aJava

Asian Journal of Animal and Veterinary Advances

Academic
Journals Inc.

www.academicjournals.com

Asian Journal of Animal and Veterinary Advances 6 (8): 789-804, 2011
ISSN 1683-9919 / DOI: 10.3923/ajava.2011.789.804

Selection of Suitable Reference Genes for Real-time Quantitative PCR Studies in Lanzhou Fat-tailed Sheep (*Ovis aries*)

[1,2]R. Zang, [2,3]J. Bai, [4]H. Xu, [1]L. Zhang, [2]J. Yang, [1]L. Yang, [2]J. Lu and [1]J. Wu

[1]College of Animal Science and Technology, Gansu Agricultural University, No. 1, New Village of Yingmeng, Anning District, Lanzhou 730070, China

[2]College of Life Science and Engineering, Northwest University for Nationalities, No. 1, New Village of Northwest, Chenguan District, Lanzhou 730030, China

[3]Gansu Engineering Research Center for Animal Cell, No. 1, New Village of Northwest, Chenguan District, Lanzhou 730030, China

[4]Experimental Center, Northwest University for Nationalities, No. 1, New Village of Northwest, Chenguan District, Lanzhou 730030, China

Corresponding Author: J. Wu, College of Animal Science and Technology, Gansu Agricultural University, No. 1, New Village of Yingmeng, Anning District, Lanzhou 730070, China Tel: 0086-13919286502

ABSTRACT

Real-time quantitative PCR (qPCR) is the most accurate method of quantifying gene expression, provided that suitable endogenous controls are used to normalize the data. To date, no reference genes have been validated for development in Lanzhou fat-tailed sheep (*Ovis aries*). We have determined the expression profiles of 7 housekeeping genes as candidate reference genes (*Actb*, *Ywhaz*, *Sdha*, *Gapdh*, *Tubb2*, *Pgk1* and *18S rRNA*) in 7 developmental stages (1, 3, 5, 7, 9, 11 and 13 months of age) and 6 tissues (omental fat, liver, tail fat, thigh muscle, subcutaneous fat (backfat above 12th and 13th rib) and perirenal fat) in Lanzhou fat-tailed sheep. The software packages geNorm, NormFinder and BestKeeper were used to evaluate the stability of potential reference genes; each produced comparable results. Initial results showed several of the candidate genes exhibited stable expression throughout development while *Actb* was identified as the least stable gene. Further analysis with geNorm, NormFinder and BestKeeper identified *Gapdh*, *Tubb2*, *Sdha* and *Ywhaz* as acceptably stable in gene expression. Comparison of diacylglycerol O-acyltransferase 1 (*Dgat1*) expression data normalized with geometric averages obtained from combinations of either *Gapdh*, *Sdha*, *Ywhaz* or *Tubb2*, *Sdha*, *Ywhaz* showed no significant differences, indicating that these two combinations are similar. The data provided in this paper may also be useful in guiding researchers performing gene expression in other species of sheep.

Key words: qPCR, housekeeping genes, reference genes, lanzhou fat-tailed sheep

INTRODUCTION

Gene expression studies are fundamental to understand the molecular basis of specific functional gene expression in specific tissues and/or organisms in mammalian development, growth and production. Real-time quantitative PCR (qPCR) is one of the most accurate methods to measure small changes in mRNA levels for individual genes. However, the quality of the results is directly related to normalization with reference genes whose expression is stable. Much work has been done on normalization using housekeeping genes with presumed stability of expression but many were

done without validation under the specific experimental conditions. According to Vandesompele *et al.* (2002), errors of up to 20-fold can result from using only one reference gene and at least two are recommended. Vandesompele *et al.* (2002) demonstrated that use of a single reference gene can lead to aberrant gene expression values and now it is widely accepted that using several reference genes for normalization is preferable. Furthermore, the expression of reference genes can vary greatly among cell types and experimental systems and many of the "classical" reference genes are unsuitable for general use (Radonic *et al.*, 2004). Accordingly, potential reference genes should be validated to assure stability of expression under specific experimental conditions. A vast number of housekeeping genes have been proposed as references for gene expression analysis (Warrington *et al.*, 2000; Laud *et al.*, 2001; Grubor *et al.*, 2004; Hein *et al.*, 2004; Garcia-Crespo *et al.*, 2005).

Housekeeping genes are typically constitutive genes that are transcribed at a relatively constant level across various conditions, such as developmental stage or tissue type and their expression is assumed to be unaffected by experimental parameters. It is unlikely that an ideal universal housekeeping gene exists. Many studies have used reference genes without proper validation of their presumed stability of expression, even though transcript levels of housekeeping genes can vary considerably. For example, two of the most commonly used reference genes in qPCR studies, glyceraldehyde-3-phosphate dehydrogenase (*Gapdh*) and α-actin (*Actb*), have been found to vary in expression by 8-fold and 7- to 22-fold, respectively, during a survey of 535 human housekeeping genes (Warrington *et al.*, 2000). The use of a single reference gene has been shown to lead to erroneous normalization of up to 3- and 6-fold in expression studies in various human tissues (Vandesompele *et al.*, 2002). Hence, the use of only one reference gene for normalization of gene expression studies should not be considered sufficient (Bustin *et al.*, 2005). The choice of suitable reference genes for normalization of qPCR data during development is by no means trivial, since they must be relatively unaffected by marked changes in transcriptional activity. Nevertheless, this task can be achieved by computational methods recently developed to assess the expression stability of candidate reference genes, namely geNorm (Vandesompele *et al.*, 2002), NormFinder (Andersen *et al.*, 2004) and BestKeeper (Pfaffl *et al.*, 2004). The geNorm uses a pairwise analysis of gene expression to identify stable reference genes. Likewise, BestKeeper performs a pairwise comparison, whereas NormFinder uses a mathematical model to estimate overall expression variation of candidate reference genes and variation between sample groups. Since there is not a single accepted method to examine gene expression stability and different statistical methods can potentially yield variable results. Validating the reference genes using several applications and identifying differences and similarities between the outputs of alternative software, makes the consensus conclusions more reliable.

Studies on selection of ovine housekeeping genes for normalization by real-time RT-PCR were carried out using geNorm software package (Garcia-Crespo *et al.*, 2005; Passmore *et al.*, 2009; Zaros *et al.*, 2010). Currently there is no information available on reference gene stability in sheep using the above mentioned three packages. In this paper, using geNorm, NormFinder and BestKeeper packages, we examined the stability of seven potential reference genes, 18S ribosomal RNA (18S rRNA), glyceraldehyde-3-phosphate dehydrogenase (Gapdh), tyrosine 3-monooxygenase (Ywhaz), α-actin (Actb), phosphoglycerate kinase 1 (*Pgk1*), α_2-Tubulin (*Tubb2*) and succinate dehydrogenase (Sdha) (Table 1) in seven developmental stages (1, 3, 5, 7, 9, 11 and 13 months of age) and six tissues, Omental Fat (OF), Liver (LI), Tail Fat (TF), Thigh Muscle (TM), Subcutaneous Fat (SF) and Perirenal Fat (PF) in Lanzhou fat-tailed sheep, one of the popular meat sheep distributed in the suburban districts of Lanzhou prefecture, in Gansu province of China.

