Reference Gene Selection for RT-qPCR Assays in Different Tissues of *Huperzia Serrata* Based on Full-Length Transcriptome Sequencing

Yanping Fu
Northwest University https://orcid.org/0000-0001-6672-1693
Fei Niu
Northwest University
Hui Jia
Northwest University
Yanli Wang
Northwest University
Yana Feng
Northwest University
Bin Guo
Northwest University
Wei He
Northwest University
Yahui Wei (weiyahui@nwu.edu.cn)
Northwest University https://orcid.org/0000-0002-1473-3585

Research

Keywords: Huperzia serrata, Huperzine A, RT-qPCR, Reference Gene

Posted Date: June 8th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-567251/v1

Abstract

Background

Huperzia serrata (*H. serrata*) produces various types of effectively lycopodium alkaloids, especially Huperzine A (HupA), which is a promising drug for the treatment of Alzheimer's disease. Numerous studies focused on the chemistry, bioactivities, toxicology, and clinical trials of HupA, however, the public genomic and transcriptomic resources are very limited for *H. serrata* research, especially for the selection of optimum reference genes.

Results

Based on the full-length transcriptome datasets and previous studies, thirteen candidate reference genes were selected in different tissue of *H. serrata*. Then, two optimal reference genes *GAPDHB* and *HisH2A* were confirmed by various analysis softwares. In order to further verify the accuracy of the two reference genes, they were used to analyze the expression patterns of four HupA-biosynthetic related genes (lysine decarboxylas, RS-norcoclaurine 6-O-methyltransferase, cytochrome P45072A1 and copper amine oxidase). The data suggested that the expression trends of HupA-biosynthetic related genes were consistent with them in transcriptome sequencing in different tissue of *H. serrata*.

Conclusions

This study screened the best reference genes *GAPDHB* and *HisH2A* in different tissues of *H. serrata*, which provides suitable normalization for analyzing the expression of HupA-biosynthetic gene in transcriptional level in *H. serrata.*

Background

Huperzia serrata (*H. serrata*) belongs to the *Huperzia* genus, Lycopodiaceae order. The whole plant of *H. serrata* has been used as a medicine in China to treat different kinds of ailments, including bruises, strains, swelling, rheumatism, schizophrenia, myasthenia gravis, and fever since 739 (during the Tang Dynasty) [10]. *H. serrata* has been widely known as a medicinal plant since Chinese scientists discovered Huperzine A (HupA) from it during the 1980s [15]. HupA is a promising candidate drug for treating Alzheimer's disease (AD), it could improve cognitive function, daily living activity, and global clinical assessment in patients with AD disease, with relatively few and mild adverse effects [34, 25]. However, *H. serrata* is scarce in nature and grows very slowly in specialized habitats. Furthermore, the HupA content is very low in *H. serrata* [20]. At present, the rapidly growing demand has put *H. serrata* resources on the brink of extinction. Although, a lot of efforts have been focused on artificial culture and tissues culture for *H. serrata* production, the results were unsatisfactory.

Now, researchers try to improve HupA content by studying the gene information of HupA biosynthesis. However, the public genomic and transcriptomic resources are very limited. Only two papers focused on transcriptomic resources [2, 36]. Real-time quantitative PCR (RT-qPCR) has been widely used in gene expression measurement in transcriptional level. Identification of suitable reference genes (RGs) is pre-requisite for RT-qPCR assays [29, 3]. Many housekeeping genes have been used as RGs under different experimental conditions, such as actin, tubulin, elongation factor (*EF*), 18S ribosomal RNA (*18S* rRNA), glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*), histone, ubiquitin and so on [21, 9]. However, there is no RG suitable to all biological systems because all previous studies have revealed alterations for housekeeping genes expression. So, researchers reach a consensus that specific RG for a given species and treatment needs to be identified firstly [3]. Unfortunately, the reportorial RGs are not suitable for the research of HupA-biosynthetic genes [35].

In present study, in order to obtain the optimal RGs for researching HupA-biosynthetic, we detected the concentration of HupA and carried out full-length transcriptome sequencing for the different tissues of *H. serrata*. Based on full-length transcriptome sequencing data and previous studies, thirteen candidate RGs were selected. Finally, *GAPDHB* and *HisH2A* stood out among the thirteen candidate RGs, and became the best combination for normalization in different tissues of *H. serrata* by the comprehensive analysis of four softwares. Using *GAPDHB* and *HisH2A* as RGs, the consistent expression trend was obtained between transcriptome sequencing and RT-qPCR when researched the expression profile of four HupA-biosynthetic genes, lysine decarboxylase (*LDC*), RS-norcoclaurine 6-O-methyltransferase (*MET*), cytochrome P45072A1 (*CYP*) and copper amine oxidase (*CAO*). This further verified that *GAPDHB* and *HisH2A* were suitable for HupA-biosynthetic gene expression normalization. This work provides suitable RGs for the subsequent research of HupA-biosynthetic in *H. serrata*.

Methods

Plant Materials

H. serrata plants were collected from Hanzhong, Shaanxi, China (107°09'/32°30') in March 2018. All materials used in this study were identified by phytotaxonomist. The plants were rinsed carefully by running water. Root, stem, and leaves were collected in liquid nitrogen and were immediately frozen at -80°C for RNA extraction. Other plant materials were dried at 60°C and powdered for determining HupA content.

HPLC parameters and conditions

HupA were extracted from the different plant tissues as previously described [19, 12]. After the plant material was dried and milled, 100 mg each of powdered plant tissues was extracted by adding 2% (2:100, w/v) aqueous tartaric acid (5 ml) for overnight and then sonicating for 2 h at 25°C. Centrifuging for 30 min at RT and the upper extraction solution was filtered into a 1.5 mL measuring flask through a 0.45 µm filter. Finally, the filtered solutions (10 µL) were injected into the HPLC system (LC-20AT, Shimadzu, Japan) for detection of HupA content. Each experiment comprised three independent biological replicates. The details are as follows: The elution conditions: (Flow rate, 0.8 mL / min; Column temperature, 28°C; Injection volume, 10 µL; Detection, the detection was performed at the wavelength of 308 nm). The mobile phase was methanol / acetonitrile / 0.08 M ammonium acetate (pH = 6) (10:30:60). Chromatography was performed on a C-18 column (Hypersil ODS2, China) of 250 × 4.6 mm dimensions and 5 µm particle size.

Candidate RGs selection and primers design

RGs for this study were selected from full-length transcriptome sequencing dates and previous studies where these were found to have stable expression in other plants by using local NCBI-blast (version 2.4.0+). Thirteen candidate RGs (MH560040 - MH560049, MZ042627-MZ042629) were selected (**Table 1**). Gene specific primers for each RG were designed using the Primer 5.0. Conserved domains of RGs were evaluated and the primer binding positions were presented. Initially, primer specificity was verified by RT-qPCR and confirmed with 2% (w/v) agarose gel electrophoresis and melting curve.

RNA extraction and cDNA synthesis

The total RNA was extracted according to the modified CTAB method [11]. RNA samples were treated with DNase I (Ambion, Waltham, MA, USA) to remove any DNA contamination. Using cDNA synthesis kit (Roche, Basel, Switzerland), first strand cDNA was prepared with 3 µg RNA as manufacturer's instructions.

RT-qPCR analysis

The RT-qPCR reactions were performed with FastStart Universial SYBR GreenMaster (Roche, Basel, Switzerland) on a CFX-96 thermocycling system (Bio-Rad, Hercules, CA, USA). Each RT-qPCR reaction was performed as described previously [24]. PCR amplifications were carried out with the following cycling conditions: one cycle at 95°C (180 s), followed by 40 cycles of denaturation at 95°C (30 s), annealing at 58°C for 10 s and extension at 72°C for 20 s [16]. Finally, melt curve analyses were done by slowly heating the PCR mixtures from 58 to 95°C. Amplification efficiencies (E) and correlation coefficients (R^2) for each primer pair were calculated by LinRegPCR program [26]. In the negative control group, RT-qPCR was performed using water instead of cDNA as the template. Three technical replicates were analyzed for each biological sample, and each experiment comprised three independent biological replicates.

