Identification of reference genes for quantitative real-time PCR in the bovine mammary gland during the lactation cycle

Massimo Bionaz and Juan J. Loor
Mammalian NutriPhysio Genomics, Department of Animal Sciences, University of Illinois, Urbana, Illinois
Submitted 10 October 2006; accepted in final form 29 January 2007

Bionaz M, Loor JJ. Identification of reference genes for quantitative real-time PCR in the bovine mammary gland during the lactation cycle. Physiol Genomics 29: 312–319, 2007. First published February 6, 2007; doi:10.1152/physiolgenomics.00223.2006.—Achieving greater understanding of the genomic influence on milk synthesis in dairy cows represents a daunting challenge. Bovine-specific microarrays have allowed for high-throughput gene expression analysis of the mammary transcriptome. However, real-time PCR (qPCR) still represents the method of choice for accurate expression profiling of small numbers of genes and verification of key microarray relationships. This method is extremely sensitive but requires data normalization to account for analytical errors. Ideally, expression of genes used as internal controls should not be affected by specific treatments or physiological state. Mammary biopsies were collected from five cows each at −15, 1, 15, 30, 60, 120, and 240 days relative to parturition for gene expression profiling. We evaluated expression of nine genes (RPS9, ACTB, GAPD, GTP, ITGβ4BP, MRPL39, RPS23, RPS15, and UXT) that could serve as internal controls in mammary tissue using qPCR. Due to gradual increases in mammary RNA concentration (μg/mg tissue) over lactation, all genes investigated experienced a dilution effect. We used pairwise comparison of expression ratios to analyze the reliability of these genes as internal controls. UXT, RPS9, and RPS15 had the most stable expression ratios across cow and time. We also assessed co-regulation among genes through network analysis. Network analysis suggested co-regulation among most of the genes examined, with MYC playing a central role. Pairwise comparison was suitable for finding appropriate internal controls in mammary gland tissue. Results showed that the geometrical average of UXT, RPS9, and RPS15 expression could be used as internal control for longitudinal mammary gene expression profiling.

cow; genomics; reference gene; RNA; dilution effect

A KEY CHALLENGE TO BE MET by dairy scientists and the dairy industry is to understand which genes control the composition of milk, how these genes are regulated, and how they might be manipulated to enhance both the manufacturing and health properties of dairy products. Genomic tools developed at the University of Illinois have allowed the evaluation of gene expression in bovine tissues on a comprehensive scale (22, 23). This, in turn, will permit identification of and will allow ascribing function to genes that are important regulators of milk component synthesis during lactation. Our group is currently embarked in projects utilizing bovine-specific microarray technology to increase understanding of the genomic influence on milk synthesis in dairy cows (http://labs.ansci.uiuc.edu/loor).

Despite the advent of bovine-specific microarrays (e.g., Ref. 8), real-time PCR (qPCR) is still the most accurate method to analyze mRNA expression of low numbers of genes and also to confirm key relationships identified by microarray analysis (15, 36). qPCR requires data normalization to obtain high precision because this technique is subject to analytical errors that introduce variation (16, 36). There are other variables that also need to be controlled (e.g., number of cells in the tissue, efficiency of reactions) (37). Data normalization can be achieved using the total amount of cells or total amount of tissue, total amount of RNA, genomic DNA, use of internal controls such as 18S or 28S ribosomal RNA, or alien molecules (16). However, all these methods present limitations (16, 27, 37). The use of internal control genes, often referred to as housekeeping genes (HKG), is the most reliable method for taking into account quantity of input RNA, sample loss during handling, and variation in the kinetics of the reverse-transcription reaction. HKG also can serve as negative controls to compare with other candidate genes that fluctuate in response to treatments or physiological state. Central to the concept of HKG for normalization is the notion that expression of a HKG should not vary between tissues or cells under investigation, or in response to experimental treatments (37).

There is increased agreement among scientists working with qPCR technology that proper evaluation of HKG should be performed before selecting the most appropriate for data normalization (16) to avoid additional variation and errors in the final data (11). Unfortunately, many studies reporting qPCR data have relied on published literature or previous experience for the selection of HKG without proper validation. The most widely used method to evaluate HKG is the statistical analysis of the raw gene expression data (i.e., not normalized) to ensure that a group effect (or treatment effect) among samples is not associated with the variation in the expression of the selected HKG. However, this approach does not take into consideration the variation generated by errors during the analytical procedure. Several statistical algorithms have been developed to properly validate HKG (2, 27, 35, 37).

