

Assessment of normalization strategies for quantitative RT-PCR using microdissected tissue samples

Heidi S Erickson^{1,*}, Paul S Albert², John W Gillespie^{1,3}, Benjamin S Wallis¹, Jaime Rodriguez-Canales¹, W Marston Linehan⁴, Sergio Gonzalez⁵, Alfredo Velasco⁶, Rodrigo F Chuaqui¹ and Michael R Emmert-Buck¹

Gene expression measurement techniques such as quantitative reverse transcriptase (qRT)-PCR require a normalization strategy to allow meaningful comparisons across biological samples. Typically, this is accomplished through the use of an endogenous housekeeping gene that is presumed to show stable expression levels in the samples under study. There is concern regarding how precisely specific genes can be measured in limited amounts of mRNA such as those from microdissected (MD) tissues. To address this issue, we evaluated three different approaches for qRT-PCR normalization of dissected samples; cell count during microdissection, total RNA measurement, and endogenous control genes. The data indicate that both cell count and total RNA are useful in calibrating input amounts at the outset of a study, but do not provide enough precision to serve as normalization standards. However, endogenous control genes can accurately determine the relative abundance of a target gene relative to the entire cellular transcriptome. Taken together, these results suggest that precise gene expression measurements can be made from MD samples if the appropriate normalization strategy is employed.

Laboratory Investigation (2007) **87**, 951–962; doi:10.1038/labinvest.3700659; published online 23 July 2007

KEYWORDS: endogenous control genes; gene expression analysis; microdissected specimens; normalization; qRT-PCR; validation

To compare gene expression levels between two biological samples, normal cells vs tumor cells for example, it is critical to have a normalization control. Historically, this has been accomplished using genes such as glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and ACTB, which are presumed to have stable expression levels in all cell types. Additional approaches for normalization include using cell number, cell content (DNA, RNA, or protein amounts), or more recently a multigene readout on expression microarrays. In each case, the normalization procedure is designed to allow investigators to obtain an accurate measurement of the relative levels of a gene of interest across different samples.

Review of the literature shows that a common set of ‘housekeeping genes’ are used as endogenous controls for quantitative reverse transcriptase (qRT)-PCR. Overall, the

majority of these mRNAs are effective normalization tools, although several groups report that variability in their expression levels can exist in different sample and cell types and needs to be considered when designing a study.^{1–8} For example, Vandesompele *et al*⁹ reported that universally stable control genes do not exist, thus necessitating the development of an accurate normalization strategy for each experimental system. Aerts *et al*,¹⁰ in a study of 26 tumor cell lines including prostate cancer DU-145 and PC-3, showed GAPDH has the highest variability in gene expression of commonly used controls, whereas 18s rRNA, GUS, and ACTB had the best expression stability.¹⁰ ACTB also showed stable expression in skeletal muscle and diabetic glomeruli; however, this was not observed in colonic and adipose tissue, reinforcing the notion that many commonly thought of ‘housekeeping’ genes are not stably expressed in certain cell

¹Pathogenetics Unit, Laboratory of Pathology and Urologic Oncology Branch, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA; ²Biometric Research Branch, Division of Cancer Treatment and Diagnosis, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA; ³SAIC-Frederick Inc., National Cancer Institute at Frederick, Frederick, MD, USA; ⁴Urologic Oncology Branch, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA; ⁵Department of Pathology, Catholic University, Santiago, Chile and ⁶Department of Urology, Catholic University, Santiago, Chile

Correspondence: Dr MR Emmert-Buck, MD, PhD, Pathogenetics Unit, Laboratory of Pathology and Urologic Oncology Branch, National Cancer Institute, National Institutes of Health, 8717 Grovemont Circle, Bethesda, MD 20892, USA. E-mail: mbuck@helix.nih.gov

*Reprint requests: HS Erickson, PhD, Pathogenetics Unit, Laboratory of Pathology and Urologic Oncology Branch, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA. E-mail: HSE@nih.gov

Research was performed at the Pathogenetics Unit, Laboratory of Pathology and Urologic Oncology Branch, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA.

Received 24 March 2007; revised 4 June 2007; accepted 5 June 2007

types and environments, thus reinforcing the need to validate endogenous controls in each tissue and experimental condition.^{2,11–13}

Beyond differences in housekeeping gene expression related to tissue type, investigators must also consider variability in gene expression related to disease state and procurement method. Tumor cells show alterations in many transcripts, particularly those related to increased growth and metabolism, a fact that must be considered when housekeeping genes are selected for studies of cancer or other proliferative processes.^{2,14–18} Additionally, many previous studies examining housekeeping gene stability utilized cells grown in culture as opposed to *ex vivo* tissue specimens. Our group and others have found that cells in culture show some expression features, such as consistently high levels of ACTB, which appear to be related to *in vitro* growth conditions and are not necessarily observed in tissue samples. Thus, ACTB may be an excellent housekeeping gene for cells grown *in vitro*, but less accurate when studying tissues. Microdissected (MD) cells add another level of complexity to quantitative gene expression analysis as the mRNA is derived from a single cell type that may have a unique expression profile, including the housekeeping genes, relative to the bulk tissue sample from which they were derived.

The objective of the present study was to evaluate methods for normalizing gene expression measurements in MD tissue samples. Cell count, total RNA, individual mRNAs, and small sets of mRNAs were each assessed for their ability to serve as internal controls.

MATERIALS AND METHODS

Tissue Specimens

Five frozen specimens of newly diagnosed and untreated human prostate carcinoma (PCa) were obtained from patients who underwent prostatectomy at Catholic University in Santiago, Chile (cases 1–5) and were anonymized before transfer to the National Cancer Institute (NCI). Tumors were evaluated by two pathologists (JWG and RFC) and assigned Gleason scores of 6 (3,3; $n = 1$), 7 (4,3; $n = 2$), and 9 (4,5; $n = 2$). Tissue specimens used in this study were selected using the NCI Pathogenetics Unit Prostate Tissue Database (data not shown). Tissue quality has previously been shown to be of excellent quality for isolation of RNA.¹⁹

Microdissection and RNA Isolation

Five cases of matched frozen nontumor prostate (benign hyperplasia) and moderate-grade prostate cancer (Gleason pattern 3,3), or high-grade prostate cancer (Gleason 4,5) or the Gleason 4 portion of a Gleason pattern 4,3 were MD in triplicate according to standard methods.^{20,21} Staining of the frozen tissue samples before laser capture microdissection (LCM) was performed by dipping the tissue sections into solutions as follows: 70% ethanol (EtOH) for 20 s, H₂O for 10 s, hematoxylin (Sigma-Aldrich, St Louis, MO, USA) for 10 s, H₂O for 10 s, 70% EtOH for 10 s, eosin (Sigma-Aldrich)

for 4 s, 95% EtOH for 10 s two times, 100% EtOH for 10 s two times, and xylenes for 20 s three times.

