



Reference gene screening for analyzing gene expression in the heart, liver, spleen, lung and kidney of forest musk deer

Xi WU¹#, Wei YANG¹#, Ziwei REN¹, Jianguo CHENG²*, Yan LUO¹*, Yin WANG¹, Zexiao YANG¹, Xueping YAO¹, Wei ZHAO¹, Yimeng LI¹ and Kexin TANG¹

¹College of Veterinary Medicine, Sichuan Agricultural University, Wenjiang 611130, Sichuan Province, China

²Sichuan Institute of Musk Deer Breeding, Dujiangyan 611830, Sichuan Province, China

ABSTRACT. The screening of reference genes for real-time quantitative PCR (qPCR) in forest musk deer (FMD) tissue is of great significance to the basic research on FMD. However, there are few reports on the stability analysis of FMD reference genes so far. In this study, We used qPCR to detect the expression levels of 11 reference gene candidates (*18S rRNA*, beta-actin [*ACTB*], glyceraldehyde-3-phosphate dehydrogenase [*GAPDH*], TATA box-binding protein [*TBP*], hypoxanthine phosphoribosyltransferase 1 [*HPRT1*], tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein zeta polypeptide [*YWHAZ*], hydroxymethylbilane synthase [*HMB*], eukaryotic translation elongation factor 1 alpha 1 [*EEF1A1*], succinate dehydrogenase complex flavoprotein subunit A [*SDHA*], peptidylprolyl isomerase B [*PPIB*], and ubiquitin C [*UBC*]) in heart, liver, spleen, lung and kidney of FMD. After removing *18S rRNA* on account of its high expression level, geNorm, NormFinder, BestKeeper and Δ Ct algorithms were used to evaluate the expression stability of the remaining genes in the five organs, and further comprehensive ranking was calculated by RefFinder. According to the results, the selected reference genes with the most stable expression in the heart of FMD are *SDHA* and *YWHAZ*, while in the liver are *ACTB* and *SDHA*; in the spleen and lung are *YWHAZ* and *HPRT1*; in the kidney are *YWHAZ* and *PPIB*. The use of common reference genes in all five organs is not recommended. The analyses showed that tissue is an important variability factor in genes expression stability. Meanwhile, the result can be used as a reference for the selection of reference genes for qPCR in further study.

KEY WORDS: expression stability, forest musk deer, reference genes

J. Vet. Med. Sci.

83(11): 1750–1759, 2021

doi: 10.1292/jvms.21-0281

Received: 16 May 2021

Accepted: 14 September 2021

Advanced Epub:

7 October 2021

Forest musk deer (FMD, *Moschus berezovskii*), an exclusive species in Asia, belongs to Moschidae, which is closest to bovines in the phylogenetic tree of Pecora [10]. China is the country with the most abundant resources of FMD [4]. Due to the scarcity of FMD, the FMD is listed as a national first-level protected animal by China and classified as an “endangered species” in the International Union for Conservation of Nature (IUCN) [7, 35]. The musk secreted by the adult male FMD has an important position in traditional Asian medicine. At the same time, musk is also used as a fragrance and widely used in perfumes [39, 41]. Therefore, musk has great economic value. The artificial breeding of FMD to obtain natural musk has developed rapidly in China. But the captive FMD generally suffers from the pressure of diseases, such as digestive tract diseases, pneumonia and purulent diseases, and so on [14, 38]. In order to deal with these diseases, it is necessary to carry out in-depth research on its pathogenesis and immune mechanism, and gene expression analysis plays an important role in this process [18, 30].

Real-time quantitative PCR (qPCR) is currently the most commonly used method for gene expression analysis, with the characteristics of fast, high sensitivity, high specificity and accuracy [3, 23, 27]. However, the accuracy of qPCR results is greatly affected by the quality of the experimental materials, the purity and concentration of RNA, the quality of reagents, and instrumental errors during the experiment [6, 19]. Therefore, in order to improve the reliability of qPCR results, specific reference genes are usually selected in the research to normalize the data of the target genes to correct and standardize the experimental results. The expression of an “ideal” reference gene will not vary with treatment or physiological state in the tissue studied. However, several studies have shown that no perfect reference gene exists because the reference gene can be influenced by species, different tissues, developmental stages and diseases, etc [2, 17, 29]. Consequently, it is becoming an essential component of qPCR to systematically evaluate common and novel reference genes derived from genomewide analyses so as to improve the reliability of research results.

*Correspondence to: Cheng, J.: 447933018@qq.com; Luo, Y.: Lycjg@163.com

#These authors contributed equally to this work.

©2021 The Japanese Society of Veterinary Science



This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial No Derivatives (by-nc-nd) License. (CC-BY-NC-ND 4.0: <https://creativecommons.org/licenses/by-nc-nd/4.0/>)

At present, reference genes for qPCR in various mammalian tissues have been evaluated, such as pigs, cows, mice and humans, and so on [11, 12, 34]. Unfortunately, because FMD is so endangered that it's hard to get fresh tissues, to the best of our knowledge, no studies have been performed on reference gene screening for qPCR data normalization in this species to date. In this study, we selected 10 candidates of reference genes: beta-actin (*ACTB*), glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*), TATA box-binding protein (*TBP*), hypoxanthine phosphoribosyltransferase 1 (*HPRT1*), tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein zeta polypeptide (*YWHAZ*), hydroxymethylbilane synthase (*HMBS*), eukaryotic translation elongation factor 1 alpha 1 (*EEF1A1*), succinate dehydrogenase complex flavoprotein subunit A (*SDHA*), peptidylprolyl isomerase B (*PPIB*), and ubiquitin C (*UBC*) to evaluate their expression stability in the five organs (heart, lung, spleen, liver and kidney) of FMD. qPCR detection method based on SYBR Green I technology was established and the transcription stability was subsequently evaluated by geNorm [31], NormFinder [1], BestKeeper [26], Δ Ct method [28] and RefFinder [37], which aimed to provide scientific basic experimental data for the selection of reference genes for qPCR of FMD organs.

MATERIALS AND METHODS

Ethics statement and sample collection

The animal experiments were approved by the National Institute of Animal Health Care and Use Committee at Sichuan Agricultural University (approval number SYXK2019-187) and all institutional and national guidelines for the care and use of laboratory animals were followed in all process of animal care, samples collection and experiments.

Samples of the FMD were collected from Sichuan Institute of Musk Deer Breeding (Dujiangyan and Maerkang, China). Five adult captive FMDs that died by accident, including 2 females and 3 males, were dissected instantly. The heart, lung, spleen, liver and kidney samples of each FMD were immediately collected and next flash-frozen in liquid nitrogen until RNA isolation.

RNA extraction and cDNA synthesis

RNA of samples were extracted using RNAiso Plus (TaKaRa, Beijing, China). Three different parts from the parenchyma of each organ were taken, then mixed according to the organs for grinding, adding 1 ml RNAiso Plus for every 50 mg of tissue sample. Nano-drop 2000 spectrophotometer (ThermoFisher, Shanghai, China) was used to detect the yield and quality of RNA. The A260/280 and A260/230 ratios ranged from 1.96 to 2.08 and from 1.90 to 2.10, respectively, which indicated the samples were of good quality. According to the manufacturer's procedures, 1 μ g of RNA from each sample was primed with 50 pM Oligo (dT)₂₃VN and 50 ng of random hexamers respectively to synthesize complementary DNA (cDNA) in a final volume of 20 μ l by reverse transcription PCR (RT-PCR) using HiScript II 1st Strand cDNA Synthesis Kit (Vazyme, Nanjing, China). The cDNA was stored at -20°C . All vessels used in the whole process are treated with 0.1% diethyl oxydiformate (DEPC) water.

