Analytical Biochemistry **299,** 63–70 (2001)
doi:10.1006/abio.2001.5369, available online at http://www.idealibrary.com on IDEAL®

# Identification and Validation of Endogenous Reference Genes for Expression Profiling of T Helper Cell Differentiation by Quantitative Real-Time RT-PCR[1]

H. K. Hamalainen,*[,2] J. C. Tubman,† S. Vikman,* T. Kyrölä,* E. Ylikoski,* J. A. Warrington,‡ and R. Lahesmaa*

**Turku Centre for Biotechnology, University of Turku and Åbo Akademi University, P.O. Box 123, FIN-20521 Turku, Finland; †Faculty of Biological Sciences, School of Biological Sciences, Dublin Institute of Technology, Kevin Street, Dublin 8, Ireland; and ‡Affymetrix, Inc., 3380 Central Expressway, Santa Clara, California 95051*

Received June 4, 2001; published online November 2, 2001

**Real-time RT-PCR method was exploited to identify endogenous reference genes in differentiating human T helper cells. When using this technology in our experimental system, finding a set of genes whose mRNA expression levels would not change appeared to be very challenging. Our initial plan to use the expression level of GAPDH in normalizing the results failed, because the mRNA expression of GAPDH underwent significant changes during the cell culture. Additional studies on the transcription of several other classical housekeeping genes led to similar results. Our second approach was to use results from an extensive survey of gene expression done by oligonucleotide micorarrays and to select another panel of genes for testing. This resulted in the identification of three genes whose expression was relatively stable in our experimental system and, therefore, suitable as endogenous reference genes in these cells. The results indicate that the expression level of a constitutively expressed gene may change during the cell culture *in vitro*, which emphasizes again the importance of carefully validating endogenous control genes for comparative quantification.** 

Measurement of RNA transcripts by real-time (kinetic) RT-PCR circumvents several obstacles known to limit the quantification potential of PCR-based methods (1, 2). Classical PCR relies on end-point analyses of the amplified product, whereas in real-time RT-PCR methods the fluorescent signal is recorded in real time from the linear range of amplification (3, 4). Measurement of fluorescence has a broader dynamic range than classical visual or autoradiography detection methods and has the potential of measuring concentration scales of up to at least 5 to 6 orders of magnitude (5, 6). These unique features of real-time PCR in combination with optimized robust amplification provide a good basis for RNA quantification.

To quantify the level of mRNA expression of a target gene, the expression level of an external reference molecule is often measured. However, the preparation and inclusion of a synthetic DNA or RNA standard (competitive RT-PCR) (7) in PCR reactions is a cumbersome effort if the expression of tens of target genes in the same sample is simultaneously followed. In addition, these approaches do not control for manipulations of the sample before the addition of a standard. Therefore, in most cases a natural choice for quantification is to measure the level of expression of an endogenous control gene exposed to the same manipulations as the target gene prior to measurements. Hence, the expression level of a reference gene in the sample can be used to normalize the mRNA expression level of target genes. For this purpose the expression of a constitutively expressed endogenous gene (housekeeping gene, maintenance gene) in the same sample is often monitored. However, several reports have indicated that differences in the expression levels of endogenous reference genes may exist between different individuals (8) or occur as a result of pathological changes (9) or cell differentiation (10).

[1] This study was supported by grants from the Academy of Finland, National Technology Agency (TEKES), and Turku University Hospital Fund.

[2] To whom correspondence and reprint requests should be addressed at Turku Centre for Biotechnology, P.O. Box 123, FIN-20520 Turku, Finland. Fax: +358-2-333 8000. E-mail: hhamalai@btk.utu.fi.

0003-2697/01 $35.00

Our studies aim at understanding the molecular mechanisms of T helper cell differentiation. To quantify gene expression of multiple genes during the lymphocyte differentiation process, it was necessary to find a set of genes whose expression would remain stable throughout the process. We first tested a set of commonly used housekeeping genes as endogenous genes and found out that their expression is subject to significant changes during differentiation. We then validated a new set of candidate genes that had been selected based on the results from a microarray screen of approximately 7000 human genes in adult and fetal tissues (11). Based on the microarray data a panel of genes whose expression is relatively stable in various human tissues was validated in our *in vitro* culture of T cell differentiation using real-time RT-PCR. This approach led to the discovery of new endogenous reference genes that were expressed at a relatively constant level in differentiating T cells.