Okuducu *et al*²² reported that methyl green causes the least fluorescent interference for qRT-PCR. We have traditionally used hematoxylin and eosin (H&E) as our stain of choice because it allows for the optimum visualization of cells of interest. Therefore, before LCM, staining of additional triplicate sections of one case was performed according to the above H&E-staining method, with the deletion of eosin and either the use of hematoxylin alone or the substitution of methyl green (MTR Scientific, Ijamsville, MD, USA) for hematoxylin. This analysis was conducted to assess if there is a statistically detectable fluorescent interference of the different tissue stains.

Approximately 10 000 dissected cells were procured for each case, which equated to ~3000 shots per tissue type (Figure 1). Samples were immediately placed into lysis buffer and stored at -80°C until RNA extraction. Total RNA isolation of each individual sample was conducted with

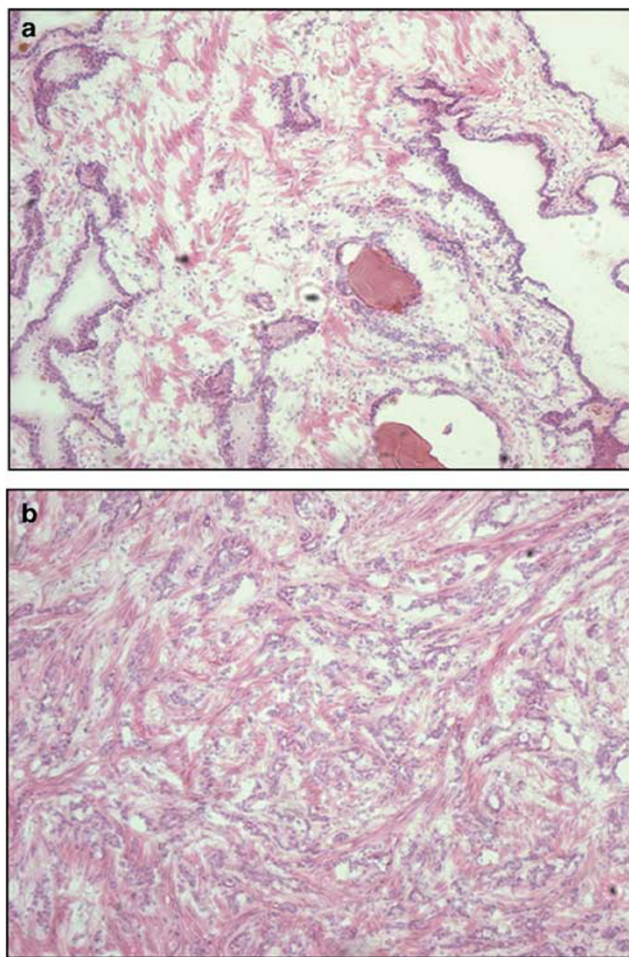


Figure 1 Representative histology from matched prostate normal (a) and PCa (b) sample areas captured by laser microdissection. From these areas, 10 000 cells were MD and used as the normalized samples as described in Materials and Methods. H&E stained and original magnification, $\times 100$.

PicoPure RNA extraction kit (Arcturus Engineering Inc., Mountain View, CA, USA) according to the manufacturer's protocol. Because DNA is often left in the sample using glass filter RNA extraction, the samples were subjected to DNase treatment for 15 min.¹⁹ Following RNA isolation, 1 μ l samples were aliquoted for RNA quantitation and qualification analysis and were used immediately. All remaining RNA samples were stored at -80°C .

RNA Quantitation and Qualitation

Quantitation of individual sample total RNA, in triplicate, was conducted using Bioanalyzer (Agilent Technologies Inc., Santa Clara, CA, USA) and NanoDrop (NanoDrop Technologies, Wilmington, DE, USA) equipment according to the manufacturer's protocol for < 10 ng total RNA quantities per microliter of the sample. Quality of individual samples was assessed with the Bioanalyzer (Agilent Technologies Inc.) and a RNA integrity number was provided by 2100 Expert software (Agilent Technologies Inc.). Qualitation of case no. 5 total RNA was also visualized via electrophoresis on a 1.0% agarose gel.

Reverse Transcription

Reverse transcription (RT) was conducted on all individual RNA samples using TaqMan RT Reagents (Applied Biosystems Inc. (ABI), Foster City, CA, USA; catalogue no. N808-0234), with random hexamers as the RT primers, according to the manufacturer's protocol. Template was used at 4 μ l volumes per reaction. All RT reactions were performed on the MJ Research PTC-200 thermocycler. The cDNA was used

immediately for all PCR. Remaining cDNA was stored at -80°C .

Housekeeping Genes

Ten housekeeping genes that are constitutively expressed in various human tissues and used in numerous previous studies were chosen as endogenous controls for gene expression analysis (Table 1). Significant reduction of the chance that genes may be coregulated was achieved by selecting genes that belong to different functional classes.⁹ Except for PGK1 and GAPDH, which are both involved in glycolysis, it is assumed that the regulation of expression of these housekeeping genes is not directly related, and that they have independent functions in cellular maintenance.²³

Quantitative Real-Time PCR (qPCR) Gene Expression Measurement

Ten commercially available optimized endogenous control primer/probe sets (Table 1) and $2 \times$ TaqMan Universal Master Mix (ABI catalogue no. 4304437) were used in singleplex qPCR according to the manufacturer's protocol. The assays for ACTB, CYPA, HPRT, GAPDH, TFRC, B2M, GUS, PGK1, and RLP are cDNA specific; whereas the primers and probes for 18s rRNA may detect genomic DNA. Negligible contamination (0.1%) of genomic DNA contamination was confirmed by the observation of a cycle threshold (C_T) of 10 comparing RT-negative with RT-positive samples.^{24,25} Stratagene M \times 3000P™ real-time PCR machine was used for all qPCR and detections. cDNA template was used at 4 μ l volumes per reaction. All qPCR assays were performed in

Table 1 Genes evaluated as endogenous controls for MD prostate N- and T-gene expression analysis

Symbol ^a (Entrez gene official symbol)	Entrez gene GeneID	Gene name	Cellular function	Map
18s (LOC100008588)	100008588	18s ribosomal RNA	Ribosome subunit	^b
ACTB (ACTB)	60	β -actin	Cytoskeletal structural protein	7p15–p12
CYPA (PPIA)	5478	Cyclophilin A	Serine–threonine phosphatase inhibitor	7p13–p11.2
HPRT (HPRT1)	3251	Hypoxanthine ribosyltransferase	Purine synthesis in salvage pathway	Xq26
GAPDH (GAPDH)	2597	GAPDH	Oxidoreductase in glycolysis and gluconeogenesis	12p13
TFRC (TFRC)	7037	Transferrin receptor	Cellular iron uptake	3q29
B2M (B2M)	567	β -2-microglobulin	β -Chain of major histocompatibility complex class I molecules	15q21–q22
GUS (Gusb)	110006	β -Glucuronidase	Lysosome exoglycosidase	7q21.11
PGK1 (PGK1)	5230	Phosphoglycerokinase 1	Glycolysis enzyme	Xq13
RLP (RPLP0)	6175	Large ribosomal protein	Transcription	12q24.2

^aGene symbol used in this paper.