Reference gene selection and primer designing

Reference genes with different functions, including *18S rRNA*, *ACTB*, *GAPDH*, *TBP*, *HPRT1*, *YWHAZ*, *HMBS*, *EEF1A1*, *SDHA*, *PPIB*, and *UBC* was selected, all of which had been tested for stability previously in bovids [15, 20, 40, 42]. The information of candidate reference genes is shown in Table 1.

Based on our previous transcriptome analysis, three commonly used reference genes (*18S rRNA*, *ACTB* and *GAPDH*) in FMD were cloned (data not shown). For other genes, to provide a set of species-specific primers for reference genes in FMD, Basic Local Alignment Search Tool (BLAST) of National Center of Biotechnology Information (NCBI) was used to obtain the sequences by comparing the homologous sequences of *Bos taurus* with published FMD whole-genome shotgun database. Primer pairs were designed by primer5.0 software and sent to Sangon Biotech (Shanghai, China) Co., Ltd. for synthesis.

Real-time quantitative PCR

For PCR efficiency, cDNA samples were 10-fold made into eight dilution series as the standard substance, and qPCR was performed using the standard substance as the template. Each dilution was amplified in triplicate PCR amplifications and plotted as mean values to generate a standard curve. The single threshold mode was used to calculate the cycle threshold (Ct) value based on the threshold crossing point of individual fluorescent trace. According to the software ABI 7500 Software v2.0.1, the melting curve, standard curve, amplification efficiency and the linear correlation between template concentration and Ct value were automatically generated. Three replicates were set for each sample, with negative control and blank control. The reaction volume was 20 μ l, including SYBR Green Mix (TaKaRa) 10 μ l, forward and reverse primers (10 pM/l) 1 μ l, cDNA template 1 μ l, ddH₂O 7 μ l. The amplification program was 95 $^{\circ}\text{C}$ for 3 min, 40 cycles of 95 $^{\circ}\text{C}$ for 10 sec and annealing 30 sec. Followed melting curve analysis was 95 $^{\circ}\text{C}$ for 10 sec and continuous fluorescent acquisition with heating samples from 65 $^{\circ}\text{C}$ to 95 $^{\circ}\text{C}$.

The expression levels of the 11 genes in the heart, lung, spleen, liver, and kidney of FMD were analyzed according to the established qPCR detection methods. A 10-fold dilution of cDNA was used as a template, PCR was performed using the same volume of cDNA sample for each gene to record the Ct value.

Data analysis

The data in this study were statistically analyzed by using SPSS 22.0, unless specified. The Ct value of each gene in the heart, lung, spleen, liver, and kidney of FMD was performed stability analysis by four classic programs, geNorm, NormFinder, BestKeeper and Δ Ct. Furthermore, the overall final ranking was provided by RefFinder. The figs. were generated by using

Table 1. Description of candidate reference genes analyzed in this study

Gene	Full gene name	Function	<i>Bos taurus</i> ^{a)}	FMD ^{b)}
<i>18S rRNA</i>	18S ribosomal RNA	Ribosomal RNA	-	MT757929
<i>ACTB</i>	Beta-actin	Cytoskeletal structural protein	-	MT767758
<i>GAPDH</i>	Glyceraldehyde-3-phosphate dehydrogenase	Carbohydrate metabolism	-	MT767759
<i>TBP</i>	TATA box-binding protein	General RNA polymerase II transcription factor	NM_001075742.1	SPDX01012420.1: 3442–4203
<i>HPRT1</i>	Hypoxanthine phosphoribosyltransferase 1	Purine synthesis in salvage pathway	NM_001034035.2	SGQJ01001359.1: 570650–572055
<i>YWHAZ</i>	Tyrosine 3-monooxygenase/tryptophan 5-Monooxygenase activation protein zeta polypeptide	Signal transduction by binding to phosphorylated serine residues on a variety of signaling molecules	NM_174814.2	SPDX01002592.1: 129239–129806
<i>HMBS</i>	Hydroxymethylbilane synthase mRNA	Heme synthesis, porphyrin metabolism	NM_001046207.1	SGQJ01001526.1: 1541500–1542902
<i>EEF1A1</i>	Eukaryotic translation elongation factor 1 alpha 1	The enzymatic delivery of aminoacyl tRNAs to the ribosome	NM_174535.2	SPDX01009820.1: 188730–190117
<i>SDHA</i>	Succinate dehydrogenase complex flavoprotein subunit A	Electron transporter in the TCA cycle and respiratory chain	NM_174178.2	SGQJ01002011.1: 1148542–1148875
<i>PPIB</i>	Peptidylprolyl isomerase B	Protein folding catalyst	NM_174152.2	SPDX01001493.1: 203348–203991
<i>UBC</i>	Ubiquitin C	Protein degradation	NM_001206307.2	SGQJ01047728.1: 1240–2158

^{a)}The accession numbers in Genbank. Gene sequences of *Bos taurus* were used to search the published forest musk deer whole-genome shotgun database to identify gene sequences not annotated in their genome. ^{b)}Forest musk deer (FMD). The accession numbers in Genbank or gene full/partial sequences location in the forest musk deer genome. Sequences of forest musk deer were used for primer design.

Graphpad Prism 9.0.0 (GraphPad Software, San Diego, CA, USA).

The standard deviation of the expression ratio of two genes that are not jointly regulated and have the same expression rate in all samples is defined as a pairwise variation by geNorm [36]. The analysis method of geNorm is to first calculate the relative quantitative data of each candidate reference gene using the $2^{-\Delta Ct} (\Delta Ct = Ct_{(sample)} - Ct_{(minimum)})$ method, and then import the relative quantitative data into the geNorm program for calculation to obtain the average pairwise variation of each gene with all other reference genes as gene-stability measure (M) [31]. The M value of a stable gene that can be used for data normalization should be lower than 1.5. The lower the M value is, the more suitable it is as the reference gene of qPCR, therefore genes with the highest M values have the least stable expression. The genes with the highest M value are ranked by stepwise elimination until a combination of two constitutively expressed reference genes were screened out [31]. geNorm can also analyze the optimal number of reference genes suitable for normalization of qPCR data [22].

NormFinder is a software by which the expression stability S for each candidate reference genes can be directly output. The calculation and analysis method of NormFinder program is similar to geNorm program. Both use $2^{-\Delta Ct}$ method to calculate the relative quantitative data of reference gene. The NormFinder program is introduced to obtain the stability value of reference gene expression. The lower the stability value of the gene, the more stable [24].

BestKeeper is an excel-based tool used to calculate the coefficients of correlation (R) standard deviation (SD) and coefficients of variation (CV) for reference genes. Unlike the previous two software, it evaluates gene expression stability using raw Ct value data. By comparing the output values, stable reference genes were identified. In general, the higher the R value, the lower the SD and CV value, the more stable the reference gene. Furthermore, the SD value of the reference gene given by BestKeeper should be lower than 1 and any studied genes with the SD higher than 1 can be considered inconsistent [33].

The ΔCt method is a simple approach that only uses an excel sheet to conduct the pairwise comparison of the Ct value. It works by comparing the relative expression of pairs and calculating the average SD for each gene, and the gene with lowest average SD is considered as the most stable reference gene [28].

After obtaining the rankings given by each algorithms, we determine the final ranking by RefFinder. RefFinder is an integration tool based on the four major computational programs. According to the rankings from each program, it assigns an appropriate weight to an individual gene and calculated the geometric mean of their weights for the overall final ranking [37].