The bovine mammary gland increases milk yield dramatically during the first few weeks of lactation. During this period characterized by close hormonal regulation, there is a set of well-studied genes involved in milk synthesis that increase dramatically in expression (25). As a consequence, the RNA concentration in mammary tissue also increases (4, 7, 25, 33). After peak lactation, milk synthesis decreases gradually until the mammary gland is induced into involution prior to entering the nonlactating or dry period. During the declining phase of milk production mammary RNA concentration also decreases (7, 33). We believe that the variation in mammary RNA concentration is an important factor to consider when an internal control gene(s) is chosen for long-term studies of
mammary gene expression, or longitudinal studies of key metabolic tissues such as liver or adipose. We argue that genes with relatively constant expression across the lactation cycle will appear as significantly downregulated due to the gradual increases in total mammary RNA concentration, i.e., there is a potential dilution effect. A number of research papers dealing with studies of mammary gland gene expression have reported data where clear dilution of the housekeeping genes was observed (e.g., 24, 28, 29, 33). Some of these have highlighted potential dilution effects (e.g., 28), but none, to our knowledge, attempted to test the appropriateness of the chosen HKG. Thus, the potential dilution of stably expressed genes appears to be an “accepted” limitation.

The objective of the current study was to examine the expression patterns of nine genes (RPS9, ACTB, GAPD, GTP, ITGB4BP, MRPL39, RPS23, RPS15, and UXT) that could serve as internal controls for longitudinal mammary gene expression studies.

**MATERIALS AND METHODS**

*Animals and sampling.* Five multiparous (3rd or greater lactation) Holstein dairy cows from the University of Illinois Dairy Cattle Research Unit were used. Percutaneous biopsies from each cow were obtained from the right or left rear quarter of the mammary gland at -15 (-13 ± 3 days), 1, 15, 30, 60, 120, and 240 days relative to parturition. The right rear quarter was biopsied at -15, 15, 60, and 240 days, and the left rear quarter was biopsied at 1, 30, and 120 days. The midsection of the rear quarters were selected for the initial biopsy. Subsequently, a different section located ~2–3 inches from the original incision site of the right or left rear quarter was selected for biopsy. Biopsies were conducted at 0700 h (post-AM milking) as described by Farr et al. (9). Briefly, after making the skin incision, we performed blunt dissection of the mammary capsule to ensure tissue obtained during the biopsy was mammary parenchyma. Immediately after removal of the biopsy instrument from the capsule, pressure was applied to the wound until visual signs of bleeding were absent. The skin incision was closed with four or five Michel clips (11 mm; Henry Schein, www.henryschein.com). The incision site was sprayed with topical antiseptic (Povidone Iodine Ointment, 10%; Taro Pharmaceuticals). Health was monitored post-surgery by recording rectal temperature, milk yield, and feed intake daily for 7 days. Surgical clips were removed 7 days postbiopsy. No incidences of mastitis were detected for any cow during the 7 days postsurgery. All procedures were conducted under protocols approved by the University of Illinois Institutional Animal Care and Use Committee.