## MATERIALS AND METHODS

### *Differentiation of CD4$^+$ Th1 and Th2 Cells*

Human CD4$^+$ T helper cells were isolated from cord blood of healthy neonates essentially as described by Sornasse *et al.* (12). Cells were grown in Yssel's medium supplemented with 1% of human AB serum (Gemini Bioproducts, Woodland, CA) and 100 U/ml human recombinant IL-2[3] (Cellular Products, U.S.A.). Cell differentiation was primed with either 2.5 ng/ml human recombinant IL-12 (R & D Systems, Minneapolis, MN) or 10 ng/ml human recombinant IL-4 (R & D Systems) for Th1 or Th2 cells, respectively. Culture medium for Th2 cells contained 10 $\mu$g/ml human recombinant $\alpha$-IL-12 (R & D Systems). Initial stimulation of the cells was carried out by plating $1.0 \times 10^6$ T cells/ml, $0.5 \times 10^6$/ml of irradiated CD32-B7 transfected mouse L fibroblasts (13), and 100 ng/ml PHA (Murex Diagnostics, Chatillon, France) in 24-well flat-bottomed plates (Lindbro, ICN Biomedicals, Inc., U.S.A.). CD4$^+$ T cells were harvested at 6, 24, and 48 h and isolated from the cultures using immunomagnetic purification (Dynabeads M-450 CD4, Dynal A.S., Oslo, Norway) prior to RNA isolation. The cells were fed every other day and split 3–4 days after stimulation. In these cell culture conditions T helper cells proliferated well and evaded the feeder cells rapidly from the culture. At day 7 T cells were restimulated and cultured further as described above for another 7 days (day 14 cells). Cell proliferation was monitored daily and cultures split to maintain cell density at approximately $0.5–2 \times 10^6$ cells/ml.

[3] Abbreviations used: IL, interleukin; PHA, phytohemagglutinin.

### *Isolation of RNA and cDNA Synthesis*

Total RNA was isolated with the GlassMAX RNA microisolation spin cartridge system (Gibco, BRL, Life Technologies). One microgram of total RNA was treated with DNase I (Amplification Grade, Gibco, BRL, Life Technologies) and the sample was divided into two. The oligo(dT)$_{12–18}$-primed first-strand cDNA synthesis was carried out in a total volume of 20 $\mu$l either with Superscript II RNase H$^-$ reverse transcriptase (+RT cDNA) or without the enzyme (−RT control) according to the manufacturer's protocols (superscript preamplification system for first strand cDNA synthesis, Gibco, BRL, Life Technologies). A total of 0.5 $\mu$l of cDNA was used for each TaqMan measurement.

### *Real-Time RT-PCR (TaqMan)*

TaqMan probes and primers (Table 1 and Ref. 14) were designed by using the computer software Primer Express (PE Biosystems, Foster City, CA), except for human $\beta$-actin, where commercially available assay reagents were used (PE Biosystems, Foster City, CA). Specificity of the PCR amplification of each primer pair was confirmed with a 4% NuSieve agarose gel (BioWhittaker Molecular Applications, Inc., U.S.A.). The PCR reactions were carried out using TaqMan universal PCR master mix (PE Biosystems) with 300 nM oligonucleotide primers (Gibco, BRL, Life Technologies) and 200 nM fluorogenic probe (MedProbe, Oslo, Norway). For the signal detection, ABI Prism 7700 sequence detector was programmed to an initial step of 2 min at 50°C and 10 min at 95°C, followed by 40 thermal cycles of 15 s at 95°C and 1 min at 60°C. Each measurement was carried out in duplicate and repeated.

