	30	27
RG1206	29	30	13	3	30
RG65560	30	21	16	14	29

Supplemental File Legends

Supplemental Table Legends

Table S1 Information of selected 82 candidate reference genes from two transcriptome sequencings

Supplemental Figure Legends

Fig.S1 Expression profile of 82 candidate reference genes under develoment stage and temperture stress of leaf.

Fig.S2 Cluster of four experiments according to ranking of expression stability.

Figures

Figure 1

Functional distribution of the 82 candidate reference genes as classified using the KEGG database.

Figure 2

Functional distribution of the 82 candidate reference genes as classified using the KEGG database.

Figure 3

Comparison of the ranking results from Delta Ct, Bestkeeper, NormFinder and GeNorm. The correlation were evaluated for the ranking results of 30 candidate reference genes in all samples, by comparative delta Ct method, Bestkeeper, NormFinder and GeNorm.

A

B

Figure 4

Venn diagram illustrating shared and unique number of candidate reference genes in four experiments. A. Ten most stable reference genes; B. Ten least stable reference genes; H,high temperature stress;L,low temperature stress;De,developmental stages;Di,disease stress.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Fig.S1.tif](#)
- [Table.S1.xlsx](#)
- [Fig.S2.tif](#)