

## Selection and use of reference genes in mouse mammary glands

L.Q. Han<sup>1,2</sup>, G.Y. Yang<sup>2</sup>, H.S. Zhu<sup>1</sup>, Y.Y. Wang<sup>1</sup>, L.F. Wang<sup>1</sup>, Y.J. Guo<sup>1</sup>,  
W.F. Lu<sup>1</sup>, H.J. Li<sup>1</sup> and Y.L. Wang<sup>1,2</sup>

<sup>1</sup>College of Animal Science and Veterinary Medicine,  
Henan Agricultural University, Zhengzhou, China

<sup>2</sup>Key Laboratory of Animal Growth and Development,  
Ministry of Agriculture, Zhengzhou, China

Corresponding author: Y.L. Wang

E-mail: ylwang2001@yahoo.cn

Genet. Mol. Res. 9 (1): 449-456 (2010)

Received November 30, 2009

Accepted January 5, 2010

Published March 16, 2010

**ABSTRACT.** Obtaining quantitative data concerning gene expression is important for understanding milk synthesis in mammary glands. Quantitative real-time PCR (qRT-PCR) is an efficient tool to calculate gene expression; however, it is necessary to find valid reference genes for normalization of qRT-PCR data. We applied the geNorm software to eight commonly used reference genes to identify the most stable and optimal genes for the mouse mammary gland. Based on this analysis, HPRT, RPL and GAPDH are the most appropriate reference genes for data normalization. We tested the expression of the  $\alpha$ -lactalbumin and fatty acid synthase genes using these three reference genes, both normalized and non-normalized. The normalized mRNA expression ratio was significantly different from the non-normalized ratio. We recommend the use of these three reference genes for the normalization of qRT-PCR data in gene expression studies of mouse mammary glands.

**Key words:** geNorm; Lactation; Quantitative real-time PCR; Reference gene

## INTRODUCTION

The mammary gland is among the most complex tissues in mammals, where more dramatic changes occur, especially in gene expression in preparation for milk synthesis in lactation. Lactation is normally associated with the end of pregnancy and around the time of nursing. The acquisition of knowledge about gene expression during lactation cycle is crucial for a better understanding of metabolism and is useful for further refinement of gene regulation in milk synthesis. Recently, most studies of mammary gland have focused on quantitative real-time polymerase chain reaction (qRT-PCR) (Pfaffl et al., 2003; Modha et al., 2004; Bernard et al., 2005; Bionaz and Looor, 2008).

qRT-PCR has been used for gene expression analysis for over a decade (Heid et al., 1996). Many factors can contribute to variability in the analysis of samples, making the results difficult to reproduce between experiments (Schmittgen and Zakrajsek, 2000). Accurate normalizations of qRT-PCR results are thus essential to allow precise comparisons between samples. The most applied approach for normalization is the use of reference genes, often referred to as housekeeping genes (HKG) (Bustin et al., 2005). However, reference genes need to be properly validated when designing quantitative gene expression studies (Schmittgen and Zakrajsek, 2000), and the use of single housekeeping genes is inappropriate for qRT-PCR studies (Tricarico et al., 2002). Vandesompele et al. (2002) described the geNorm software to identify the most stable reference genes in a given set of tissue samples, and to determine the optimal number of genes required for reliable normalization of qRT-PCR data. geNorm has been applied to the identification of reference genes in bovine and swine mammary gland (Bionaz and Looor, 2007; Tramontana et al., 2008), but not in mouse mammary gland.

In the present study, the expression stability of eight commonly used reference genes was compared in mouse mammary gland. The aim of this study was to identify the most stable, or a combination of the most stable reference genes, in mouse mammary gland from pregnancy to lactation. This could enable us to assess the suitability of the genes for normalization and to apply the best number of reference genes for studies of gene expression.

## MATERIAL AND METHODS

### Animals and sampling

KM mice were maintained in accordance with the laboratory animal central of Henan province and following procedures approved by the Animal Care and Use Committee of the University of Zhengzhou for Medical Sciences. Animals were housed individually in polycarbonate cages under 24°C, 40% humidity, and a 12-h light-dark cycle. Animals were provided food and water *ad libitum*. P1 was identified as the first day when a postcoital plug was observed. Similarly, L1 was identified as the first day when litters were present. The litter number was normalized to 8 on L3. Mammary glands from six mice were taken for the time series at P18, L6, L12 and L18. The excised tissue was frozen in liquid nitrogen and stored at -80°C until RNA extraction.