*RNA extraction and PCR.* Biopsy tissue was weighed (~0.5–1.0 g) and immediately subjected to RNA extraction with ice-cold TRIzol (Invitrogen) as described in Loo et al. (22, 23). RNA integrity was assessed by electrophoretic analysis of 28S and 18S rRNA subunits. Genomic DNA was removed from RNA with DNase (Qiagen) using RNase Mini Kit columns (Qiagen). RNA concentration was measured with a NanoDrop ND-1000 spectrophotometer (www.nanodrop.com). The purity of RNA (A260/A280) was measured with a NanoDrop ND-1000 spectrophotometer (www.nanodrop.com). RNA integrity was measured with a NanoDrop ND-1000 spectrophotometer (www.nanodrop.com). Genomic DNA was removed from RNA with DNase (Qiagen) using RNase Mini Kit columns (Qiagen). RNA concentration was measured with a NanoDrop ND-1000 spectrophotometer (www.nanodrop.com). The purity of RNA (A260/A280) was measured with a NanoDrop ND-1000 spectrophotometer (www.nanodrop.com). RNase Free Water (Qiagen) was added. The mixture was incubated at 65°C for 5 min, kept on ice for 3 min. A total of 6 μl of Master Mix composed of 4 μl 5X First-Strand Buffer, 1 μl 0.1 M DDT, 0.5 μl (100 U) of SuperScript III RT (Invitrogen), and 0.5 μl of RNase Inhibitor (Promega) was added. The reaction was performed in an Eppendorf Mastercycler Gradient using the following temperature program: 25°C for 5 min, 50°C for 60 min, and 70°C for 15 min. cDNA was then diluted 1:4 with DNase/RNase-free water. We combined 4 μl of diluted cDNA with 6 μl of a mixture composed of 5 μl 1X SYBR Green master mix (Applied Biosystems), 0.4 μl each of 10 μM forward and reverse primers, and 0.2 μl DNase/RNase-free water in a MicroOptic Optical 384-Well Reaction Plate (Applied Biosystems). Each sample was run in triplicate and a seven-point relative standard curve plus the nontemplate control were used (User Bulletin #2, Applied Biosystems). The twofold standard curve was made using cDNA from an RNA pool of known amounts of RNA from five bovine tissues. The reactions were performed in an ABI Prism 7900 HT SDS instrument (Applied Biosystems) using the following conditions: 2 min at 50°C, 10 min at 95°C, 40 cycles of 15 s at 95°C, and 1 min at 60°C. The presence of a single PCR product was verified by the dissociation protocol using incremental temperatures to 95°C for 15 s plus 65°C for 15 s. Complete details regarding amplification of target genes can be found at http://docs.appliedbiosystems.com/pe/bio/docs/04364014.pdf.

Data were calculated with the 7900 HT Sequence Detection Systems Software (version 2.2.1, Applied Biosystems) using a seven-point standard curve, as required for the evaluation of gene stability. Myelocytomatosis viral oncogene homolog (MYC), lactalbumin (LALBA), and estrogen receptor 1 (alpha) (ESR1) expression data were normalized using the geometric mean of the three most stable genes among the ones tested as internal controls (37).

*Selection of genes and primer design.* Primers for the nine selected genes and MYC, LALBA, and ESR1 are shown in Suppl. Table S1. (The online version of this article contains supplemental material.) Commonly used HKG such as β-actin (ACTB) and gyceraldehyde-3-phosphate-dehydrogenase (GAPD) were selected. Ribosomal protein S15 (RPS15) was chosen because it had been previously used as HKG in a bovine mammary gland longitudinal study (21). Ribosomal protein S23 (RPS23) also was previously used by various authors as HKG (e.g., Ref. 2). Ribosomal protein S9 (RPS9) was found to be stable in a time course experiment with bovine liver conducted in our laboratory (17). The remaining four genes were observed to be stable in a microarray data set from a longitudinal study of the bovine mammary gland transcriptome (Biomaz M and Loor JJ, unpublished). Among that data set containing expression patterns for over 13,000 bovine genes, integrin-B4 binding protein (ITGB4BP), GTPase-like protein (GTP or MTGI), mitochondrial ribosomal protein L39 (MRPL39), and ubiquitously expressed transcript isofrom 2 (UX1) appeared to behave as a suitable HKG (assessed with GeneSpring GX 7, Agilent Technologies). Primers were designed with Primer Express software fixing the amplicon length to 100–150 bp with low specificity. PCR products were run in a 3% agarose gel (Invitrogen) stained with ethidium bromide (Suppl. Fig. S1). Efficiency of PCR amplification for each gene using the standard curve method is reported in Suppl. Table S1. Although the efficiency of amplification might vary when using SYBR Green versus a commercial kit to test the specificity of a given primer, the capacity of the primer to bind (i.e., specificity) the DNA sequence should not be affected. PCR products were sequenced at the Core DNA Sequencing Facility of the Roy J. Carver Biotechnology Center at the University of Illinois,
Urbana-Champaign (Suppl. Table S2). Aliquots of all PCR products are stored to enable additional confirmatory sequencing of PCR products in the future if required.