RESULTS

Amplification efficiency and primer specificity of qPCR

The specific primer sequences are shown in Table 2. In the melting curve analysis of 11 primer pairs, it was observed that the standard single peak of the melting curve appeared at 79–86°C, but no peak appeared at other positions, indicating that the primer

Table 2. The primers and their annealing temperature (T_a), PCR efficiency, slope, and regression coefficient of selected candidate reference genes

Gene	Primer sequence forward/reverse (5'-3')	T_a (°C)	PCR efficiency (%)	Slope	Regression coefficient (r^2)	Product length (bp)
<i>18S rRNA</i>	CGGTCCCCAACTTCTT/CGGACATCTAAGGGCATCA	58	101.4	-3.289	0.997	86
<i>ACTB</i>	GAATCCTGCGGCATTACG/TCTTCATCGTGCTGGGTGC	58	102.9	-3.254	0.994	172
<i>GAPDH</i>	CTCAACGGGAAGCTCACTG/CTTGGCAGGTTTCTCCAGG	57	97.1	-3.395	1.000	93
<i>TBP</i>	AGTTTGAGCCTCTACAGTGACG/GAAGCCTTTGAGAATCTACCC	58	94.2	-3.469	0.999	282
<i>HPRT1</i>	GAGGATTTGGAGAGGGTGTTTAT/ATAGCCCCCTTGAGCACA	58	109.9	-3.105	0.994	129
<i>YWHAZ</i>	ATCCCCAACGCTTCACAA/TTGGTATGCTTGCTGTGACTG	58	107.7	-3.151	0.998	135
<i>HMBS</i>	TGGGTGAAACACAACAGCATC/CCACCACGGGGGACAAGA	58	101.3	-3.292	0.997	208
<i>EEF1A1</i>	GTTTGCCTCTCCAGGATGTCTA/ATTGCCACGACGGACATCT	58	99.3	-3.338	0.998	233
<i>SDHA</i>	CAATCAAAAAGCGGTTACGA/GACTGACTGTGCCACTGTTCCC	58	93.3	-3.494	0.998	123
<i>PPIB</i>	GCAGCACGGAGCCCTATGG/CGATGGTCGGGACAAGCCT	58	92.5	-3.517	0.999	183
<i>UBC</i>	CAGGCTCTCTCCACCA/TTGGGAGTTGTGCGTTA	58	90.6	-3.571	0.999	91

pairs has high specificity. Using qPCR amplification by gradient dilution of cDNA samples as the template, the standard curve of 11 genes was drawn. The amplification efficiency of the standard curves is all above 90% shown in Table 2. The regression coefficients of the standard curve are also more than 0.99, showing good linear correlation. Therefore, all the primer pairs of candidate reference genes meet the requirements of expression analysis using quantitative analysis.

Expression profiling of candidate reference genes

The expression of 11 candidate reference genes in heart, liver, spleen, lung and kidney was assayed by SYBR Green I qPCR detection methods. In the obtained results, *18S rRNA* showed high abundance expression with an average Ct value of 10.59, which was significantly higher than the mRNA expression levels of other genes. In view of the absence in purified mRNA samples and the high abundance compared to mRNA transcripts, which caused errors easily in qPCR analysis, it was excluded from our further study. The distribution of Ct values except *18S rRNA* was displayed by boxplot shown in Fig. 1. All Ct values ranged from 18.65 to 34.45, which provided an overview of the variation in the expression of the candidate reference genes. According to the overall expression level, the remaining 10 candidate reference genes can be divided into two class: a moderate abundant class (*ACTB*, *GAPDH*, *YWHAZ*, *EEF1A1*, *SDHA* and *PPIB* with mean Ct values between 18 and 25) and low abundant class (*TBP*, *HPRT1*, *HMBS* and *UBC* with mean Ct values greater than 25). The expression level of *HMBS* with an average Ct value of 30.96 was the lowest among them. Overall, the maximum and minimum Ct values of candidate reference genes have the smallest difference in the heart by roughly observing. Although the Ct value of each gene was different, the distribution of Ct value of the five organs was similar in general.

geNorm analysis

The expression data of 10 candidate reference genes in heart, lung, spleen, liver and kidney of FMD were all analyzed using geNorm. After progressively removing the genes with the highest M values, the final ranking revealed that the optimal reference genes differed for each organ. According to the results of geNorm analysis shown in Fig. 2, *HPRT1* and *YWHAZ* were the most stably expressed genes in heart; *ACTB* and *HMBS* were the most stable genes in liver; *ACTB* and *GAPDH* were the most stable genes in spleen; *HPRT1* and *PPIB* were the most stable genes in lung; *ACTB* and *SDHA* were the most stable genes in kidney. Simultaneously, the stability of reference genes in five organ combinations of FMD was evaluated by geNorm, and the two most stable reference genes screened for use in combination were *TBP* and *HMBS*, followed by *YWHAZ*, *HPRT1*, *EEF1A1*, *ACTB*, *PPIB*, *UBC*, *GAPDH* and *SDHA*. In addition, it can be seen that the average expression stability values were all less than the threshold value of 1.5 after eliminating the reference gene with the worst stability for the first time.

NormFinder analysis

NormFinder can directly output the most stable expression reference gene. According to the judgment criterion that the lower the stability value is, the more stable the gene expression is, *SDHA*, *ACTB*, *YWHAZ*, *HPRT1*, *YWHAZ* and *HMBS* are considered as the most stable genes in heart, liver, spleen, lung, kidney and combined group by NormFinder respectively. It can be found that for the most stable genes, NormFinder results showed a difference from the geNorm. However, for the least stable genes, NormFinder analysis showed that *SDHA* was the most unstable gene in the heart of FMD, *UBC* was the most unstable gene in the liver, spleen, lung, and kidney, and *SDHA* was the most unstable gene in the combination group, which was completely consistent with the results of geNorm. In addition, NormFinder also outputs a stability value for the optimal combination of the two genes when group identifiers are included. Therefore, when all the data were inputted and each organ was taken as a sample group, NormFinder produced the best combination of two genes, *HPRT1* and *YWHAZ*, with a stability value of 0.276. The result is shown in Fig. 3.