### Selection of genes and primer design

A total of eight genes, most commonly used as housekeeping genes for normalization,

were selected for gene expression analysis (Table 1). Primer sequences were designed with the Primer Express Software (Applied Biosystems). The primer sequence, correlation of stand curve and reaction efficiencies for the genes were submitted to the RT Primer Database [<http://medgen.ugent.be/rtprimerdb/>] and are shown in Table 1.

**Table 1.** Primer information for qRT-PCR amplification.

Gene symbol (GenBank number)	Primer sequence (5'-3')	Product length (bp)	Correlation	Efficiency	Function
ACTB (NM-007393)	CATCCGTAAGACCTCTATGCCAAC ATGGAGCCACCGATCCACA	171	0.99	1.99	Cytoskeletal structural protein
GAPDH (NM-001001303)	TGTGTCCGTCGTGGATCTGA TTGCTGTTGAAGTCGCAGGAG	150	0.99	1.95	Glycolytic pathway enzyme
HPRT (NM-013556)	CCTAAGATGAGCGCAAGTTGAA CCACAGGACTAGAACCTGCTAA	86	0.99	2.03	Metabolic salvage of nucleotides
B2M (NM-009735)	TGCTACTCGGCGCTTCAGTC AGGCGGTGGAAGTGTGTTAC	200	0.97	2.02	Major histocompatibility complex
UBC (NM-016723)	AGCCAGTGTACCACCAAG ACCCAAGAACAAGCACAAGG	97	0.98	1.97	Protein degradation
RPL13A (NM-009438)	CCTAAGATGAGCGCAAGTTGAA CCACAGGACTAGAACCTGCTAA	168	0.98	1.93	Structural component of ribosomal subunit
TUBULIN (NM-011653)	FAAGGAGGATGCTGCCAATAA GCTGTGGAAAACCAAGAAGC	135	0.99	2.01	Cytoskeletal structural protein
CYC (NM-025567)	CCAGGTATACAAGCAGGTGTGCTC CATCATTAGGGCCATCCTGGAC	140	0.98	1.98	Electron transport pathway
LA (NM-010679)	AAGTAGTGAGTCCCGAGTTC GGCTTCCAGTAGTCGATTCCTT	132	0.97	1.99	Whey protein
FAS (NM-007988)	AGCACTGCCTTCGGTTCAGTC AAGAGCTGTGGAGGCCACTTG	94	0.99	1.97	Fatty acid synthesis enzyme

## RNA extraction

All RNA samples from mouse mammary gland were extracted using the Pure Yield TMRNA Midiprep System (Promega) according to the instruction manual. RNA integrity was assessed by electrophoretic analysis of 28S and 18S rRNA subunits. RNA quality and quantity were measured with a NanoDrop ND-1000 spectrophotometer. The average A260/A280 ratio was 2.1 and the average A260/A230 was 2.11.

## qRT-PCR

Each cDNA was synthesized by M-MLV reverse transcriptase (RT, Promega). For the first-strand cDNA synthesis, 2 µg total RNA and 1 µL Oligo (dT) 15 primer (0.5 µg/µL) were added to each sample before incubation at 70°C for 5 min. Samples were cooled on ice immediately and subjected to a mixture consisting of 5 µL 5X M-MLV RT-buffer, 1.25 µL dNTP (10 mM), 0.75 µL RNAsin (40 U/µL), 1 µL M-MLV RT (200 U/µL), and nuclease-free water to 25 µL. Samples were mixed, and then incubated for 60 min at 42°C, and finally for 15 min at 70°C. cDNA was then diluted 1:10 with nuclease-free water and stored at -20°C. Real-time PCR was performed using 2 µL cDNA diluted

with 18  $\mu\text{L}$  of a mixture composed of 10  $\mu\text{L}$  2X SYBR green master mix (TOYOBO), 0.25  $\mu\text{L}$  each of 10  $\mu\text{M}$  forward and reverse primers, and 7.5  $\mu\text{L}$  nuclease-free water in a MicroAmp Optical 96-well reaction plate. A 10-fold standard curve was made with cDNA from a pool of known amounts of RNA from tissues. The reactions were performed in Eppendorf Mastercycler realplex instrument using the following conditions: 2 min at 95°C, 40 cycles of 15 s at 95°C, 15 s at 60°C and 45 s at 72°C. The presence of a single PCR product was verified by the dissociation protocol.