**HKG stability evaluation.** HKG were submitted to statistical analysis using a MIXED model of SAS (release 8.0; SAS Inst., Cary, NC) with repeated measures to evaluate the effect of time relative to parturition on mRNA expression. Compound symmetry was the most appropriate covariate structure used for repeated measures analysis. The model included the fixed effect of time (−15, 1, 15, 30, 60, 120, and 240 days relative to parturition) and the random effect of cow. Statistical correlations between genes were performed using the PROC CORR procedure of SAS.

Gene expression stability was evaluated using the geNorm software (http://medgen.ugent.be/~jvdesomp/genorm/) following the procedures of Vandesompele et al. (37). Briefly, stability (M = gene-stability measure) refers to the constancy of the expression ratio between two non-co-regulated genes among all samples tested. The more stable the expression ratio among two genes, the more likely that the genes are appropriate internal controls, i.e., two ideal control genes should have an identical expression ratio in all samples regardless of experimental conditions, cell, and/or tissue type. The lower the M value, the higher the stability. Once the more stable genes are selected, the normalization factor (NF) is calculated using a geometrical average. geNorm also performs an analysis to test how many genes should be used to normalize expression data (37). A minimum of three genes should be used for normalization. However, the software is able to determine the utility of including more than two genes for normalization by calculating the pairwise variation (V) between NF obtained using n genes (best references) and NF obtained using n + 1 genes (addition of an extra less stable reference gene) (NF_{n+1}). A large decrease in the pairwise variation indicates that addition of the subsequent more stable gene (i.e., with lowest M value) has a significant effect and should be included for calculation of NF (37).

**Milk yield.** Milk yield was electronically recorded twice a day (each milking) during the entire lactating period, i.e., ≥240 individual measurements for each cow. To take into consideration the daily variation in milk yield as well as the effect of tissue biopsy on milk production, data from the last 5 days before biopsy were averaged with data from 4 days postbiopsy, disregarding production data from both the day of biopsy and the first day postbiopsy. This was deemed necessary because the biopsy procedure typically leads to a reduction in milk yield for the first two or three milkings (9).

**RESULTS AND DISCUSSION**

Primer for HKG were highly specific as shown by a single band when PCR products (amplicons) were run in an agarose gel (Suppl. Fig. S1), and by the presence of a single peak in the dissociation curve after the qPCR reaction (Suppl. Fig. S2). Efficiency of PCR reactions averaged 114% (Suppl. Table S1). All HKG had a similar pattern of expression in mammary tissue with a significant time effect (P < 0.05). Responses were characterized by decreased expression from day −15 relative to parturition to a nadir at 60–120 days postpartum (Fig. 1). Subsequently, expression patterns increased gradually until reaching prepartum (day −15) values at 240 days.

The relationship between the average expression pattern of HKG, RNA concentration (µg/mg tissue), and milk yield is reported in Fig. 2. Results for genes with a stable expression likely are affected by a dilution effect particularly when the metabolic capacity of the lactating mammary gland to synthesize milk components signals a dramatic increase in the expression of other genes (e.g., CSN3, BTN1A1; Refs. 3, 25). The existence of

---

Fig. 1. mRNA expression patterns of potential housekeeping genes (HKG) during the lactation cycle (data after transformation of CT values using standard curve).