Data analysis of gene expression stability

Three software algorithms: geNorm [29], NormFinder [1], BestKeeper [23] and RefFinder [27] were applied to determine the stability of RGs. For geNorm and NormFinder, the raw Ct values were converted into the relative quantities using the formula $2^{-\Delta Ct}$ (ΔCt = each corresponding Ct value - lowest Ct value). M value, which respents average expression stability, was calculated in geNorm. The candidate RGs showing a higher M value (M>1.5) are not considered for normalization studies [29]. geNorm software was also used to confirm the numbers of RGs with pair wise variation (Vn/Vn+1, n refers to the RGs number) [29]. NormFinder provides the stability value for each gene,

which is a direct measure of the estimated expression variation enabling evaluation of the systematic error introduced when using the gene for normalization [1]. For BestKeeper, the raw Ct values and amplification efficiencies were used to calculate the coefficient of variation (CV) and standard deviation (SD). The most stable genes are determined to be those which exhibit the lowest CV and SD (CV ± SD). The comprehensive ranking order was recommended on the basis of geometric mean (GM) by RefFinder [37].

Validation of RGs

The primer of four Hup A-biosynthetic genes *LDC* (GO914645) *MET* (GO914756), *CYP* (GO914428) and *CAO* (JN247732) [17, 33, 28] were designed using the Primer 5.0. The combination of the top two best ranked RGs and worst ranked RGs were used to standardize the expression of two target genes. The target gene expression data was normalized using the geometric mean values calculated for the RG pairs [29]. Relative expression level and fold change were determined using the comparative $2^{-\Delta\Delta Ct}$ method [22]. One-way analysis of variance was performed using SPSS software (Version 16.0, SPSS Inc., Chicago, IL, USA).

Results And Discussion

The Hup A content analyses

HPLC–UV was performed to detect the HupA content in *H. serrata*. Typical chromatograms from HupA standard and three tested samples are shown in **Figure S1**, indicating that HupA has good peak shape and is well separated from different tissues. A linear relationship exists between the peak area (measured at 308 nm) and the concentration of HupA in the sample injected into the HPLC. The results showed that there was obvious difference for the HupA content in different tissues. The highest HupA content (72 μg/g) was found in the leaves of *H. serrata*. The lowest content (19 μg/g) of HupA was found in root tissues of *H. serrata* (**Figure 1**).

The screening of candidate RGs

In consideration of the varied difference of Hup A concentrations in different tissues, the root, stem and leaf samples were collected and proceeded the full-length transcriptome sequencing by Nanopore. After assembly, 43,443 unigenes were retrieved. CPM（counts per million） is the index for measuring the expression of unigenes. Based on the CPM value and reported literatures, ten traditional RGs and three new RGs candidate were choosed. The three new RGs candidate had stable expression in full-length transcriptome sequencing. They were annotated hypothetical or uncharacterized proteins by NCBI Nr database, furthermore, they were not used as RGs before. The three new RGs candidate as following: ONT.10684 represented the high expression level (CPM over 100), EVM0022608 was the middle level (CPM 29-34), and EVM0017784 was the low level (CPM less than 5). The detail information of total thirteen candidate RGs was showed in Table 1.

Verification of the primer specificity and RT-qPCR amplification efficiency

The primer information of thirteen candidate RGs was given in **Table 2**. Each primer pair was designed except the conserved domains to ensure the specificity (**Figure S2**). Initially, the agarose gel electrophoresis yielded a specific fragment of expected size (**Figure S3A, S3B and S3C**). Further, the melting curve analysis in the RT-qPCR reaction showed the single peak for each primer pair indicating an absence of non-specific product amplification (**Figure S4**). For all primer pairs, the amplification efficiencies were spanning from 90.4% to 103.6%, and the correlation coefficient (R^2) were greater than 0.990 (**Table 2**). Taken together, these results indicated each primer pair was specificity and the RT-qPCR assays were highly efficient.

Expression profiles of candidate RGs

The expression profiles of RT-qPCR products for all experimental samples are shown in **Figure2**. The results illustrated that the mean Ct values of all RGs ranged from 24.04 to 29.43. Lower Ct value indicates the higher expression abundance, conversely means the lower expression profiles. *EF1dt* and *UBQ1* were highly expressed with mean Ct values between 24.04 and 24.08 while *EFTS* was the least expressed gene on account of its highest mean Ct value (29.43). All candidate genes showed expression variability in different samples as evident from a wide range of Ct values. Genes such as *GAPDHB* and *EFTS* showed relatively smaller variation (< 2 cycles), while others like *UBQ11* had the highest expression variation (3.07 cycles). The results indicated that there was still variable expression even for relative stable housekeeping genes.

geNorm analysis

To identify the most stable RG, geNorm algorithm calculated the average expression stability values (M values) of each RG. As **Figure3** shown, each M value was less than 1.5, which suggested the appropriateness of all RGs for normalization consideration in different tissues of *H. serrata*. Concretely analyzing, *EF1dt*, *HisH2A* and *GAPDHB* were the most stable genes in each *H. serrata* samples, while *HisH3.3* and *EFTS* were the least stable genes in each *H. serrata* samples.

NormFinder analysis

NormFinder evaluates each RG according to the stability value. Lower stability value indicates more stable gene expression, and vice versa. As shown in **Table 3**, *GAPDHB* and *HisH2A* were obviously stable in all samples, and *EFTS* (highest stability value = 0.210) was the least stable gene. For the root samples, *HisH2A* and *a-tub3* were most stable, and the *EFTS* (stability value = 0.506) still was the least stable gene. Whereas *GAPDHB* and *HisH2A* were the most stable gene and *HisH3.3* (stability value = 0.358) was the least one in stem. In leaf tissues, the most stable RGs were *EF1dt*, *HisH2A* and, *GAPDHB*, meanwhile, the least stable RG was *Actin7* (stability value = 0.364). Overall, with NormFinder analysis, *GAPDHB*, *HisH2A* and *EF1dt* were the most stable genes, while *EFTS* and *HisH3.3* were the least stable genes in different tissues of *H. serrata*.

BestKeeper analysis

The stability standard deviation (SD) and its relationship to the BestKeeper index were considered as two important evaluation criteria in BestKeeper analysis [23]. The results showed that each RG had a SD value < 1.0, which indicated that the candidate RGs were relatively stable for RT-qPCR normalization. In present, *GAPDHB*, *EF1dt* and *HisH2A* were the top three ranked genes with lowest CV ± SD values in all samples, stem and leaf tissues (**Table 4**). In root samples, the top three ranked genes were *HisH2A*, *UBQ11* and *a-tublin*. Rather, *HisH3.3* was deemed to the least RG with the highest CV ± SD value (27.35 ±0.21 and 28.66 ±0.20) in all samples and stem tissues, while *EFTS* (in root tissues) and *Actin7* (in leaf samples) showed the least stable expression. Taken together, with BestKeeper analysis, *GAPDHB*, *EF1dt* and *HisH2A* were the most stable genes, while *HisH3.3* and *EFTS* were the least stable genes in different tissues of *H. serrata*.

RefFinder analysis

Athough the results (geNorm analysis, NormFinder and BestKeeper analysis) were similar, it was not strictly consistent. Therefore, we performed overall evaluate using RefFinder to recommend a comprehensive ranking of the most stable genes in diverse tissues (**Table 5**). In root tissues, the final ranking calculations based on the RefFinder found *HisH2A* (GM = 1.67), *GAPDHB* (3.33) and *a-tub3* (3.33) were the best genes. For stem samples, the top three stable RGs were *GAPDHB* (1.67), *EF1dt* (2.00) and *HisH2A* (3.67), while *EF1dt* (1.33), *HisH2A* (2.00) and *GAPDHB* (2.33) in leaf samples. Across all samples, the top three stable RGs were *GAPDHB* (1.67), *HisH2A* (1.67) and *EF1dt* (2.33). On the other hand, *HisH3.3* and *EFTS* were ranked as the two least stable genes (**Table 5**).