### Gene expression stability analysis

Analysis of the gene expression stability and determination of the optimal number of reference genes were performed using the geNorm software. geNorm Visual Basic Application for Microsoft Excel is available at [[http://medgen.ugent.be/\\_jvdesomp/genorm/](http://medgen.ugent.be/_jvdesomp/genorm/)].

### Relative mRNA expression ratio normalization

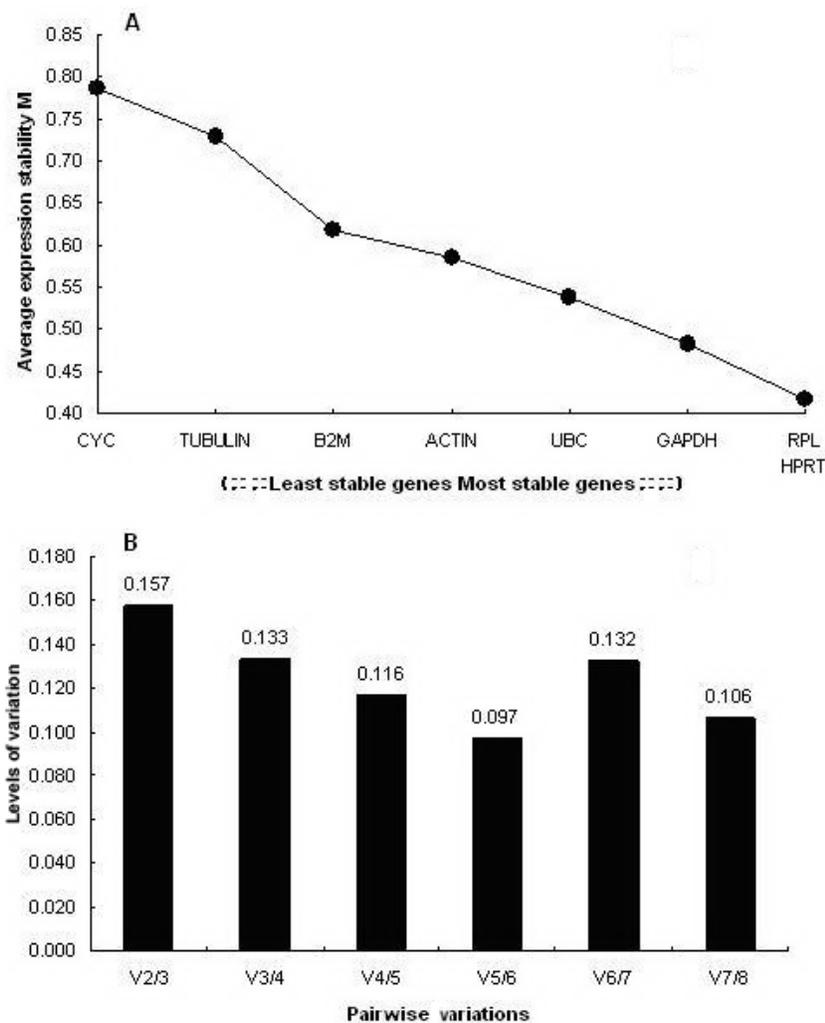
The target genes  $\alpha$ -lactalbumin (LA) and fatty acid synthase (FAS) were used to measure the change in mRNA expression ratio between non-normalized and normalized with the geometric average of three reference genes (HPRT, RPL, GAPDH). To calculate the relative expression ratio (R) of the target gene, the formula described by Pfaffl (2001) and Vandesompele et al. (2002), which included correction for PCR efficiency (E) and three reference genes (REF), was used:  $\text{Ratio} = E_{\text{target}} \Delta C_{\text{t target}} (\text{control-sample}) / \text{GEOMEAN} (E_{\text{ref}} \Delta C_{\text{t ref}} (\text{control-sample}))$ . The data were analyzed by GLM in SPSS10.0 to evaluate the effect of time and normalization on gene expression.