Fig. 2. RNA concentration (SE = 0.3 µg/mg tissue), milk production (SE = 3.4 kg/day), and geometrical average expression of all HKG during the lactation cycle.
a dilution effect is reinforced by the expression pattern of LALBA, which codes for the regulatory subunit of the lactose synthase heterodimer (Fig. 3 and Table 1). Mammary RNA concentration increased \((P < 0.05)\) from \(-15\) prepartum through 30 days postpartum and then decreased gradually \((P > 0.10)\) as reported by others \((4, 33)\). RNA concentration had a weak correlation with milk yield (Table 1) similar to results observed by Capuco et al. \((4)\). A positive relationship between RNA synthesis and milk yield has been observed in mice, rabbits, and sheep \((7)\). We observed a significant negative correlation between RNA concentration and all HKG, as well as MYC, and a positive correlation with LALBA (Table 1). It is possible that losses of RNA occurred during the extraction protocol, which required centrifugations and transfer of supernatant. However, despite potential shortcomings during RNA extraction, these data suggest that at the onset of copious milk synthesis there is a battery of genes whose expression increases dramatically and leads to greater RNA concentration in mammary tissue. Thus, mRNA for genes with stable expression is diluted because the starting amount of RNA used for RT is the same across all samples. This could explain the statistically significant time effect found for all HKG tested and the negative correlations among those genes and both RNA concentration and mRNA for genes with stable expression is diluted because the starting amount of RNA used for RT is the same across all samples. This could explain the statistically significant time effect found for all HKG tested and the negative correlations among those genes and both RNA concentration and mRNA for genes with stable expression is diluted because the starting amount of RNA used for RT is the same across all samples. This could explain the statistically significant time effect found for all HKG tested and the negative correlations among those genes and both RNA concentration and milk yield, MYC, LALBA, and genes tested as internal controls.

**Table 1.** Pearson correlations between RNA concentration, milk yield, MYC, LALBA, and genes tested as internal controls

<table>
<thead>
<tr>
<th>Gene</th>
<th>Milk Yield</th>
<th>LALBA</th>
<th>MYC</th>
<th>RNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACTB</td>
<td>-0.23</td>
<td>-0.63</td>
<td>0.33</td>
<td>-0.40</td>
</tr>
<tr>
<td>GAPD</td>
<td>-0.12</td>
<td>-0.55</td>
<td>0.32</td>
<td>-0.41</td>
</tr>
<tr>
<td>GTP</td>
<td>-0.16</td>
<td>-0.64</td>
<td>0.74</td>
<td>-0.40</td>
</tr>
<tr>
<td>MRPL39</td>
<td>-0.03</td>
<td>-0.59</td>
<td>0.69</td>
<td>-0.35</td>
</tr>
<tr>
<td>RPS9</td>
<td>-0.28</td>
<td>-0.85</td>
<td>0.80</td>
<td>-0.54</td>
</tr>
<tr>
<td>RPS15</td>
<td>-0.22</td>
<td>-0.77</td>
<td>0.90</td>
<td>-0.45</td>
</tr>
<tr>
<td>RPS23</td>
<td>-0.33</td>
<td>-0.75</td>
<td>0.70</td>
<td>-0.37</td>
</tr>
<tr>
<td>UXT</td>
<td>-0.28</td>
<td>-0.81</td>
<td>0.78</td>
<td>-0.45</td>
</tr>
<tr>
<td>MYC</td>
<td>-0.36</td>
<td>-0.70</td>
<td>0.97</td>
<td>-0.43</td>
</tr>
<tr>
<td>ESR1</td>
<td>0.25</td>
<td>0.22</td>
<td>-0.27</td>
<td>-0.11</td>
</tr>
<tr>
<td>Milk Yield</td>
<td>0.47</td>
<td>-0.35</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>LALBA</td>
<td>-0.71</td>
<td>0.56</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Comparisons for the best internal controls are underlined. ACTB, β-actin; GAPD, glyceraldehyde-3-phosphate-dehydrogenase; GTP, GTPase-like protein; ITGβ4BP, integrin-β4 binding protein; MRPL39, mitochondrial ribosomal protein L39; RPS9, ribosomal protein S9; UXT, ubiquitously expressed transcript isoform 2; MYC, myelocytomatosis viral oncogene homolog; ESR1, estrogen receptor 1; LALBA, lactalbumin. Data normalized by the geometric average of UXT, RPS9, and RPS15. Data not normalized. *P < 0.10, **P < 0.05, ***P < 0.01.