Optimal Number of RGs for Normalization

Though a single and stable RG is sufficient for quantifying gene expression, the use of more than one RG for effective normalization of gene expression data is suggested [29]. Based on the geNorm software, the optimal number of RGs needed for normalization was determined by pairwise variation (Vn/n+1). In our data, the all pairwise variation of V2/3 values were lower than 0.15 (**Figure 3**), which suggested that the combination of the two most stable RGs was optimum for normalization.

Together with RefFinder analysis, *GAPDHB* and *HisH2A* were the best combination for normalization in different tissues of *H. serrata*.

RG Validation

To demonstrate the utility of identified stable RGs, four HupA-biosynthetic related genes *LDC*, *MET CYP* and *CAO* were selected in *H. serrata*. For the purpose of comparison, expression values of target genes were normalized with respect to the most stable gene pair (*GAPDHB* and *HisH2A*) and the least stable gene pair (*EFTS* and *HisH3.3*) in *H. serrata* different tissues. When normalized using the most stable genes, the transcription levels of *LDC*, *MET*, *CYP* and *CAO* were (over 2 –fold) in the tissues of leaf, stem and root were compared, and the expression trend was consistent with that of transcriptome sequencing dates (**Figure 4)**. By contrast, when normalized using the

least stable genes, the transcription level of *MET* and *CYP* were not up-regulated (less than 2 –fold) in stem and leaf tissues. The transcription level of *LDC* was down-regulated (0.77 –fold) in stem tissues, and the *CAO* was down-regulated (0.67 –fold) in leaf tissues. This expression trend was not consistent with that of transcriptome sequencing dates. In all, the expression of the most stable gene pair was more reliable than the least stable gene pair.

Discussion

H. serrata has received extensive concern due to produce biologically active lycopodium alkaloids, especially HupA [7]. HupA was found to possess potent acetylcholine esterase inhibition (AChEI) and had been clinically exploited for the treatment of Alzheimer's disease. More studies are focused on the isolation and identification of compounds and endophytic bacteria [31], but little on the transcriptional level. Especially for the selection of optimum reference genes, little research has been reported [35]. So, in present study, we screened and selected the optimal RGs based on full-length transcriptome sequencing and previous researches in different tissues of *H. serrata*. By the analysis of four softwares and the verification of four HupA-biosynthetic genes (*LDC*, *MET*, *CYP* and *CAO*), we obtained two optimal reference genes *GAPDHB* and *HisH2A* for studying HupA-biosynthetic related genes. This study provides suitable normalization for analyzing the expression of HupA-biosynthetic gene. In addition, we found the expression trend of HupA-biosynthetic genes were similar with the trend of HupA content in different tissues of *H. serrata*. This result will provide the information for further studying the biosynthesis and transportation of HupA.

In general, the expression level of RGs should be constant stable in any physical conditions. However, there is no RG suitable to all biological systems. We had to screen the most suitable RGs for studying the HupA-biosynthetic. Based on the CPM value of transcriptome sequencing and reported literatures [32, 8, 14, 5], ten traditional RGs and three new RGs candidate were chosen (Table 1**)**. Especially for three new RGs candidate, they had stable expression in full-length transcriptome sequencing but they were not used as RGs before.

The primer specificity is the primary condition of RT-qPCR. The ideal primers which should cross intron regions to availably avoid genomic contamination in cDNA samples and cannot be set in conservative domain. Firstly, each primer pair was designed except the conserved domains (**Figure S2**). Subsequently, the products of each primer pair were detected by agarose gel electrophoresis (**Figure S3**) and melting curves (**Figure S4**). The results indicated that there were no primer dimers and non-specific amplification for each primer pair. Furthermore, the E value of PCR varied from 90.4–103.6%, and all of the R^2 were greater than 0.990 (Table 2), which were similar to previous literatures [16]. In conclusion, these results indicated each primer pair was specificity and the RT-qPCR assays were highly efficient.

Based on the analysis of four softwares (geNorm, NormFinder, BestKeeper and RefFinder), *GAPDHB* and *HisH2A* stood out among the thirteen candidate RGs, and became the best combination for normalization in different tissues of *H. serrata* (Table 5). Many studies have shown that *GAPDH* was most often as relatively stable internal RGs in different tissues and under a variety of experimental conditions [6, 24, 14]. Although histone and elongation factor were be reported as most stable RGs in other species [24], *HisH3.3* and *EFTs* were the most instable RGs in different tissues of *H. serrata.* These results suggested that the traditional RG may not suitable for all samples. Interestingly, the performance between the homologous genes was obviously different. *HisH3.3* and *EFTs* were more instable than *HisH2A* and *EF1dt*. Similar findings can be found that the expression level of *Actin2/7* was more stable than *Actin11* in diverse tissues of soybean [13]. These results also stated clearly that the expression level and stability of RGs from the same gene family may be different in the same samples. Taken together, the results further proved the necessity for screening suitable RGs in different tissues of *H. serrata*.

A proposed biosynthesis pathway for HupA and related lycopodium alkaloids was reported [18]. However, only two enzymes, *LDC* and *CAO* have been proved to participate in the biosynthesis of HupA [33, 4, 28]. Three enzymes *RS*-norcoclaurine 6-*O*-methyltransferase (*MET*) and cytochrome P45072A1 (*CYP*) [17, 33], type III polyketide synthase (*PKS*) [30], have been described to be possible involvement of the biosynthesis of HupA. In order to verify the accuracy of the stable RGs identified in this paper, four HupA-biosynthetic genes, *LDC*, *MET*, *CYP* and *CAO* were selected. The results showed that when the combination of stable RGs (*GAPDHB* and *HisH2A*) was used, the consistent expressions trend of *LDC, MET, CYP and CAO* were obtained between transcriptome sequencing and RT-qPCR (Fig. 5**and** Table 1). Conversely, the use of the most unstable RGs (*HisH3.3* and *EFTS*) may lead to declinational results (Fig. 5**and** Table 1). The results further verified that *GAPDHB* and *HisH2A* were suitable for gene expression normalization, especially for HupA-biosynthetic genes. In addition, we tested the content of HupA in different tissues. The results indicated the content of HupA in root was obviously lower than that in stem or leaf (Fig. 1), which was consistent with the previous study [31, 18]. Therefore, in order to protect the wild resources, we suggested picking the aboveground parts instead of uprooting the whole plan when digging *H. serrata*. In addition, we found the expression trend of HupA-biosynthetic genes were similar with the trend of HupA content in different tissue of *H. serrata*, which indicated that the biosynthesis of HupA may be in stem and leaf. This result will provide the information for further studying the biosynthesis and transportation of HupA.

Conclusions

In the present study, based on full-length transcriptome sequencing data and the analysis of four softwares, we obtained two optimal reference genes *GAPDHB* and *HisH2A* from thirteen candidate reference genes in different tissue of *H. serrata*. The expression patterns of four HupA-biosynthetic related genes *LDC*, *MET*, *CYP* and *CAO* further verified that *GAPDHB* and *HisH2A* were suitable for gene expression normalization. This work provides suitable RGs for the subsequent research of HupA-biosynthetic and transportation in *H. serrata.*

Declarations

Acknowledgements

The authors are particularly grateful to Kai Yao, Youfeng Yang, Xin liu and other members of our laboratory for their effective comments in the process of experiment.

Authors' contributions

YF designed the experiment and analyzed data. FN performed the partial experiment, prepared the first draft. HJ, YW and YF contributed reagents/materials and detected the content of HupA. BG provided plant materials used for experiments and discussed data. WH revised the paper. YW edited the graphs and revised the paper. All the authors approved the final draft.

Funding

This work was supported by Scientific Research from Shaanxi Provincial Department of Education [16JK1756] and Natural Science Foundation of China (31702159) to YF, Key Research and Development Plan Project of Shaanxi Province (2018ZDXM-SF-016) to YW.

Availability of data and materials

All data generated or analysed during this study are included in this published article (and its Additional file 1).