Table 1: Housekeeping genes as candidate reference genes evaluated in this study

Gene symbol	Gene name	Accession No.	Function
18S rRNA	18S Ribosomal RNA	EU861215	Ribosome assembly
Gapdh	glyceraldehyde-3-phosphate dehydrogenase	AF022183	Oxidative phosphorylation
Pgk1	phosphoglycerate kinase 1	NM_001142516	Glycolysis
Sdha	succinate dehydrogenase	AY970969	Citric acid cycle
Tubb2	α2-Tubulin	GQ338157	Microtubule structure
Ywhaz	tyrosine 3-monooxygenase	AY970970	Cell signaling
Actb	α-Actin	NM_001009784	Cytoskeleton structure

MATERIALS AND METHODS

Sheep husbandry and sample collection: Lanzhou fat-tailed lambs used in this study were produced by artificial insemination in Gansu Huajia Animal Husbandry Company. Sheep were raised in confinement and fed pelleted feed and hay, with supplementary green forage from May to October. Thirty one-month-old lambs were randomly divided into 6 groups in March 2008. According to the Regulations for the Administration of Affairs Concerning Experimental Animals in China, animals were properly anaesthetized and sacrificed by exsanguination at corresponding months of 1, 3, 5, 7, 9, 11, 13 and samples of liver, muscle, omental fat, subcutaneous fat, perirenal fat and tail fat were collected from each animal. Samples were flash frozen in liquid nitrogen at-196°C and then kept in a -80°C freezer after transport to the laboratory until total RNA extraction.

RNA extraction and cDNA synthesis: Approximately 100 mg liver and muscle tissue were used for total RNA extraction with TRIzol reagent (Invitrogen) according to manufacturer instructions. About 150-200 mg of fat tissue were used for total RNA extraction with modified method of TRIzol reagent (Invitrogen). Frozen fat tissue was ground in a mortar three times. Homogenized tissue was then transferred to a homogenizing tube, 0.4 mL Trizol added and homogenization was continued until the homogenate was agranular and transparent. The homogenate was transferred to a 2 mL Eppendorf tube, 0.8 mL Trizol added, then mixed thoroughly and let stand at room temperature for 5 min. The homogenate was centrifuged at 12,000 rpm, 25°C, for 5 min, liquid removed and remaining homogenate transferred to a new Eppendorf tube. Chloroform was added at 200 μL of chloroform per 1 mL of Trizol, the tube inverted several times and vortexed at least 15 sec. The solution was held at ambient temperature for 15 min, centrifuged at 12,000 rpm at 4°C for 10 min and the upper aqueous phase was transferred to a clean tube. Total RNA was precipitated with isopropanol and washed with 75% ethanol. Total RNA was treated with the gDNA wipeout buffer supplied with the QuantiTect reverse transcription kit (Qiagen) to remove traces of genomic DNA contamination. Assessment of RNA quality was performed on a 1.2% agarose gel containing Ethidium Bromide (EB) DNA gel stain and photographed with the Bio-BEST 140E (SIM) imaging system (China). RNA samples were then quantified using a Nanodrop spectrophotometer (China). All samples had absorbance ratios of greater than 1.9 at 260/280 nm which indicated high purity RNA. Synthesis of cDNA was done with the superscript™ III first-strand synthesis system for RT-PCR (Invitrogen) using a 96 well PCR plate from Techne (Barloworld Ltd, Stone, USA). The resulting single stranded cDNA products were quantified using the Nanodrop spectrophotometer and then diluted 50-fold with deionized water prior to using them as templates for the qPCR reactions.

Primer design: For qPCR, primers of *Gapdh*, *Pgk1* and *tubb2* were designed using Primer 3 software (Rozen and Skaletsky, 2000) and sequences were validated against the Ovine genome. Primers of *Sdha, Ywhaz, Actb* reported by Garcia-Crespo *et al.* (2005) were used. For the 18S rRNA gene, ovine sequences were not available and therefore, multiple sequence alignments of these genes from different animal species (Bos taurus, Mus musculus, Rattus norvegicus and Homo sapiens) were done using the program Aligning X (Vector NTI 8.0 suite, Informax Inc., North Bethesda, MD, USA) to identify conserved regions for primer design. For normalization, the primers of diacylglycerol O-acyltransferase 1 (*Dgat1*) gene were also designed based on the sheep *Dgat1* sequence (GenBank accession no. EU178818). Four concentrations of primers (100, 200, 300 and 500 nM) were evaluated and formation of primer-dimers was assessed by melting curve analysis. Thus, only those concentrations of primers which showed dimer-free reactions were used for the final analysis. Primers were synthesized from Takara Bio, Dalian, China (Table 2).

Real-time quantitative PCR (qPCR): Gene amplifications by qPCR were performed with a SLAN thermocycler (Hongshi, Shanghai, China). Each 25 μL reaction in a 96-well plate was comprised of 4 μL of 50x diluted cDNA template, 1 μL of each primer pair at 10 and 12.5 μL of SYBR Premax Tag containing DNA polymerase, buffer, dNTP and SYBR® Green I (Promega). 96-well plates were sealed with adhesive optical film (Promega) and after an initial denaturation step of 15 min at 95°C, 45 cycles of amplification were performed according to the following thermo cycling profiles: denaturation for 10 sec at 95°C, annealing for 20 sec at 60°C and extension for 20 sec at 72°C. Fluorescence data were acquired during last step. A dissociation protocol with a gradient from 65 to 97°C was used to investigate the specificity of the qPCR reaction and the presence of primer dimers. Gene expression levels were recorded as C_T values that corresponded to the number of cycles at which the fluorescence signal can be detected above a threshold value, arbitrarily set to 0.3. The C_T value is therefore inversely correlated to the initial amount of DNA present in the PCR reaction. All samples were run in three replications and minus reverse

Table 2: Nucleotide sequences of primers and amplification parameters for real-time PCR (forward primers are listed first)

Gene	Forward and reverse primers (5'→3')	[C][a]	Amplicon size (bp)	Time (sec)[b]	T_m (°C)[c]	Slope	R^{2d}	E (%)[e]
18S rRNA	CCCTGTAATTGGAATGAGTCCACTT	300	100	25	85	-3.3313	-0.999	99.61
	ACGCTATTGGAGCTGGAATTACC							
Gapdh	GTTCCACGGCACAGTCAAGG	500	116	25	87.7	-3.2851	-0.998	101.56
	ACTCAGCACCAGCATCACCC							
Pgk1	TTCGCTTTGTCATTGTTTCACC	300	84	20	81.7	-3.3136	-0.998	100.35
	ATCACCAACTTCATCTCCCTCC							
Sdha	CATCCACTACATGACGGAGCA	200	90	20	82	-3.2518	-0.994	103.01
	ATCTTGCCATCTTCAGTTCTGCTA							
Tubb2	ACACTGTGGTTGAGCCCTACA	300	120	20	84	-3.3606	-0.994	98.41
	GCTTGAGGGTGCGGAAA							
Ywhaz	TGTAGGAGCCCGTAGGTCATCT	300	102	25	79	-3.1935	-0.997	105.65
	TTCTCTCTGTATTCTCGAGCCATCT							
Actb	ATGCCTCCTGCACCACCA	300	125	25	85	-3.3854	-0.995	97.42
	GCATTTGCGGTGGACGAT							
Dgat1	GGACACAGACAAGGACGGAGA	300	144	25	87.7	-3.1732	-0.997	106
	ATCAGCATCACCACACACCAA							

[a] Primer coucentrations in nM. [b] Annealing and extension time. [c] Theoretical amplicon melting temperature calculated with Primer Express software (Applied Biosystems). [d] Correlation coefficient. [e] PCR efficiency

transcriptase and no template controls were included in all plates, along with a positive plate control. Five-point standard curves of a 5-fold dilution series (1:1-1:625) from pooled cDNA were used for PCR efficiency calculation. The PCR efficiency (E) is given by the equation (Radonic *et al.*, 2004):

$$E = (10^{-\frac{1}{m}} - 1) \times 100$$

where, m is the slope of linear regression model fitted over log-transformed data of the input cDNA concentrations versus C_T values.

Evaluation of expression stability: Evaluation of expression stability was done using three independent statistical applications: geNorm (Vandesompele *et al.*, 2002), NormFinder (Andersen *et al.*, 2004) and BestKeeper (Pfaffl *et al.*, 2004). The geNorm is a Microsoft Excel application that determines the expression stability of reference genes based on overall pairwise comparisons between them. In brief, the principle behind this algorithm is that expression ratio of two ideal reference genes is identical in all samples, regardless of the experimental conditions or cell type. The software calculates a stability value which is inversely correlated to gene expression stability and ranks the reference genes accordingly. Stepwise exclusion of the gene with the highest stability index, assuming that the genes are not co-regulated, results in a combination of two reference genes that have the most stable expression across the tested samples. BestKeeper is based on pairwise comparisons of raw Cycle Threshold (C_T) values of each gene. The genes showing least variation are incorporated into a BestKeeper index which can be used in very much the same way as a single reference gene in normalization of gene expression. Stable reference genes shows a strong correlation with the BestKeeper index. NormFinder is also an application for Excel but, unlike geNorm and BestKeeper, it uses a model-based approach to determine the optimal reference genes. In this mathematical model a separate analysis of the sample subgroups and estimation of both intra- and inter-group variation in expression levels are included into the calculation of a gene stability value. NormFinder also calculates the best combination of two genes for a two-gene normalization factor and its corresponding stability value.

Statistical analyses: Differences in expression levels of *Actb, Ywhaz, Sdha, Gapdh, Tubb2, Pgk1* and *18S rRNA* with developmental stage and different tissues were examined by one-way ANOVA with Holm-Sidak post-hoc tests. When the data did not meet the normality and/or equal variance requirements, a Kruskal-Wallis one-way ANOVA on ranks with a Dunn's test for post-hoc comparison was performed instead. The SigmaStat statistical package (Systat software, London, UK) was used for all analyses. Significance levels were set at $p<0.05$.

RESULTS

All primers used to amplify the candidate reference genes generated single amplicons, as demonstrated by the presence of single bands in agarose gels following electrophoresis (data not shown) and by the single-peak melting curves of the PCR products. Further analysis of the dissociation curves confirmed the absence of primer dimers or other products resulting from non-specific amplification. Efficiency of PCR reactions ranged from 99.4% for *Sdha* and *Tubb2* to 105.65% for *Ywhaz* (Table 2).

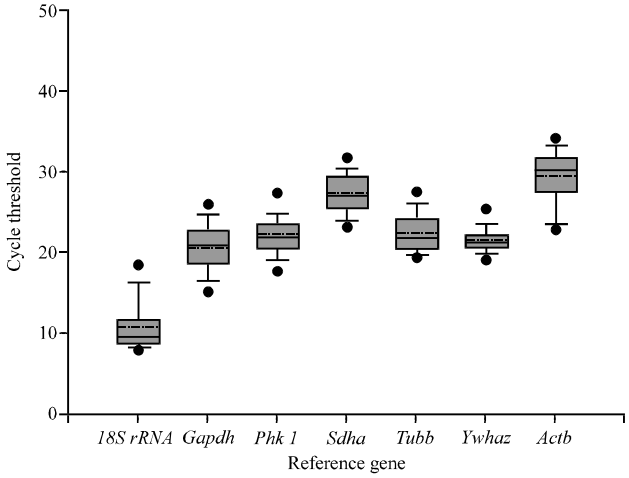

Fig. 1: Overall expression patterns of potential reference genes in development in Lanzhou fat-tailed sheep (*Ovis aries*). The raw cycle threshold (C_T) qPCR data of each reference gene in all samples (n = 258) are represented in a box-and-whisker diagram. Mean values are represented by a dashed line. The 5th and 95th percentiles are indicated by the dots below and above each box, respectively

Reference gene stability

Developmental expression profiles of candidate reference genes: Analysis of the raw expression levels of C_T value across all samples (7 developmental stages, 3 independent samples per stage) identified some variation among candidate reference genes (Fig. 1). *Ywhaz* showed the lowest global variability, whereas *Actb* had the highest variation in expression levels. The candidate reference gene 18S rRNA, was highly expressed (mean C_T = 10.89) while other candidate reference genes were expressed at moderate levels, with mean C_T values of 20.60, 21.63, 22.10, 22.38, 27.28 and 29.74 for *Gapdh, Ywhaz, Pgk1, Tubb2, Sdha* and *Actb*, respectively (Fig. 1).

A more detailed analysis of individual gene expression levels at each developmental stage revealed that all selected genes were expressed throughout the selected developmental stages from the age of one month to the age of thirteen months, albeit with different expression patterns (Fig. 2). Apart from *Pgk1* ($p = 0.432$), all other candidate reference genes showed overall changes in expression with developmental stage ($p<0.05$). It should be noted that *Pgk1* has the highest intra-group variation, despite its relatively stable expression throughout development (Fig. 2c). As shown in Fig. 2, expression profiles were generally characterized by lesser transcriptional levels (i.e., higher C_T values) at the age of one month and greater at the ages of five and seven months. Since transcriptional activity varies substantially during development, it is evident that a straightforward statistical analysis of raw C_T values is not suitable to select the best reference genes for normalization of these qPCR data. The most evident increase in expression levels throughout development was detected for *18S rRNA*: the difference in mean C_T values for this gene between the age of one month and the age of five month was about 6.38 which corresponds to an expression increase of approximately 638-fold (Fig. 2a).

Expression profiles of potential reference genes in different tissues: Averaged over developmental stage, the individual gene expression profiles in different tissues, as indicated by raw C_T values, are shown in Fig. 3. Apart from *18S rRNA* ($p = 0.212$), all other candidate reference genes showed significant overall changes in their expression in different tissue ($p<0.05$). Although