The raw data for each single gene (after transformation of CT values) were analyzed using geNorm. Results of the pairwise comparison are reported in Fig. 4. Genes showing higher stability were UXT and RPS9, followed by RPS15. UXT is a ubiquitously expressed protein playing an essential role in normal centrosomal biogenesis and cell survival \((38)\). RPS9 is a ribosomal protein essential for the assembly of 30S ribosomal protein complex \((5)\). Interestingly, we previously demonstrated that RPS9 has high stability in bovine liver across different dietary treatments and physiological states \((17)\). Similar to the mammary gland, the periparturient liver undergoes drastic physiological adaptations at the level of gene expression \((22, 23)\). Thus, our overall data suggest that RPS9 is the most suitable internal control gene studied to date in bovine mammary and liver tissue and should be useful in longitudinal studies of their transcriptome. We believe, however, that the...
use of a single gene as internal control should be avoided as suggested by others (27, 37).

geNorm analysis of the optimal number of HKG to use as potential internal controls for normalization is reported in Fig. 5. It appeared that three genes (V2/3) were appropriate for data normalization. The cut-off was reduced from 0.15 to 0.10, as suggested previously (37), because of the lower variation in stability observed in the genes. Addition of a 4th gene (V3/4) would increase reliability in normalization but it was not justified given the lower cut-off value used. Among all genes tested, the least stable genes during the lactation cycle were ACTB and GAPD. These results are not surprising because a number of authors previously found them as unsuitable internal controls (11, 16, 37). Normalization of the remaining six genes by geNorm using the geometrical mean of UXT, RPS9, and RPS15 indicated a significant time effect for GAPD, MRPL39, and ITGB4BP, and a tendency for a time effect in the case of ACTB (Suppl. Fig. S3). These results suggest a differential transcriptional regulation for these genes during the lactation cycle, rendering them unreliable as internal controls in longitudinal experiments of mammary gene expression.

The absence of co-regulation among the genes used in the pairwise comparison appears to be crucial (37). To examine potential relationships among the genes chosen for study, network analysis was applied using Ingenuity Pathways Analysis (www.ingenuity.com, Redwood City, CA). This is a web-based application that enables the discovery, visualization, and exploration of interaction networks. The software relies on currently known relationships (i.e., published manuscripts) among human, mouse, and rat genes. Figure 6 depicts the results of a comprehensive analysis to find potential networks among the nine genes under investigation. The relationships shown do not pertain to temporal expression patterns but, rather, depict known networks encompassing gene expression regulation (activation/inhibition). We found evidence for co-regulation among most of the genes under investigation.

The transcriptional regulator MYC is crucial in breast cancer (19) and plays some functions in the mammary gland (14). MYC directly regulates the expression of GAPDH, ACTB, RPS9, RPS15, and RPS23. ACTB expression is repressed by the action of MYC (32), while other genes are upregulated. A dilution effect was observed with MYC, as shown by correlations (Table 1). However, normalized data of MYC were characterized by a large decrease in gene expression between the nonlactating period (~15 days) and the onset of lactation. This pattern suggests marked downregulation of this protein in mammary gland due to a change in physiological state (i.e., nonlactating to lactating), which also was observed previously (34). Decreased MYC mRNA could potentially affect the expression of RPS9, RPS15, RPS23, and GAPD as shown by others (12, 13). However, we were unable to find a correlation between the normalized expression of MYC and those genes (Suppl. Table S3).

ESR1 is another nuclear receptor with an interesting role in regulating the genes under investigation. ESR1 expression (Fig. 3) decreased numerically from late prepartum through the onset of lactation then increased (P < 0.001) exponentially through 240 days postpartum. Estrogen, through the binding with ESR1, functions as an inhibitor of MYC (18) and protects TP53 (tumor protein 53) from being deregulated (20). Estrogen also was shown to directly upregulate the expression of EGFR (epidermal growth factor receptor) in HeLa cells (31). From the gene network analysis, it appears that the type I receptor tyrosine kinases (EGFR, ERBB2, ERBB3) play important roles as well (Fig. 6). They repress ACTB and stimulate the expression of ITGB4BP in murine and human tumor cells (1). ITGB4BP stimulates murine mast cells to synthesize interleukin-2 (26), while interleukin-2 activates GAPDH transcription in murine T-lymphocytes (30). ITGB4BP regulation of GAPDH occurs in immune cells, and we found a significant correlation between the normalized data of the two genes (Suppl. Table S3). TP53 directly upregulates GAPDH in rat cerebellar granule cells (6) and downregulates ACTB in primary rat embryo fibroblasts (10).