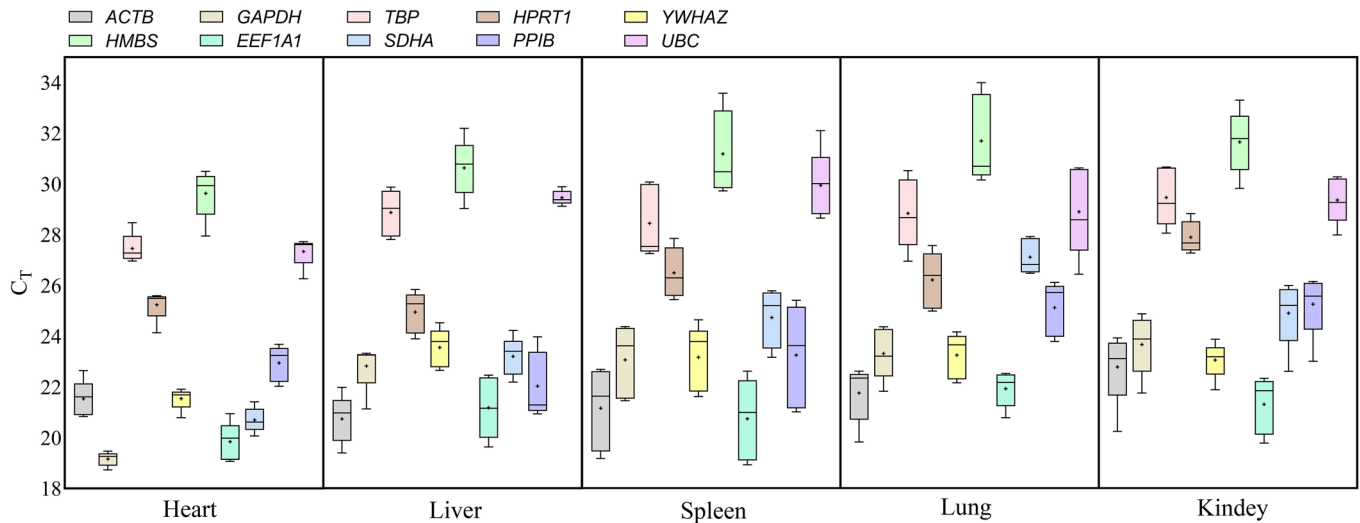


Fig. 1. The distribution of reference gene expression levels in each organ. Means and medians are indicated by asterisks and lines, respectively. The boxes encompass the 25th to the 75th percentiles. Whisker caps denote the maximum and minimum values.

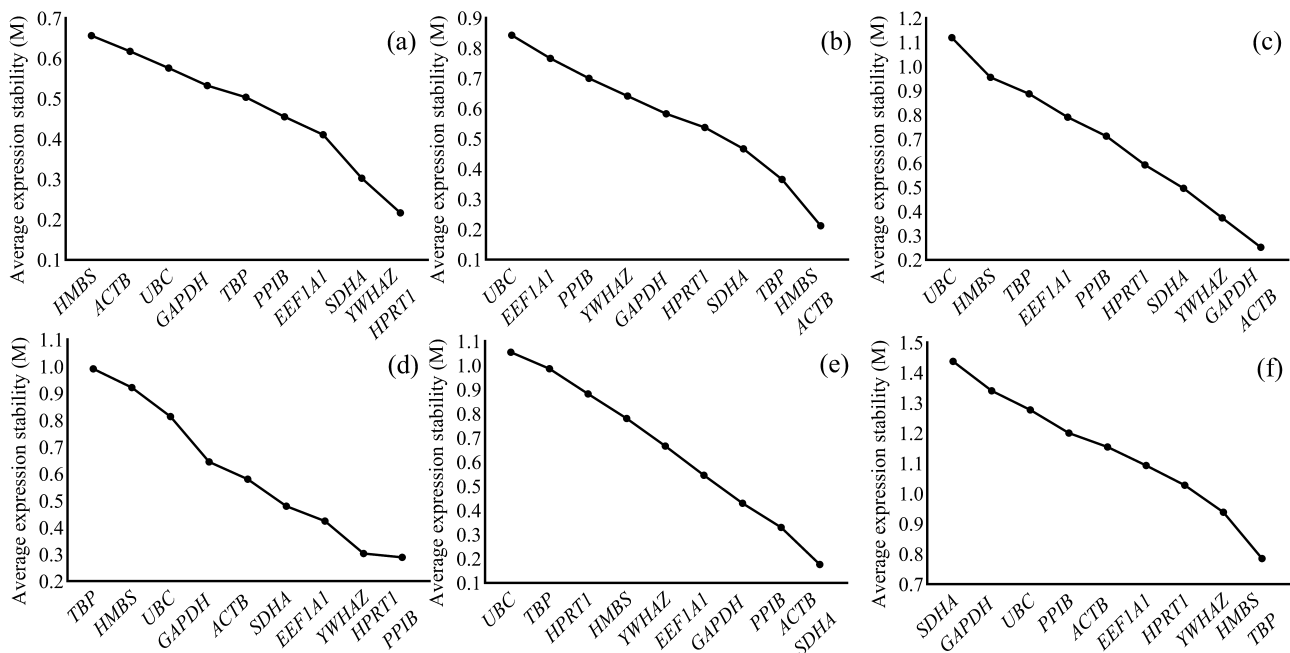


Fig. 2. Gene expression stabilities and rankings of reference genes as calculated by geNorm. (a), (b), (c), (d), (e) and (f) represent the rank-order of gene expression stability of heart, liver, spleen, lung, kidney, and combined group, respectively.

BestKeeper analysis

BestKeeper showed the results with good consistency and certain significance ($P < 0.05$) after repeated paired correlation analysis and correlation analysis of candidate reference genes. All SD, CV and R values calculated by BestKeeper were shown in Fig. 4. According to the criteria, *GAPDH* (SD=0.21), *UBC* (SD=0.22), *HPRT1* (SD=0.81), *EEF1A1* (SD=0.55), *YWHAZ* (SD=0.47) were ranked by BestKeeper as the most stable gene in heart, liver, spleen lung and kidney respectively. For the combined group, *YWHAZ* was the most stable gene with the SD value was 0.93, and SD value of other genes in the group were all greater than 1.

ΔC_t method analysis

The expression stability of reference genes was also evaluated via the average SD calculated by ΔC_t method. According to the principle that the smaller the average SD value, the less expression variability, the results in Table 3 showed that the most stable expression genes in the heart, liver, spleen, lung, kidney and the combined group were *SDHA*, *ACTB*, *YWHAZ*, *YWHAZ*, *YWHAZ* and *YWHAZ*, respectively (average SD of 0.50, 0.67, 0.83, 0.74, 0.84 and 1.24).

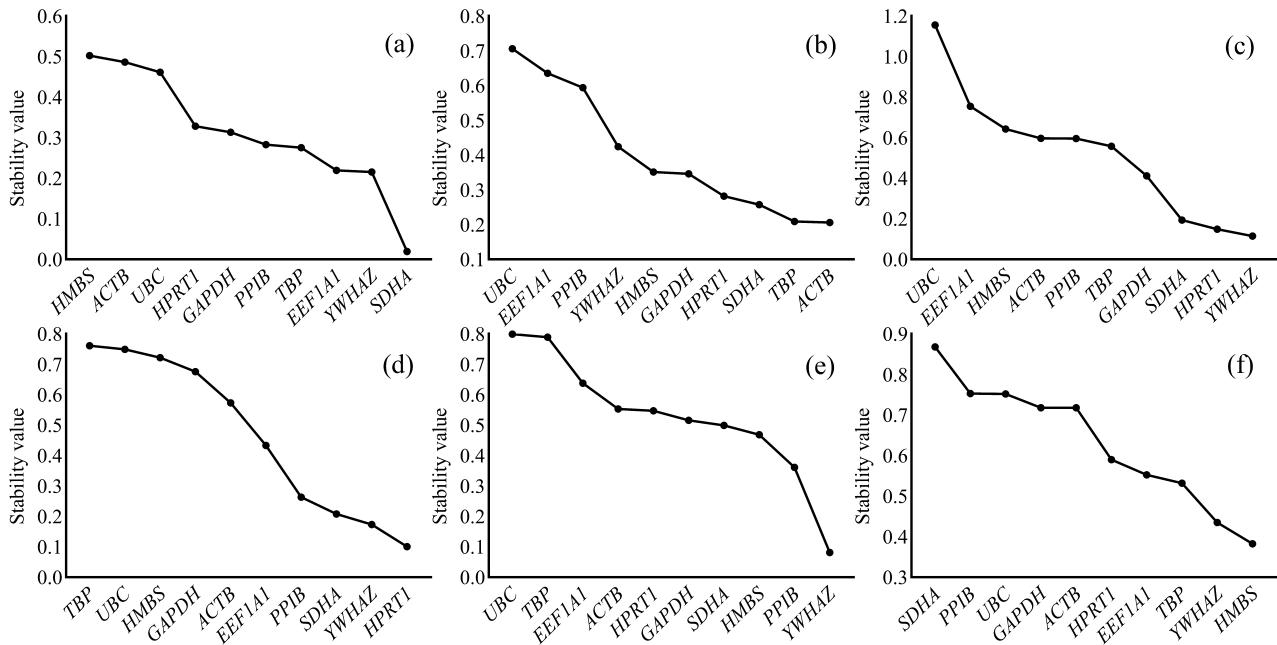


Fig. 3. Gene expression stabilities and rankings of reference genes as calculated by NormFinder. (a), (b), (c), (d), (e) and (f) represent the rank-order in NormFinder of gene expression stability of heart, liver, spleen, lung, kidney, and combined group, respectively.