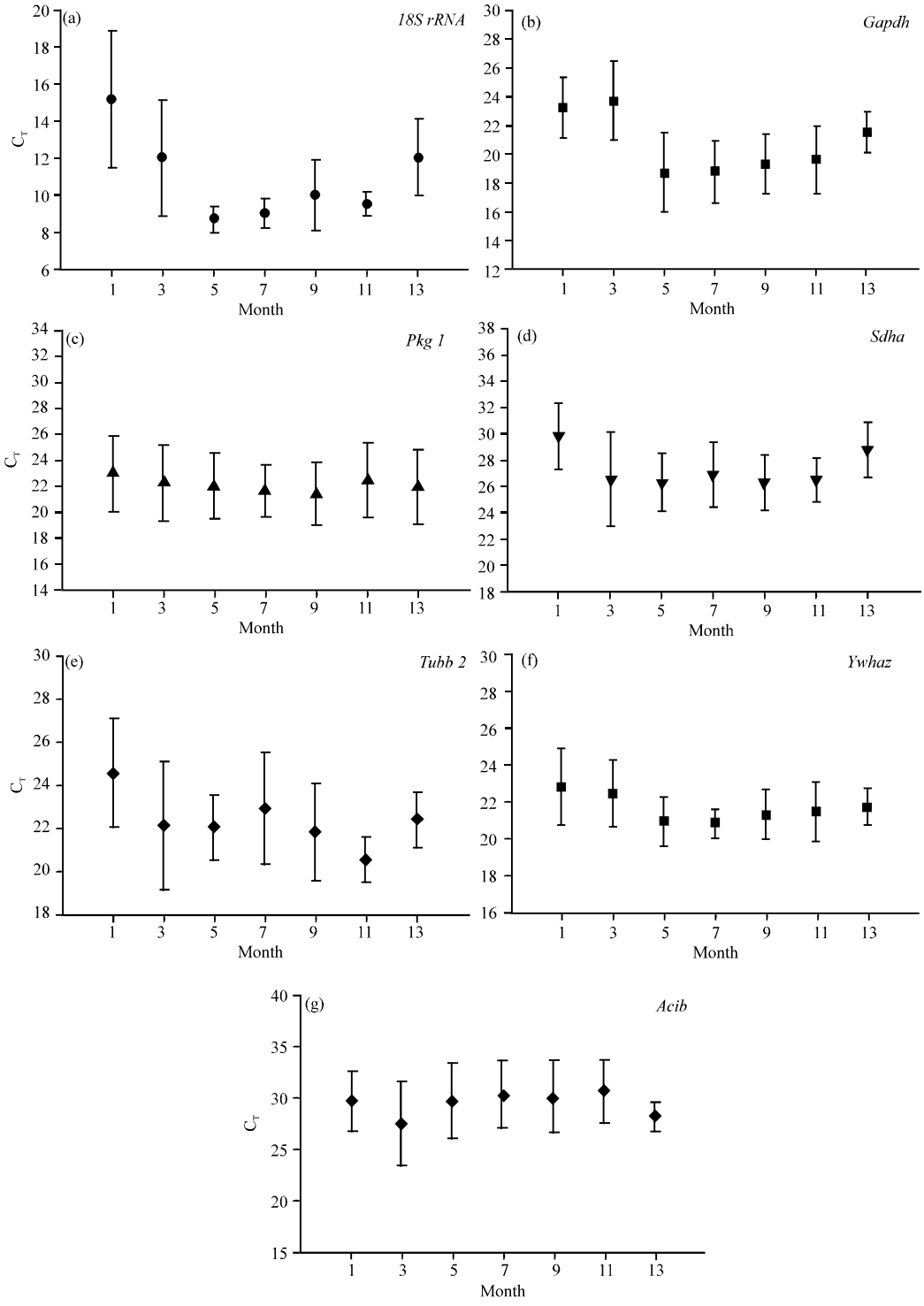

Fig. 2 (a-g): Individual expression profiles of potential reference genes. Transcript levels of *Gapdh* (a), *Pgk1* (b), *Sdha* (c), *Tubb2* (d), *Ywhaz* (e), *18S rRNA* (f) and *Actb* (g) in seven developmental stages (age of 1, 3, 5, 7, 9, 11 and 13 months) in Lanzhou fat-tailed sheep were determined by qPCR. Data are shown as raw cycle threshold (C_T) values and represented as Mean±SE (n = 6)

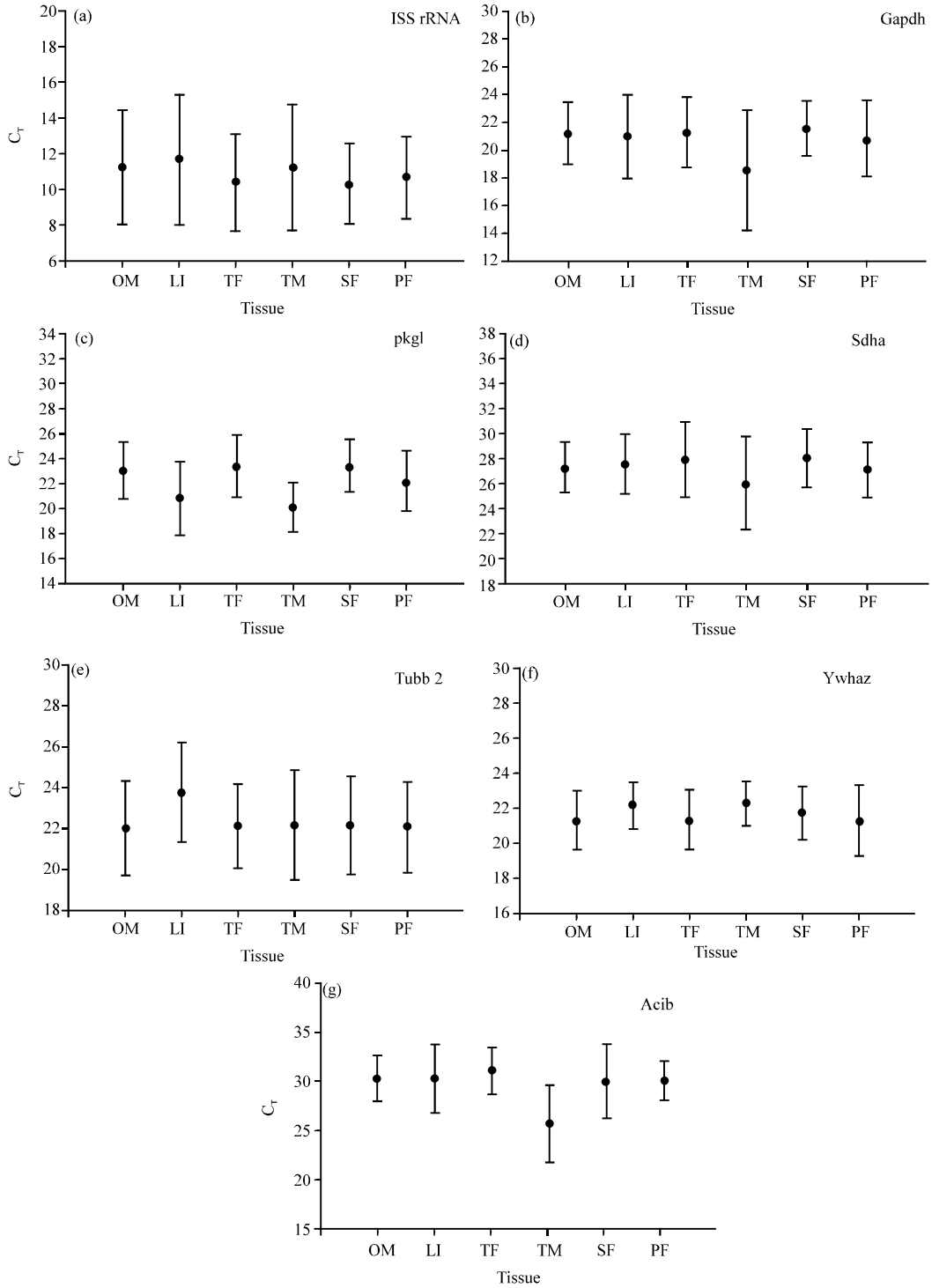

Fig. 3 (a-g): Individual expression profiles of potential reference genes. Transcript levels of *Gapdh* (a), *Pgk1* (b), *Sdha* (c), *Tubb2* (d), *Ywhaz* (e), *18S rRNA* (f) and *Actb* (g) in six tissues of OF, LI, TF, TU, SF and PF in Lanzhou fat-tailed sheep were determined by qPCR. Data are shown as raw cycle threshold (C_T) values and represented as Mean±SE (n = 6). Abbreviations: omental fat (OF), Liver (LI), Tail Fat (TF), Thigh Muscle (TM), Subcutaneous Fat (SF) and Perirenal Fat (PF)

Table 3: Inter-gene relations and correlations between the reference genes and the *BestKeeper* index

	18S rRNA	*Gapdh*	*Pgk1*	*Sdha*	*Tubb2*	*Ywhaz*	*Actb*
Gapdh	0.723	-	-	-	-	-	-
Pgk1	0.140*	0.504	-	-	-	-	-
Sdha	0.705	0.662	0.386	-	-	-	-
Tubb2	0.651	0.676	0.421	0.733	-	-	-
Ywhaz	0.617	0.592	0.379	0.421	0.700	-	-
Actb	0.071*	0.502	0.389	0.268	0.377	0.110*	-
BestKeeper index	0.775	0.907	0.608	0.819	0.857	0.694	0.524

Pairwise correlation analyses were performed based on the cycle threshold values of the six housekeeping genes. Pearson's correlation coefficients (r) are shown. *Indicates correlations below the significance threshold ($p>0.05$)

18S rRNA expression was relatively stable across the different tissues, it had the highest intra-group variation (Fig. 3a). The expression of the seven reference genes was stable in adipose tissues OF, TF, SF and PF but not in LI and TM.

Analyses of gene expression stability by BestKeeper: Using the initial statistics in Fig. 1 produced by BestKeeper, the genes were ranked in the following order from most to least stable: *Gapdh>Tubb2>Pgk1>Sdha>18S rRNA>Ywhaz>Actb* when all developmental stages were examined. Due to its high variabihty ($\sigma = 2.70$ C_T), *Actb* can be classified as a clearly unstable gene and therefore, unsuitable for normalization. Nevertheless and for comparison purposes, all genes were included in the calculation of the BestKeeper index. Apart from *Pgk1* with *18S rRNA* and *Actb* with *18S rRNA*, *Actb* and *Ywhaz*, the seven housekeeping genes correlated well with each other (Table 3). Particularly strong inter-gene correlations were found for *18S rRNA/Gapdh* ($r = 0.723$), *18S rRNA/Sdha* ($r = 0.705$), *Gapdh/Tubb2* ($r = 0.676$), *Sdha/Tubb2* ($r = 0.733$) and *Tubb2/Ywhaz* ($r = 0.700$) which indicates that these gene pairs have very similar overall expression patterns. Repeated pairwise correlation and regression analyses between the candidate reference genes and the *Best-Keeper* index showed that all genes contributed significantly to the index (Table 3). The largest correlations with the *BestKeeper* index were observed for *Gapdh* ($r = 0.907$, $p = 0.001$), *Tubb2* ($r = 0.857$, $p = 0.001$) and *Sdha* ($r = 0.819$, $p = 0.001$), thus identifying these three genes as the most reliable reference genes for normalization. Conversely, *Actb* was found to be the least suitable gene ($r = 0.524$, $p = 0.001$), despite showing a regular mean expression level throughout the growth and development (Fig. 2g and 3g). In summary, the BestKeeper pairwise analysis sorted the candidate reference genes in the following series of decreasing stability: *Gapdh>Tubb2>Sdha>18S rRNA>Ywhaz>Pgk1>Actb*. The expression stability of each reference gene in each developmental stage analyzed by BestKeeper application also indicated that *Gapdh* is the most stable reference gene, followed by *Tubb2* and *Ywhaz* (Table 4). Averaged over the seven developmental stages, the individual gene expression stability from the BestKeeper application in different tissues are shown in Table 5. Because *18S rRNA* was the most stable gene in 3 (omental, liver and perirenal fat) of the 6 tissues, it was classified as the best candidate reference gene, followed by *Ywhaz* and *Gapdh* (Table 5). The least stable gene in 4 tissues (omental, liver, tail and perirenal fat) was *Actb*.

Analyses of gene expression stability by geNorm: Most to least stable, as follows: *Gapdh>Ywhaz>Tubb2>Sdha>18S rRNA>Pgk1>Actb*. Pairwise comparisons performed by geNorm to identify the best combination of genes for calculation of a normalization factor revealed that *Tubb2* and *Yhwaz* were the best pair (Fig. 4a. There were also differences in the relative stability of gene expression The stability of gene expression each developmental stage was also assessed by

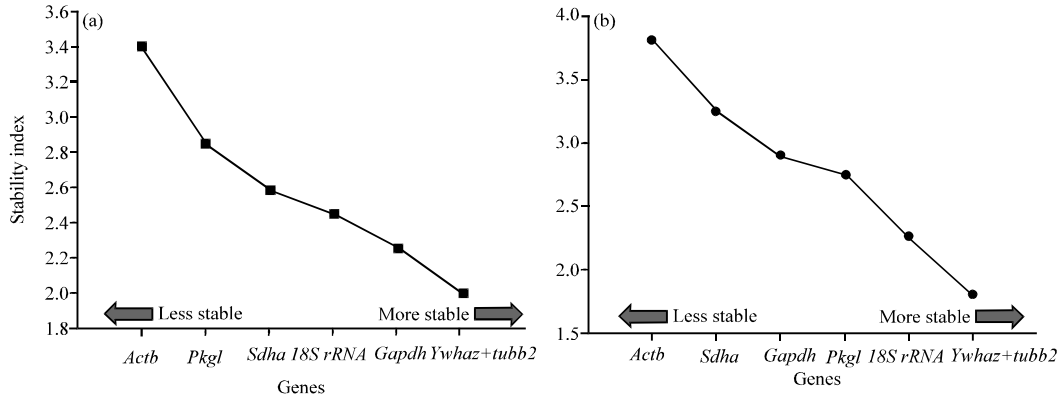

Fig. 4: Ranking of reference genes according to their expression stability in development (a) and in different tissues (b). Average expression stability values were calculated by geNorm. Expression stability of the reference genes is inversely correlated to their stability index. Both algorithms identify *Actb* reference gene as the least suitable gene for normalization of qPCR data. But the most stable reference gene pairs are *Ywhaz/Tubb2* and *Ywhaz/Gadph*, respectively

Table 4: Expression stability of reference gene during development analyzed by *BestKeeper* application

Developmental stage (Month)	*18S rRNA*	*Gapdh*	*Pgk1*	*Sdha*	*Tubb2*	*Ywhaz*	*Actb*
1	0.942	0.902	0.611	0.652	0.939	0.744	0.536
3	0.923	0.920	0.487	0.769	0.882	0.923	0.508
5	0.305	0.970	0.940	0.841	-0.053**	0.531	0.838
7	0.292	0.937	0.613	0.804	0.554	-0.421	0.889
9	0.529	0.930	0.312	0.870	0.923	0.358	0.777
11	0.504	0.938	0.675	0.806	0.139	0.505	0.709
13	0.777	0.843	0.331	0.783	0.877	0.346	0.475

Stability data for reference gene expression were determined with the *BestKeeper* application. The stability values are in proportion to gene expression stability.

Table 5: Expression stability of reference gene in different tissue analyzed by *BestKeeper* applications

Tissue	*18S rRNA*	*Gapdh*	*Pgk1*	*Sdha*	*Tubb2*	*Ywhaz*	*Actb*
Omental fat	0.946	0.823	0.787	0.728	0.877	0.904	0.324
Liver	0.949	0.908	0.568	0.836	0.664	0.536	0.060
Tail fat	0.724	0.809	0.284**	0.788	0.689	0.828	-0.062
Thigh muscle	0.942	0.959	0.194	0.875	0.879	0.653	0.791
Subcutaneous fat	0.667	0.676	0.408	0.779	0.472	0.782	0.522
Perirenal fat	0.943	0.901	0.897	0.674	0.722	0.900	0.225

Stability data for reference gene expression were determined with the *BestKeeper* application. The stability values are in proportion to gene expression stability

geNorm (Table 6). There were differences in the relative stability of gene expression among developmental stages but *Gapdh* was classified as the best candidate reference gene, since it was the most stable gene in 3 (age of 1, 7 and 9 months) of the 7 developmental stages, followed by *Tubb2*, *Sdha* and *Ywhaz* which were most stable at one developmental stage (Table 6). The overall least stable gene was *Actb*, rated as least stable in 4 stages (age of 3, 5, 9 and 11 months). When the data from all stages were combined and analyzed simultaneously, the candidate reference genes were ordered from the between different tissues when the data from all stages were combined

Table 6: Indices of expression stability of seven reference genes during development of Lanzhou fat-tailed sheep

Development	geNorm							NormFinder						
stage (month)	*18S rRNA*	*Gapdh*	*Pgk1*	*Sdha*	*Tubb2*	*Ywhaz*	*Actb*	*18S rRNA*	*Gapdh*	*Pgk1*	*Sdha*	*Tubb2*	*Ywhaz*	*Actb*
1	2.769	1.649	2.17	1.929	1.798	1.947	2.224	0.155	0.035	0.078	0.065	0.026	0.064	0.082
3	2.567	2.429	3.584	3.157	2.552	2.383	4.346	0.099	0.034	0.092	0.069	0.061	0.03	0.08
5	2.384	2.158	2.111	2.18	2.915	2.284	3.219	0.078	0.066	0.044	0.031	0.082	0.055	0.076
7	2.336	2.1	2.401	2.438	2.953	2.583	2.857	0.088	0.056	0.067	0.046	0.084	0.076	0.059
9	2.757	1.961	2.961	1.999	2.043	2.56	3.114	0.118	0.04	0.078	0.032	0.04	0.06	0.071
11	2.056	2.198	2.913	2.013	2.305	2.327	3.074	0.054	0.054	0.063	0.036	0.057	0.06	0.063
13	2.213	1.682	3.17	2.141	1.581	1.948	1.954	0.109	0.026	0.106	0.031	0.023	0.046	0.042

Stability data for reference gene expression were determined with the *geNorm* and *NormFinder* applications. The stability values are inversely correlated to gene expression stability

Table 7: Indices of expression stability of seven reference genes in different tissues during development in Lanzhou fat-tailed sheep

	geNorm							NormFinder						
Tissue	*18S rRNA*	*Gapdh*	*Pgk1*	*Sdha*	*Tubb2*	*Ywhaz*	*Actb*	*18S rRNA*	*Gapdh*	*Pgk1*	*Sdha*	*Tubb2*	*Ywhaz*	*Actb*
Omental fat	2.204	2.031	2.085	1.973	1.866	1.707	3.032	0.136	0.055	0.05	0.054	0.039	0.034	0.102
Liver	3.156	3.005	3.294	2.682	2.593	2.813	4.629	0.187	0.06	0.09	0.049	0.063	0.059	0.134
Tail fat	3.224	3.292	3.83	3.766	3.123	2.718	3.881	0.165	0.055	0.103	0.069	0.052	0.038	0.107
Thigh muscle	3.329	4.362	3.731	3.742	3.021	3.13	3.731	0.146	0.101	0.118	0.056	0.054	0.092	0.067
Subcutaneous fat	2.885	2.994	2.544	2.835	2.662	2.269	3.772	0.137	0.055	0.062	0.035	0.082	0.036	0.1
Perirenal fat	2.553	3.3	2.993	3.091	2.375	2.296	2.832	0.106	0.051	0.041	0.051	0.062	0.033	0.097

Stability data for reference gene expression were determined with the *geNorm* and *NormFinder* applications. The stability values are inversely correlated to gene expression stability

and analyzed simultaneously. The best candidate reference gene was *Ywhaz*, since it was the most stable gene in 4 (omental fat, tail fat, subcutaneous fat and perirenal fat) of the 6 tissues (Table 7). The least stable gene was *Actb* with the least stability in 4 tissues (omental fat, liver, tail fat and subcutaneous fat). Combining data from all stages and all tissues, the candidate reference genes were ordered from the most to least stable, as follows: *Ywhaz>Tubb2>18S rRNA>Pgk1>Gapdh>Sdha>Actb*. Pairwise comparisons performed by geNorm to identify the best combination of genes for calculation of a normalization factor revealed that *Tubb2* and *Ywhaz* were the best pair (Fig. 4b).

Analyses of gene expression stability by NormFinder: The NormFinder analysis of individual developmental stage was considerably different from the corresponding geNorm output (Table 6). The most stable gene was found to be *Sdha* (age of 5, 7, 9 and 11 months). The gene that showed the greatest variability was *18S rRNA* (age of 1, 3, 7, 9 and 13 months) (Table 6). When the data from all stages were combined and analyzed simultaneously, the candidate reference genes were ordered from the most to least stable, as follows: *Sdha>Ywhaz>Tubb2>Gapdh>Pgk1>Actb>18S rRNA*. *Sdha* and *18S rRNA* was found to be the most and least stable housekeeping genes, with stability index of 0.0459 and 0.1867, respectively. Pairwise comparisons performed by NormFinder to identify the finest combination of genes for calculation of a normalization factor revealed that *Gapdh* and *Tubb2* were the best pair (Fig. 5a). According to results from NormFinder, there were also differences in the relative stability of gene expression amongst different tissues when the data from all stages were combined and analyzed simultaneously, the best candidate reference gene was *Ywhaz* , since it was the most stable gene in 3 (omental, tail fat and perirenal fat) of the 6 tissues (Table 7). The least stable gene was *18S rRNA* in all 6 tissues. Taking the data from all stages and all tissues together, the candidate reference genes were ordered from the most to least stable, as follows: *Sdha>ywhaz>Tubb2>Actb>Pgk1>18S rRNA*. *Sdha* and *18S rRNA* was found to be the most and least stable housekeeping genes, with stability index of 0.0459 and 0.1867, respectively. Nevertheless, *Gapdh* and *Ywhaz* were selected as the best combination of genes for a two-gene normalization factor (Fig. 5b).

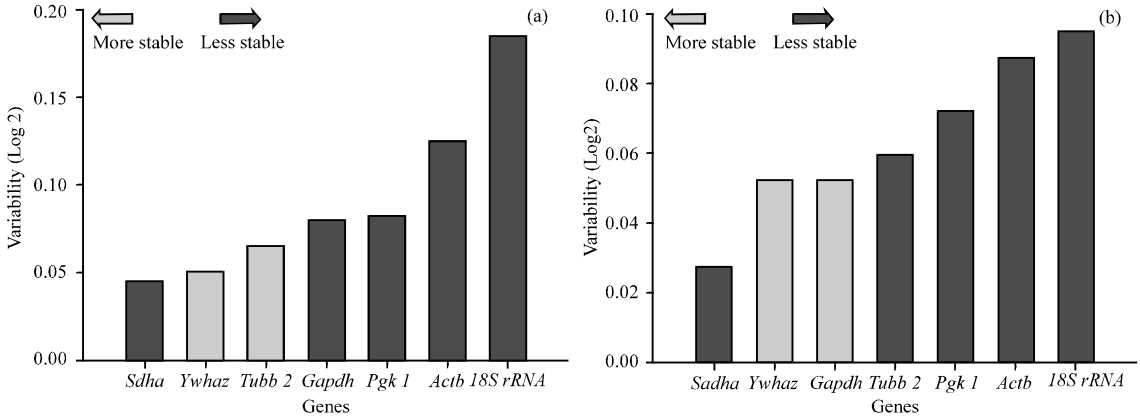

Fig. 5: Ranking of reference genes according to their expression stability in development (a) and in different tissues (b). Average expression stability values were calculated by NormFinder. Expression stability of the reference genes is inversely correlated to their stability index. Both algorithms identify *18S rRNA* and *Sdha* reference gene as the least suitable gene and the most stable gene for normalization of qPCR data. But the most stable reference gene pairs are *Ywhaz/Tubb2* and *Ywhaz/Gapdh* that are shaded in grey color, respectively

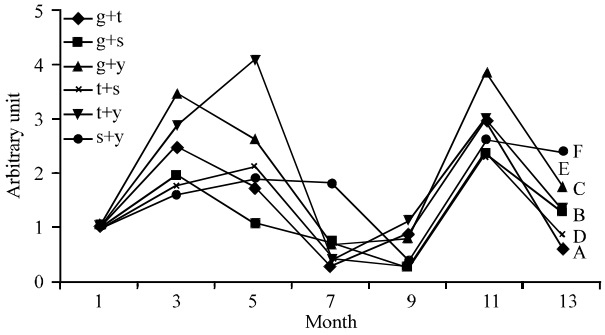

Fig. 6: Normalisation of *Dgat1* mRNA expression to various combinations of the geometric average for two genes. Data shown are all calculated relative to the age of one month, so that one month values are equal to 1 arbitrary. Six normalization factors were derived by calculating the geometric averages of the following gene combinations: (A) *Gapdh, Tubb2* (rhombus, blue); (B) *Gapdh, Shda* (rectangle, red); (C) *Gapdh, Ywhaz* (triangle, yellow); (D) *Tubb2, Shda* (cross green); (E) *Tubb2, Ywhaz* (cross, purple); (F) *Shda, Ywhaz* (rhombus, purple). Values shown are the mean normalized value±SE (n = 9)

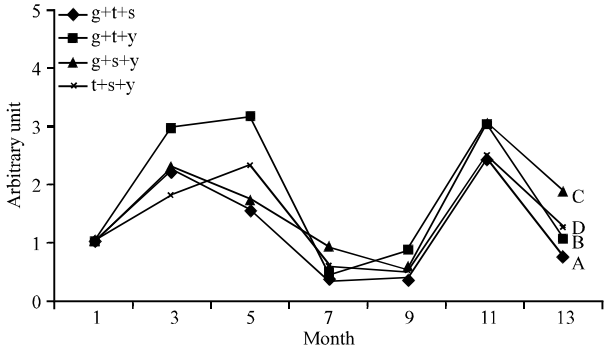

Fig. 7: Normalisation of *Dgat1* mRNA expression to various combinations of the geometric average for three genes. Data shown are all calculated relative to the age of one month, so that one month values are equal to 1 arbitrary. Six normalization factors were derived by calculating the geometric averages of the following gene combinations: (A) *Gapdh, Tubb2, Shda* (Rhombus, blue); (B) *Gapdh, Tubb2, Ywhaz* (rectangle, red); (C) *Gapdh, Sdha, Ywhaz* (triangle, yellow); (D) *Tubb2, Sdha, Ywhaz* (cross, blue). Values shown are the mean normalized value±SE (n = 9)

Normalization: Based on the results from the three methods of analysis, four genes, *Gapdh, Tubb2, Sdha* and *Ywhaz* were consistently stable. In order to assess the stability of the normalization factors obtained, we first compared the normalized expression of diacylglycerol O-acyltransferase 1 (*Dgat1*) to various combinations of the geometric average of two genes (Fig. 6). Six normalization factors were derived by calculating the geometric averages of the following gene combinations: A: *Gapdh, Tubb2*; B: *Gapdh, Shda*; C: *Gapdh, Ywhaz*; D: *Tubb2, Shda*; E: *Tubb2, Ywhaz*; F: *Shda, Ywhaz*. We found there were significant differences ($p<0.05$) in *Dgat1* gene expression at age of month 5 (A vs E), month 9 (A vs B, A vs. D, C vs F, D vs. E) and month 13 (A vs F, D vs. F). Most significant differences (ANOVA $p<0.01$) in month 7 (A vs. F, D vs F, E vs. F) and in month 9 (A vs. B, B vs. E, A vs. F). We therefore examined normalization using geometric average of three genes (Fig. 7). Four normalization factors were derived by calculating

the geometric averages of the following gene combinations: A: *Gapdh, Tubb2, Shda*; B: *Gapdh, Tubb2, Ywhaz*; C: *Gapdh, Sdha, Ywhaz*; D: *Tubb2, Sdha, Ywhaz*. Apart from significant difference between A and B in *Dgat1* gene expression in month 9, no significant differences were observed ($p>0.05$) when the normalized data at each developmental stage were compared between the different normalization factors, indicating that using a three gene normalization factors calculating the geometric average of combination either *Gapdh, Sdha, Ywhaz* or *Tubb2, Sdha, Ywhaz*.

DISCUSSION

In this study, we examined the expression of seven candidate reference genes for normalization of real-time quantitative PCR data in seven developmental stages and six tissues in Lanzhou fat-tailed sheep. The identification of genes with stable expression in all samples of an experiment is crucial as it is necessary to normalize for variability between samples introduced during the production of the cDNA (Bustin, 2000; Le Bail *et al*., 2008). As a universal reference gene with stable expression in all experimental system is not available, suitable reference genes for each experiment need to be determined.

To identify stably expressed genes, analysis packages such as geNorm and BestKeeper perform a pairwise comparison of gene expression across the various samples in an experiment. Therefore it is crucial that genes used are not co-regulated or present on the same pathway, as co-regulated genes will likely have similar expression patterns and would therefore appear to be stably expressed in any biological experiment. For this reason we chose genes involved in a number of different biological processes (Table 1), such as glycolysis (*Pgk1*), citric acid cycle (*Sdha*), ribosome assembly (*18S rRNA*), cytoskeleton structure (*Actb*), oxidative phosphorylation (*Gapdh*), microtubule structure (*Tubb2*) and cell signaling (*Ywhaz*).

Based on the M values obtained in geNorm (Fig. 3), the stability index from NormFinder (Fig. 4) and the descriptive statistics produced by BestKeeper (Table 3-5), it would appear that several of the genes used in this study are suitable for normalization of gene expression data in Lanzhou fat-tailed sheep. But, ranking of genes for stability was not the same from each of the three software packages. The least stable gene identified by geNorm and BestKeeper was *Actb*, whereas the least stable gene was *18S rRNA* in NormFinder analysis. The genes *Actb* and *18S rRNA* have been used as reference genes in a number of studies using different tissues (Laud *et al*., 2001; Grubor *et al*., 2004; Hein *et al*., 2004; Huggett *et al*., 2005), however their higher inter and intra-group variation make them unsuitable for normalization in Lanzhou fat-tailed sheep.

Results obtained from geNorm identify *Tubb2* and *Ywhaz* as the more stable genes when all time points were examined, however, the M values obtained for *Gapdh* and *Sdha* are quite similar (Fig. 3a) and thus these genes are also likely to be suitable for normalization. The same set of gene was found to be stable when all tissues were examined (Fig. 3b). The values for the pairwise variation between two sequential normalization factors were above the threshold value of 0.15 (data not shown). However, practical considerations must be taken into account when selecting the number of reference genes required to minimize variation in the normalization factor. For example, it might be impractical to use multiple reference genes in experiments involving a large number of samples but only a few target genes, or when limited amounts of RNA are available.

The most stable genes identified for all time points by NormFinder ranked in descending order were *Sdha>ywhaz>Tubb2>Actb>Pgk1>18S rRNA*. The similar stability indices obtained for *Tubb2, Ywhaz, Gapdh* and *Sdha*, identified all of these genes as suitable for normalization. In this

analysis, *18S rRNA* was ranked as the worst reference gene. Vandesompele *et al.* (2002) criticize the use of *18S rRNA* as a housekeeping gene due to its high abundance making baseline subtraction difficult. Also, transcription of rRNA and mRNA occur via RNA polymerase I and II, respectively which may lead to imbalances in the two mRNA fractions as reported by Solanas *et al.* (2001). Present result is agreement with this viewpoint.

Similar to the results of geNorm and NormFinder, BestKeeper analysis revealed the most stable genes to be *Gapdh, Tubb2, Sdha* (Table 3). The BestKeeper pairwise analysis sorted the candidate reference genes in the following series of decreasing stability: *Gapdh>Tubb2>Sdha>18S rRNA>Ywhaz>Pgk1>Actb*. Interestingly, *Actb* which has been used as a reference gene in numerous studies (Grubor *et al.*, 2004; Hein *et al.*, 2004), was found to be the least stable gene in this analysis, This apparent discrepancy is due to the different expression pattern of *Actb* when compared to the other potential reference genes, since the BestKeeper (Pfaffl *et al.*, 2004) and geNorm (Vandesompele *et al.*, 2002) algorithms assume that the expression ratio of two ideal reference genes is identical in all samples. *18S rRNA* was found to be the most stable gene in tissues in this analysis. As the *18S rRNA* is highly abundant and account for the vast majority of RNA, it is unsurprising that *18S rRNA* is found to be stable across the samples as equal amounts of RNA were reverse transcribed.

Vandesompele *et al.* (2002) and Bower and Johnston (2009) recommend using the geometric average of three reference genes for accurate normalization. To assess the suitability of the reference gene candidates, we first normalized the expression of *Dgat1* to combinations of the geometric average of two reference genes from *Gapdh, Tubb2, Sdha* and *Ywhaz* (Fig. 6). We found significant differences in *Dgat1* expression in the age of month 5, 9 and 13 and the most significant differences in *Dgat1* expression in the age of month 7 and 9 when comparing results from different combinations of reference genes. However, when three genes were used, apart from the combinations between *Gapdh, Tubb2, Sdha* and *Gapdh, Tubb2, Ywhaz* in month 9, there were no significant differences between any other of the combinations of reference genes (Fig. 7), indicating the best choices of combinations either *Gapdh, Sdha, Ywhaz* or *Tubb2, Sdha, Ywhaz* are suitable for normalization when the geometric average of three genes is used.

CONCLUSION

This study identified two sets of reference genes of *Gapdh, Sdha, Ywhaz* and *Tubb2, Sdha, Ywhaz* that can be used for normalization of expression data in the development and in the different tissues of Lanzhou fat-tailed sheep. The data provided in this paper may also be useful in guiding researchers performing gene expression in other species of sheep, *Ovis aries*.

ACKNOWLEDGMENTS

This work was partially funded by science and technology research program of Lanzhou Municipality (2006-2-59), agricultural bio-technology research and development plan of Gansu Province (GNSW-2009-13), Research program of State Ethnic Affairs Commission of China (2009-158-09XB02). We appreciate Dr. Michael Brown for the review of manuscript.

REFERENCES

Andersen, C.L., J.L. Jensen and T.F. Orntoft, 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.

Bower, N.I. and I.A. Johnston, 2009. Selection of reference genes for expression studies with fish myogenic cell cultures. BMC Molecular, Biol., 10: 80-80.

Bustin, S.A., 2000. Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. J. Mol. Endocrinol., 25: 169-193.

Bustin, S.A., V. Benes, T. Nolan and M.W. Pfaffl, 2005. Quantitative real-time RT-PCR-a perspective. J. Mol. Endocrinol, 34: 597-601.

Garcia-Crespo, D., R.A. Juste and A. Hurtado, 2005. Selection of ovine housekeeping genes for normalisation by real-time RT-PCR; analysis of PrP gene expression and genetic susceptibility to scrapie. BMC Vet. Res., 1: 3-3.

Grubor, B., J.M. Gallupa, R. Ramirez-Romero, T.B. Bailey, E.C. Crouchd, K. Brogdene and M.R. Ackermann, 2004. Surfactant protein D expression in normal and pneumonic ovine lung. Vet. Immunol. Immunopathol., 101: 235-242.

Hein, W.R., T. Barber, S.A. Cole, L. Morrison and A. Pernthaner, 2004. Long-term collection and characterization of afferent lymph from the ovine small intestine. J Immuno. Methods, 293: 153-168.

Huggett, J., K. Dheda, S. Bustin and A. Zumla, 2005. Real-time RT-PCR normalisation: Strategies and considerations. Genes and Immunity, 6: 279-284.

Laud, K., L. Hornez, I. Gourdou, L. Belair, A. Arnold,J.P. Peyrat and J. Djiane, 2001. Expression of BRCA1 gene in ewe mammary epithelial cells during pregnancy: Regulation by growth hormone and steroid hormones. Eur J Endocrinol, 145: 763-770.

Le Bail, A., S.M. Dittami, P.O. de Franco, S. Rousvoal, M.J. Cock, T. Tonon and B. Charrier, 2008. Normalisation genes for expression analyses in the brown alga model Ectocarpus siliculosus. BMC. Mol. Biol., 9: 75-75.

Passmore, M., M. Nataatmadja and J.F. Fraser, 2009. Selection of reference genes for normalization of real-time RT-PCR in brain-stem death injury in *Ovis aries.*, BMC Mol. Biol., 10: 72-72.

Pfaffl, M.W., A. Tichopad, C. Prgomet and T.P. Neuvians, 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.

Radonic, A., S. Thulke, I.M. Mackay, O. Landt, W. Siegert and A. Nitsche, 2004. Guideline to reference gene selection for quantitative real-time PCR. Biochem. Biophys. Res. Commun., 313: 856-862.

Rozen, S. and H. Skaletsky, 2000. Primer3 on the WWW for general users and for biologist programmers. Methods, Mol. Boil., 132: 365-386.

Solanas, M., R. Moral and E. Escrich, 2001. Unsuitability of using ribosomal RNA as loading control for Northern blot analyses related to the imbalance between messenger and ribosomal RNA content in rat mammary tumors. Anal. Biochem., 288: 99-102.

Vandesompele, J., K. De Preter, F. Pattyn, B. Poppe, N. van Roy, A. De Paepe and F. Speleman, 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol., 3: 1-11.

Warrington, J.A., A. Nair, M. Mahadevappa and M. Tsyganskaya, 2000. Comparison of human adult and fetal expression and identification of 535 housekeeping/maintenance genes. Phys. Genomics., 2: 143-147.

Zaros, L.G., L.L. Coutinho, L.H. Sider, H.R. de Medeiros and M.R.M. das Neves *et al.*, 2010. Evaluation of reference genes for real-time PCR studies of Brazilian Somalis sheep infected by gastrointestinal nematodes. Genetics, Mole. Biol., 33: 486-490.