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests

References

1. Andersen, C. L., Jensen, J. L. and Ørntoft, T. F. 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.
2. André, S., Sun, Y., Song, J., Wu, Q., Sun, C., Li, Y., Luo, H. and Chen, S. Comparison of 454-ESTs from *Huperzia serrata* and *Phlegmariurus carinatus* reveals putative genes involved in lycopodium alkaloid biosynthesis and developmental regulation. BMC Plant Biology 2010, 10(1):209.
3. Bansal, R., Mittapelly, P., Cassone, B. J., Mamidala, P., Redinbaugh, M. G. and Michel, A. Recommended Reference Genes for Quantitative PCR Analysis in Soybean Have Variable Stabilities during Diverse Biotic Stresses. Plos One 2015, 10(8):e0134890.
4. Bunsupa, S., Katayama, K., Ikeura, E., Oikawa, A., Toyooka, K., Saito, K. and Yamazaki, M. Lysine decarboxylase catalyzes the first step of quinolizidine alkaloid biosynthesis and coevolved with alkaloid production in leguminosae. Plant Cell 2012, 24(3):1202-1216.
5. Chen, C., Wu, J., Hua, Q., Tel-Zur, N. and Qin, Y. Identification of reliable reference genes for quantitative real-time PCR normalization in pitaya. Plant Methods 2019, 15(70).
6. Chen, X., Mao, Y., Huang, S., Ni, J. and Wu, L. Selection of Suitable Reference Genes for Quantitative Real-time PCR in Sapium sebiferum. Frontiers in Plant Science 2017, 8.

7. Christenhusz, M. J. M., Zhang, X. C. and Schneider, H. A linear sequence of extant families and genera of lycophytes and ferns. Phytotaxa 2011, 19(1):7-54.
8. Dudziak, K., Sozoniuk, M., Szczerba, H., Kuzdraliński, A., Kowalczyk, K., Brner, A. and Nowak, M. Identification of stable reference genes for qPCR studies in common wheat (*Triticum aestivum* L.) seedlings under short-term drought stress. Plant Methods 2020, 16.
9. Fei, X., Shi, Q., Yang, T., Fei, Z. and Wei, A. Expression Stabilities of Ten Candidate Reference Genes for RT-qPCR in *Zanthoxylum bungeanum* Maxim. Molecules 2018, 23(4):802.
10. Ferreira, A., Rodrigues, M., Fortuna, A., Falcão, A. and Alves, G. Huperzine A from Huperzia serrata: a review of its sources, chemistry, pharmacology and toxicology. Phytochemistry Reviews 2016, 15:51-85.
11. Gasic, K., Hernandez, A. and Korban, S. S. RNA extraction from different apple tissues rich in polyphenols and polysaccharides for cDNA library construction. Plant Molecular Biology Reporter 2004, 22(4):437-438.
12. Ishiuchi, K. i., Park, J. J., Long, R. M. and Gang, D. R. Production of huperzine A and other Lycopodium alkaloids in Huperzia species grown under controlled conditions and in vitro. Phytochemistry 2013, 91:208-219.
13. Jian, B., Liu, B., Bi, Y., Hou, W., Wu, C. and Han, T. Validation of internal control for gene expression study in soybean by quantitative real-time PCR. BMC Molecular Biology 2008, 9:59.
14. Liang, WX, Zou, XX, Carballar-Lejarazu, LJ, Sun, WH, Yuan and XY. Selection and evaluation of reference genes for qRT-PCR analysis in *Euscaphis konishii* Hayata based on transcriptome data. Plant Methods 2018, 14(42).
15. Liu, J. S., Yu, C. M., Zhou, Y. Z., Han, Y. Y., Wu, F. W., Qi, B. F. and Zhu, Y. L. Study on the Chemistry of Huperzine A and B. Acta Chimica Sinica 1986, 44:1035-1040.
16. Liu, X., Guan, H., Song, M., Fu, Y., Han, X., Lei, M., Ren, J., Guo, B., He, W. and Wei, Y. Reference gene selection for qRT-PCR assays in*Stellera chamaejasme* subjected to abiotic stresses and hormone treatments based on transcriptome datasets. PeerJ 2018, 6.
17. Luo, H., Sun, C., Li, Y., Wu, Q., Song, J., Wang, D., Jia, X., Li, R. and Chen, S. Analysis of expressed sequence tags from the *Huperzia serrata* leaf for gene discovery in the areas of secondary metabolite biosynthesis and development regulation. Physiol Plant 2010, 139(1):1-12.
18. Ma, x. and Gang, D. The Lycopodium Alkaloids. Natural Product Reports 2005, 21(6):752-772.
19. Ma, X., Tan, C., Zhu, D. and Gang, D. R. Is there a better source of huperzine A than *Huperzia serrata*? Huperzine A content of Huperziaceae species in China. J Agric Food Chem 2005, 53(5):1393-1398.
20. Ma, X., Tan, C., Zhu, D., Gang, D. R. and Xiao, P. Huperzine A from Huperzia species-an ethnopharmacolgical review. Journal of Ethnopharmacology 2007, 113(1):15-34.
21. Nguyen, D. Q., Eamens, A. L. and Grof, C. P. L. Reference gene identification for reliable normalisation of quantitative RT-PCR data in *Setaria viridis*. Plant Methods 2018, 14(24).
22. Pfaffl, W. A new mathematical model for relative quantification in real-time RT–PCR. Nucleic Acids Research 2001, 29(9):2002-2007.
23. Pfaffl, W., Tichopad, A., Prgomet, C. and Neuvians, P. 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.
24. Plant MethodsZhuang, H., Fu, Y., He, W., Wang, L. and Wei, Y. Selection of appropriate reference genes for quantitative real-time PCR in *Oxytropis ochrocephala* Bunge using transcriptome datasets under abiotic stress treatments. Frontires in Plant Science 2015, 6:475.
25. Qian, Z. M. and Ke, Y. Huperzine A: Is it an Effective Disease-Modifying Drug for Alzheimer's Disease? Frontiers in Aging Neuroscience 2014, 6:216.
26. Ruijter, J. M., Ramakers, C., Hoogaars, W. M. H., Karlen, Y., Bakker, O., Hoff, M. J. B. V. D. and Moorman, A. F. M. Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. Nucleic Acids Research 2009, 37(6):e45.
27. Silver, N., Best, S., Jiang, J. and Thein, S. L. Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR. BMC Molecular Biology 2006, 7(1):33.
28. Sun, J., Morita, H., Chen, G., Noguchi, H. and Abe, I. Molecular cloning and characterization of copper amine oxidase from *Huperzia serrata*. Bioorganic & Medicinal Chemistry Letters 2012, 22(18):5784-5790.
29. Vandesompele, J., Preter, K. D., Pattyn, F., Poppe, B., Roy, N. V., Paepe, A. D. and Speleman, F. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biology 2002, 3(7):34.
30. Wang, J., Wang, X., Liu, X., Li, J., Shi, X., Song, Y., Zeng, K., Zhang, L., Tu, P. and Shi, S. Synthesis of Unnatural 2-Substituted Quinolones and 1,3-Diketones by a Member of Type III Polyketide Synthases from *Huperzia serrata*. Organic Letters 2016, 18(15):3550-3553.
31. Wu, S., Fan, Z. and Xiao, Y. Comprehensive relative quantitative metabolomics analysis of lycopodium alkaloids in different tissues of *Huperzia serrata*. Synthetic & Systems Biotechnology 2017, 56(4):698.

32. Wu, Y., Tian, Q., Huang, W., Liu, J., Xia, X., Yang, X. and Mou, H. Identification and evaluation of reference genes for quantitative real-time PCR analysis in *Passiflora edulis* under stem rot condition. Molecular Biology Reports 2020, 47:2951-2962.
33. Xu, B., Lei, L., Zhu, X., Zhou, Y. and Xiao, Y. Identification and characterization of L-lysine decarboxylase from *Huperzia serrata* and its role in the metabolic pathway of lycopodium alkaloid. Phytochemistry 2017, 136:23-30.
34. Yang, G., Wang, Y., Tian, J. and Liu, J. P. Huperzine A for Alzheimer's disease: a systematic review and meta-analysis of randomized clinical trials. Plos One 2013, 8(9):e74916.
35. Yang, M., Wu, S., You, W., Jaisi, A. and Xiao, Y. Selection of Reference Genes for Expression Analysis in Chinese Medicinal Herb *Huperzia serrata*. Frontiers in Pharmacology 2019, 10:44.
36. Yang, M., You, W., Wu, S., Fan, Z., Xu, B., Zhu, M., Li, X. and Xiao, Y. Global transcriptome analysis of *Huperzia serrata* and identification of critical genes involved in the biosynthesis of huperzine A. BMC Genomics 2017, 18(1):245.
37. Zhang, B., Liu, J., Yuan, G., Chen, X. and Gao, X. Selection and evaluation of potential reference genes for gene expression analysis in greenbug (*Schizaphis graminum* Rondani). Journal of Integrative Agriculture 2018, 17(9):2054-2065.