^b18s rRNA is a ribosomal subunit and not a gene that is mapped to a specific chromosome; therefore, no map location is given.

triplicate after RT. Cycling conditions consisted of one cycle of 50°C for 2 min followed by 95°C for 10 min, and then 50 cycles of 95°C for 15 s followed by 60°C for 1 min. Controls consisting of total human prostate RNA (~12 ng/μl; Ambion, Austin, TX, USA) were positive in all runs, and controls consisting of sterile molecular grade water were negative in all runs. Relative quantitation analysis of gene expression data was conducted according to the $2^{-\Delta\Delta CT}$ method.^{26,27}

Statistical Analysis

Analyses were carried out on the original C_T scale. The ‘gold standard’ normalization factor was the average of all 10 endogenous control genes (Table 1). The variation for particular genes and the average values across all 10 genes was examined using a variance components model.²⁸ This approach allowed decomposition of the sources of variation into components attributable to individual sample, microdissection replicate, and qPCR technical replicate. Specifically, we fit the model:

$$(\text{Gene specific or Average } C_T \text{ value})_{ijk} = \beta + \tau_i + \delta_{ij} + \varepsilon_{ijk},$$

where i denotes the i th sample, j denotes the j th MD sample, and k denotes the qPCR technical replicate. In the analysis, there are five individual samples, with three microdissections for each, and then three qPCR technical replicates for each microdissection. The parameter β reflects the average value, while the random effects τ_i , δ_{ij} , and ε_{ijk} measure variation due to individual sample, microdissections within individual sample, and qPCR technical replication, respectively. Variances of τ_i , δ_{ij} , and ε_{ijk} , denoted as σ_τ^2 , σ_δ^2 , and σ_ε^2 , characterizing the different sources of variation for these three components, are estimated from the model. Ninety-five percent tolerance intervals demonstrating the natural variation in single measurements were constructed based on estimates obtained from the variance components model. These tolerance intervals can be interpreted as bounds on likely ranges for the average C_T values. As an illustration, for a single MD sample with a single qPCR technical replicate across tumor samples, the 95% tolerance interval was computed as $\pm 1.96\sqrt{\sigma_\tau^2 + \sigma_\delta^2 + \sigma_\varepsilon^2}$. The 95% tolerance interval for a single MD sample with a single qPCR technical replicate within a tumor sample was computed as $\pm 1.96\sqrt{\sigma_\delta^2 + \sigma_\varepsilon^2}$.

The difference between using a single endogenous control gene or a pair of endogenous control genes as a normalization factor, relative to using all 10 endogenous control genes (‘gold standard’) was also assessed. For evaluating each of the 10 individual endogenous control genes or averages of any two control genes (ie, 45 possible combinations of two endogenous control genes), we fit the variance components models:

$$(\text{Gene } C_T \text{ value} - \text{Average } C_T \text{ value})_{ijk} = \beta + \tau_i + \delta_{ij} + \varepsilon_{ijk},$$

where the gene C_T value was either the C_T value for each of the 10 genes individually, or an average of the C_T values for each of the 45 possible pairs of the 10 housekeeping genes.

The average C_T value is the average value over all 10 genes. The parameter β measures the average difference, while the random effects τ_i , δ_{ij} , and ε_{ijk} measure variation in the difference due to individual sample, microdissections within individual sample, and qPCR technical replication, respectively. Variances of τ_i , δ_{ij} , and ε_{ijk} , denoted as σ_τ^2 , σ_δ^2 , and σ_ε^2 , characterizing the different sources of variation in the difference for these three components, are estimated from the model. Estimates of the model parameters were used to compute 95% tolerance intervals for normalization factors based on only one or two genes vs the average of all 10 genes. These intervals were constructed for a single measurement (ie, one particular qPCR technical replicate) as well as for the more practical situation of the average of three qPCR technical replicates. These tolerance intervals can be interpreted as bounds on likely values of the difference in C_T value of using only a single or two genes as compared with the average C_T value across all 10 control genes. Genes were ranked according to the length of the interval. These intervals were constructed separately for normal and tumor tissue. Further, we also considered the common situation in which both normal and tumor tissue are collected from the same individual (ie, paired normal and tumor tissue).

We examined whether there was a significant effect of staining on C_T values using a linear mixed model.²⁹ A comparison across stains was made with an F -test.

Untransformed and transformed data means and s.d. are presented as means (s.d.). All P values correspond to two-sided tests with $P \leq 0.05$ considered as statistically significant.

RESULTS

Three strategies for gene expression normalization were evaluated using cells MD from radical prostatectomy specimens; cell count during LCM, total RNA measurement, and the use of endogenous housekeeping genes. However, before this analysis the basic parameters of the gene expression measurement system were assessed. First, it was determined if stains used for visualizing tissues interfere with subsequent qRT-PCR. Three stains were tested: (1) H&E, (2) hematoxylin alone, and (3) methyl green. An F -test revealed no statistically significant C_T value difference for the three stains showing P values of 0.21, 0.08, and 0.18 for ACTB, HPRT, and GUS, respectively (Table 2). Therefore, H&E was used as

Table 2 Mean difference, relative to H&E, for hematoxylin and methyl green staining for ACTB, HPRT, and GUS

Stain	Endogenous control gene		
	ACTB	HPRT	GUS
Hematoxylin	-0.024 (0.035)	-0.050 (0.042)	-0.049 (0.040)
Methyl green	0.047 (0.035)	0.068 (0.042)	0.038 (0.040)

Means and s.e. are on log base 2 scale.

the tissue stain for the remainder of the study. Second, the basic characteristics of qRT-PCR applied to MD samples were examined. Complementary DNA (cDNA) was produced from each of the 10 samples (one normal epithelium and one tumor dissection from five cases) and used to measure the levels of 10 housekeeping genes. A dilution series analysis of positive control RNA (human prostate total RNA) demonstrated that a minimum of 0.02 ng/ μ l total RNA was necessary for the assay. For each gene, C_T values for the positive control RNA and all dissected samples were consistent with a within run s.d. of $<0.3 C_T$, establishing that C_T values of all 10 endogenous control genes can be reproducibly determined in dissected cells, and that PCR inhibitors were not present in them. Varying levels of expression were demonstrated for each of the 10 evaluated endogenous control housekeeping genes in the normal and tumor samples (Table 3).