## RESULTS

Eight reference genes were amplified in the tissues examined, and single-band products of appropriate size were obtained by agarose gel electrophoresis (Supplementary Figure 1). All qRT-PCR assays produced a single peak in the melting curve (Supplementary Figure 2).

The gene expression stability analysis, using the geNorm software, ranked the genes based on their average expression stability value (M) (Figure 1A). The most stable reference genes were identified by stepwise exclusions of the least stable gene and recalculating the M values: the lower the M value, the more stable the gene under consideration. The genes HPRT and RPL were found to be the two most stable genes with the lowest M value, followed by the genes GAPDH, UBC, ACTB, B2M, TUBULIN, in their order of appearance. CYC is the least stable gene observed with the highest M value in mammary gland (Figure 1A).

The levels of variation (V) in reference gene stability are presented in Figure 1B. Vn values were calculated by stepwise inclusion of more reference genes until the (n + 1)th gene had no significant contribution to the newly calculated normalization factor. In this case, the V of two reference genes (V2/3) was the highest (0.158), and five reference genes (V5/6) had the lowest value (0.097). Vandesompele et al. (2002) proposed 0.15 as a cut-off value for the variation below which the inclusion of an additional reference gene is not required. It appears that three reference genes is the best (V3/4 < 0.15, Figure 1B).



**Figure 1.** Gene expression stability analysis of the eight reference genes using the geNorm software. **A.** Average expression stability values (M), starting from the least stable gene at the left, ending with the two most stable genes on the right. **B.** Pairwise variation (V), starting with the two most stably expressed genes on the left, with the inclusion of a 3rd, 4th, 5th gene, etc., moving to the right.

There are significant differences of time and normalization effect on the gene expression ratio (Table 2). The normalized ratio was significantly upregulated by the stage of lactation ( $P < 0.05$  or  $P < 0.01$ ), but the non-normalized ratio was only significantly upregulated at L18 ( $P = 0.001$  for LA,  $P = 0.020$  for FAS). From the significance test of normalization, the normalized ratio is significantly different from non-normalized ratio ( $P = 0.030$  for LA,  $P = 0.022$  for FAS; Table 2).

**Table 2.** Relative gene expression ratio of  $\alpha$ -lactalbumin (LA) and fatty acid synthase (FAS).

Time	Non-normalized				Normalized with 3 HKG			
	LA		FAS		LA		FAS	
	Ratio <sup>a</sup>	P <sup>b</sup>	Ratio	P	Ratio	P	Ratio	P
P18	1.75 ± 2.01	-	1.49 ± 1.52	-	1.19 ± 0.47	-	1.18 ± 0.48	-
L6	8.39 ± 6.01	0.228	3.42 ± 1.95	0.274	13.93 ± 9.90	0.041	5.11 ± 1.26	0.024
L12	10.36 ± 5.19	0.079	3.19 ± 1.71	0.376	21.53 ± 7.44	0.001	6.47 ± 1.82	0.002
L18	20.43 ± 8.22	0.001	4.82 ± 1.91	0.020	21.88 ± 8.85	0.001	5.61 ± 3.69	0.010
P <sup>c</sup>	-	-	-	-	0.030	-	0.022	-

<sup>a</sup>Ratio of LA and FAS was calculated in mouse mammary tissue at L6, L12, L18 (lactation 6, 12, 18 day). HKG = housekeeping gene. Values are reported as means ± SD and were expressed as fold change relative to P18 (pregnant 18 day). <sup>b</sup>P: Significance test of time effect on gene expression ratio. The data of P18 were set as control group. <sup>c</sup>P: Significance test of normalization effect on gene expression ratio. The non-normalized data were set as control group.

## DISCUSSION

Despite the advent of microarrays, qRT-PCR is still the most accurate method to analyze mRNA expression of small numbers of genes (Hembruff et al., 2005). As shown previously, the classical approach for selecting appropriate reference genes is important. The so-called HKG always regulates basic and ubiquitous cellular functions. To our knowledge, the eight selected HKG have independent functions in cellular maintenance, and the regulation of their expression is not assumed to be directly related. Only ACTB and TUBULIN share an identical biochemical process in the cell, namely cytoskeletal (Table 1).

The geNorm software is a statistical algorithm that determines HKG stability measure (M) and pairwise variation (V). This analysis relies on the principle that the expression ratio of two ideal reference genes is identical in all samples, regardless of the experimental condition or cell type (Vandesompele et al., 2002). Thus, assessment of HKG based on pairwise comparisons appears to be a more reliable method for selection of HKG, because it takes into account a potential dilution effect on gene expression (Bionaz and Looor, 2007). Our results showed that HPRT and RPL were the two most stable genes, while the most commonly used reference gene ACTB was not a suitable HKG for mouse mammary gland. To determine how many reference genes should be used, geNorm shows the pairwise variation  $V_n / V_{n+1}$ . A large variation means that the added gene has a significant effect and should probably be included for calculation of the normalization factor (Goossens et al., 2005). Our results showed that three genes were appropriate for data normalization. The cut-off was reduced from 0.158 (V2/3) to 0.133 (V3/4 < 0.15) (Figure 1B). The addition of a 4th gene (V4/5) would increase reliability in normalization, but it was not justified given the lower cut-off value used.

We used two target genes (LA and FAS) to determine the effect of normalization on gene expression. By means of a thorough analysis, the results showed that the normalized ratio was significantly different from non-normalized data. Lisowski et al. (2008) found that the use of a single gene for normalization may lead to errors, and that it is important to use multiple control genes. Considering the variable mRNA levels and wide variability of the conventional qRT-PCR procedure, we clearly point out that the use of three reference genes is the appropriate solution for the normalization problem of real-time PCR.

## CONCLUSIONS

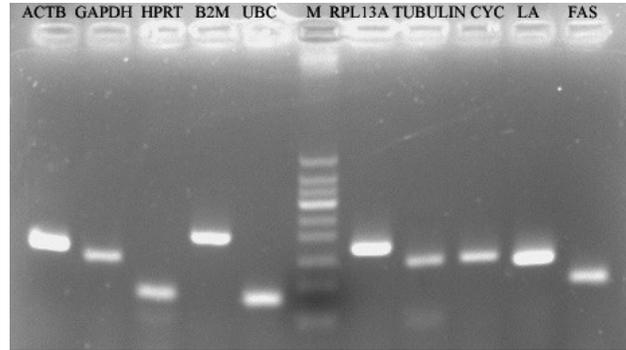
In summary, our findings confirm that the use of reference genes is important for normalizing qRT-PCR data. We recommend the use of HPRT, RPL and GAPDH for the normalization of qRT-PCR data for a broad range of different gene expression studies in mouse mammary gland.

## ACKNOWLEDGMENTS

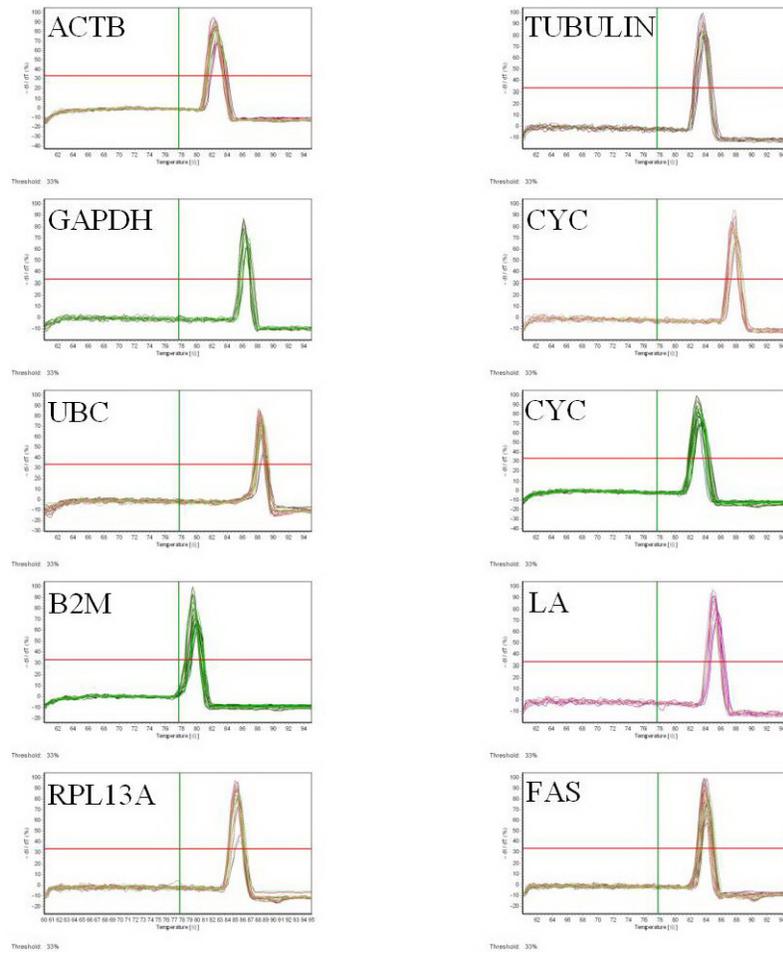
Research supported by Chinese National Programs for Science and Technology Development (#2006BAD04A03-10).