Tables

Table 1. Description of candidate RGs and target genes.

Unigene gene ID	Accession number	Length (bp)	Gene symbol	Gene name	Homolog locus	Root_ CPM	Stem_ CPM	Leaf_ CPM	E value[a]
EVM0025399	MH560040	1041	*Actin4*	actin related protein 4	KM496528	24.88	39.46	30.90	1e-147
EVM0014845	MH560041	1704	*Actin7*	actin-related protein 7-like	XM_024546532.1	20.90	24.57	31.48	0
EVM0006291	MH560042	2444	*EF1dt*	elongation factor 1-delta 2-like	XM_004507134.3	205.15	193.46	160.29	1e-78
EVM0021957	MH560043	3747	*EFTS*	elongation factor Ts family protein	NM_119050.3	1.79	4.28	2.53	0
EVM0027572	MH560044	2137	*a-tub3*	tubulin alpha-3 chain	XM_020240093.1	16.55	18.42	23.00	0
EVM0033890	MH560045	1698	*GAPDHB*	glyceraldehyde-3-phosphate dehydrogenase B	NM_001302308.1	36.73	44.16	35.66	0
EVM0008093	MH560046	1008	*HisH3.3*	histone H3.3 isoform X1	XM_017375185.1	77.28	62.50	75.11	2e-91
EVM0031551	MH560047	1098	*HisH2A*	histone H2A-III-like	XM_024521943.1	64.68	81.59	50.83	3e-37
EVM0033477	MH560048	906	*UBQ1*	ubiquitin-NEDD8-like protein RUB1	XM_007048546.2	65.34	89.70	52.61	2e-89
EVM0000551	MH560049	948	*UBQ11*	ubiquitin 11	NM_001203752.2	314.46	421.33	369.47	2e-73
ONT.10684	MZ042629	817	10684	uncharacterized protein	XP_010251846	117.85	116.38	111.58	2e-17
EVM0022608	MZ042627	2036	22608	hypothetical protein	PTQ36569	34.00	29.033	32.86	5e-99
EVM0017784	MZ042628	1605	17784	uncharacterized protein	XP_002972660	3.15	2.55	3.68	0
EVM0027909	GO914645	637	*LDC*	Lysine decarboxylase	GO914645	2.20	47.53	53.52	9e-128
EVM0017005	GO914756	1682	MET	(RS)-norcoclaurine 6-O-methyltransferase	GO914756	2.82	13.46	10.48	0
EVM0022835	GO914428	1975	CYP	Cytochrome P450 72A1	GO914428	0.30	21.94	30.97	0
EVM0024797	JN247732	2640	*CAO*	copper amine oxidase	JN247732	4.11	13.86	39.70	0

a. E value represents high homology of candidate RGs with stable reference genes in other plants using the local Blast program.

Table 2. Selected candidate RGs and target genes, primers, and amplicon characteristics.

Name	Sequence 5' to 3'	Amplicon size (bp)	Product Tm (℃)	E (%)*	R^{2*}
Actin4-RT-F	TGTCTCTAAAGTTTCTTGTAGCACC	174	77.5	99.6	0.997
Actin4-RT-R	GCACAGCGGACAAGACTCTG				
Actin7-RT-F	AACCCTTATCTGTAGGCTTCTTG	139	78.5-79	95.7	0.998
Actin7-RT-R	TCATACACTGCACGTCAGGTAG				
EF1dt-RT-F	GACTGAGCAAATAAGAGGGG	213	75.5-76	91.7	0.997
EF1dt-RT-R	CGATATTGCGGCTTTTAACA				
EFTS-RT-F	AAGTATCCTAATACAGGGTTGG	158	75-75.5	92.5	0.997
EFTS-RT-R	AAAATTTGGTCTATCGCG				
a-tub3-RT-F	AAATCCAAACAATATGTATGAACAA	210	80-80.5	94.8	0.997
a-tub3-RT-R	GCCAAGGGTTTCAATCTTCTA				
GAPDHB-RT-F	GCAAAGTATATGAAGATTAGGCTC	161	76.5-77.5	91.3	0.998
GAPDHB-RT-R	GCGTCCACCAACGAACA				
HisH3.3-RT-F	AACTTGGTCTTGCTATGAAACTAAC	212	77-77.5	90.4	0.996
HisH3.3-RT-R	CAAGTCAAGAAACTCAACACGA				
HisH2A-RT-F	CTTCCTGCTTTCATCACTTT	270	75.5-76	92.2	0.994
HisH2A-RT-R	GTCCCCAAGCCTTACATT				
UBQ1-RT-F	GTGGCGGTCACTTATAGAGAG	130	73-73.5	97.3	0.997
UBQ1-RT-R	CTTCTGTAGTTCTGACATCAGTAAA				
UBQ11-RT-F	TCTGAAATGTCGCTTATCCG	174	76.5-77	97.4	0.995
UBQ11-RT-R	TCTGTTGGCGTCATTTGTTAG				
10684-RT-F	GCGCTTGATAAGTCACATGCTAC	253	76	94.3	0.995
10684-RT-R	GAAAAAAAAGATTGCCATAATAAGG				
22608-RT-F	TGGCGAATTTAGAGGGCAAT	198	77.5	95.1	0.998
22608-RT-R	CCTCAAGCCCATTTTAATTTCTCT				
17784-RT-F	AATATGCCCACAGGGTCACC	212	75	93.8	0.996
17784-RT-R	GTTGACCACTTGGCTTCCTACC				
LDC-F	GTATAGCCAATTACTCTATCCTCC	155	73-73.5	91.6	0.997
LDC-R	GTAACACCCATCCATTGTAGC				
MET-F	GCTCATCTCTGGGACATGG	237	80.5	98.2	0.996
MET-R	ACCCGACACTGAATCCTCTAT				
CYP-F	AAGAGTCAGGCTCTACTGTGC	151	78-79	91.4	0.998
CYP-R	GCAAGGAAGAACGTCGAGA				
CAO-F	CAAAGCTTGGAATTACGCTT	174	81	94.7	0.997
CAO-R	TTATATGTCTTGGCTAGTGTAAATG				

* E represents amplification efficiencies and R^2 represents correlation coefficient of RGs, respectively.

Table 3. Expression stability of candidate RGs as calculated by Normfinder.

Rank	All		Root		Stem		Leaf	
	Gene name	Stability value	Gene name	Stability value	Gene name	Stability value	Gene name	Stability value
1	*GAPDHB*	0.030	*HisH2A*	0.042	*GAPDHB*	0.045	*EF1dt*	0.014
2	*HisH2A*	0.044	*a-tub3*	0.046	*HisH2A*	0.096	*HisH2A*	0.017
3	*EF1dt*	0.060	*GAPDHB*	0.086	*EF1dt*	0.131	*GAPDHB*	0.024
4	*22608*	0.069	*Actin4*	0.108	*22608*	0.156	*22608*	0.064
5	*a-tub3*	0.077	*22608*	0.108	*10684*	0.174	*UBQ1*	0.083
6	*10684*	0.077	*UBQ1*	0.109	*a-tub3*	0.189	*10684*	0.089
7	*Actin4*	0.078	*10684*	0.113	*UBQ11*	0.228	*17784*	0.089
8	*17784*	0.094	*UBQ11*	0.116	*17784*	0.241	*Actin4*	0.090
9	*UBQ1*	0.104	*17784*	0.151	*Actin7*	0.257	*a-tub3*	0.165
10	*UBQ11*	0.109	*EF1dt*	0.167	*Actin4*	0.280	*HisH3.3*	0.217
11	*Actin7*	0.160	*Actin7*	0.211	*EFTS*	0.281	*UBQ11*	0.224
12	*HisH3.3*	0.166	*HisH3.3*	0.287	*UBQ1*	0.348	*EFTS*	0.305
13	*EFTS*	0.210	*EFTS*	0.506	*HisH3.3*	0.358	*Actin7*	0.364

Table 4. Expression stability of candidate RGs as calculated by BestKeeper.

Rank	All			Root			Stem			Leaf		
	Gene	CV[b]	SD[b]	Gene	CV	SD	Gene	CV	SD	Gene	CV	SD
1	*GAPDHB*	7.48	0.07	*HisH2A*	4.99	0.05	*GAPDHB*	5.36	0.05	*GAPDHB*	8.38	0.07
2	*HisH2A*	9.71	0.09	*UBQ11*	6.14	0.06	*EF1dt*	10.46	0.09	*EF1dt*	9.59	0.09
3	*EF1dt*	11.83	0.10	*a-tub3*	7.96	0.07	*HisH2A*	10.83	0.09	*HisH2A*	10.19	0.09
4	*a-tub3*	12.44	0.11	*22608*	8.06	0.167	*a-tub3*	15.35	0.12	*a-tub3*	10.63	0.09
5	*22608*	12.67	0.24	*UBQ1*	8.28	0.07	*22608*	15.61	0.21	*22608*	10.69	0.23
6	*10684*	12.97	0.18	*GAPDHB*	8.57	0.08	*UBQ11*	16.02	0.14	*10684*	10.73	0.18
7	*UBQ1*	13.07	0.11	*10684*	9.17	0.21	*10684*	16.21	0.20	*Actin4*	10.78	0.10
8	*17784*	13.14	0.19	*17784*	9.84	0.19	*Actin7*	16.86	0.13	*17784*	11.68	0.22
9	*UBQ11*	13.19	0.11	*Actin4*	9.94	0.09	*17784*	16.95	0.17	*UBQ1*	12.17	0.10
10	*Actin4*	14.44	0.12	*Actin7*	13.63	0.11	*EFTS*	17.20	0.15	*EFTS*	14.85	0.13
11	*Actin7*	16.84	0.13	*EF1dt*	14.68	0.12	*Actin4*	18.36	0.14	*UBQ11*	16.04	0.13
12	*EFTS*	20.13	0.17	*HisH3.3*	25.81	0.19	*UBQ1*	20.46	0.16	*HisH3.3*	19.47	0.17
13	*HisH3.3*	27.35	0.21	*EFTS*	26.12	0.20	*HisH3.3*	28.66	0.20	*Actin7*	20.31	0.16

b. CV and SD represent standard deviation and coefficient of variation, respectively.

Table 5. Expression stability ranking of the ten candidate RGs by RefFinder.

Rank	All		Root		Stem		Leaf	
	Gene	GM[c]	Gene	GM	Gene	GM	Gene	GM
1	*GAPDHB*	1.67	*HisH2A*	1.67	*GAPDHB*	1.67	*EF1dt*	1.33
2	*HisH2A*	1.67	*GAPDHB*	3.33	*EF1dt*	2.00	*HisH2A*	2.00
3	*EF1dt*	2.33	*a-tub3*	3.33	*HisH2A*	3.67	*GAPDHB*	2.33
4	*a-tub3*	4.33	*22608*	4.00	*22608*	4.67	*22608*	5.00
5	*22608*	4.67	*UBQ11*	6.33	*a-tub3*	5.67	*a-tub3*	6.00
6	*10684*	6.67	*EF1dt*	7.33	*UBQ11*	5.67	*Actin4*	7.33
7	*17784*	8.00	*UBQ1*	7.33	*Actin7*	6.00	*UBQ1*	7.67
8	*UBQ1*	8.67	*17784*	7.67	*10684*	7.00	*10684*	7.67
9	*UBQ11*	8.67	*10684*	8.00	*17784*	8.33	*17784*	8.00
10	*Actin4*	9.33	*Actin4*	8.67	*Actin4*	10.33	*UBQ11*	10.00
11	*Actin7*	9.33	*Actin7*	9.67	*EFTS*	11	*Actin7*	10.00
12	*EFTS*	12.33	*EFTS*	11	*UBQ1*	11.67	*EFTS*	11.33
13	*HisH3.3*	12.67	*HisH3.3*	12	*HisH3.3*	13	*HisH3.3*	11.67

c. GM represents geometric mean.

Figures

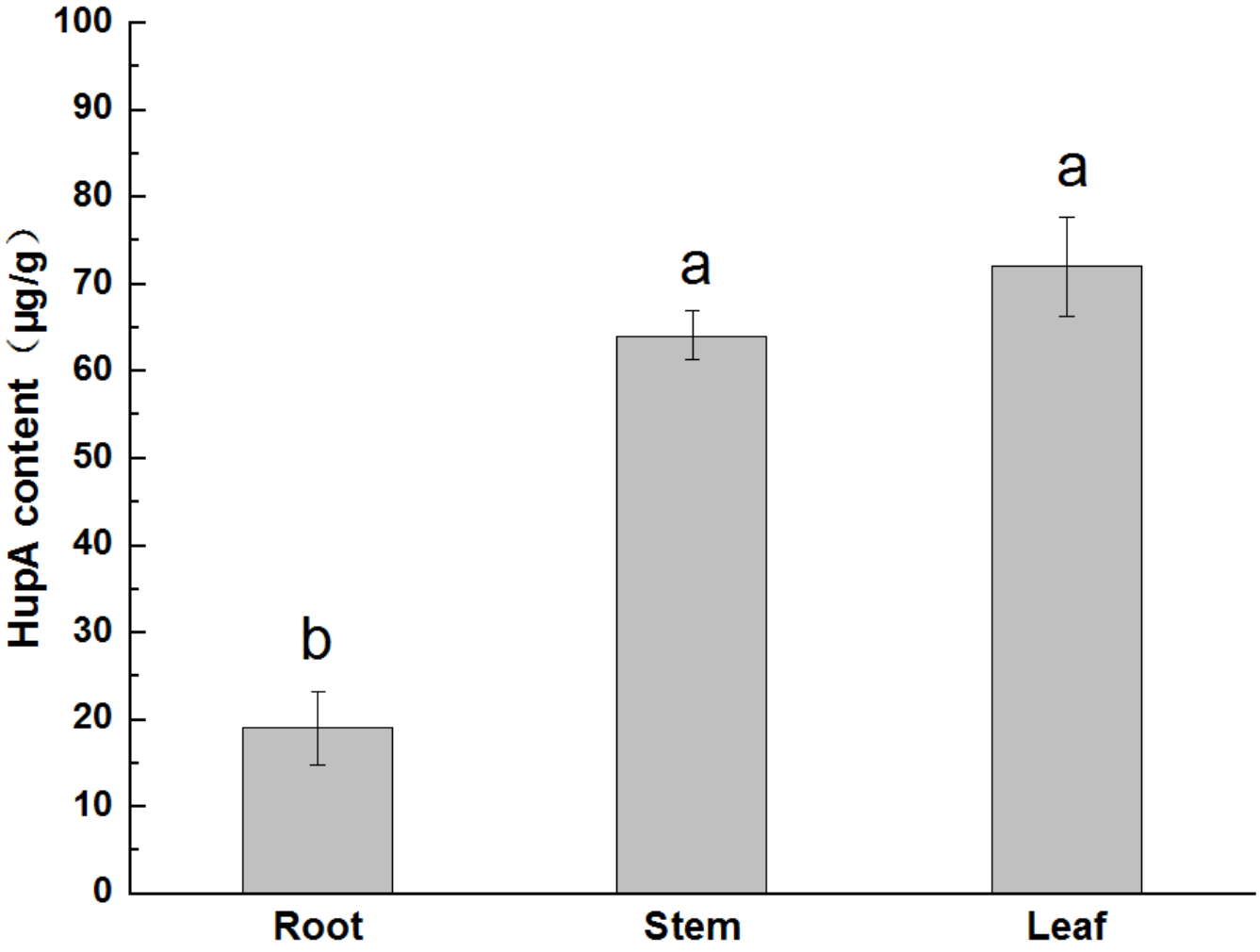

Figure 1

The content of Hup A in different tissues of H. serrata. The mean and standard deviation were calculated using the data from three independent biological replicates.

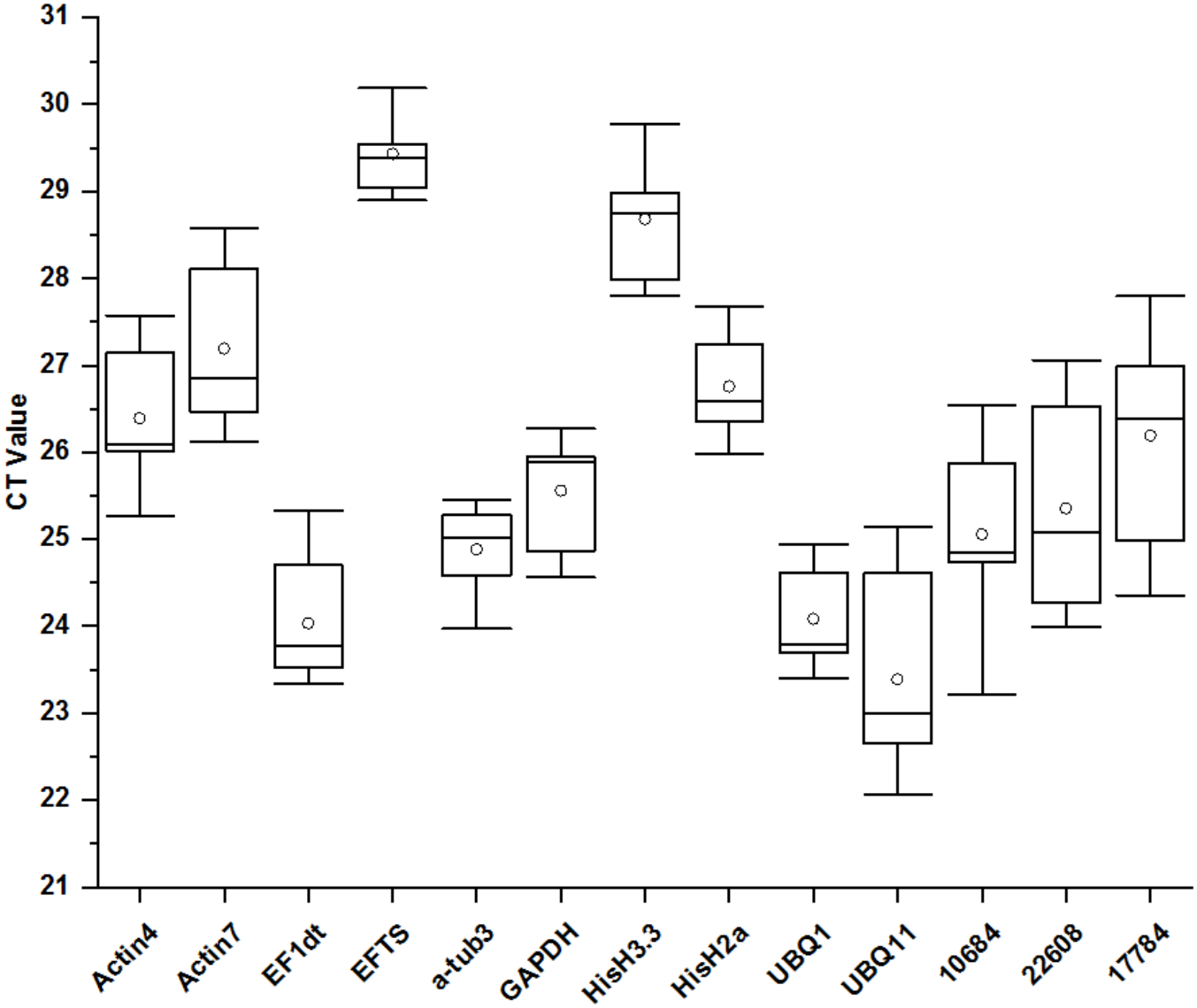

Figure 2

Distribution of Ct values for ten candidate reference genes in different tissues of H. serrata. Lines across the boxes denote the medians. The box represents the 25th and 75th percentile. The top and bottom whisker caps depict the maximum and minimum values, respectively. The white dots represent mean Ct values.

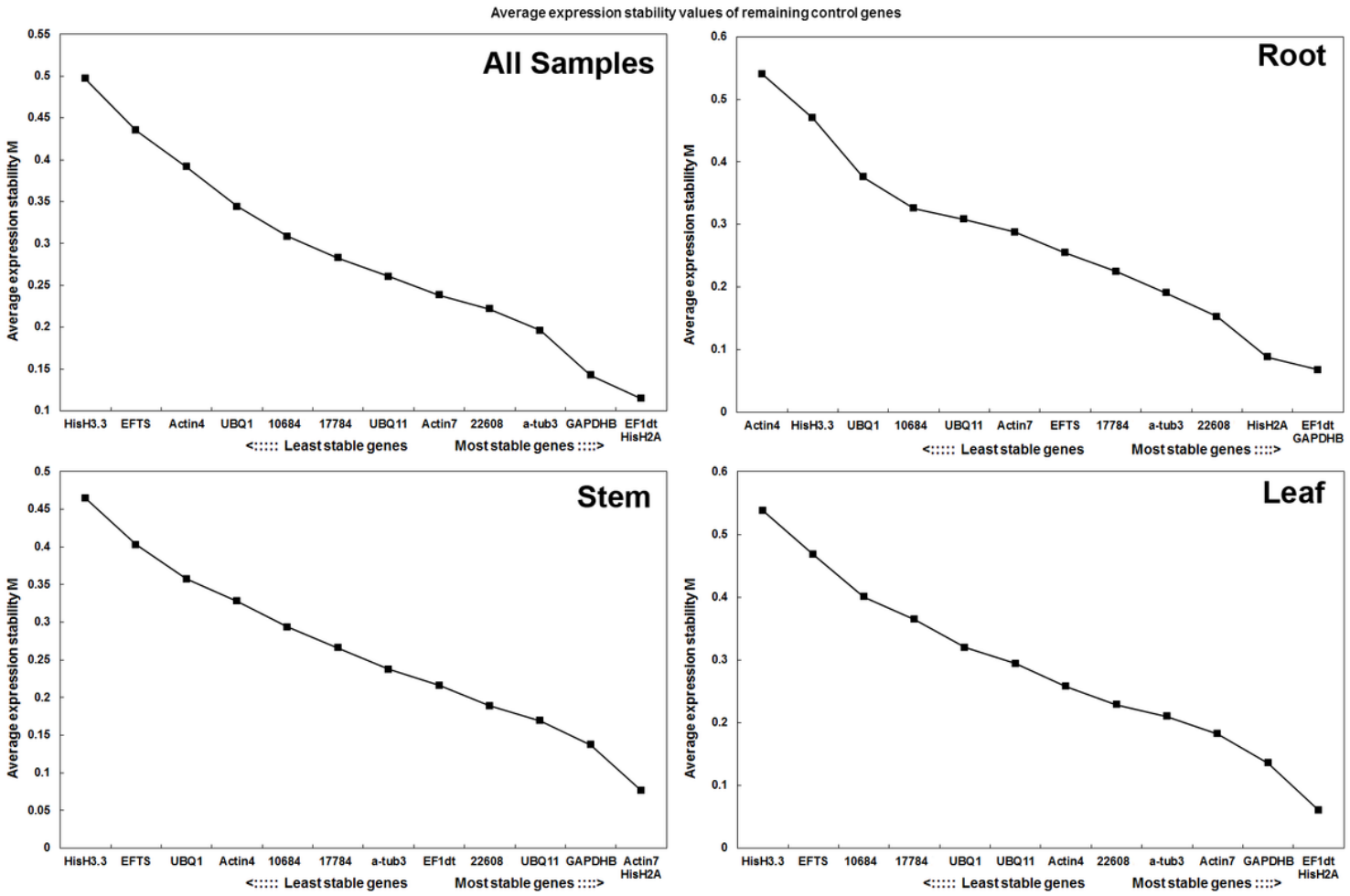

Figure 3

Average expression stability value (M) and ranking of the thirteen candidate reference genes analyzed by geNorm. (A) All samples.(B)Root. (C) Stem. (D)Leaf. The least stable genes are listed on the left, while the most stable genes are exhibited on the right.

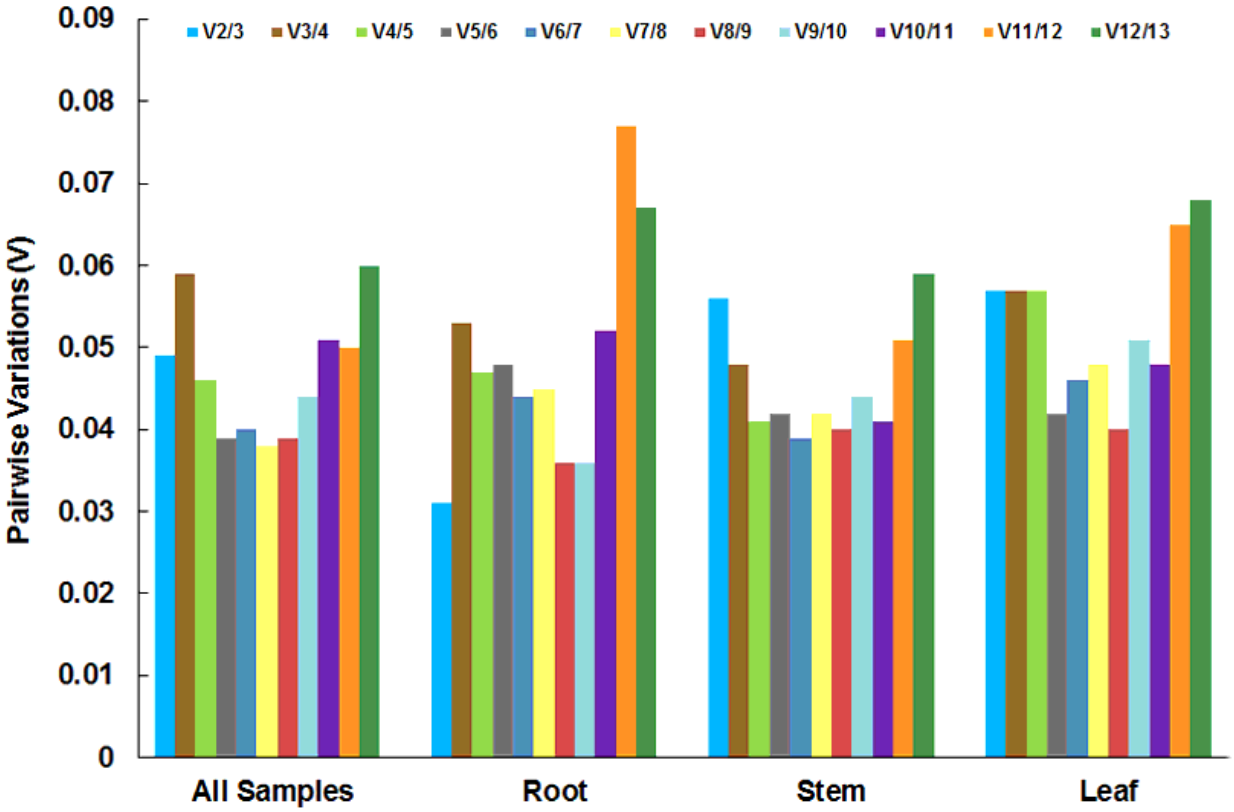

Figure 4

Pairwise variation (Vn/n+1) of ten candidate reference genes calculated by geNorm. The cut-off value to determine the optimal number of RGs for qRT-PCR normalization is 0.15.

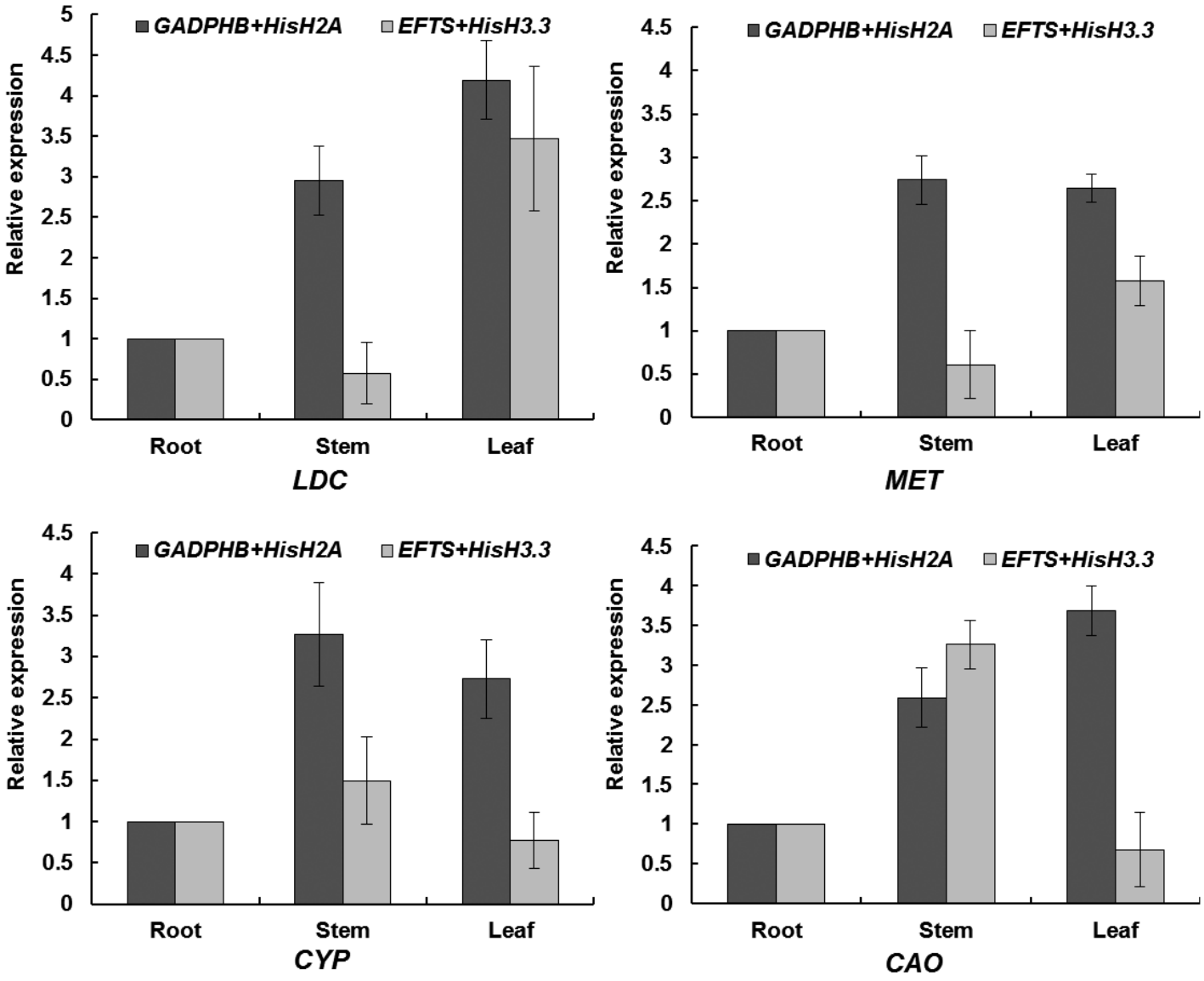

Figure 5

Relative expression levels of four HupA-biosynthetic genes in different tissues normalized by the most stable and unstable combination. The expression level of LDC, MET, CYP, CAO. The relative expression was calculated using the comparative threshold method (2-ΔΔCT). The transcription levels of target gene in the root were set to one. Bars represent the standard error from three biological replicates.

Supplementary Files

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

- Additionalfile1.docx