Table 3 Average gene expression level (s.d.) of all 10 endogenous controls measured by C_T values

Endogenous control gene	Tissue type N or T	Average gene expression level per tissue type C_T value (s.d.) ^a	Average gene expression level per gene ^b C_T value (s.d.)
18s	N	18.95 (3.21)	19.00 (2.72)
	T	19.05 (2.28)	
CYPA	N	29.24 (2.94)	28.95 (2.44)
	T	28.65 (2.03)	
ACTB	N	33.28 (4.10)	33.09 (3.32)
	T	32.90 (2.59)	
HPRT	N	35.68 (2.92)	35.61 (2.10)
	T	35.55 (1.65)	
GAPDH	N	33.69 (3.84)	33.43 (3.03)
	T	33.16 (2.34)	
TFRC	N	35.96 (3.52)	36.16 (2.60)
	T	36.35 (1.92)	
B2M	N	27.85 (2.65)	28.08 (2.16)
	T	28.33 (1.72)	
GUS	N	31.85 (3.11)	31.63 (2.42)
	T	31.41 (1.96)	
PGK1	N	31.74 (2.71)	31.88 (2.05)
	T	32.02 (1.53)	
RLP	N	28.90 (2.81)	28.49 (2.08)
	T	28.08 (1.42)	

N = normal prostate tissue; T = tumor prostate tissue.

^as.d. for a single measurement with a single MD and qPCR replication. The s.d. was estimated using a variance components model. Specifically, the s.d. is the square root of the sum of the sources of variation due to individual sample, MD sample within individual sample, and qPCR replication.

^bAverage of normal and tumor C_T value at each qPCR replicate.

LCM cell count was first evaluated as a normalization method by determining if consistent, reproducible amounts of RNA could be reliably procured from tissue samples. The following experimental strategy was employed. Three 5 μ M thick serial histological sections were prepared from each of the 10 tissue blocks, 10 000 cells were dissected from identical areas of each of the three slides, 10 housekeeping genes were measured in all of the samples, and the intra-case variability in recovered RNA amounts was determined. In other words, we evaluated if the three serial histological sections within each case could produce identical amounts of RNA based on matched LCM cell count. Since the three recut sections contained cells that were immediately adjacent to each other within the tissue block, we made the assumption that the housekeeping gene levels within each case were constant. Thus, the only variables were the dissection procurement process and the RNA purification procedure. The data shown in Figure 2, and Tables 4 and 5, indicate that cell count during dissection was not a reliable means of normalizing expression data. Sizable variation was observed within the three intra-case dissections, from ± 1.12 to 1.74 C_T values in normal tissues and ± 1.07 to 1.84 C_T values in tumor tissues. These data imply that a comparison of 10 000 dissected cells from sample A against 10 000 dissected cells from sample B can produce an artifactual difference of up to 1.84 C_T values using cell count as the normalization strategy. Since changes in gene expression of 1.5–2 C_T values (~ 3 - to 4-fold) are considered biologically important,³⁰ the use of dissection cell count as a means to compare samples is limited only to those studies that do not require precise measurements. However, we did find that cell count was useful in producing 'ballpark' RNA input levels within the range that can be successfully analyzed by qRT-PCR, and thus serves as a useful first step when analyzing dissected samples.

The use of total RNA measurement as an internal comparator for qRT-PCR was next evaluated. LCM was performed on three serial sections from each of the 10 tissue blocks as described above, total RNA was recovered and measured using the NanoDrop method, all 10 housekeeping genes were analyzed by qRT-PCR in the 30 samples, and the intra-case reproducibility of total RNA measurement was calculated. Since, we had established that LCM cell count was not a reliable means to normalize samples, the total RNA data were compared to housekeeping gene levels to determine if the intra-case RNA values were reproducible. Again, we made the assumption that housekeeping gene levels were constant in the three serial recut slides prepared from each tissue block. The MD samples yielded approximately 16 μ l of 8–10 ng/ μ l average of total RNA, which is similar to our previous MD sample collection for microarray studies (data not shown). The total RNA quantity per sample was at the low end of the detection limits for quantitation by NanoDrop (cases 1–4: normal samples mean = 8.4 ng/ μ l; median = 5.4 ng/ μ l; range = 2.0–26.8 ng/ μ l. Tumor samples: mean = 10.8 ng/ μ l; median = 6.4 ng/ μ l; range = 1.6–41.5 ng/ μ l).

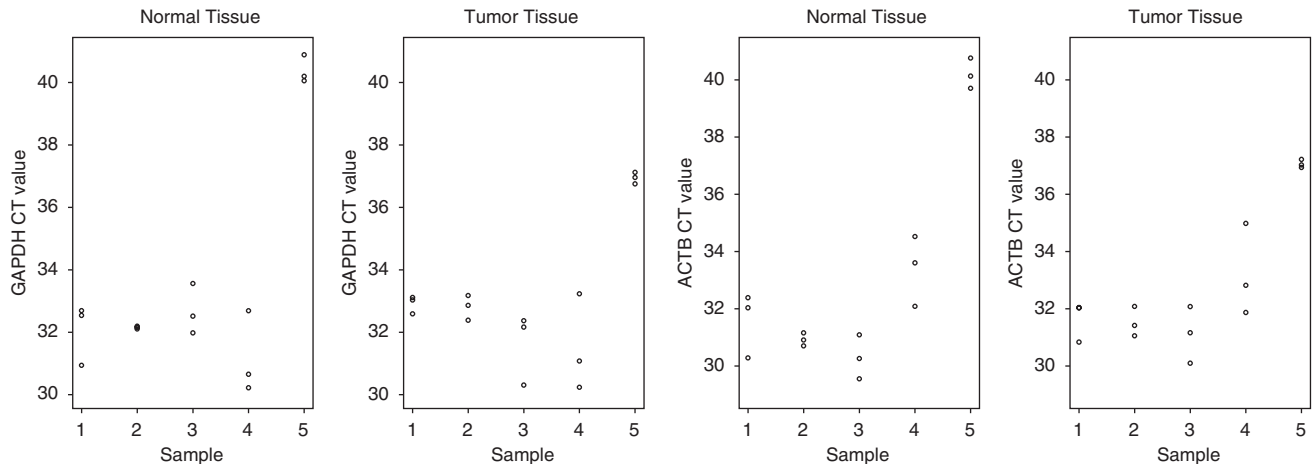


Figure 2 Variation in C_T values for normal and tumor tissue across LCM replicates. C_T value variation for ACTB and GAPDH is representative of all 10 housekeeping genes. Each observation is the average of three qPCR technical replicates.

Table 4 Ninety-five percent tolerance intervals showing LCM variation within a tumor or normal sample

Gene symbol	Normal	Tumor
18S	$\pm 1.30^a$	± 1.67
CYPA	± 1.31	± 1.71
ACTB	± 1.69	± 1.80
HPRT	± 1.74	± 1.83
GAPDH	± 1.64	± 1.74
TFRC	± 1.40	± 1.84
B2M	± 1.27	± 1.49
GUS	± 1.12	± 1.51
PGK1	± 1.18	± 1.27
RLP	± 1.34	± 1.07

The intervals assume a single MD replication with three qPCR replicates.

^aValues in units of C_T .

The data shown in Figure 3 indicate that total RNA quantity measurements are not reproducible across LCM samples and cannot serve as an effective normalization strategy for MD samples. The large variance generated by the RNA analysis technique alone precludes accurate measurement of small and moderate gene expression differences among samples. Similar to using cell count, however, we did find that total RNA is a useful parameter for calibrating ballpark RNA input levels within the range that can be successfully analyzed by qRT-PCR.

We then evaluated whether housekeeping genes could be used as an internal control for dissected samples. In this approach, the amount of target transcript per cell or per RNA content is not determined as with LCM cell count or RNA measurement; rather, the amount of target transcript relative to

the overall transcriptome of a sample is calculated, much like an expression array experiment where input amounts of RNA are normalized between samples based on multiple housekeeping genes or overall mRNA readout. However, unlike expression arrays, it is impractical to use a large set of housekeeping genes for normalization in qRT-PCR samples, due to the time and resources involved; and more importantly, the limited amount of starting material. Thus, it was determined if one or two housekeeping genes could serve as a surrogate measurement for the transcriptome of dissected cells.

To perform this assessment, we started with the assumption that the average of the 10 housekeeping genes selected for the study was indeed reflective of the overall transcriptome of the dissected samples. The gene set was intentionally chosen to include diverse biological functions to increase this likelihood, even in tumors where many metabolic activities are increased. The strategy employed was to measure the 10 housekeeping genes in three separate dissected samples from each case, and statistically evaluate how accurately each individual gene, or small sets of genes, tracked with the average of the 10.

However, to complete this analysis the types of tissue samples that are typically used in qRT-PCR studies first needed to be considered. The two general scenarios are paired samples (diseased and normal samples from the same patient) or unpaired samples (diseased and normal samples from different patients). When the 10 housekeeping genes were initially analyzed, a strong positive correlation was observed between the paired normal and tumor samples from the five cases, indicating that the two study designs need to be separately considered (Figure 4).

For unpaired data, 95% tolerance intervals corresponding to a ΔC_T for a single tumor sample (one tumor sample, one LCM sample, and the average of three qPCR replicates) were constructed using the results from the variance components

Table 5 Average qRT-PCR replicates housekeeping gene expression C_T values for MD replicates per case and tissue type

Case ID no.	LCM replicate	18s N	18s T	CYPA N	CYPA T	ACTB	ACTB	HPRT	HPRT T	GAPDH N	GAPDH T
<i>(a) 18S, CYPA, ACTB, HPRT, and GAPDH</i>											
1	1	18.15	18.29	28.49	28.48	32.39	32.01	36.14	35.21	32.54	33.04
1	2	16.67	17.13	27.01	26.66	30.28	30.80	34.00	35.70	30.94	32.59
1	3	17.50	18.02	27.97	27.50	32.04	32.01	35.16	36.62	32.69	33.11
1	Average	17.44	17.81	27.82	27.54	31.57	31.61	35.10	35.84	32.06	32.91
2	1	17.24	18.10	27.30	27.45	30.91	31.37	34.41	35.89	32.19	32.86
2	2	17.15	17.97	27.13	27.67	30.71	31.03	33.72	35.94	32.16	32.39
2	3	17.38	18.88	26.88	28.34	31.16	32.05	34.47	36.08	32.11	33.18
2	Average	17.26	18.32	27.10	27.82	30.93	31.48	34.20	35.97	32.15	32.81
3	1	15.46	17.15	27.06	27.33	29.56	31.16	34.63	34.93	31.98	32.37
3	2	15.49	15.50	27.41	26.07	30.26	30.15	34.99	33.77	32.52	30.31
3	3	16.78	17.62	28.08	28.38	31.09	32.06	35.91	36.07	33.56	32.17
3	Average	15.91	16.76	27.52	27.26	30.30	31.12	35.18	34.93	32.69	31.62
4	1	20.37	18.83	30.04	27.73	33.61	31.87	33.77	32.28	30.65	30.24
4	2	19.38	20.09	28.42	28.55	32.09	32.79	31.95	33.39	30.22	31.08
4	3	21.23	21.30	30.56	30.05	34.52	34.95	34.65	35.39	32.69	33.23
4	Average	20.33	20.07	29.67	28.78	33.41	33.20	33.45	33.69	31.19	31.52
5	1	23.30	21.91	33.73	31.73	39.70	37.07	39.92	37.32	40.05	36.75
5	2	24.04	22.83	34.20	32.12	40.76	37.24	40.43	37.25	40.20	36.96
5	3	24.18	22.06	34.38	31.77	40.28	36.97	41.00	37.44	41.03	37.11
5	Average	23.84	22.27	34.10	31.87	40.25	37.09	40.45	37.34	40.43	36.94

Table 5 Continued

Case ID no.	LCM replicate	TFRC N	TFRC T	B2M N	B2M T	GUS N	GUS T	PGK1 N	PGK1 T	RLP N	RLP T
<i>(b) TFRC, B2M, GUS, PGK1, and RLP</i>											
1	1	35.39	36.21	29.13	29.32	32.17	32.03	32.02	32.32	29.26	28.28
1	2	33.38	36.61	27.66	28.16	30.72	30.84	30.40	31.60	27.72	28.14
1	3	34.86	37.34	28.50	29.25	31.53	32.06	31.58	32.39	29.89	28.52
1	Average	34.54	36.72	28.43	28.91	31.47	31.64	31.34	32.10	28.96	28.31
2	1	34.06	36.21	24.65	26.43	28.79	30.30	29.28	30.09	26.45	26.53
2	2	34.09	35.49	24.70	26.40	28.68	30.42	29.01	30.08	26.11	26.82
2	3	34.25	36.45	24.99	26.89	28.97	31.13	29.23	30.80	26.14	27.20
2	Average	34.13	36.05	24.78	26.57	28.81	30.62	29.17	30.33	26.24	26.85
3	1	34.40	35.54	27.05	28.08	30.30	31.74	30.89	33.26	28.31	28.48
3	2	34.82	34.04	26.64	27.19	30.09	29.47	31.41	31.44	28.05	27.25
3	3	35.39	36.64	28.01	28.77	31.30	31.19	31.63	32.69	28.70	28.73
3	Average	34.87	35.41	27.23	28.01	30.56	30.80	31.31	32.46	28.35	28.15
4	1	34.36	33.06	26.53	26.10	31.61	28.52	31.06	30.18	27.42	26.04
4	2	33.08	34.53	26.34	27.70	30.56	29.75	29.74	31.23	26.77	27.09
4	3	35.12	35.92	27.96	28.43	32.09	30.35	31.18	31.97	28.31	27.67
4	Average	34.19	34.50	26.94	27.41	31.42	29.54	30.66	31.13	27.50	26.94
5	1	41.70	38.87	31.28	30.61	37.10	34.46	35.61	34.29	32.85	30.30
5	2	42.09	39.33	32.14	31.40	37.15	34.54	36.73	34.13	33.47	30.12
5	3	42.45	39.06	32.15	30.34	36.68	34.29	36.23	33.81	34.06	30.01
5	Average	42.08	39.09	31.86	30.79	36.98	34.43	36.19	34.08	33.46	30.14

N = normal; T = tumor.

analyses on tumor samples, and the genes were ranked in terms of the size of the interval (Table 6). The data indicate that as few as two housekeeping genes can effectively serve as a surrogate of the overall transcriptome. For example, using the average of *CYPA* and *TFRC*, the interval is within ± 0.54 C_T s of the average C_T value across all 10 housekeeping genes 95% of the time; whereas, using *ACTB* alone, the interval is within only ± 2.33 C_T s of this average 95% of the time.

For paired data, we formed 95% tolerance interval around the $\Delta\Delta C_T$ s (one LCM replicate and three qPCR replicates; Table 7). In contrast to unpaired data, only one gene is required for normalization, with *PGK1* and *RLP* tracking the average of the 10 housekeeping genes best (± 0.83 C_T s and ± 0.93 C_T s of the average C_T value across all 10 housekeeping genes 95% of the time, respectively).

For both unpaired and paired data, the effects of using three LCM replicates vs one dissection, and the effects of

using three qPCR replicates as opposed to just one qPCR for each sample on the variation in the average C_T value across the 10 housekeeping genes were evaluated. The three

Table 6 Ninety-five percent tolerance intervals (TI) on the difference between a single gene or the average of two genes and the average of all 10 housekeeping genes for unpaired samples

Gene(s)	TI	Single or pair ranking
18s	$\pm 2.585^a$	10
<i>CYPA</i>	± 1.552	6
<i>ACTB</i>	± 2.330	9
<i>HPRT</i>	± 2.091	8
<i>GAPDH</i>	± 1.535	5
<i>TFRC</i>	± 1.375	2
<i>B2M</i>	± 1.456	4
<i>GUS</i>	± 1.244	1
<i>PGK1</i>	± 1.795	7
<i>RLP</i>	± 1.437	3
<i>CYPA</i> and <i>TFRC</i>	± 0.518	1
<i>CYPA</i> and <i>GUS</i>	± 0.537	2
<i>CYPA</i> and <i>RLP</i>	± 0.613	3
<i>GAPDH</i> and <i>RLP</i>	± 0.646	4
<i>GAPDH</i> and <i>PGK1</i>	± 0.648	5

The intervals were estimated using tumor samples from different patients. Tolerance intervals are based on one MD replicate and the average of three qPCR technical replicates per sample. Individual genes and pairwise combinations of two genes were ranked according to the length of the 95% tolerance interval.

^aValues in units of C_T .

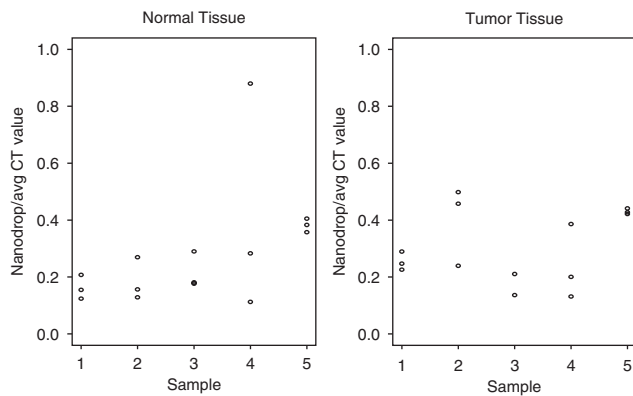


Figure 3 Variation in total RNA quantitation measurements (NanoDrop) relative to average C_T values across the 10 housekeeping genes. Each observation is the average of three qPCR technical replicates for each LCM replicate.

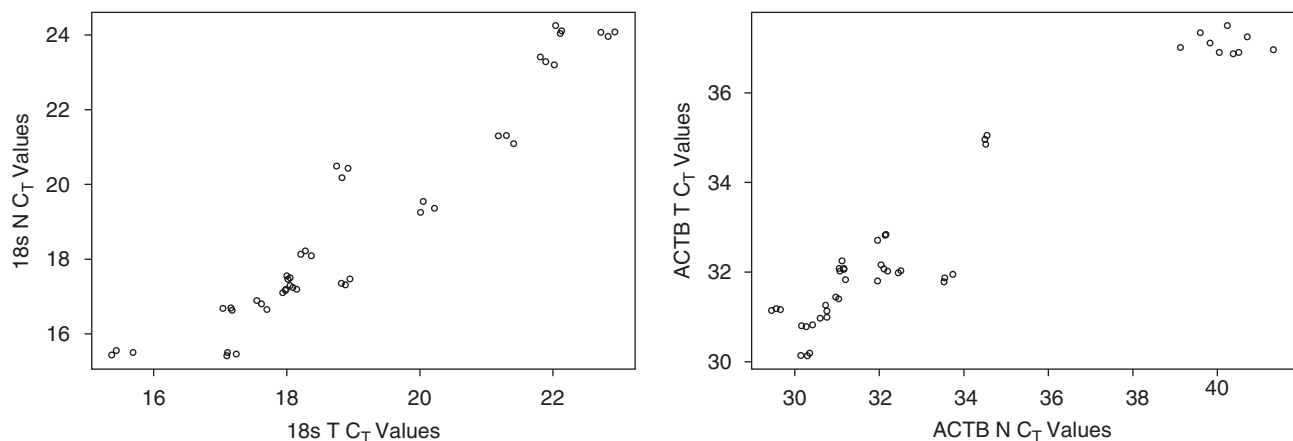


Figure 4 Correlation of paired tumor (T) and normal (N) samples (ie, T and N from the same patient). Each observation is a single qPCR technical replicate for each LCM replicate.

Table 7 Ninety-five percent Tolerance Intervals (TI) on the difference between a single gene or the average of two genes and the average of all 10 housekeeping genes for paired samples

Gene(s)	TI	Single or pair ranking
18s	$\pm 1.062^a$	5
CYPA	± 1.042	3
ACTB	± 1.050	4
HPRT	± 1.388	7
GAPDH	± 1.712	9
TFRC	± 1.696	8
B2M	± 1.211	6
GUS	± 1.771	10
PGK1	± 0.832	1
RPL	± 0.926	2
<hr/>		
18s and HPRT	± 0.426	1
CYPA and HPRT	± 0.565	2
18s and TFRC	± 0.587	3
CYPA and ACTB	± 0.662	4
CYPA and PGK1	± 0.667	5

The intervals were estimated using tumor and normal samples from the same patient. Tolerance intervals are constructed based on one MD replicate and the average of three qPCR technical replicates per sample. Individual genes and pairwise combinations of two genes were ranked according to the size of the tolerance interval.

^aValues in units of C_T .

potential sources of variation were decomposed into components attributable to; (1) sample to sample variability, (2) LCM replicate intra-case variability, and (3) qPCR replicate variability for each dissection. Using tumor tissue (unpaired case), the three components were estimated as 2.5, 0.56, and 0.01, respectively. These data indicate that the majority of variance is from true biological differences among the samples, and that the qPCR measurements within each dissection are highly stable. However, the intra-case variation among LCM samples is not trivial, suggesting there are advantages in performing multiple dissections from each case in a study. Results were similar for the paired situation (data not shown).

Analyzing the gene expression data using three separate normalization techniques revealed an interesting finding in one patient specimen in the study. Both the normal and tumor samples from case no. 5 appeared to contain approximately 100 times more RNA than the other specimens (Figure 5). This unexpected and anomalous result was supported by four separate experimental methods, and was observed consistently in multiple independent dissections. First, the initial NanoDrop readings from the 10 000 dissected

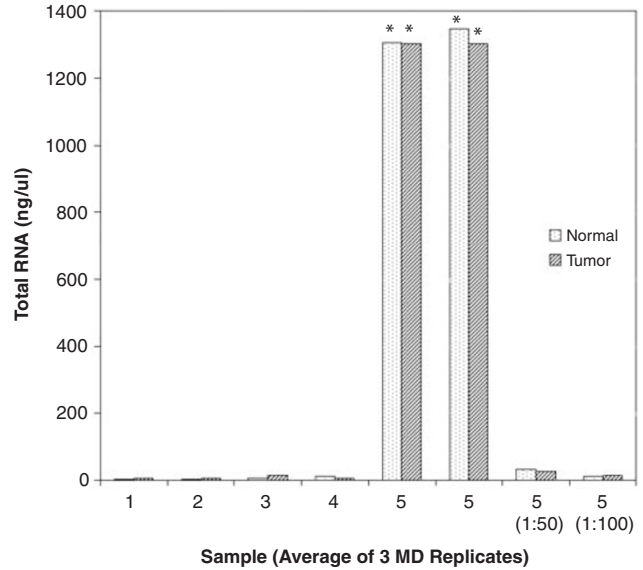


Figure 5 Total RNA concentrations per case sample (ie, average of three MD replicates) tissue type (normal (N) or tumor (T)). Case no. 5 was dissected two times and total RNA from the second set of dissections was diluted 1:50 and 1:100. Average total RNA concentrations (ng/ μ l) per case tissue type: 1N = 4.84, 1T = 7.71; 2N = 5.26, 2T = 7.43; 3N = 6.33, 3T = 16.25; 4N = 12.66, 4T = 7.00; 5^aN = 1305.50, 5^aT = 1301.63; 5^bN = 1345.37, 5^bT = 1302.50; 5^bN 1:50 = 31.27, 5^bT 1:50 = 27.81; 5^bN 1:100 = 13.71, 5^bT 1:100 = 14.37. *a* = first set of three MD replicates. *b* = second set of three MD replicates. * = qRT-PCR did not work (no C_T value) due to 100 times higher template concentration causing template inhibition of qRT-PCR.

normal and from the 10 000 dissected tumor cells showed significantly elevated RNA levels compared to the other four cases, in the range of 26 000 ng as opposed to 260 ng. Second, the BioAnalyzer chip that is routinely used to assess RNA quality in the samples showed an unusual pattern, an apparent large smear of RNA extending from the top to the bottom of the lane, with barely discernible 28S and 18S rRNA bands. This pattern is not consistent with degraded RNA where a smear of small size fragments at the bottom of the lane is typically observed, along with diminished or absent rRNA 28S and 18S bands. Third, the same samples were analyzed on a standard agarose gel and showed an identical pattern as the Agilent BioAnalyzer. Finally, the qRT-PCR failed for each dissection of case no. 5, even though the protocol was identical to that used for the other eight dissected samples that reliably amplified all 10 housekeeping gene transcripts. However, after diluting the normal and tumor samples from case no. 5 100-fold, each gene was consistently amplified, and expression data similar to the other four cases in the study was obtained.

DISCUSSION

The present study examined the use of three separate strategies for normalizing qRT-PCR measurements of MD cell samples. The data show that neither LCM cell count nor total RNA are precise enough to serve as internal standards

for most studies, although as a practical matter both of these methods are useful in calibrating 'ballpark' mRNA amounts. However, endogenous housekeeping genes can serve as a surrogate marker of the cellular transcriptome and provide a precise internal control for expression measurements. Of interest is the finding that as few as one or two endogenous housekeeping genes can be used for MD material. With high probability, a combination of two housekeeping genes was within approximately 0.5 C_T values of the average across all 10 genes. For paired normal and tissue studies, a single housekeeping gene was within one C_T value. Since MD samples typically produce limited amounts of mRNA, the requirement for only one or two genes for normalization will facilitate measurement of a larger number of genes of interest in future studies.

There are several implications of these data for expression analysis of dissected samples. If cell count is utilized as an internal standard, then one can calculate the number of transcripts on a per cell basis and this measurement is independent of the overall transcriptome. Biologically, this information may be uniquely important if the intent of a study is to determine how the absolute amount of mRNA is related to the normal or pathological process being investigated. In other words, if cell count is employed as the normalization standard, an increase in a gene of interest in one sample relative to another indicates that the number of transcripts per cell in that sample is elevated.

Alternatively, comparing gene expression levels using the cellular transcriptome (either RNA measurement or endogenous housekeeping genes as a surrogate) as the normalization standard provides different information. Since the denominator (RNA amount, transcriptome levels) in the ratio of 'gene of interest to internal control' is not necessarily stable across dissected samples, the absolute change in the amount of mRNA per cell cannot be calculated. Rather, the relative level of the gene to the amount of RNA is determined and compared among samples. This analysis can produce a different result than when using cell count. For example, if dissected sample A contains 50% more mRNA of a gene of interest than sample B, but also contains 50% more mRNA overall (including 50% more of the endogenous housekeeping genes), then one would conclude that the gene of interest showed no change between the samples when in fact it was elevated by 50% in absolute amount. On the basis of the results of the present study, however, it appears these issues are of less concern when using MD samples as only endogenous housekeeping genes are an effective normalization tool, and cell count does not provide enough precision to be useful for most studies.

By employing tissue microdissection and multiple independent methods to normalize gene expression levels, we discovered a putatively unique RNA state in one case in the study. Both the normal and tumor cells from case no. 5 appeared to show approximately 100-fold more mRNA than the other eight dissected samples. This finding was verified by

four experimental approaches and in multiple repeat dissections. Since all five specimens in the study were procured at the same institution in an identical manner, were received together at the NIH in the same shipment, and were processed side-by-side in our laboratory and received the same active DNase treatment, there is no obvious rationale for this difference. Histologically, the tumor from this unusual case showed an overall Gleason score of $5 + 4 = 9/10$, and displayed a variety of aggressive patterns of growth, including solid sheets of cells, cords, focal neuroendocrine appearance, and areas of comedonecrosis, a variant that is classified as Gleason 5.³¹ In addition, highly aggressive basaloid carcinoma type areas were identified.

Case no. 5 highlights the unique information that is provided by each internal normalization method. Relying only on endogenous housekeeping genes, it is possible that this unusual finding would have been missed. Typically, we would have calibrated the input RNA using NanoDrop and then subsequent comparison of genes of interest to endogenous control genes would have appeared within the normal range. In other words, the change in apparent absolute mRNA content would not have been readily observed. However, since LCM cell count was used to calibrate general input amounts at the outset of the study, it was immediately clear that case no. 5 was unusual and appeared to contain significantly elevated global mRNA level on a per cell basis.

The cause of the anomalous mRNA results in case no. 5 is not yet known, nor is the relationship to the high-grade cancer observed in this specimen, if any. Analysis of the patient's history, family history, and clinical condition are underway and may reveal an important biological phenomenon. Alternatively, this finding may represent a previously unknown technical artifact associated with gene expression studies of MD patient samples, DNase-resistant mRNA for example, or other artifact related to tissue processing. If so, this problem will need to be uncovered and understood to eliminate this error in future expression studies of dissected samples. Our current interpretation of the data from case no. 5 is that the result is artifactual; however, to date attempts to uncover a technical, non-biological basis for this finding have not been successful.

In summary, three methods for normalizing gene expression data were evaluated in MD tissue samples; cell count during MD, total RNA, and endogenous housekeeping genes. The results indicate that housekeeping genes are the most useful internal standard and they are capable of enabling precise gene expression measurements of dissected cell samples.

ACKNOWLEDGEMENT

We thank Lauren C Georgevich for contribution to the literature review. This research was supported by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research.

1. Simon R, Korn E, McShane L, *et al*. Design and Analysis of DNA Microarray Investigations. Springer: New York, 2004.

2. Murphy RM, Watt KK, Cameron-Smith D, *et al*. Effects of creatine supplementation on housekeeping genes in human skeletal muscle using real-time RT-PCR. *Physiol Genomics* 2003;12:163–174.
3. Khimani AH, Mhashilkar AM, Mikulskis A, *et al*. Housekeeping genes in cancer: normalization of array data. *Biotechniques* 2005;38:739–745.
4. Warrington JA, Nair A, Mahadevappa M, *et al*. Comparison of human adult and fetal expression and identification of 535 housekeeping/maintenance genes. *Physiol Genomics* 2000;2:143–147.
5. Ross DH, Scherf U, Eisen MB, *et al*. Systematic variation in gene expression patterns in human cancer cell lines. *Nat Genet* 2000;24:227–235.
6. Thellin O, Zorzi W, Lakaye B, *et al*. Housekeeping genes as internal standards: use and limits. *J Biotechnol* 1999;75:291–295.
7. Suzuki T, Higgins PJ, Crawford DR. Control selection for RNA quantitation. *Biotechniques* 2000;29:332–337.
8. Bustin SA. Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *J Mol Endocrinol* 2000;25:169–193.
9. Vandesompele J, De Preter K, Pattyn F, *et al*. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 2002;3:research0034.
10. Aerts JL, Gonzales MI, Topalian SL. Selection of appropriate control genes to assess expression of tumor antigens using real-time RT-PCR. *Biotechniques* 2004;36:84–86, 88, 90–91.
11. Biederman J, Yee J, Cortes P. Validation of internal control genes for gene expression analysis in diabetic glomerulosclerosis. *Kidney Int* 2004;66:2308–2314.
12. Tsuji N, Kamagata C, Furuya M, *et al*. Selection of an internal control gene for quantitation of mRNA in colonic tissues. *Anticancer Res* 2002;22:4173–4178.
13. Gorzelnik K, Janke J, Engeli S, *et al*. Validation of endogenous controls for gene expression studies in human adipocytes and preadipocytes. *Horm Metab Res* 2001;33:625–627.
14. Mahoney DJ, Carey K, Fu MH, *et al*. Real-time RT-PCR analysis of housekeeping genes in human skeletal muscle following acute exercise. *Physiol Genomics* 2004;18:226–231.
15. Pogue-Geile KL, Greenberger JS. Effect of the irradiated microenvironment on the expression and retrotransposition of intracisternal type A particles in hematopoietic cells. *Exp Hematol* 2000;28:680–689.
16. Barnard GF, Staniunas RJ, Bao S, *et al*. Increased expression of human ribosomal phosphoprotein P0 messenger RNA in hepatocellular carcinoma and colon carcinoma. *Cancer Res* 1992;52:3067–3072.
17. Henry JL, Coggin DL, King CR. High-level expression of the ribosomal protein L19 in human breast tumors that overexpress erbB-2. *Cancer Res* 1993;53:1403–1408.
18. Vaarala MH, Porvari KS, Kyllonen AP, *et al*. Several genes encoding ribosomal proteins are over-expressed in prostate-cancer cell lines: confirmation of L7a and L37 over-expression in prostate-cancer tissue samples. *Int J Cancer* 1998;78:27–32.
19. Best CJM, Gillespie JW, Yi Y, *et al*. Molecular alterations in primary prostate cancer after androgen ablation therapy. *Clin Cancer Res* 2005;11:6823–6834.
20. Gillespie JW, Best CJ, Bichsel VE, *et al*. Evaluation of non-formalin tissue fixation for molecular profiling studies. *Am J Pathol* 2002;160:449–457.
21. Emmert-Buck MR, Bonner RF, Smith PD, *et al*. Laser capture microdissection. *Science* 1996;274:998–1001.
22. Okuducu AF, Janzen V, Hahne JC, *et al*. Influence of histochemical stains on quantitative gene expression analysis after laser-assisted microdissection. *Int J Mol Med* 2003;11:449–453.
23. de Kok JB, Roelofs RW, Giesendorf BA, *et al*. Normalization of gene expression measurements in tumor tissues: comparison of 13 endogenous control genes. *Lab Invest* 2005;85:154–159.
24. Schmid H, Cohen CD, Henger A, *et al*. Validation of endogenous controls for gene expression analysis in microdissected human renal biopsies. *Kidney Int* 2003;64:356–360.
25. Cohen CD, Frach K, Schlondorff D, *et al*. Quantitative gene expression analysis in renal biopsies: a novel protocol for a high-throughput multicenter application. *Kidney Int* 2002;61:133–140.
26. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2⁻(Delta Delta C(T)) Method. *Methods* 2001;25:402–408.
27. Fink L, Seeger W, Ermert L, *et al*. Real-time quantitative RT-PCR after laser-assisted cell picking. *Nat Med* 1998;4:1329–1333.
28. Searle S, Casella G, McCulloch C. *Variance Components*. Wiley Press: New York, 1992.
29. Pinheiro J, Bates DM. *Mixed-effects Models in S and S-PLUS*. Springer: New York, 2000.
30. Xu LL, Sun C, Petrovics G, *et al*. Quantitative expression profile of PSGR in prostate cancer. *Prostate Cancer Prostatic Dis* 2006;9:56–61.
31. Epstein JI, Allsbrook Jr WC, Amin MB, *et al*. The 2005 International Society of Urological Pathology (ISUP) consensus conference on Gleason grading of prostatic carcinoma. *Am J Surg Pathol* 2005;29:1228–1242.