## REFERENCES

- Bernard L, Leroux C, Bonnet M, Rouel J, et al. (2005). Expression and nutritional regulation of lipogenic genes in mammary gland and adipose tissues of lactating goats. *J. Dairy Res.* 72: 250-255.
- Bionaz M and Loor JJ (2007). Identification of reference genes for quantitative real-time PCR in the bovine mammary gland during the lactation cycle. *Physiol. Genomics* 29: 312-319.
- Bionaz M and Loor JJ (2008). Gene networks driving bovine milk fat synthesis during the lactation cycle. *BMC Genomics* 9: 366.
- Bustin SA, Benes V, Nolan T and Pfaffl MW (2005). Quantitative real-time RT-PCR - a perspective. *J. Mol. Endocrinol.* 34: 597-601.
- Goossens K, Van Poucke M, Van Soom A, Vandesompele J, et al. (2005). Selection of reference genes for quantitative real-time PCR in bovine preimplantation embryos. *BMC Dev. Biol.* 5: 27.
- Heid CA, Stevens J, Livak KJ and Williams PM (1996). Real time quantitative PCR. *Genome Res.* 6: 986-994.
- Hembruff SL, Villeneuve DJ and Parissenti AM (2005). The optimization of quantitative reverse transcription PCR for verification of cDNA microarray data. *Anal. Biochem.* 345: 237-249.
- Lisowski P, Pierzchala M, Goscik J, Pareek CS, et al. (2008). Evaluation of reference genes for studies of gene expression in the bovine liver, kidney, pituitary, and thyroid. *J. Appl. Genet.* 49: 367-372.
- Modha G, Blanchard A, Iwasio B, Mao XJ, et al. (2004). Developmental changes in insulin-like growth factor I receptor gene expression in the mouse mammary gland. *Endocr. Res.* 30: 127-140.
- Pfaffl MW (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* 29: e45.
- Pfaffl MW, Wittmann SL, Meyer HH and Bruckmaier RM (2003). Gene expression of immunologically important factors in blood cells, milk cells, and mammary tissue of cows. *J. Dairy Sci.* 86: 538-545.
- Schmittgen TD and Zakrajsek BA (2000). Effect of experimental treatment on housekeeping gene expression: validation by real-time, quantitative RT-PCR. *J. Biochem. Biophys. Methods* 46: 69-81.
- Tramontana S, Bionaz M, Sharma A, Graugnard DE, et al. (2008). Internal controls for quantitative polymerase chain reaction of swine mammary glands during pregnancy and lactation. *J. Dairy Sci.* 91: 3057-3066.
- Tricarico C, Pinzani P, Bianchi S, Paglierani M, et al. (2002). Quantitative real-time reverse transcription polymerase chain reaction: normalization to rRNA or single housekeeping genes is inappropriate for human tissue biopsies. *Anal. Biochem.* 309: 293-300.
- Vandesompele J, De Preter K, Pattyn F, Poppe B, et al. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* 3: 0034.1-0034-11.



Supplementary Figure 1. Agarose gel electrophoresis of PCR products.



Supplementary Figure 2. Dissociation curve with single peak of qRT-PCR assays.