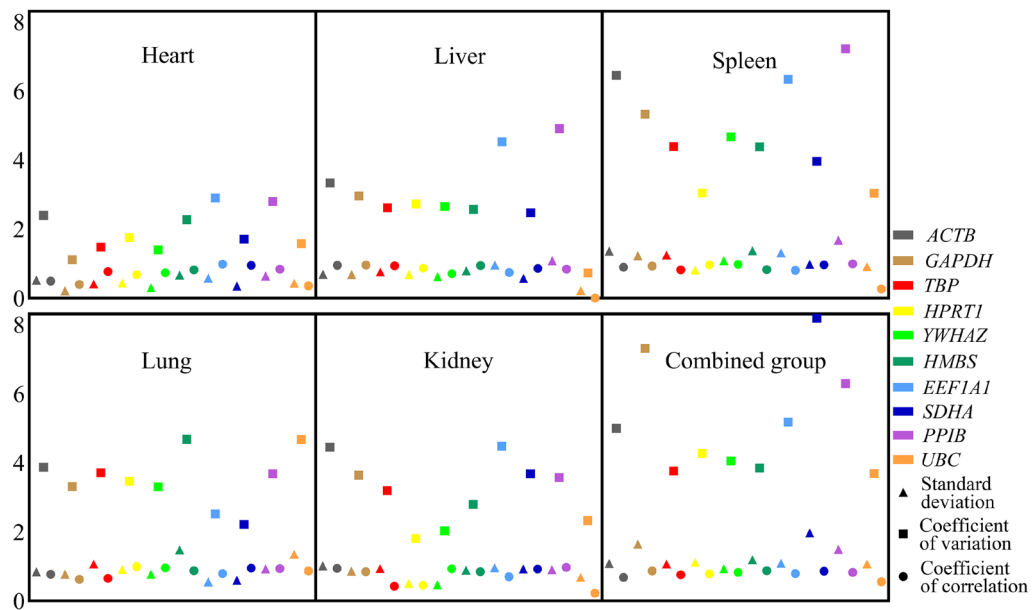


Fig. 4. Reference genes stability related parameters generated by BestKeeper. BestKeeper determine the most stable gene by repeated pairwise correlation analysis. Different colors represented different genes, and different symbols represented standard deviation (SD), coefficient of variation (CV) and coefficient of correlation (R) generated by BestKeeper respectively. Their positions on the vertical axis correspond to the values. Within each correlation appeared in the Figure, the probability P -value <0.05 .

Table 3. The average standard deviation (SD) calculated by Δ Ct method

Gene	ACTB	GAPDH	TBP	HPRT1	YWHAZ	HMBS	EEF1A1	SDHA	PPIB	UBC
Heart	0.81	0.65	0.62	0.64	0.57	0.81	0.58	0.50	0.65	0.79
Liver	0.67	0.79	0.69	0.74	0.86	0.76	1.05	0.72	1.00	1.15
Spleen	1.09	0.94	1.13	0.89	0.83	1.21	1.28	0.88	1.16	1.78
Lung	1.03	1.12	1.27	0.74	0.74	1.23	0.92	0.81	0.79	1.26
Kidney	1.00	0.99	1.30	1.07	0.84	1.03	1.14	0.94	0.91	1.33
Combined group	1.47	1.54	1.31	1.33	1.24	1.25	1.34	1.83	1.50	1.59

The optimal number of reference genes

The optimal number of reference genes needed for accurate normalization was determined by geNorm on the basis of the average pairwise variation ($V_{n/n+1}$) value. geNorm software takes 0.15 as the default value, below this value, inclusion of an additional reference gene offers little further improvement to the data normalization, that is, when $V_{n/n+1}$ is less than 0.15, it does not need to use the number of reference genes $\geq n+1$ as parameters. As shown in Fig. 5, the value of $V_{2/3}$ fell below 0.15 for heart, liver, spleen, lung and kidney groups. It meant that for qPCR analysis in a single organ, the optimal number of reference genes was 2, and there was no need to select a third candidate gene as the internal reference, which could meet the requirements of accuracy. However, in the combined group, the value of $V_{2/3}$ was greater than 0.15, and it was not until $V_{8/9}$ that the value was less than 0.15, indicating that two reference genes were insufficient for accurate normalization among samples. Therefore, when detecting the expression level of target genes in five organs, the optimal number of reference genes was 8.

Comprehensive ranking of optimal reference genes

The results of our analysis of the reference genes stability using different algorithms were broadly similar, but the rankings were slightly different. Therefore, we used RefFinder to calculate the comprehensive ranking of the four algorithms. The results of the calculation of candidate genes stability in the analysis are shown in Table 4. According to the comprehensive ranking of the four algorithms, the selected reference genes with the most stable expression in heart, liver and lung were *SDHA*, *ACTB* and *HPRT1*, respectively. In addition, *YWHAZ* was the most stable reference gene in spleen, kidney and combined group. With the optimal number of reference genes considered, the recommended genes in heart are *SDHA* and *YWHAZ*, while in liver are *ACTB* and *SDHA*; in spleen and lung are *YWHAZ* and *HPRT1*; in kidney are *YWHAZ* and *PPIB*; in combined group are *YWHAZ*, *TBP*, *HMBS*, *HPRT1*, *EEF1A1*, *ACTB*, *UBC* and *PPIB*. These results suggest that tissue is an important variation factor affecting gene expression stability. Thus, this demonstrates the necessity of using reference genes based on tissue analysis.

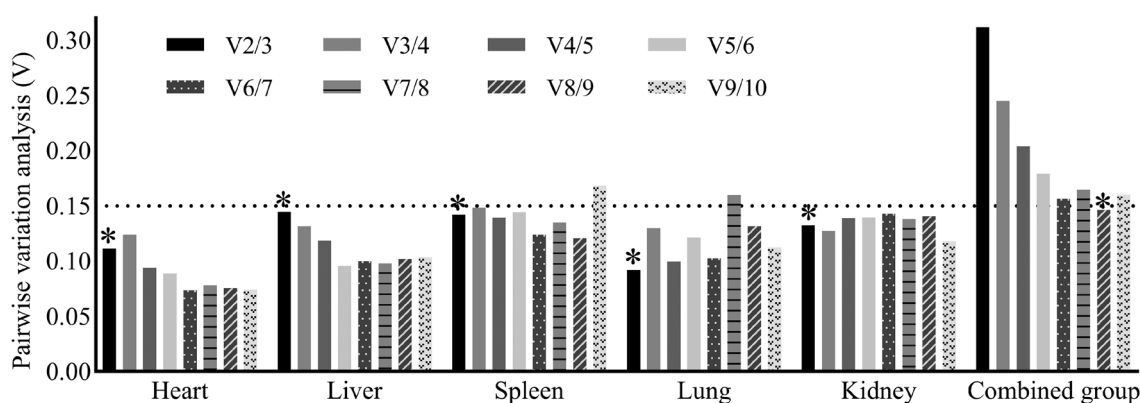


Fig. 5. Determination of the optimum number of genes required for normalisation. The dotted line indicates the cut-off value of 0.15 for pairings of pairwise variation analysis value (V). When $V_{n/n+1} < 0.15$, n is the optimal number of reference genes. The optimal number of reference genes in each group was marked with an asterisk.