The networks described in this study suggest a more complicated level of co-regulation of the genes under investigation because MYC is not the only gene involved. We uncovered a potential effect of estrogen and type I receptor tyrosine kinases on ACTB expression. The estrogen network can explain to a certain extent the downregulation of MYC. Interestingly, RPS9 and RPS15, two of the three genes with the most stable expression based on pairwise comparison, also can be sub-

![Fig. 4. Average stability of expression ratio values (M = gene-stability measure) of remaining genes tested during pairwise comparison. Values are reported as stepwise exclusion of the least stable control gene (figure was generated by geNorm).](http://physiolgenomics.physiology.org/)

![Fig. 5. Determination of the optimal number of control genes for normalization (figure generated by geNorm). Y-axis, pairwise variation V (V_{n+1}) between the normalization factors NF_n and NF_{n+1}; X-axis, comparison between the use of n or n + 1 genes to calculate the normalization factor, i.e., V2/3 is the comparison between the use of 2 or 3 HKG to calculate the normalization factor (see MATERIALS AND METHODS and Ref. 36 for detailed explanation).](http://physiolgenomics.physiology.org/)
jected to co-regulation. We could not determine known relationships encompassing UXT, MPRL39, and GTP (MTG1).

The presence of co-regulation between RPS9 and RPS15 might be seen as a problem in applying the pair-wise comparison method because co-regulation could add bias to the calculations, i.e., stability is evaluated on a gene by gene basis using pair-wise comparisons. The absence of co-regulation between MYC and UXT and the high correlation between UXT and the two ribosomal proteins (RPS9 and RPS15, $r = 0.89$ and 0.90, $P < 0.001$), along with the larger negative correlation between RNA concentration and UXT, RPS9, and RPS15 (Table 1), suggest a lack of differential regulation of these genes during the lactation cycle. We conclude that even if a potential co-regulation exists between some of the best HKG it was not apparent in our study. This provided reasonable confidence about the chosen genes for calculation of the normalization factor. We evaluated potential co-regulation after the expression analysis was performed, which is not a
Toolbox

318 MAMMARY REFERENCE GENES

drawback because all genes tested as potential HKG were previously assumed to be reasonably stable. Furthermore, time had a statistically significant effect on their expression (Fig. 1). Thus, if we used a statistical approach to select HKG, none of the genes studied would have been suitable. The potential for an effect of dilution on the expression of stable genes during the lactation cycle prompted the re-evaluation of the genes tested after results were obtained. Thus, our findings demonstrated that the method developed by Vandesompele et al. (37) was more suitable than a simple statistical analysis of time effects.

Summary and Conclusions

Results showed that changes in RNA concentration in the bovine mammary gland lead to a dilution effect of genes that have stable mRNA expression across the lactation cycle. This hypothesis is supported by correlation analysis between the genes under investigation and total mammary RNA. As a result, we conclude that the evaluation of a suitable internal control gene requires more than a simple statistical analysis. The pair-wise comparison method was suitable for selection of HKG. Several genes were tested as potential internal controls and UXT, RPS9, and RPS15 were found to be the best for normalization of other gene expression data. The results also uncovered co-regulation among several genes (GAPD, ACTB, RPS9, RPS15, RPS23, and ITGβ4B5). More importantly, no evidence was found of co-regulation among UXT and the ribosomal proteins RPS9 and RPS15. A potentially critical role was uncovered for MYC and ESR1 in regulating some genes that have been previously used as internal controls.

In conclusion, this study showed that several genes must be evaluated as potential internal controls for qPCR studies of the bovine mammary gland particularly when the effect of physiological state on tissue genomics is of interest. The geometrical average of the data from RPS9, RPS15, and UXT can be used for normalization in future studies of this type or to verify important relationships that arise from bovine mammary microarray studies. Lastly, caution should be taken when interpreting qPCR data from studies lacking proper evaluation of HKG.

ACKNOWLEDGMENTS

We gratefully acknowledge Dr. Robert Collier (University of Arizona) for providing the primer sequence for LALBA and Nicole Janovick Guretzky (University of Illinois) for providing the sequence for RPS9. The help from the staff of the University of Illinois Dairy Research and Teaching Unit for animal care also was instrumental.

REFERENCES

4. Capuco AV, Wood DL, Baldwin R, Meleod K, Pape MJ. Mammary cell number, proliferation, and apoptosis during a bovine lactation: