Screening and validation of suitable internal reference genes for gene expression analysis in Rubia cordifolia L. under different abiotic stresses

Wenwen Li
Henan University of Chinese Medicine
Ke Wang
Henan University of Chinese Medicine
Mingzhu Xiao
Henan University of Chinese Medicine
Conglong Lian
lianc100@163.com
Henan University of Chinese Medicine
Suiqing Chen
Henan University of Chinese Medicine

Research Article

Keywords: Rubia cordifolia L., RT-qPCR, internal reference genes, gene expression, abiotic stress

Posted Date: June 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9575819/v1

Additional Declarations: No competing interests reported.

Abstract

Rubia cordifolia L. is routinely employed as a traditional Chinese medicinal herb in clinical practice. As research on *R. cordifolia* advances, the identification of stable internal reference genes has become essential for gene expression studies. A comprehensive literature search reveals that the screening of internal reference genes for *R. cordifolia* has been reported neither domestically nor internationally. This study employed quantitative real-time PCR (RT-qPCR) to assess the expression stability of several candidate internal reference genes in *R. cordifolia*. The identification of internal reference genes that are stably expressed across different tissues and under various stress conditions in *R. cordifolia* has laid an important foundation for gene expression studies in this plant. This study profiled the expression stability of 12 candidate internal reference genes in *R. cordifolia* across various tissues and under multiple abiotic stresses (NaCl, dehydration, SA, MeJA, ABA), leveraging a suite of algorithms (GeNorm, NormFinder, BestKeeper, ΔCt, and RefFinder) for comprehensive evaluation.Analysis of the 12 candidate internal reference genes was performed using RefFinder, a computational tool that calculates a comprehensive stability ranking based on weighted geometric means. The collective analysis of candidate internal reference genes using geNorm, Normfinder, BestKeeper, and ΔCt methods identified *hnRNP, UBQ, HIS3.3, PP2A*, and *PTBP2* as the five most stable genes. In contrast, *ACT7* and *EF-1α* were found to be the least stable. Consistent with the findings from the four individual algorithms, RefFinder also identified *hnRNP, HIS3.3, PTBP2, PP2A*, and *RPL5* as the top five most stable genes. Moreover,to further validate the reliability of the selected internal reference genes, we examined the expression patterns of the mevalonate kinase (*MVK)* gene across a range of tissues and under a variety of treatment conditions, Using an inappropriate internal reference genes as a control can lead to significant bias in the results. This study systematically evaluated the expression stability of internal reference genes in *R. cordifolia*, as a fundamental step toward reliable functional genomics research in this species. The findings pave the way for accurately elucidating the roles of its genes.

1. Introduction

Rubia cordifolia L. is a perennial climbing herb and a member of the *Rubia* genus in the Rubiaceae family, its primary distribution encompasses regions of Northeast, North, and Northwest China, as well as northern Sichuan and the Qamdo area of Tibet. It is often found in sparse forests, forest edges, shrublands, or grasslands[1]. The medicinal material of *R. cordifolia* is the dried roots and rhizomes, it has the functions of cooling the blood to arrest bleeding and activating blood circulation to resolve stasis. As a traditional medicinal plant, the dried roots and rhizomes of *R. cordifolia* are commonly used as medicinal parts, which possess the effects of cooling blood to stop bleeding and activating blood circulation to remove stasis. Relevant prescriptions have been widely applied in the clinical treatment of diseases[2–3]. In recent years, with the in-depth development of modern research, the chemical constituents and pharmacological effects of Rubia cordifolia L. have been gradually clarified, and its main chemical constituents include anthraquinones and their glycosides, cyclopeptides, polysaccharides, etc[4]. Apart from its medicinal value, Beyond its pharmacological relevance, *R. cordifolia* has been extensively exploited in non-pharmaceutical sectors, particularly as a natural pigment for dyes, cosmetics, and food products, therefore, it possesses significant medicinal and economic value[5].

In the current field of scientific research, Domestic and international scholars have conducted research and discussions on the pharmacological effects and chemical constituents of the Rubia plant to a certain extent[6]. With the advancement of molecular pharmacognosy, researchers have placed growing emphasis on exploring the biosynthetic pathways and regulatory mechanisms of the major bioactive constituents in *R. cordifolia*, with a particular focus on the functional characterization of its key genes.[6]. In the process of gene expression analysis, selecting stable internal reference genes is particularly crucial, as the stability of gene expression directly impacts the reliability and accuracy of experimental results[7]. However, up to now, no research on gene screening in *R. cordifolia* has been reported. This study aims to conduct an in-depth analysis of the expression stability of internal reference genes in *R. cordifolia* using real-time quantitative polymerase chain reaction (RT-qPCR) technology. The objective is to identify internal reference genes that exhibit stable expression across different tissues of *R. cordifolia* and under various abiotic stresses.

RT-qPCR has become an essential technique for gene expression quantification research due to its high sensitivity, high specificity, excellent reproducibility, and high throughput capabilities[8]. However, the reliability of RT-qPCR data is highly dependent on the normalization of target transcripts using stably expressed internal reference genes[9]. In theory, internal reference genes should maintain constant expression levels across different tissues, treatment conditions, and developmental stages to ensure the accuracy and comparability of quantitative results. However, extensive research indicates that the expression stability of internal reference genes exhibits significant species specificity and condition dependence. The most suitable endogenous internal reference genes vary not only between different plant species but also within the same species under different stress conditions or developmental stages. The use of unstable internal reference genes can lead to systematic bias in the expression profile of target genes. Therefore, the systematic screening and validation of optimal internal reference genes, tailored to different tissues and processing conditions within a specific experimental system, is a fundamental prerequisite for ensuring the accuracy and reproducibility of RT-qPCR results[10].

According to previous reports, the primary computational tools for RT-qPCR data normalization include: ΔCt, geNorm, NormFinder, and BestKeeper. To systematically evaluate the expression stability of candidate internal reference genes, this study comprehensively employed four commonly used algorithms—ΔCt, geNorm, NormFinder, and BestKeeper—to quantify the stability values of internal reference genes across all samples. Based on these calculations, the optimal single gene or gene combination was selected for RT-qPCR data normalization[11].

Plants are rich in structurally diverse secondary metabolites, and their content varies significantly across different tissues and under various stress conditions. The accumulation of secondary metabolites itself can constitute a stress signal, which further influences the expression of genes involved in related synthetic pathways within plants[12]. Therefore, analyzing the expression patterns of key genes involved in secondary metabolite synthesis across different tissues and under various stress treatments is crucial for elucidating their regulatory networks[13]. For *R. cordifolia*, a medicinal plant rich in anthraquinones, there is a lack of systematic reports on reliable selection criteria for suitable candidate internal reference genes. This study provides the first comprehensive assessment of the expression stability of 12 candidate internal reference genes in different tissues (roots, stems, leaves) of *R. cordifolia*

under various stress treatments, with the aim of identifying optimal internal reference genes or their combinations for the normalization of RT-qPCR data in *R. cordifolia*.[10]. This work not only provides a standardized basis for analyzing the functional gene expression of *R. cordifolia*, but also lays a methodological foundation for subsequent in-depth studies on the regulatory mechanisms of secondary metabolic pathways.

2. Materials and Methods

2.1 Processing and Collection of Plant Materials

The *R. cordifolia* seeds utilised in this experiment were obtained from one-year-old *R. cordifolia* plants cultivated in the Medicinal Plant Garden of Henan University of Chinese Medicine. The plants were cultivated for a period of three months within a light-controlled incubator set at 24–26℃ and 70% relative humidity, under a 16 h/8 h dark photoperiod (120 μmol m^{-2} s^{-1}). At the conclusion of this period, the roots, stems and leaves were harvested (hereafter referred to as the 'DT group'). During the vegetative growth stage, three months after germination, the *R. cordifolia* seedlings were subjected to various abiotic stress treatments. For the simulated salt stress treatment, uniformly growing *R. cordifolia* seedlings were thoroughly irrigated with a 200 mM NaCl solution, and leaves were harvested at 0, 4, 8 and 12 h post-treatment. For the dehydration stress treatment, *R. cordifolia* seedlings were cultivated in a controlled environment and then subjected to a process of dehydration. The seedlings were then desiccated using filter paper and positioned on a laboratory bench. Leaf samples were collected at 0, 2, 4 and 8h following the commencement of the dehydration process. The two abiotic stress samples were designated as the ND group. For the salicylic acid (SA), methyl jasmonate (MeJA) and abscisic acid (ABA) treatments (referred to as the 'SMA group'), 200 μmol SA, 200 μmol MeJA and 300 μmol ABA were uniformly sprayed onto the *R. cordifolia* leaves. Leaves from the various treatment groups were harvested at 0, 4, 8 and 12 h after treatment. All collected samples were immediately frozen in liquid nitrogen and subsequently stored at − 80℃ for subsequent analysis.The plant material of *R. cordifolia* was formally identified by Professor Suiqing Chen (Henan University of Chinese Medicine). The voucher specimen was deposited in the Herbarium of Henan University of Chinese Medicine, Zhengzhou, China.The voucher specimen accession number is HUCM-RC-0401.

2.2 Selection of Candidate internal reference genes, Primer Design and RT-qPCR

Common endogenous genes were selected as internal controls for gene expression in order to identify candidate internal reference genes in *R. cordifolia*. Utilising the NCBI Local BLAST tool in BioEdit 7.1.9, a BLASTx search was conducted against the assembled *R. cordifolia* transcriptome database with an E = 10^{-50}. Select 12 candidate internal reference genes (*ACT7, TUB8, UBQ, HIS3.3, eIF, EF-1α, PP2A, TBP, RPL5, hnRNP, PTBP2, PEPC*), investigate the impact of internal reference genes selection on RT-qPCR normalization. Use the NCBI Primer Design Tool (https://www.ncbi.nlm.nih.gov/tools/primer-blast/) to design specific primer pairs for each Candidate internal reference genes.

2.3 Total RNA isolation, cDNA synthesis, RT-qPCR and amplification efficiency

Total RNA was extracted from all samples using the FastPure Universal Plant Total RNA Isolation Kit(Vazyme Biotech Co., Ltd., Nanjing, China), in accordance with the manufacturer's instructions. The quantity and purity of the RNA were determined using a NanoDrop 2000 spectrophotometer to measure the optical density ratio OD_{260}/OD_{280} (1.8–2.2) and the absorption ratio OD_{260}/OD_{230} (1.7–2.0). The integrity of the total RNA was then assessed on a 1.0% agarose gel. Following the adjustment of each RNA sample to the same concentration, 800 ng of RNA was prepared for cDNA synthesis. The latter was performed using the ToloScript All-in-one RT EasyMix for qPCR kit（TOLOBIO, Shanghai, China） with random oligomers. In this study, the RT-qPCR experimental procedures were carried out in accordance with previously published protocols and MIQE guidelines. RT-qPCR amplification was performed using the 2 × Q3SYBR qPCR Master Mix (Universal) kit（TOLOBIO, Shanghai, China）. The total reaction volume was 20 μL, containing 100 ng of cDNA as template, 10 μL of 2 × TransStart Top Green qPCR SuperMix, 0.4 μL of 50 × Passive Control Dye II, 0.4 μL each of specific forward and reverse primers (10 μM), and nuclease-free water. The reaction conditions were as follows: The first stage of the process was an incubation at 94℃ for 30s. This was followed by 45 cycles of 5s at 94℃ and 30s at 60℃. In order to analyse the efficiency of the PCR, equal volumes of each cDNA sample were utilised. Subsequently, a standard curve was determined using mixed cDNA templates that had been diluted 2.5-fold, 5-fold, 10-fold and 100-fold. Thereafter, a linear regression model was established using the following formula: The following equation is to be used: $E = -10^{(-1/slope)}$. Each experiment comprised three technical replicates.

2.4 Stability Analysis of Candidate internal reference genes

Gene expression stability analysis was performed using QuantStudio 5 software to assess the Ct values of 12 internal reference genes across 23 distinct samples. Four algorithms were employed to assess the stability of 12 Candidate internal reference genes: ΔCt[14], NormFinder[15], BestKeeper[16], geNorm[17], and RefFinder.

2.5 Internal reference genes Stability Validation

In order to validate the reliability of the selected internal reference gene and to determine whether the use of different internal reference genes would affect the normalisation of target gene expression data, MVKwas selected for validation analysis. MVK is one of the key enzymes in the mevalonate (MVA) metabolic pathway. It is responsible for catalysing the conversion of 3-hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA) into mevalonate (MVA). This pathway is a key metabolic pathway in the biosynthesis of anthraquinone compounds[18]. Consequently, the expression of *MVK* gene in *R. cordifolia* across different tissues and under various stress conditions was analysed. The expression was normalised using the most stable internal reference gene identified via the RefFinder algorithm, followed by validation analysis. The relative expression levels of *MVK* were calculated using the $2^{-\Delta\Delta Ct}$ method. For each biological sample, three technical replicates were performed.

3. Results

3.1 Selection and Amplification Specificity of Candidate Internal Reference Genes

To evaluate the potential of 12 common candidate internal reference genes, previously reported in the literature (*ACT7, TUB8, UBQ, HIS3.3, eIF, EF-1, PP2A, TBP, RPL5, hnRNP, PTBP2, PEPC*), as reliable reference genes for Rubia across different tissue types and abiotic stresses, these candidates were combined with the *R. cordifolia* transcriptomic data.The sequences of these 12 candidate reference genes were retrieved from the *R. cordifolia* transcriptome data, and their primer sequences and relevant annotations are listed in Table 1.To verify the accuracy and specificity of the primers, we performed agarose gel electrophoresis. The results revealed a single DNA band within the 70–300 bp range for each primer combination, consistent with the predicted PCR product sizes (Fig. 1A).Additionally, all 12 amplification reactions targeting candidate internal reference genes produced unimodal RT-qPCR melting curves, with specific melting temperature ranges from 79.69°C (*hnRNP*) to 89.55°C (*PTBP2*) (Fig. 1B).The sequences of 12 candidate reference genes are listed in Table S1.These two results indicate that the primer pairs used in our study exhibit good specificity. Moreover, the PCR efficiency of the candidate internal reference genes ranged from 1.7030 to 2.0424, with *PEPC* exhibiting the lowest amplification efficiency, whereas *TBP* demonstrated the highest. In summary, the primers designed in this study are well-suited for RT-qPCR analysis of gene expression under the tested conditions.

Table 1
Gene Description, Primer Sequences, Amplicon Length, and PCR Efficiency of Candidate candidate internal reference genes in *R. cordifolia*.

Gene	Gene annotation	Primer sequence of forward (5'–3')	primer sequence of reverse (5'–3')	Amplicon length (bp)	Tm (℃)	PCR efficiency
ACT7	Actin7	GGGTCAATCCTTGCATCCCT	GCATCTCAGCATCCATCTGC	160	80.64	1.7853
TUB8	Tubulin	GTTGTGGAAAGGCCGTTTGT	TAACGCAACCTGACAAGGCA	245	81.34	1.9968
UBQ	Ubiquitin	GGCCGTACTCTCGCGGAC	CACAGCGACACTCAGAGACC	89	82.03	1.8732
HIS3.3	Histone superfamily protein	TGAGGCTTACCTTGTTGGGC	CATGATGGTCACACGCTTGG	73	80.5	1.9941
eIF	eukaryotic translation initiation factor 3 subunit D-like, partial	CCAAAGTTCAATCCGCGGTG	GTCGAGCGGTTGAGGTTGTA	149	89.25	1.8307
TBP	TATA binding protein	GATCCATTTCGCCTCCCTGT	TCTTGGCTCCGGTACAAACC	348	81.2	2.0424
PP2A	serine/threonine-protein phosphatase 2A activator isoform X2	CCATACAATTTCCACGCGCC	GGAAGTTGGCGAGAGTAGGG	88	82.89	1.9552
EF-1α	elongation factor 1-alpha	TCTTATCGCCTCCCTCTGCT	TCACGTTCAGCCTTGAGCTT	247	81.44	1.8727
RPL5	60S ribosomal protein L5	ACTGGTCTGTTACTTGCCCG	CTCTGTTGCCGGTAGTGGTT	172	81.23	1.8303
hnRNP	Heterogeneous nuclear ribonucleoprotein	CAGTAAGCCGAATGCCTCCA	GGGCTACTGGGCCATTTCTT	141	79.69	1.8101
PTBP2	Polypyrimidine tract-binding protein homolog 2	TCCGGTAGGATTTGTGGTGC	CTCCTCAGTACGTTCCCACG	119	89.55	1.8414
PEPC	Phosphoenolpyruvate carboxykinase	GAAGAAATCGGCTCCGACGA	CCGCCGGTGAGAGATTGTAA	278	83.86	1.7030

3.2 Expression Profile Analysis of Candidate Internal Reference Genes in *R. cordifolia*

The cycle threshold (Ct) values were used to determine the expression levels of 12 candidate internal reference genes across all 23 samples,and were also displayed as box plots (Fig. 2A).The Ct values are listed in Supplementary Table S2. The results indicate that the *UBQ* gene exhibited the highest expression level, with the lowest average Ct value of 20.38. The average Ct values of the other candidate internal reference genes ranged from 21.51 to 28.51, with *TUB8* exhibiting the highest Ct value of 28.51, indicating the lowest expression level among all tested samples (Fig. 2) .

3.3 Analysis of Expression Stability of Candidate Internal Reference Genes in *Rubia cordifolia*

Four primary computational algorithms were employed in this study to screen for suitable candidate internal reference genes: the comparative ΔCt method, BestKeeper, NormFinder, and geNorm.

3.3.1 comparative ΔCt method

In the comparative ΔCt method, the stability of candidate internal reference genes was evaluated using the mean standard deviation (STDEV) index.The candidate internal reference genes exhibited low STDEV values, suggesting that they have stable expression across different samples. According to the results derived from this method, *TUB8* was identified as the most stable internal reference gene, exhibiting the lowest STDEV values across the total sample set, salicylic acid (SA) treatment, and SMA treatment (Fig. 3A, E, I) .*UBQ* was identified as the most stable internal reference gene under methyl jasmonate (MeJA) treatment and in the ND group (a combination of NaCl and dehydration treatments) (Fig. 3F, H) .Meanwhile, *PP2A* and *hnRNP* were the most stable under dehydration and NaCl treatments, respectively (Fig. 3C, D).

3.3.2 Normfinder Analysis

As another algorithm for determining stable internal reference genes, NormFinder evaluates the stability-based variation across total sample subsets.According to NormFinder results, the stability of candidate internal reference genes increases as the stability value (SV) decreases. Among all *R.*

cordifolia and ABA samples, *PTBP2* was ranked as the internal reference gene with the lowest SV values, at 0.64 and 0.2, respectively.The genes with the lowest expression stability were *EF-1α* and *ACT7*, which presented the highest stability values (SVs) of 3.08 and 1.35, respectively (Fig. 4A,G). For the combined SA and SMA sample group, *RPL5* exhibited the optimal expression stability and was identified as the most reliable internal reference gene, with SVs of 0.19 and 0.32, respectively.Conversely, *TUB8* showed the lowest expression stability with the highest SV values of 2.67 and 2.53 (Fig. 4E,I). Among the samples treated with DT, dehydration, NaCl, MeJA, and ND, *HIS3.3,EIF,hnRNP,EF-1α* and *UBQ* were ranked as the most stably expressed genes in each corresponding treatment group (Fig. 4B,C,D,F,H).

3.3.3 BestKeeper Analysis

As an Excel-based algorithm, BestKeeper screens stable internal reference genes across samples and evaluates the standard deviation (SD) and coefficient of variation (CV) of each candidate gene.A smaller SD and a lower CV indicate greater gene stability. According to BestKeeper analysis, *TUB8* was the most stable internal reference gene across all *R. cordifolia*, SA, and SMA composite groups (Fig. 5A, E, I).In NaCl and ABA samples, *hnRNP* was the most stable internal reference gene (Fig. 5D, G).In DT, dehydrated, MeJA-treated, and ND samples, *HIS3.3, PP2A, UBQ*, and *PTBP2* ranked first and were identified as the most stable genes (Fig. 4B, C, F, H).Among all *R. cordifolia* samples, dehydrated samples, NaCl samples, and ND samples, *EF-1α* was the least stable gene. In DT samples, MeJA-treated samples, and ABA samples, *ACT7* was the least stable gene.

3.3.4 GeNormrAnalysis

The gene expression stability measure M for an internal reference gene is calculated by GeNorm as the average pairwise variation between that gene and all other measured internal reference genes. According to GeNorm, the two most stable internal reference genes are determined based on their M-values. The M-value is negatively correlated with the stability of candidate internal reference genes

a smaller M-value indicates higher stability, whereas a larger M-value indicates lower stability.Among all *R. cordifolia*, ABA, and SMA composite samples, the most stably expressed genes were *hnRNP* and *PTBP2*, with M-values of 0.033, 0.013, and 0.021, respectively. In DT, SA-treated, NaCl, dehydrated, and ND samples, *PP2A, PTBP2, ACT7, TUB8, EIF*, and *UBQ* ranked first and were identified as the most stable genes. In MeJA-treated samples, *EF-1α* and *RPL5* were identified as the most stable genes (Table 2). GeNorm is also used to determine the optimal number of internal reference genes required for accurate normalization.The pairwise variation ($V_{n/n+1}$) between the normalization factors NF_n and NF_{n+1} was calculated. When $V_{n/n+1} > 0.15$, the number of genes required for normalization is n + 1. When $V_{n/n+1} \leq 0.15$, the number of genes required for normalization is n. In this study, all pairwise variation $v_{2/3}$ values fell below the 0.15 threshold, thus requiring normalization using only two internal reference genes (Table 3).

3.4 Use RefFinder to comprehensively rank internal reference genes

Considering the differences in stability analysis among the three programs, RefFinder was employed to reanalyze the 12 candidate internal reference genes using a weighted geometric mean.Integrating the results from GeNorm, NormFinder, BestKeeper, and the ΔCt method for candidate internal reference genes revealed that *hnRNP, UBQ, HIS3.3, PP2A*, and *PTBP2* ranked among the top five most stable genes, whereas *ACT7* and *EF-1α* were identified as the least stable genes.

RefFinder identified the top five most stable genes as *hnRNP, HIS3.3, PTBP2, PP2A*, and *RPL5*, whereas the least stable internal reference genes were *ACT7* and *EF-1α*. These results are largely consistent with those obtained from the four aforementioned methods (Table 4).

3.5 Validation of the selected candidate internal reference genes

Analysis of different *R. cordifolia* samples revealed that *hnRNP* and *HIS3.3* were not the most stably expressed internal reference genes during data normalization. Therefore, to validate the reliability of the two most stable candidate internal reference genes identified by RefFinder (Table 4), we used them both individually and in combination as internal controls to analyze the expression patterns of MVK in different Rubia root samples.These results were then compared with those obtained using *hnRNP* or HIS3.3 alone for data normalization. As shown in Fig. 6, identical expression patterns were observed when using combinations of the two or three most stable genes, or when using *hnRNP* or *HIS3.3* as a single internal reference gene to normalize RT-qPCR data.However, when data were normalized using *hnRNP* or *HIS3.3* alone as internal controls, MVK expression levels differed (Fig. 6).Considering both cost and experimental feasibility, we recommend using *hnRNP* alone for gene expression normalization across different tissues, dehydration treatments, and NaCl-stressed *R. cordifolia*. samples.For samples treated with SA and ABA, gene expression was normalized using *hnRNP* or *HIS3.3*. For samples treated with MeJA, gene expression was normalized using *hnRNP, HIS3.3*, or *UBQ*.

Table 2
m-values for expression stability of 12 candidate candidate internal reference genes calcu

Stability rank	Total		DT		SA		MeJA		ABA		Dehydration	
	Gene	Values	Gene	Values	Gene	Values	Gene	Values	Gene	Values	Gene	V
1	*hnRNP/PTBP2*	0.033	*PP2A/TUB8*	0.005	*ACT7/PP2A*	0.022	*EF-1α/RPL5*	0.007	*hnRNP/PTBP2*	0.013	*EIF/PTBP2*	0
2	*PP2A*	0.045	*UBQ*	0.014	*EIF*	0.030	*TUB8*	0.017	*EF-1α*	0.018	*RPL5*	0
3	*EIF*	0.050	*TBP*	0.016	*PTBP2*	0.033	*PTBP2*	0.019	*HIS3.3*	0.020	*HIS3.3*	0
4	*HIS3.3*	0.065	*HIS3.3*	0.017	*hnRNP*	0.037	*PP2A*	0.020	*UBQ*	0.023	*PEPC*	0
5	*RPL5*	0.075	*PTBP2*	0.018	*RPL5*	0.041	*UBQ*	0.021	*RPL5*	0.027	*TBP*	0
6	*ACT7*	0.084	*hnRNP*	0.019	*HIS3.3*	0.056	*HIS3.3*	0.022	*TUB8*	0.030	*PP2A*	0
7	*TUB8*	0.093	*RPL5*	0.025	*PEPC*	0.063	*EIF*	0.024	*TBP*	0.033	*UBQ*	0
8	*PEPC*	0.107	*PEPC*	0.030	*TBP*	0.069	*PEPC*	0.026	*PP2A*	0.036	*hnRNP*	0
9	*UBQ*	0.117	*EF-1α*	0.041	*EF-1α*	0.082	*TBP*	0.027	*EIF*	0.040	*EF-1α*	0
10	*TBP*	0.135	*EIF*	0.050	*UBQ*	0.092	*hnRNP*	0.028	*PEPC*	0.043	*ACT7*	0
11	*EF-1α*	0.155	*ACT7*	0.061	*TUB8*	0.107	*ACT7*	0.033	*ACT7*	0.054	*TUB8*	0

Table 3
Pairwise variation in geNorm determines the optimal number of internal reference genes

Pairwise variations	Total (values)	DT (values)	SA (values)	MeJA (values)	ABA (values)	Dehydration (values)	NaCl (values)	ND (values)	SMA (values)
$V_{2/3}$	0.0162	0.0060	0.0107	0.0072	0.0068	0.0071	0.0265	0.0095	0.0128
$V_{3/4}$	0.0122	0.0039	0.0084	0.0045	0.0047	0.0087	0.0239	0.0104	0.0115
$V_{4/5}$	0.0163	0.0030	0.0082	0.0037	0.0050	0.0072	0.0126	0.0090	0.0132
$V_{5/6}$	0.0143	0.0030	0.0082	0.0032	0.0054	0.0053	0.0082	0.0085	0.0118
$V_{6/7}$	0.0135	0.0029	0.0133	0.0031	0.0050	0.0061	0.0064	0.0113	0.0123
$V_{7/8}$	0.0138	0.0053	0.0102	0.0034	0.0047	0.0054	0.0043	0.0098	0.0127
$V_{8/9}$	0.0158	0.0049	0.0094	0.0032	0.0048	0.0059	0.0046	0.0099	0.0108
$V_{9/10}$	0.0143	0.0081	0.0130	0.0028	0.0051	0.0070	0.0035	0.0116	0.0125
$V_{10/11}$	0.0179	0.0083	0.0125	0.0025	0.0050	0.0083	0.0038	0.0184	0.0120
$V_{11/12}$	0.0203	0.0091	0.0150	0.0046	0.0090	0.0089	0.0047	0.0264	0.0145

Table 4
Comprehensive Ranking of candidate internal reference genes Using RefFinder

Stability rank	Total		DT		SA		MeJA		ABA		Dehydration		NaCl		N
	Gene	Values	Gene	Values	Gene	Values	Gene	Values	Gene	Values	Gene	Values	Gene	Values	G
1	*PTBP2*	1.32	*PTBP2*	1.68	*hnRNP*	1.86	*EF-1α*	1.57	*PTBP2*	1.41	*PTBP2*	1.57	*UBQ*	1.86	*P*
2	*hnRNP*	2	*HIS3.3*	1.86	*RPL5*	1.93	*RPL5*	2.45	*UBQ*	2.11	*EIF*	1.68	*PP2A*	2	*U*
3	*PP2A*	3.66	*UBQ*	2.45	*PTBP2*	3.41	*PTBP2*	3	*hnRNP*	2.63	*RPL5*	3.08	*PTBP2*	2.78	*P*
4	*EIF*	3.66	*hnRNP*	3.13	*EIF*	4.23	*TUB8*	4.76	*EF-1α*	3.83	*PP2A*	3.31	*hnRNP*	3.6	*h*
5	*HIS3.3*	4.68	*PP2A*	5.44	*ACT7*	5.21	*UBQ*	5.27	*HIS3.3*	3.98	*HIS3.3*	5.32	*HIS3.3*	3.87	*E*
6	*TUB8*	5.32	*TUB8*	6.64	*PP2A*	5.26	*PEPC*	5.48	*RPL5*	5.05	*UBQ*	5.42	*ACT7*	4.9	*H*
7	*RPL5*	5.96	*PEPC*	7.09	*HIS3.3*	5.79	*hnRNP*	5.66	*PP2A*	7.24	*TBP*	7.97	*EIF*	6.19	*A*
8	*ACT7*	7.24	*EF-1α*	7.21	*TUB8*	6.45	*PP2A*	6.19	*TBP*	8.65	*PEPC*	8.1	*TUB8*	7.74	*T*
9	*UBQ*	8.49	*TBP*	7.48	*PEPC*	7.14	*HIS3.3*	7.02	*PEPC*	9.38	*hnRNP*	8.45	*TBP*	9.24	*R*
10	*PEPC*	9.49	*RPL5*	9.43	*TBP*	9.24	*EIF*	8.05	*TUB8*	9.46	*TUB8*	8.49	*RPL5*	9.74	*T*
11	*TBP*	11	*EIF*	10.74	*EF-1α*	10.24	*TBP*	10.74	*EIF*	9.97	*EF-1α*	10.47	*PEPC*	11	*P*
12	*EF-1α*	12	*ACT7*	12	*UBQ*	11.24	*ACT7*	12	*ACT7*	12	*ACT7*	10.74	*EF-1α*	12	*E*

4. Discussion

Real-time quantitative polymerase chain reaction (RT-qPCR) is a precise and sensitive technique for detecting gene expression. To ensure the reliability of RT-qPCR results, the selection of appropriate internal reference genes is crucial. In recent years, with the in-depth advancement of research on gene expression in medicinal plants, internal reference gene screening has been carried out in a variety of medicinal plant species. Studies have shown that the expression stability of internal reference genes may change significantly in different tissues or under different experimental conditions. Therefore, it is necessary to evaluate the stability of internal reference genes under different experimental conditions to ensure the accuracy of RT-qPCR data normalization.

RT-qPCR has been widely applied in the field of functional gene expression analysis in medicinal plants, and the selection of suitable internal reference genes for data normalization is a prerequisite for ensuring the reliability of RT-qPCR results [19]. In recent years, with the continuous in-depth study of molecular biology in medicinal plants, a growing number of studies have confirmed that there is no universal internal reference gene applicable to all species and all experimental conditions, and its expression stability changes significantly with species differences, tissue types, and external stress conditions [20].

For example, in *Isodon rubescens*, *GAPDH + 18S + eIF* was identified as the optimal internal reference gene combination for different tissues; *UBQ + Apt + HIS* and *Cycl + UBQ + PP2A* were the optimal internal reference gene combinations for dehydration and NaCl stress, respectively; and *GAPDH + 18S + eIF*, *eIF + UBQ + PP2A*, and *TUB + Cycl + UBQ* were sequentially the optimal internal reference gene combinations for SA, MeJA, and ABA treatments [10]. In *Rubia yunnanensis*, *hnRNP*, *TBP*, and *RPL5* were the optimal internal reference gene combinations for all samples and different tissues; *hnRNP + eIF* was the optimal internal reference combination for drought (PEG) treatment; *EF-1α + UBCE* was the optimal internal reference combination for salt (NaCl) treatment; *RPL5 + MDH* was the optimal internal reference combination for MeJA treatment; *TBP + UBCE* was the optimal internal reference combination for SA treatment; *TBP + RPL5* was the optimal internal reference combination for oxidative and heavy metal treatments; *RPL5 + hnRNP* was the optimal internal reference combination for heat stress; and *hnRNP + F-box* was the optimal internal reference combination for cold stress [11]. This is consistent with the results of a transcriptome study on *Rubia yunnanensis*, which also used *hnRNP* as an internal reference gene for MeJA-treated hairy roots, confirming the stability and applicability of *hnRNP* in *Rubia yunnanensis* under MeJA stress [21]. In *Panax ginseng*, *GAPDH/30S RPS20, CYP/60S RPL13*, and *CYP/QCR* were the optimal internal reference gene combinations for roots, stems, and leaves, respectively[22]. In *Metasequoia glyptostroboides*, *ACT2*, *HIS*, and *TATA* showed stable expression in different tissues, while *EF-1α*, *HIS*, and *TATA* were suitable for abscisic acid-treated samples [23].

However, to date, there have been no reports on the screening of internal reference genes in *R. cordifolia*.Therefore, the systematic evaluation of the expression stability of candidate internal reference genes in different tissues and under abiotic stress conditions of *R. cordifolia*, and the screening of suitable internal reference genes and standardization protocols, are of great theoretical and practical value for the analysis of functional gene expression, the study of the synthesis mechanism of medicinal active components, and the research on the stress adaptation mechanism of *R. cordifolia*. In this study, 12 widely used internal reference genes in the literature were selected as candidates, and these 12 selected internal reference genes have also been widely applied in other plants. For instance, in the euhalophyte *Suaeda altissima*, *ACT7*, *PP2A*, and *EF-1α* showed stable expression under long-term salt stress and salt shock, among which *PP2A* and *EF-1α* had the highest stability at different salt concentrations [24]. In *Indocalamus tessellatus*, *PP2A* exhibited high stability under drought (dehydration) stress, while *EIF* family genes were stably expressed under salt stress [25]. In *ramie*, *UBQ*, *EF-1α*, and *TUB* were confirmed to be stable internal reference genes across multiple tissues and abiotic stresses (including hormone treatments), among which *EF-1α* and *TUB* were particularly suitable for root tissue normalization [26].

Five statistical algorithms, namely ΔCt, BestKeeper, NormFinder, GeNorm, and RefFinder, were used to systematically analyze the stability of the 12 selected internal reference genes in different tissues and under abiotic stresses (NaCl, dehydration, SA, MeJA, and ABA) of *R. cordifolia*. To avoid the experimental

complexity caused by too many treatment groups, the samples were further integrated into three groups: different tissue group (DT group), salt stress and dehydration treatment group (ND group), and SA, MeJA, and ABA treatment group (SMA group), so as to accurately select internal reference genes under different environmental conditions. In addition, an overall analysis was performed on all samples. The results of the four methods (ΔCt, BestKeeper, NormFinder, and GeNorm) were integrated by the RefFinder algorithm, and it was found that hnRNP, HIS3.3, and PTBP2 were suitable for the overall sample analysis; HIS3.3 and PTBP2 were stably expressed in different tissues; while PP2A/EIF/TUB8, HIS3.3/ACT7, TUB8/RPL5/PTBP2, UBQ/EF-1α/RPL5, and HIS3.3/PTBP2/hnRNP were recommended for dehydration, NaCl, SA, MeJA, and ABA treatments, respectively.

Therefore, to verify the reliability of the two most stable candidate reference genes identified by RefFinder, they were used as internal controls individually and in combination to analyze the expression pattern of mevalonate kinase (MVK) in different samples of Rubia cordifolia. These results were compared with those normalized using hnRNP or HIS3.3 alone. When two or three of the most stable gene combinations, or hnRNP or HIS3.3 alone, were used to normalize RT-qPCR data, the results showed the same expression pattern. However, the expression level of MVK differed when only hnRNP or HIS3.3 was used as the internal control for data normalization. Considering the cost and operation process comprehensively, we recommend using hnRNP alone for normalizing gene expression in different tissues, dehydration, and NaCl-treated samples of Rubia cordifolia. For samples under SA and ABA treatments, either hnRNP or HIS3.3 can be used for normalization. For samples under MeJA treatment, hnRNP, HIS3.3, or UBQ alone should be used for normalization.

Conclusions

In this study, we conducted the first comprehensive analysis of stable reference gene selection for RT-qPCR analysis of target gene expression in different tissues or under abiotic stresses (NaCl, dehydration, SA, MeJA, ABA) in Rubia. Five algorithms (ΔCt, BestKeeper, NormFinder, GeNorm, RefFinder) were used to evaluate 12 candidate housekeeping genes, with RefFinder yielding more reliable results.For the pooled sample, *hnRNP*, *HIS3.3* and *PTBP2* were recommended; *HIS3.3* and *PTBP2* showed stable expression across different tissues. For individual treatments: *PP2A/EIF/TUB8* for dehydration, *HIS3.3/ACT7* for NaCl, *TUB8/RPL5/PTBP2* for SA, UBQ/EF-1α/RPL5 for MeJA, and *HIS3.3/PTBP2/hnRNP* for ABA. Validation confirmed consistent expression patterns of the MVK gene when normalized by the most stable reference genes (single or combined), though individual use caused expression level differences.Considering cost and simplicity, we propose: *hnRNP* alone for different tissues, dehydration and NaCl treatments; *hnRNP* or *HIS3.3* alone for SA and ABA treatments; *hnRNP, HIS3.3* or *UBQ* alone for MeJA treatments. These results will improve the accuracy of RT-qPCR-based target gene quantification in Rubia and support future gene expression studies in this species.

Declarations

Ethical approval.

During the collection of plant and seed samples, we strictly abided by the policy statements of the International Union for Conservation of Nature (IUCN) regarding research involving endangered species, as well as the relevant guiding principles stipulated by the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES). Additionally, this study did not involve any research content related to humans or animals.

Data availability

The sequence data generated in this study have been submitted to the NCBI GenBank database under the accession numbers PZ337664–PZ337675. The data will be publicly released after manuscript acceptance and completion of NCBI official processing. All sequence data are also listed in Supplementary Table S1. All other data generated or analysed during this study are included in this article and its supplementary files.

Material availability

The seeds of Rubia cordifolia L. were collected in Zhengzhou City, Henan Province, China. We declare that *R. cordifolia* is not listed as a protected plant species in China. Therefore, the collection of its seeds and other related plant materials is legally permitted.

Consent for publication

Not applicable.

Funding Declaration

This study was funded by the High-Level Key Discipline Construction Project of Traditional Chinese Medicine of the National Administration of Traditional Chinese Medicine (Grant No. SATCM-RH[2023]85).

Author contributions

C.L. and S.Q.C. conceived and designed the research framework; W.W.L. and K.W. prepared the samples and performed the experiments; W.W.L., M.Z.X. and K.W. analyzed the data; W.W.L. and K.W. wrote the paper; C.L. and S.Q.C. revised the manuscript; S.Q.C. acquired the funding. All authors have read and approved the final manuscript.

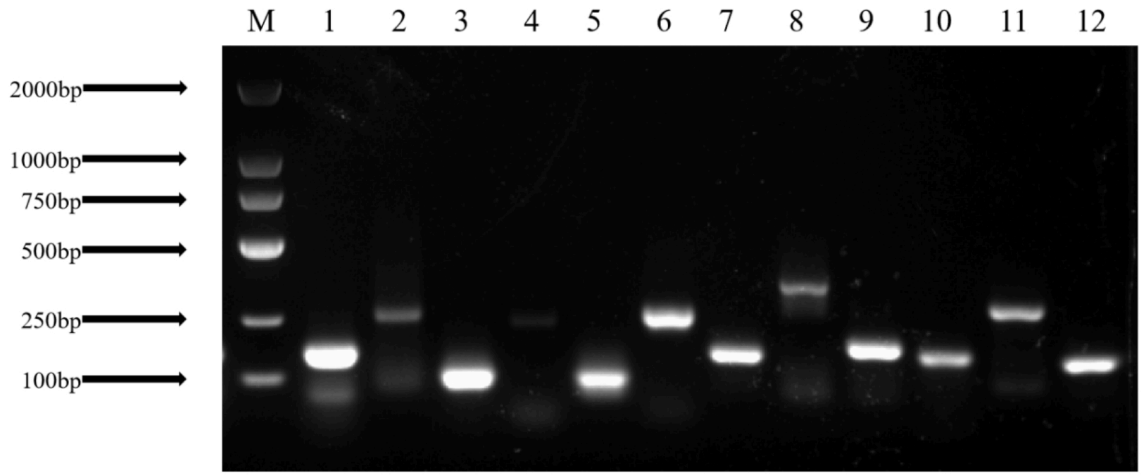
D.Competing interests

The authors declare no competing interests.

References

1. Li Wenwen G, Kaihua L, Conglong et al. Research on Quality Inspection Methods and Grading Standards for Rubia Seed [J]. Seed,2024,43(5): 146–156.(In Chinese).
2. Tang B, Ma L, Ma C. Spectrofluorimetric study of the betacyclodextrin-rubidate complex and determination of rubidate by beta-CD-enhanced fluorimetry [J]. Talanta. 2002;58(5):841–8.
3. Junyong L. Du Yixin. Technical Approaches for Expanding the Initial Resources of *Rubia cordifolia* L. [J]. Shanghai Agricultural Sci Technol, 2015(06): 101–2.
4. Liang P, Yao L, Fei L et al. Effects of Exogenous Methyl Jasmonate and Gibberellic Acid on Growth, Enzyme Activities, and Major Active Compound Content in *Rubia cordifolia* L. [J]. Chinese Herbal.
5. Liang P, Tao S, Yuanyuan W, et al. Research on Quality Grading Standards for Shaanxi-Produced Rubia Seed [J]. Cent South J Pharm. 2017;15(09):1224–7.
6. Zhang Yufang Z, Lijuan ZY. Selection and Application of internal reference genes in Gene Expression Studies [J]. Acta Physiol Plant Sinica 2014, 50(08): 1119–2510.13592/j.cnki.ppj.2014.0201
7. Cécile P, Fanny P, Jean-Marie H, et al. Validation of a set of internal reference genes to study response to herbicide stress in grasses[J]. BMC Res Notes. 2012;5(1):18.
8. Xu S, Liuyang L, ,Yashe L et al. internal reference genes selection for quantitative real-time PCR (qRT-PCR) expression analysis in Galium aparine L. [J].PloS one,2020,15(2):e0226668.
9. Keertan, D, F J H, A S B, et al. Validation of candidate internal reference genes for normalizing RNA expression in real-time PCR. [J] BioTechniques. 2004;37(1):112.
10. Conglong L, Bao Z, ,Jingfan Y, et al. Validation of suitable internal reference genes by various algorithms for gene expression analysis in Isodon rubescens under different abiotic stresses.[J]. Sci Rep. 2022;12(1):19599–19599.
11. Shanyong Y, Qianwen L, Xuejia Z, et al. Selection and Validation of Appropriate internal reference genes for Quantitative RT-PCR Analysis in Rubia yunnanensis Diels Based. on Transcriptome Data.[J].BioMed research international; 2020. p. 20205824841.
12. Niranjana HM, Sebastian KJ, Yoeup KP, et al. Anthraquinone Production from Cell and Organ Cultures of Rubia Species: An Overview[J]. Metabolites. 2022;13(1):39–39.
13. Hasan SQ, Ibrahim T. The Genetic Regulation of Secondary Metabolic Pathways in Response to Salinity and Drought as Abiotic Stresses[J]. Appl Sci. 2021;11(15):6668–6668.
14. Silver N, Best S, Jiang J, Thein SL. Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR. BMC Mol Biol. 2006;7:1–9.
15. Andersen CL, Jensen JL, Ørntoft 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. Can Res. 2004;64:5245–50.
16. Pfaffl MW. A new mathematical model for relative quantification in real-time RT–PCR. Nucleic Acids Res. 2001;29:e45.
17. Etschmann B, Wilcken B, Stoevesand K, Schulenburg VD, A., Sterner-Kock A. Selection of reference genes for quantita-tive real-time PCR analysis in canine mammary tumors using the GeNorm algorithm. Vet Pathol. 2006;43:934–42.
18. Xu Zhenxiang L, Meijie M. Recent Advances in Regulatory Strategies for Metabolic Branches of the Methylopentanoate Pathway [J]. Bioprocess Eng. 2020;18(01):44–51.
19. J K L,D T S. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method.[J]. Methods (San Diego, Calif.), 2001, 25 (4) : 402-8.
20. Jo V, Katleen PD, Filip P, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. [J] Genome biology. 2002;3(7):RESEARCH0034.
21. Rongfei Z, Yuanyuan M, Lingyun C, et al. De Novo Transcriptome Analysis Reveals Putative Genes Involved in Anthraquinone Biosynthesis in Rubia yunnanensis[J]. Genes. 2022;13(3):521–521.
22. Liu J, Wang Q, Sun M, et al. Selection of reference genes for quantitative real-time PCR normalization in Panax ginseng at different stages of growth and in different organs. [J]. PLoS ONE. 2018;9(11):e112177.
23. Li R, et al. Selection of the reference genes for quantitative gene expression by RT-qPCR in the desert plant Stipagrostis pennata. Sci Rep. 2021;11:21711.
24. Khramov ED, Nedelyaeva IO, Konoshenkova OA, et al. Identification and selection of reference genes for analysis of gene expression by quantitative real-time PCR in the euhalophyte Suaeda altissima (L.) Pall. [J]. Commun Integr Biol. 2024;17(1):2372313.
25. Hu X, Zhou C, Pan J, et al. Screening and Validation of Stable Reference Genes for Real-Time Quantitative PCR in Indocalamus tessellatus (Munro) P. C. Keng Under Multiple Tissues and Abiotic Stresses[J]. Forests. 2025;16(10):1607–1607.
26. Yongting Y, Gang Z, Yikun C, et al. Selection of Reference Genes for qPCR Analyses of Gene Expression in Ramie Leaves and Roots across Eleven Abiotic/Biotic Treatments. [J]. Sci Rep. 2019;9(1):20004.

Figures



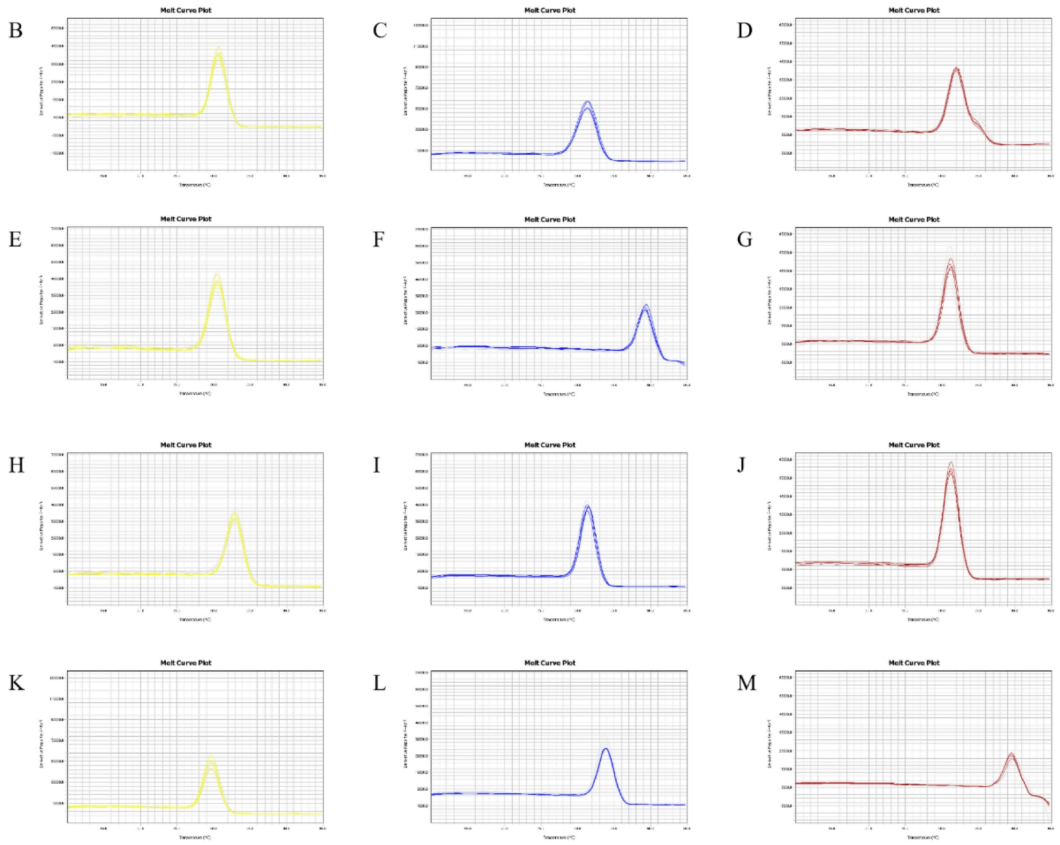


Figure 1

RT-PCR and RT-qPCR Amplification Specificity. (A) Agarose gel (2.0%) electrophoresis shows that the single PCR amplification products of 12 genes (1-12: *ACT7, TUB8, UBQ, HIS3.3, PP2A, EF-1α, eIF, TBP, RPL5, hnRNP, PTBP2, PEPC)* match their expected sizes; M denotes the 2000 bp DNA marker. (B-M) Melting curves for the 12 genes (*ACT7, TUB8, UBQ, HIS3.3, eIF, TBP, PP2A, EF-1α, RPL5, hnRNP, PTBP2, PEPC)* exhibited unimodal profiles.

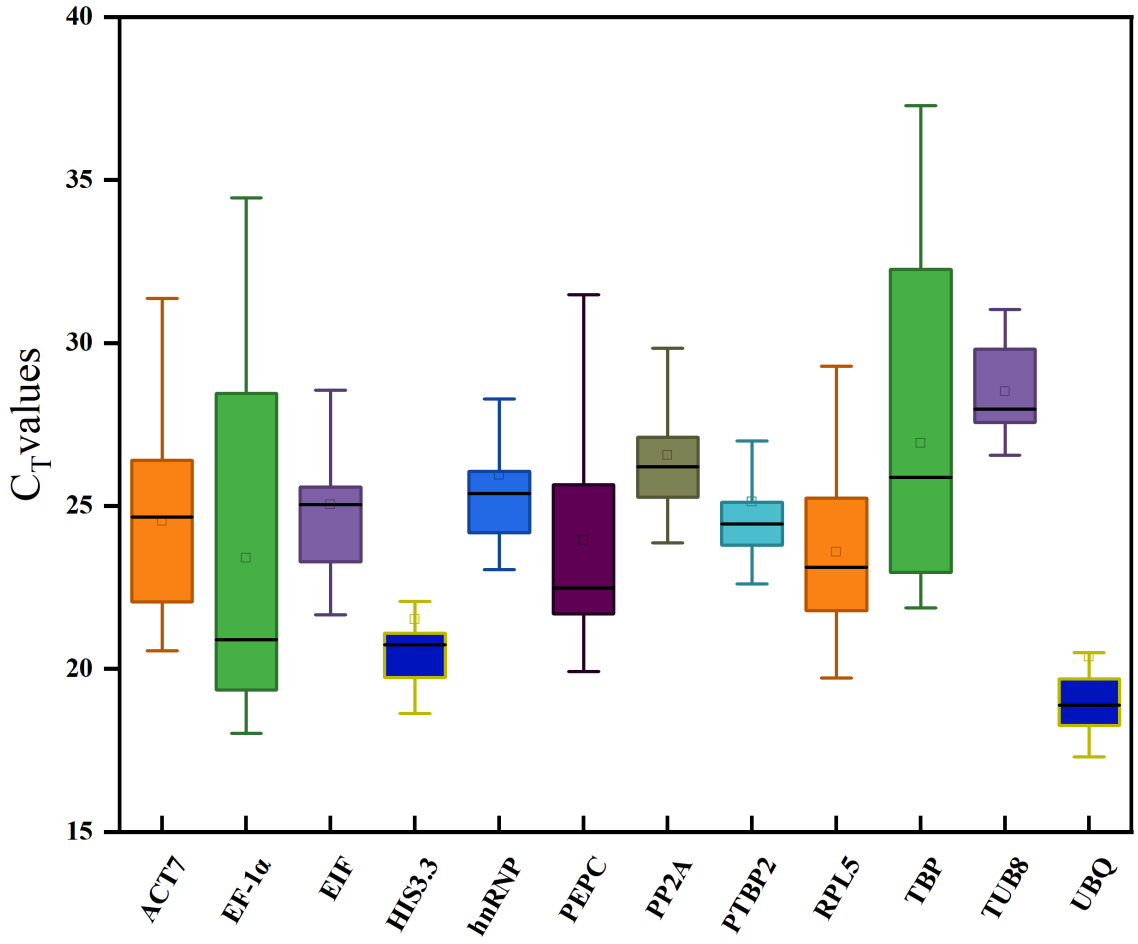


Figure 2

Expression stability of the 12 candidate internal reference genes across all 23 samples. (A) Box plots of the Ct values for each candidate internal reference genes. The plot depicts the mean (hollow square) , median (central line), maximum and minimum values (whiskers, vertical lines) , and the 25th and 75th percentiles (box).

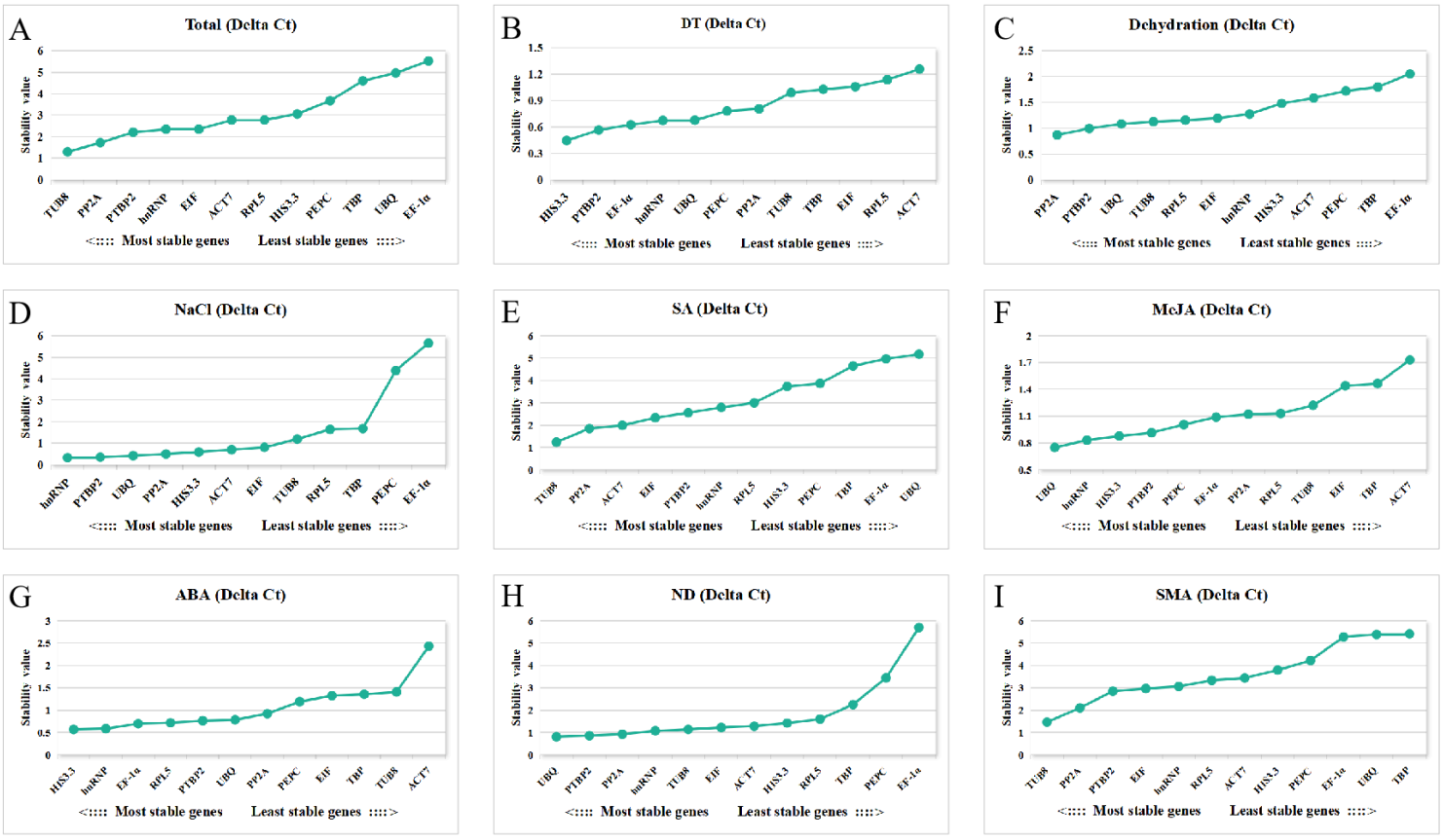


Figure 3

Ranking of candidate internal reference genes based on the ΔCt method. (A) Total (A) Total sample collection. (B) DT: Collection of samples from different tissues. (C) Dehydrated samples. (D) NaCl-treated samples. (E) SA-treated samples. (F) MeJA-treated samples. (G) ABA-treated samples. (H) ND: Collection of samples treated with NaCl and dehydration. (1) SMA: Collection of samples treated with SA, MeJA, and ABA.

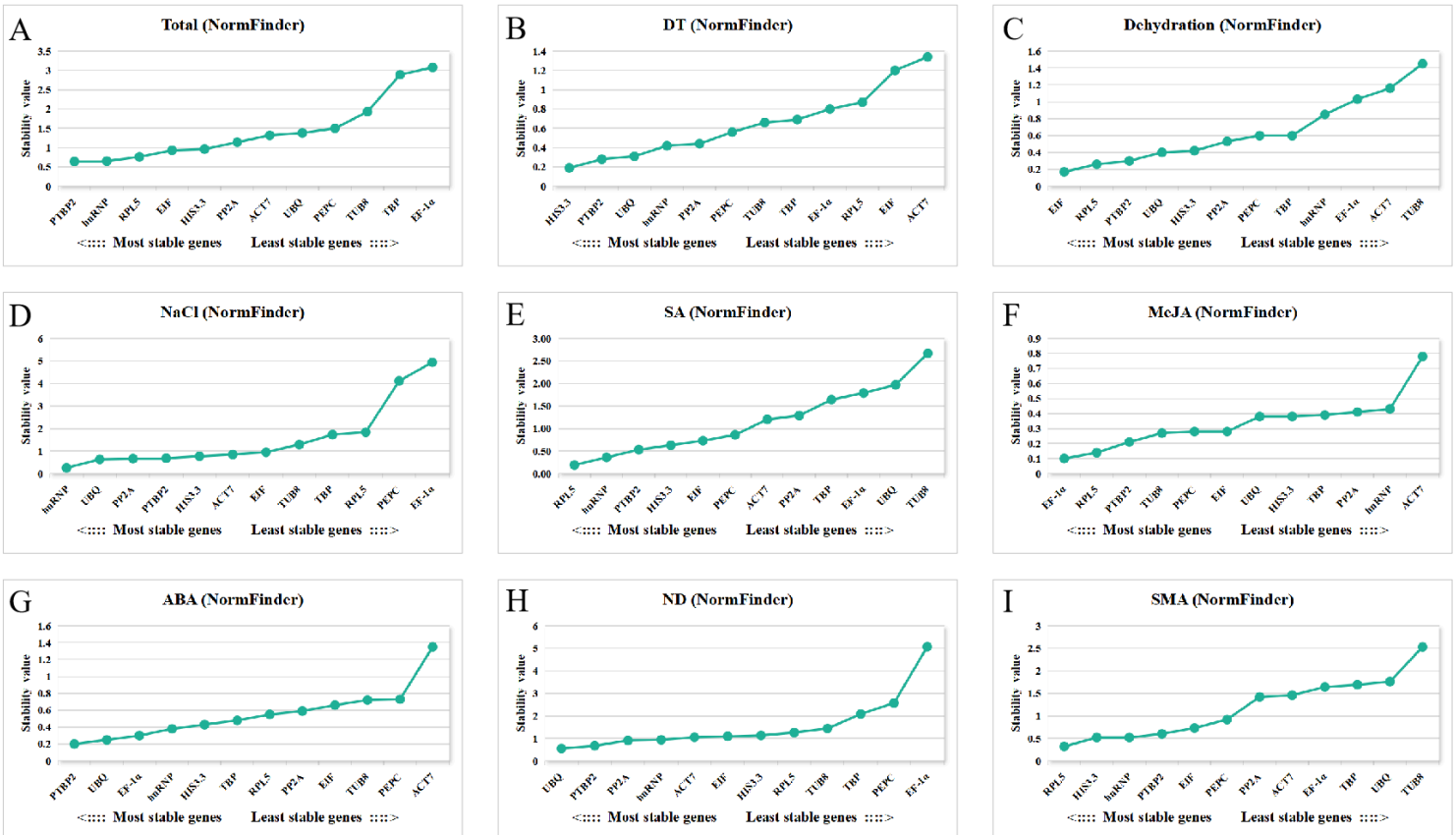


Figure 4

Ranking of candidate internal reference genes using the Normfinder method. (A) Total (A) Total sample collection. (B) DT: Collection of samples from different tissues. (C) Dehydrated samples. (D) NaCl-treated samples. (E) SA-treated samples. (F) MeJA-treated samples. (G) ABA-treated samples. (H) ND: Collection of samples treated with NaCl and dehydration. (1) SMA: Collection of samples treated with SA, MeJA, and ABA.

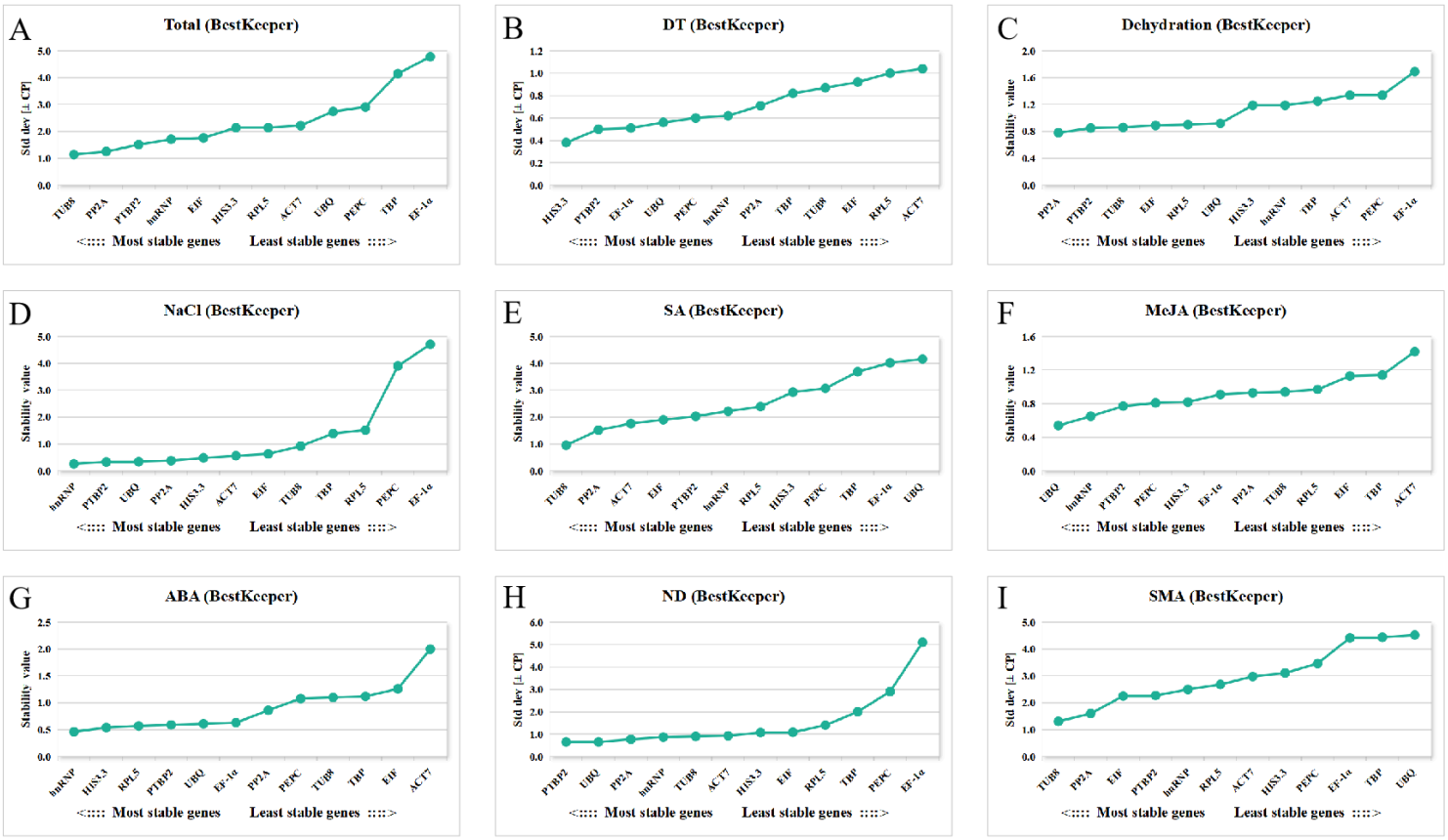


Figure 5

Ranking of candidate internal reference genes using the BestKeeper method. (A) Total (A) Total sample collection. (B) DT: Collection of samples from different tissues. (C) Dehydrated samples. (D) NaCl-treated samples. (E) SA-treated samples. (F) MeJA-treated samples. (G) ABA-treated samples. (H) ND: Collection of samples treated with NaCl and dehydration. (1) SMA: Collection of samples treated with SA, MeJA, and ABA.

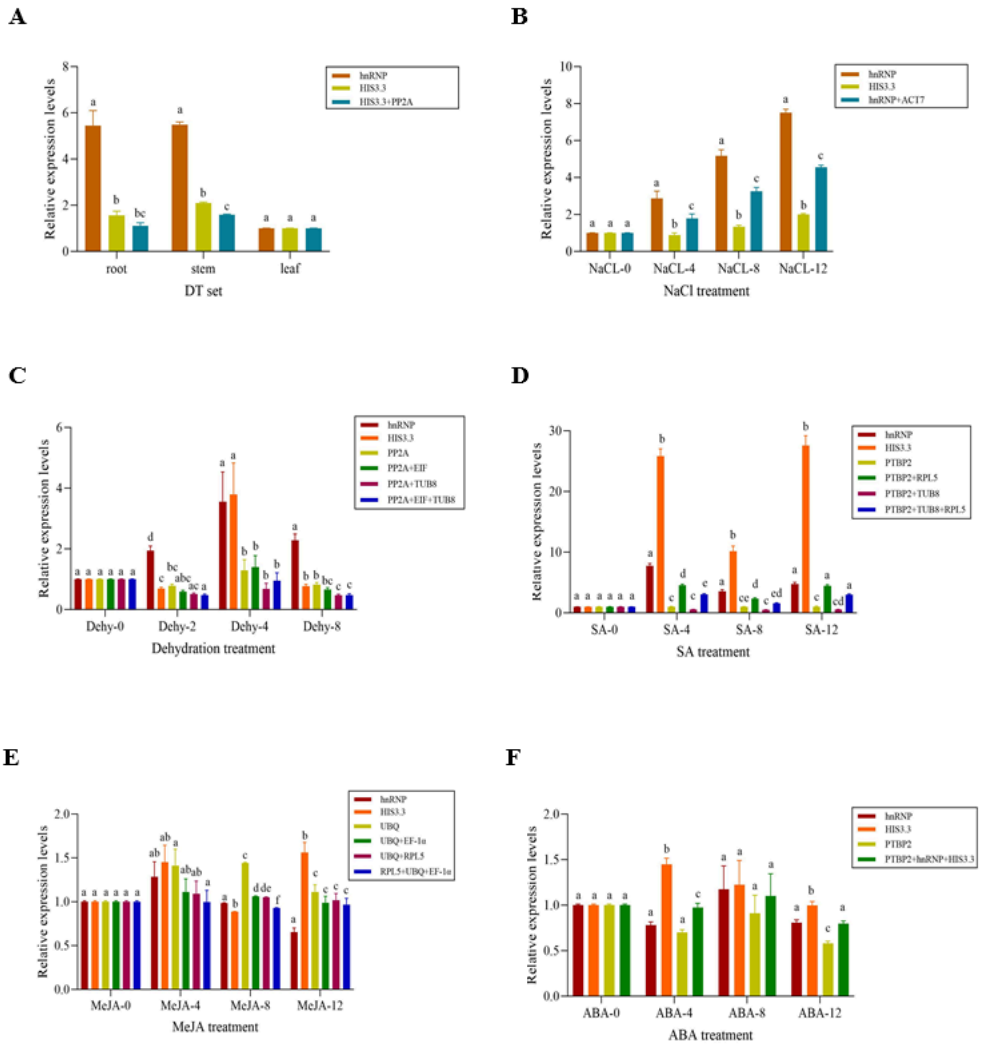


Figure 6

Relative expression levels of MVK in plants. Expression of MVK in different tissues (A) , under NaCl (B) , dehydration (C) , SA (D) , MeJA (E) , and ABA (F) treatments. Genes were normalized to *hnRNP*or *HIS3.3*, or to a combination of the two or three highest-ranked genes defined by comprehensive analysis. All values represent means. Different letters on the bar charts indicate significant differences, $p < 0.05$.

Supplementary Files

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

- Additionalfile1.docx
- Additionalfile2.xlsx
- Agarosegel2.0electrophoresis.pdf