Table 4. The comprehensive ranking of the expression stability

	Recommended comprehensive ranking									
	1	2	3	4	5	6	7	8	9	10
Heart	<i>SDHA</i> (1.73)	<i>YWHAZ</i> (1.86)	<i>EEF1A1</i> (3.72)	<i>HPRT1</i> (3.81)	<i>GAPDH</i> (3.81)	<i>TBP</i> (4.43)	<i>PPIB</i> (6.59)	<i>UBC</i> (7.33)	<i>ACTB</i> (8.91)	<i>HMBS</i> (9.21)
Liver	<i>ACTB</i> (1.57)	<i>SDHA</i> (2.91)	<i>TBP</i> (3.03)	<i>HMBS</i> (3.94)	<i>HPRT1</i> (4.47)	<i>GAPDH</i> (5.18)	<i>UBC</i> (5.62)	<i>YWHAZ</i> (5.66)	<i>PPIB</i> (8.46)	<i>EEF1A1</i> (9.00)
Spleen	<i>YWHAZ</i> (1.86)	<i>HPRT1</i> (2.34)	<i>SDHA</i> (2.91)	<i>GAPDH</i> (2.99)	<i>ACTB</i> (4.09)	<i>TBP</i> (6.16)	<i>UBC</i> (6.69)	<i>PPIB</i> (7.09)	<i>EEF1A1</i> (7.94)	<i>HMBS</i> (8.49)
Lung	<i>HPRT1</i> (1.86)	<i>YWHAZ</i> (2.06)	<i>PPIB</i> (3.03)	<i>EEF1A1</i> (3.16)	<i>SDHA</i> (3.31)	<i>ACTB</i> (5.73)	<i>GAPDH</i> (6.09)	<i>HMBS</i> (8.71)	<i>UBC</i> (8.74)	<i>TBP</i> (9.46)
Kidney	<i>YWHAZ</i> (1.57)	<i>PPIB</i> (2.91)	<i>SDHA</i> (3.03)	<i>GAPDH</i> (4.23)	<i>ACTB</i> (4.33)	<i>HMBS</i> (5.01)	<i>HPRT1</i> (5.09)	<i>EEF1A1</i> (7.33)	<i>UBC</i> (7.4)	<i>TBP</i> (8.74)
Combined group	<i>YWHAZ</i> (1.32)	<i>TBP</i> (2.28)	<i>HMBS</i> (2.30)	<i>HPRT1</i> (4.43)	<i>EEF1A1</i> (5.00)	<i>ACTB</i> (5.42)	<i>UBC</i> (6.00)	<i>PPIB</i> (7.24)	<i>GAPDH</i> (8.49)	<i>SDHA</i> (10.00)

The value after the gene name in the table represented the geometric average of weights calculated by RefFinder for each algorithm after the appropriate weights had been assigned. The reference genes ultimately recommended for use in different organs were bolded.

DISCUSSION

FMD is listed in CITES Appendix II, and it is a Class I key protected wild animal in China [7]. With the continuous expansion of the demand for natural musk in the international market, the FMD breeding industry has developed rapidly, and the research on FMD has been further deepened. At present, gene expression analysis is the most commonly used method in molecular biology research. In the process of in-depth study of FMD, the quantitative techniques for mRNA expression, including RNA blotting, gene microarray and qPCR, etc., all require the selection of reference genes to correct the expression results of target genes. The ideal reference genes should be stable expression in the studied tissue, cell and experimental conditions. But in fact, such an absolutely stable gene does not exist [21, 42]. Therefore, in order to obtain meaningful experimental data, researchers should select appropriate reference genes according to their special tissues and experimental conditions.

There have been many reports related to mammalian reference genes. Pérez *et al.* [25] screened the reference genes in bovine muscle tissue and showed that the three genes *SF3A1*, *EEF1A1* and *HMBS* are the most stable expression. Lisowski *et al.* [20] evaluated the expression stability of six internal reference genes in cattle liver, kidney, pituitary and thyroid. The results have showed that there are different optimal reference genes in different tissues. The most stable genes in the liver tissue of cattle are *ACTB* and *TBP*, while in the kidney are *GAPDH* and *YWHAZ*; in the pituitary are *GAPDH* and *SDHA*; in the thyroid are *TBP* and *HPRT1*. Zhu *et al.* [42] reported the expression of reference genes in the two muscles of *Capra hircus* at different stages of development. The results have showed that *PPIB* and *HMBS* are the most stable reference genes in different developmental stages of skeletal muscle. Zhang *et al.* [40] completed the evaluation of reference genes in different tissues of Boer goat. The results showed that *18S rRNA*, *ACTB*, *GAPDH*, *HMBS*, *TBP*, and *HPRT1* can only be used as reference gene in appropriate tissues. Finally, *18S rRNA*, *TBP* and *HMBS* were considered as the best reference combination for calibrating gene expression in goat tissues. More and more research indicated that the choice of reference gene should be changed with different experimental conditions and research types. Therefore, before using qPCR to analyze the expression level, screening suitable reference genes for research subjects is critical.

In the present study, *geNorm*, *BestKeeper* and ΔCt method are the main classical algorithms used to screen reference genes. Although *geNorm* has the function of screening the optimal number of reference genes, it is susceptible to identify co-regulated genes as optimal reference genes as they would show a constant ratio. *NormFinder* instead uses a model based approach to analyze the variance in the expression data. It takes into account both intra-group and inter-group variation which makes it more robust against co-expressed genes. *BestKeeper* employing parametric methods, but heterogeneous variance between groups of differently expressed genes invalidates the use of Pearson correlation coefficient [26]. ΔCt method is the simplest to use but suitable for some instances where parameters cannot be quantified [28].

Taking advantage of these four programs, we not only evaluated the stability of candidate genes in heart, liver, spleen, lung and kidney, but also evaluated the overall stability of candidate genes in five organs. However, the results of each algorithm are not completely consistent as well, and other similar studies have also reflected this finding [8, 9, 32]. Since each algorithm has its own advantages and disadvantages, it is impossible to uniformly decide which is better and it is difficult to decide which to adopt likewise. Therefore, we chose the *ReFinder* integration tool to calculate the final ranking based on these four algorithms. From the results, *UBC* had never ranked first among the five organs, and had always ranked in the bottom four, so it was considered unstable and unsuitable for data normalization processing of the five organs of FMD. This was consistent with a previously reported study [13]. At the same time, their results showed that *ACTB* was one of the three most stably expressed genes in different liver sample types. According to the results of Lisowski, *et al.* [20], *ACTB* was also one of the two most stably expressed genes in liver, which jointly confirmed the reliability with our results of *ACTB* as an reference gene in liver. And in his study, the expression stability of *YWHAZ* in the kidney ranked first, which was also consistent with our results. Besides, in the previous study by Zhang, *et al.* [40], except *18S rRNA*, the top three reference gene stability in different tissues of goats, *YWHAZ*, *TBP* and *HMBS*, also ranked in the top three in our combined group. In addition to the above, from the results of five methods it can be also seen that the relatively stable gene in each organ due to expression variations increases in all five organs consideration and shows instability, such as *SDHA*. However, genes that vary widely in each organ presented relatively stable results in the combined group due to their lesser tissue-specific expression differences, such as *HMBS*. Therefore, we recommend selecting appropriate genes for reference in individual organs rather than using common reference genes for all organs.

When selecting reference genes, the optimum number of reference genes should also be taken into consideration. The use of just a single reference gene may result in a more than 6-fold erroneous normalization and it is therefore recommended to use more than one reference gene [31]. From the results of *geNorm*, two reference genes can be used directly in the qPCR analysis of a single organ, while in combined group, more reference genes need to be added to achieve the desired accuracy. Therefore, it is more appropriate to carry out quantitative analysis in a single organ for data normalization as far as possible under convenient circumstances.

Furthermore, the close expression level of the reference gene and the target gene is also one of the selection indicators [16]. In general, a good reference gene should have an expression level that is relatively targeted in a similar range to the gene. For example, we know that the Ct value is inversely proportional to the level of gene transcription, that is, higher Ct values correspond to lower expression abundance. From the obtained reference gene expression profile, it can be seen that *HMBS* is suitable for the reference of low abundance genes. Perhaps *18S rRNA* is suitable for the reference of high abundance genes, however, it was excluded from our further study. On the one hand, there is no poly (A) tail follow the rRNA terminal, so using Oligo-dT primers only for reverse transcription is not feasible. On the other hand, *18S rRNA* was highly expressed compared to the target mRNA

transcript, making it difficult to deduct the baseline value in the analysis of qPCR data. Therefore, selecting *18S rRNA* as the reference gene may increase the result error [5].

Overall, in this study, the expression stability of reference genes in FMD heart, lung, liver, spleen and kidney was evaluated by geNorm, NormFinder, BestKeeper and ΔCt method, and the optimal reference gene in the combination of all five organs was given, which can be used as a reference for the selection of reference genes for qPCR of FMD organs. The analyses showed that tissue is an important variability factor in genes expression stability. Consequently, for other specific samples not shown in this study, reference genes should also be re-selected before data standardization to reliably interpret the data. In the future, our study will further develop on the basis of this study to screen more suitable reference genes for FMD, so as to provide more theoretical basis for artificial breeding and disease prevention of FMD.

CONFLICT OF INTEREST. Authors declare that they have no conflict of interest.

ACKNOWLEDGMENTS. This work was funded by the Science and Technology Achievements Transfer Project in 2020 from Scientific Research Institutions of Science and Technology Department of Sichuan Province, Sichuan, China (No. 2020JDZH0024). We thank the Animal Quarantine Laboratory of the College of Veterinary Medicine of Sichuan Agricultural University (Wenjiang, China) and Sichuan Institute of Musk Deer Breeding (Dujiangyan, China) for providing experimental conditions for the experiment. We extend our thanks to the hard work of the experimental contributors.

REFERENCES

1. Andersen, C. L., Jensen, J. L. and Ørntoft, T. F. 2004. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res.* **64**: 5245–5250. [Medline] [CrossRef]
2. Brym, P., Ruś, A. and Kamiński, S. 2013. Evaluation of reference genes for qRT-PCR gene expression studies in whole blood samples from healthy and leukemia-virus infected cattle. *Vet. Immunol. Immunopathol.* **153**: 302–307. [Medline] [CrossRef]
3. Bustin, S. A. 2002. Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. *J. Mol. Endocrinol.* **29**: 23–39. [Medline] [CrossRef]
4. Cai, Y., Yang, J., Wang, J., Yang, Y., Fu, W., Zheng, C., Cheng, J., Zeng, Y., Zhang, Y., Xu, L., Ren, Y., Lu, C. and Zhang, M. 2020. Changes in the population genetic structure of captive forest musk deer (*Moschus berezovskii*) with the increasing number of generation under closed breeding conditions. *Animals (Basel)* **10**: 255. [Medline] [CrossRef]
5. Cappelli, K., Felicetti, M., Capomaccio, S., Spinsanti, G., Silvestrelli, M. and Supplizi, A. V. 2008. Exercise induced stress in horses: selection of the most stable reference genes for quantitative RT-PCR normalization. *BMC Mol. Biol.* **9**: 49. [Medline] [CrossRef]
6. Chapman, J. R. and Waldenström, J. 2015. With reference to reference genes: a systematic review of endogenous controls in gene expression studies. *PLoS One* **10**: e0141853. [Medline] [CrossRef]
7. Fan, Z., Li, W., Jin, J., Cui, K., Yan, C., Peng, C., Jian, Z., Bu, P., Price, M., Zhang, X., Shen, Y., Li, J., Q, W. and Yue, B. 2018. The draft genome sequence of forest musk deer (*Moschus berezovskii*). *Gigascience* **7**: giy038. [Medline] [CrossRef]
8. Gao, Z., Deng, W. and Zhu, F. 2019. Reference gene selection for quantitative gene expression analysis in black soldier fly (*Hermetia illucens*). *PLoS One* **14**: e0221420. [Medline] [CrossRef]
9. Gong, Z. K., Wang, S. J., Huang, Y. Q., Zhao, R. Q., Zhu, Q. F. and Lin, W. Z. 2014. Identification and validation of suitable reference genes for RT-qPCR analysis in mouse testis development. *Mol. Genet. Genomics* **289**: 1157–1169. [Medline] [CrossRef]
10. Hassanin, A. and Douzery, E. J. 2003. Molecular and morphological phylogenies of ruminantia and the alternative position of the moschidae. *Syst. Biol.* **52**: 206–228. [Medline] [CrossRef]
11. Hildyard, J. C. W., Finch, A. M. and Wells, D. J. 2019. Identification of qPCR reference genes suitable for normalizing gene expression in the mdx mouse model of Duchenne muscular dystrophy. *PLoS One* **14**: e0211384. [Medline] [CrossRef]
12. Jacob, F., Guertler, R., Naim, S., Nixdorf, S., Fedier, A., Hacker, N. F. and Heinzelmänn-Schwarz, V. 2013. Careful selection of reference genes is required for reliable performance of RT-qPCR in human normal and cancer cell lines. *PLoS One* **8**: e59180. [Medline] [CrossRef]
13. Janovick-Guretzy, N. A., Dann, H. M., Carlson, D. B., Murphy, M. R., Loor, J. J. and Drackley, J. K. 2007. Housekeeping gene expression in bovine liver is affected by physiological state, feed intake, and dietary treatment. *J. Dairy Sci.* **90**: 2246–2252. [Medline] [CrossRef]
14. Jie, H., Zhang, P., Xu, Z., Mishra, S. K., Lei, M., Zeng, D., Zhao, G. and Li, D. 2019. microRNA and other small RNA sequence profiling across six tissues of Chinese forest musk deer (*Moschus berezovskii*). *BioMed Res. Int.* **2019**: 4370704. [Medline] [CrossRef]
15. Kaur, R., Sodhi, M., Sharma, A., Sharma, V. L., Verma, P., Swami, S. K., Kumari, P. and Mukesh, M. 2018. Selection of suitable reference genes for normalization of quantitative RT-PCR (RT-qPCR) expression data across twelve tissues of riverine buffaloes (*Bubalus bubalis*). *PLoS One* **13**: e0191558. [Medline] [CrossRef]
16. Kozera, B. and Rapacz, M. 2013. Reference genes in real-time PCR. *J. Appl. Genet.* **54**: 391–406. [Medline] [CrossRef]
17. Li, C., Hu, L., Wang, X., Liu, H., Tian, H. and Wang, J. 2019. Selection of reliable reference genes for gene expression analysis in seeds at different developmental stages and across various tissues in *Paeonia ostii*. *Mol. Biol. Rep.* **46**: 6003–6011. [Medline] [CrossRef]
18. Li, L., Wang, B. B., Ge, Y. F. and Wan, Q. H. 2014. Major histocompatibility complex class II polymorphisms in forest musk deer (*Moschus berezovskii*) and their probable association with purulent disease. *Int. J. Immunogenet.* **41**: 401–412. [Medline] [CrossRef]
19. Li, Y., Wang, K., Chen, L., Zhu, X. and Zhou, J. 2016. Quantification of mRNA levels using real-time polymerase chain reaction (PCR). *Methods Mol. Biol.* **1406**: 73–79. [Medline] [CrossRef]
20. Lisowski, P., Pierzchała, M., Gościak, J., Pareek, C. S. and Zwierzchowski, L. 2008. Evaluation of reference genes for studies of gene expression in the bovine liver, kidney, pituitary, and thyroid. *J. Appl. Genet.* **49**: 367–372. [Medline] [CrossRef]
21. Luchsinger, C., Arias, M. E., Vargas, T., Paredes, M., Sánchez, R. and Felmer, R. 2014. Stability of reference genes for normalization of reverse transcription quantitative real-time PCR (RT-qPCR) data in bovine blastocysts produced by IVF, ICSI and SCNT. *Zygote* **22**: 505–512. [Medline] [CrossRef]

22. Mehta, R., Birerdinc, A., Hossain, N., Afendy, A., Chandhoke, V., Younossi, Z. and Baranova, A. 2010. Validation of endogenous reference genes for qRT-PCR analysis of human visceral adipose samples. *BMC Mol. Biol.* **11**: 39. [Medline] [CrossRef]
23. Nolan, T., Hands, R. E. and Bustin, S. A. 2006. Quantification of mRNA using real-time RT-PCR. *Nat. Protoc.* **1**: 1559–1582. [Medline] [CrossRef]
24. Pereira-Fantini, P. M., Rajapaksa, A. E., Oakley, R. and Tingay, D. G. 2016. Selection of Reference Genes for Gene Expression Studies related to lung injury in a preterm lamb model. *Sci. Rep.* **6**: 26476. [Medline] [CrossRef]
25. Pérez, R., Tupac-Yupanqui, I. and Dunner, S. 2008. Evaluation of suitable reference genes for gene expression studies in bovine muscular tissue. *BMC Mol. Biol.* **9**: 79. [Medline] [CrossRef]
26. Pfaffl, M. W., Tichopad, A., Prgomet, C. and Neuvians, T. P. 2004. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper—Excel-based tool using pair-wise correlations. *Biotechnol. Lett.* **26**: 509–515. [Medline] [CrossRef]
27. Rebouças, E. D., Costa, J. J. D., Passos, M. J., Passos, J. R. D., Hurk, R. V. D. and Silva, J. R. V. 2013. Real time PCR and importance of housekeeping genes for normalization and quantification of mRNA expression in different tissues. *Braz. Arch. Biol. Technol.* **56**: 143–154. [CrossRef]
28. Silver, N., Best, S., Jiang, J. and Thein, S. L. 2006. Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR. *BMC Mol. Biol.* **7**: 33. [Medline] [CrossRef]
29. Song, H., Mao, W., Duan, Z., Que, Q., Zhou, W., Chen, X. and Li, P. 2020. Selection and validation of reference genes for measuring gene expression in *Toona ciliata* under different experimental conditions by quantitative real-time PCR analysis. *BMC Plant Biol.* **20**: 450. [Medline] [CrossRef]
30. Sun, X., Cai, R., Jin, X., Shafer, A. B. A., Hu, X., Yang, S., Li, Y., Qi, L., Liu, S. and Hu, D. 2018. Blood transcriptomics of captive forest musk deer (*Moschus berezovskii*) and possible associations with the immune response to abscesses. *Sci. Rep.* **8**: 599. [Medline] [CrossRef]
31. Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A. and Speleman, F. 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* **3**: RESEARCH0034. [Medline] [CrossRef]
32. Wang, B., Duan, H., Chong, P., Su, S., Shan, L., Yi, D., Wang, L. and Li, Y. 2020. Systematic selection and validation of suitable reference genes for quantitative real-time PCR normalization studies of gene expression in *Nitraria tangutorum*. *Sci. Rep.* **10**: 15891. [Medline] [CrossRef]
33. Wang, Q., Ishikawa, T., Michiue, T., Zhu, B. L., Guan, D. W. and Maeda, H. 2012. Stability of endogenous reference genes in postmortem human brains for normalization of quantitative real-time PCR data: comprehensive evaluation using geNorm, NormFinder, and BestKeeper. *Int. J. Legal Med.* **126**: 943–952. [Medline] [CrossRef]
34. Wang, S., Wang, B., He, H., Sun, A. and Guo, C. 2018. A new set of reference housekeeping genes for the normalization RT-qPCR data from the intestine of piglets during weaning. *PLoS One* **13**: e0204583. [Medline] [CrossRef]
35. Wang, Y. and Harris, R. 2015. *Moschus berezovskii*. IUCN 2015: e.T13894A103431781.
36. Wei, L., Miao, H., Zhao, R., Han, X., Zhang, T. and Zhang, H. 2013. Identification and testing of reference genes for Sesame gene expression analysis by quantitative real-time PCR. *Planta* **237**: 873–889. [Medline] [CrossRef]
37. Xie, F., Xiao, P., Chen, D., Xu, L. and Zhang, B. 2012. miRDeepFinder: a miRNA analysis tool for deep sequencing of plant small RNAs. *Plant Mol. Biol.* **80**: 75–84. [Medline] [CrossRef]
38. Yang, W., Wang, W. Y., Zhao, W., Cheng, J. G., Wang, Y., Yao, X. P., Yang, Z. X., Yu, D. and Luo, Y. 2020. Preliminary study on the role of novel LysR family gene *kp05372* in *Klebsiella pneumoniae* of forest musk deer. *J. Zhejiang Univ. Sci. B* **21**: 137–154. [Medline] [CrossRef]
39. Yang, Q. S., Meng, X. X., Xia, L. and Feng, Z. J. 2003. Conservation status and causes of decline of musk deer (*moschus spp.*) in china. *Biol. Conserv.* **109**: 333–342. [CrossRef]
40. Zhang, Y., Zhang, X. D., Liu, X., Li, Y. S., Ding, J. P., Zhang, X. R. and Zhang, Y. H. 2013. Reference gene screening for analyzing gene expression across goat tissue. *Asian-Australas. J. Anim. Sci.* **26**: 1665–1671. [Medline] [CrossRef]
41. Zhao, W., Tian, Q., Luo, Y., Wang, Y., Yang, Z. X., Yao, X. P., Cheng, J. G., Zhou, X. and Wang, W. Y. 2017. Isolation, identification, and genome analysis of lung pathogenic *klebsiella pneumoniae* (LPKP) in forest musk deer. *J. Zoo Wildl. Med.* **48**: 1039–1048. [Medline] [CrossRef]
42. Zhu, W., Lin, Y., Liao, H. and Wang, Y. 2015. Selection of reference genes for gene expression studies related to intramuscular fat deposition in *Capra hircus* skeletal muscle. *PLoS One* **10**: e0121280. [Medline] [CrossRef]