### *Calculations*

The $C_T$ value is defined as the number of PCR cycles required for the fluorescence signal to exceed the detection threshold value (5, 6). This threshold is set to the log linear range of the amplification curve and kept constant (0.05) for data analysis throughout the study. The correlation between the $C_T$ value and the fold difference in the concentration was determined individually for each probe and primer set. This was carried out by measuring the expression in a cDNA sample diluted two- or fourfold over the entire detection range. With the PCR efficiency of 100% the $C_T$ values of two separate genes can be compared ($\Delta C_T$) and the fold difference calculated from:

$$\text{Fold Difference} = 2^{|C_T1 - C_T2|} = 2^{|\Delta C_T|}$$

The average standard deviation for all 12 detections used in this study and collectively for all measurements was 0.19 (mean CV 14%).

## RESULTS

### *Criteria for the Detection Design of Endogenous Reference Genes*

When selecting the endogenous reference genes and designing their real-time RT-PCR detection, the following features were originally considered as preferential: constitutive expression during various stages of differentiation, no processed pseudogenes, known structures of intron–exon junction and low sequence variability between different individuals. We routinely treat our RNA samples with DNase I and control its success by analyzing the minus RT samples after cDNA synthesis. Therefore, a detailed knowledge of pseudogenes is not needed. However, with the current status of human sequence information the requirement of known intron–exon structures is difficult to meet. Additionally, alternative splicing of human genes is common (15) and the existence of a variety of splicing sites at the site of detection would disturb the RNA quantification. Therefore, to avoid ambiguous quantification, we selected the exons of coding sequences, if known, as sites of detection when primers and probes were designed. Some commonly used housekeeping genes (HPRT, $\beta_2$-microglobulin) were excluded from the study, because of the suggestive evidence in the literature that their expression might be affected by the cell culture conditions used here (16, 17). The ribosomal subunits 18S and 28S were excluded mainly for two reasons: (i) oligo(dT) was used for RT priming and (ii) their expression level is very high and therefore differs from the level of expression of most of the target genes we study. In our first set of experiments we included the following *bona fide* endogenous genes: ADA (adenosine deaminase), $\beta$-actin, GAPDH, L32, and Rb (retinoblastoma). The second set of experiments was carried out using a panel of candidate genes whose level of expression was reported to be relatively similar in multiple human tissues (11). This panel of genes included aldolase, c-yes, EF-1-$\alpha$, MLN51, phosphomannomutase, proteasome subunit Y, and UbcH5B.

### *Specificity and Linearity of the Real-Time RT-PCR Detection*

Specificity of the PCR amplification for each primer pair (Table 1) was analyzed on an agarose gel and confirmed to give a single amplified band of a correct size (less than 150 bp). The expression of each gene was measured in a human T lymphocyte cDNA sample, which had been serially diluted by two- or fourfold. To create a standard curve over the entire detection range, five to eight dilutions of T cell cDNA were usually required. For all 12 genes studied, the average correlation coefficient of detection was 0.989 (SD = 0.005, range 0.982–0.997). Furthermore, the mean slope value [$\log_{10}$ (cDNA dilution)] versus $C_T$ value] for all genes tested was −3.644 (SD = 0.239, range −4.034 to −3.270). The ideal detection is characterized as 100% efficient PCR (slope −3.3); hence the average efficiency of detection for all 12 primer pairs in the present study was 110% (SD = 7%, range 99–122%).

### *Endogenous Reference Genes during T Helper Cell Differentiation*

*Panel 1: Commonly used housekeeping genes.* Real-time RT-PCR was used to measure the expression of commonly used housekeeping genes at various stages of lymphocyte differentiation. Naive $CD4^+$ T cells were isolated from cord blood drawn from three neonates, and three sets of Th1 or Th2 cultures were established. The expression of ADA, $\beta$-actin, GAPDH, L32, and Rb was quantified in $CD4^+$ T cells at time 0 and at various stages of differentiation (6 h, 1 day, 2 days, 7 days, and 14 days). The expression of other genes was compared to the expression of ADA. Among the tested genes the expression of ADA was chosen as a reference, because the $C_T$ values of ADA exhibited the smallest change between the $C_T$ values of the samples at zero and 14 days. Expression levels of the genes studied changed significantly (Fig. 1). The gene expression varied from 3.4-fold (Rb) up to 7.1-fold (GAPDH) when the expression at zero was compared to that at day 14. The maximal changes for each gene were 11.4-fold for $\beta$-actin (at 48 h), 17.0-fold for GAPDH (at 48 h), 5.0-fold for L32 (at 14 days), and 3.4-fold for Rb (at 14 days). With the exception of L32, the expression of $\beta$-actin, GAPDH, and Rb increases (inverse correlation) during the T cell differentiation. The changes in the expression for the panel 1 genes could be not accounted for by differences in the starting amount of total RNA or due to a failure of the cDNA synthesis, because in all T cell culture samples originating from different individuals a similar trend of change in the expression was consistently observed (data not shown). An independent experiment in which samples of pooled RNAs were measured gave essentially similar results (not shown).

*Panel 2: Selection of genes expressed at similar levels in various human tissues.* In the original panel of housekeeping genes tested, only ADA showed some potential to act as an endogenous reference gene. However, due to the relativity of the comparative method, relying on the expression of only one reference gene was considered unreliable. Among the other genes, none fulfilled the criteria set for a good endogenous reference gene. Therefore, another panel of genes was selected. Based on the results of a microarray screen (11), the expression of aldolase, c-yes, EF-1-$\alpha$, MLN51, phosphomannomutase, proteasome subunit Y, and UbcH5B among several other genes had been demonstrated to differ less than 4-fold in several human

**TABLE 1**

TaqMan Primers and Probes Used in This Study

| Gene (human) | Accession no. | 1) 5′-(FAM)-probe-(TAMRA)-3′<br>2) 5′-primer 1-3′<br>3) 5′-primer 2-3′ | Amplicon size (bp) |
|---|---|---|---|
| ADA | X02994 | 1) 5′-CGGCATGGCCGCGTCACTT-3′<br>2) 5′-CCATTTCTGCACACACGTATACC-3′<br>3) 5′-TGGCCAGGGCACATAATCA-3′ | 65 |
| Aldolase | X12447 | 1) 5′-TTTCCATGATGGCGAGGGCTGAG-3′<br>2) 5′-TAACGGGCCAGAACATTGG-3′<br>3) 5′-TGGCGTTGTGTGCTGAAGAT-3′ | 80 |
| c-yes | M15990 | 1) 5′-TGCCGTGCCCTCAGGGCTG-3′<br>2) 5′-AGAACAAGTGGAGCGAGGATACA-3′<br>3) 5′-CATCAATTCATGGAGGGATTCTG-3′ | 70 |
| EF-1-$\alpha$ | J04617 | 1) 5′-AGCGCCGGCTATGCCCCTG-3′<br>2) 5′-CTGAACCATCCAGGCCAAAT-3′<br>3) 5′-GCCGTGTGGCAATCCAAT-3′ | 59 |
| L32 | X03342 | 1) 5′-ACATGCTGCCCAGTGGCTTCCG-3′<br>2) 5′-CAACATTGGTTATGGAAGCAACA-3′<br>3) 5′-TGACGTTGTGGACCAGGAACT-3′ | 80 |
| MLN51 | X80199 | 1) 5′-AGGCCTGTGGAAGCTGGTGGGC-3′<br>2) 5′-CAAGGAAGGTCGTGCTGGTT-3′<br>3) 5′-ACCAGACCGGCCACCAT-3′ | 64 |
| Phosphomannomutase | U86070 | 1) 5′-TGCAGGAAGGCGGCCACCTC-3′<br>2) 5′-CGATCTGCACTCTACTTCGTAGCT-3′<br>3) 5′-CGGCTCGCCAGAAAATTG-3′ | 69 |
| Proteasome subunit Y | D29012 | 1) 5′-TGGTCCCAGTGGAAACTTCTCGGCT-3′<br>2) 5′-CAAACTGCACGGCCATGATA-3′<br>3) 5′-GAGGCATTCACTCCAGACTGG-3′ | 69 |
| Retinoblastoma | M15400 | 1) 5′-ATCTCAGGACCTTGGTGGACACTGTGT<br>2) 5′-ACCTCAAACAAGGAAGAGAAATGAG-3′<br>3) 5′-ACATCTGTGAGAGACAATGAATCCA-3′ | 85 |
| UbcH5B | U39317 | 1) 5′-TGATCTGGCACGGGACCCTCCA-3′<br>2) 5′-TGAAGAGAATCCACAAGGAATTGA-3′<br>3) 5′-CAACAGGACCTGCTGAACACTG-3′ | 72 |

*Abbreviations:* FAM, 6-carboxyfluorescein; TAMRA, 6-carboxytetramethylrhodamine.

tissues. Additionally, the expression of GAPDH was measured to control and confirm a similar trend of GAPDH expression in both sets of experiments (panel 1 and panel 2, not shown). From the second panel of candidate genes, the expression of MLN51 remained most stable throughout the T cell differentiation process; therefore, its expression was used for normalization. The expression of aldolase, c-yes, EF-1-$\alpha$, MLN51, phosphomannomutase, proteasome subunit Y, and UbcH5B in reference to MLN51 at various stages of T cell differentiation is presented in Fig. 2. The overall trend for aldolase, phosphomannomutase, and proteasome subunit Y expression was an increase in expression, whereas the expression of c-yes decreased. The maximal changes for each gene relative to MLN51 (no change) and the zero sample were 20.4-fold for aldolase (at 48 h and 7 days), 3.3-fold for phosphomannomutase (at day 14), 5.8-fold for proteasome subunit Y (at day 7), and 3.4-fold for c-yes (at 24 h). Most importantly, in addition to MLN51 the expression of EF-1-$\alpha$ and UbcH5B remained relatively constant during the entire differentiation period.

### *Relative Quantification with Endogenous Reference Genes EF-1-α, MLN51, and UbcH5B*

The expression of EF-1-$\alpha$, MLN51, and UbcH5B was found to be relatively similar in naive T cells and in T cells differentiated to Th1 or Th2 cells *in vitro* for various periods of time. Moreover, these genes were expressed at similar levels independent of the donor both in Th1 and Th2 cells (data not shown). The standard curves for EF-1-$\alpha$ ($y = -3.270x + 21.55$), MLN51 ($y = -3.849x + 30.67$), and UbcH5B ($y = -3.732x + 26.07$) are almost parallel (Fig. 3). However, PCR detection efficiency for EF-1-$\alpha$, MLN51, and UbcH5B differs from the 100% theoretical efficacy by $-1$, 17, and 13%, respectively. How does the difference in the PCR detection influence the quantification of gene expression? For example, if the difference in the PCR detection efficiency is 10% on average, it means that the expression of a target gene in any two samples equally normalized to a given reference gene can be quantitatively compared within a 2-fold accuracy as long as their $C_T$ values do not differ by more than 10

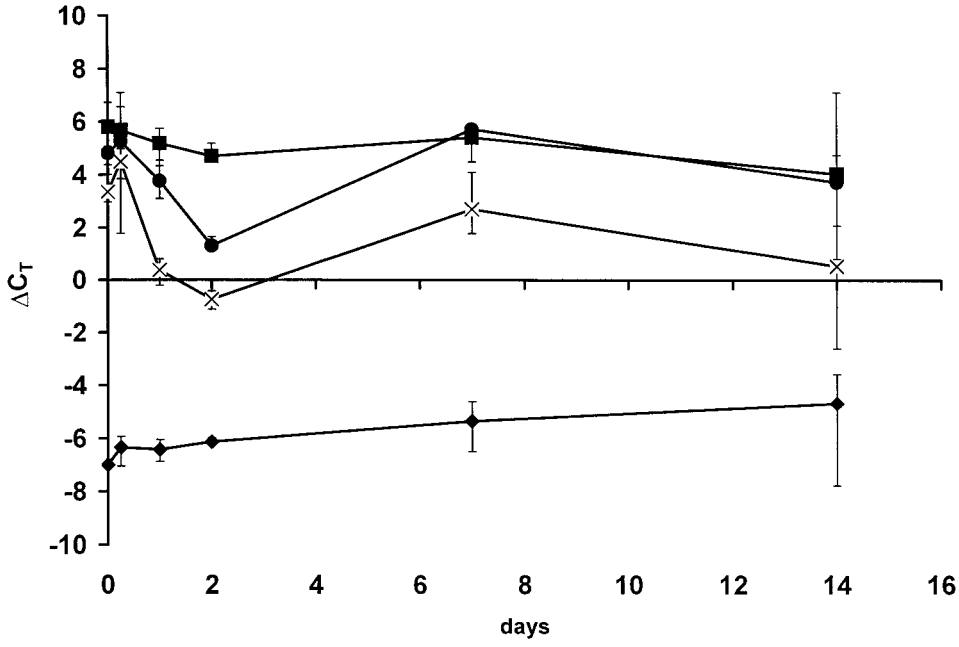


**FIG. 1.** Expression of housekeeping genes at various stages of T cell differentiation relative to ADA (adenosine deaminase). Changes in the $\Delta C_T$ values in the course of the T cell differentiation are shown for $\beta$-actin (●), GAPDH (×), L32 (◆) and Rb (retinoblastoma, ■). Expression for each gene was measured at time zero, 6 h, 1 day, 2 day, 7 day, and 14 day of cell culture. Each data point measured individually from cultured T cell samples represents a mean of $\Delta C_T$ values from three Th1 and three Th2 cell cultures derived from different individuals. In both Th1 and Th2 cell cultures similar trends of gene expression were observed. For each $\Delta C_T$ value the average range is marked (bar).

PCR cycles (Table 2). The suggested limit of 10 PCR cycles (with the error of 10%) for $C_T$ values translates to a difference of 1024-fold in the concentration of the target transcript. Such a difference in the concentration of a target gene expression between two separate samples is relatively rare, at least within our experimental settings with cultured T cells to date (see also www.HuGeIndex.org). Moreover, at least to some degree an error that indicates an increase in PCR efficiency might result from the manual pipetting of sequential dilutions, an error that in fact does not take place when actual (undiluted) samples are measured.

For routine expression profiling of $CD4^+$ T cells cultured in conditions that mimic Th1 or Th2 cell differentiation, the measurement of even one of the endogenous reference genes might be satisfactory. However, it seems reasonable to suggest that under other experimental manipulations the expression of all three could be measured and their mean $C_T$ values used for normalization. The formula for normalization would then be expressed as $\Delta C_T = C_{T(\text{target})} - C_{T(\text{mean of EF-1-}\alpha\text{, MLN51, and UbcH5B})}$.

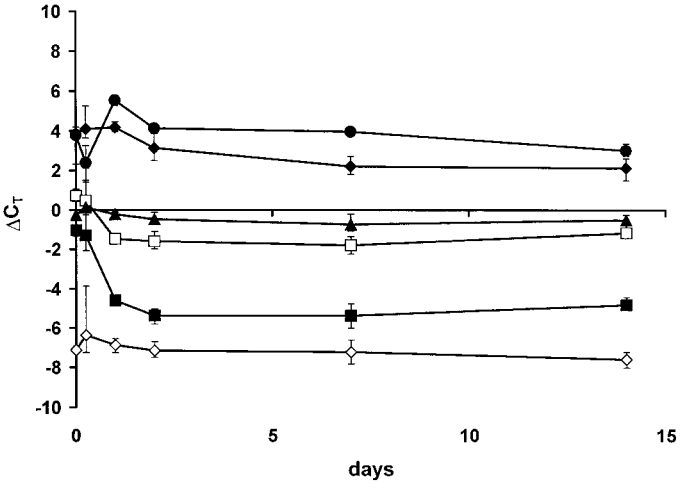


**FIG. 2.** Expression of endogenous reference candidate genes at various stages of T cell differentiation relative to MLN51. Changes in the $\Delta C_T$ values in the course of the T cell differentiation are shown for aldolase (■), c-yes (●), EF-1-$\alpha$ (◇), phosphomannomutase (◆), proteasome subunit Y (□), and UbcH5B (▲). Expression for each gene was measured at time zero, 6 h, 1 day, 2 day, 7 day, and 14 day of cell culture. The curve marked with squares represents the expression of aldolase, curve with circles expression of c-yes, with open diamonds EF-1-$\alpha$, with diamonds phosphomannomutase, with open squares proteasome subunit Y, and with triangles UbcH5B. Each value measured individually from cultured T cell samples represents a mean of $\Delta C_T$ values from three Th1 and three Th2 cell cultures derived from different individuals. In both Th1 and Th2 cell cultures similar trends of gene expression were observed. For each $\Delta C_T$ value the average range is marked (bar).

## DISCUSSION

Recent advances in molecular genetic technologies include development of RT-PCR methods as well as new approaches for large-scale expression profiling with high-density microarrays (6, 18, 19). While a mi-

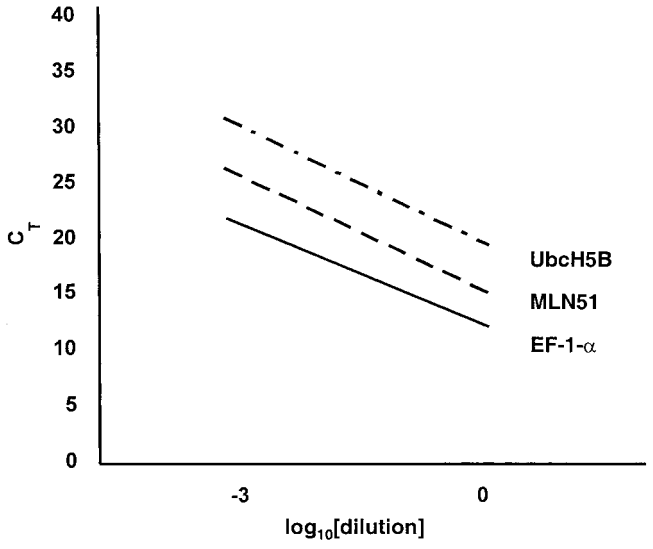


**FIG. 3.** Comparison of the standard curves for endogenous reference genes EF-1-$\alpha$, MLN51, and UbcH5B. The standard curves were created by measuring the expression in a series of dilutions of T cell cDNA with real-time RT-PCR. To illustrate parallelism, a linear regression curve of dilutions up to $10^3$-fold is shown. The curves for EF-1-$\alpha$ ($y = -3.270x + 21.55$), MLN51 ($y = -3.849x + 30.67$), and UbcH5B ($y = -3.732x + 26.07$) are presented. For each curve y denotes for the $C_T$ value and $x$ for the $\log_{10}$ (dilution).

croarray analysis can simultaneously detect and compare the expression of thousands of genes, real-time (kinetic) RT-PCR can easily be adopted for subsequent confirmatory and follow-up studies.

To quantitatively measure changes at the RNA level with RT-PCR methods, the RNA first needs to be isolated and converted into cDNA. Due to the manipulations, separate measurements of the expression level of an internal control gene in the same sample is often a prerequisite for RNA quantification. In practice, however, the expression level of a reference gene (housekeeping/maintenance) may vary. This concern was also raised in our studies when temporal changes in the expression levels of genes in human $CD4^+$ T cells were to be studied. We first selected a panel of commonly used housekeeping genes and studied the expression of ADA, $\beta$-actin, GAPDH, L32, and Rb in differentiating $CD4^+$ T cells. The expression level for commonly used housekeeping genes studied did change and this was most pronounced with $\beta$-actin (up to 11.4-fold) and GAPDH (up to 17.0-fold). The results indicated that none of the tested genes could act as a proper endogenous reference gene.

Oligonucleotide microarrays were used to study the average transcript levels of housekeeping genes (maintenance genes) in several human fetal and adult tissues (11). In a study of ~7000 genes screened with mRNA extracted from eleven tissues from multiple individuals, 47 transcripts were identified that had average expression levels that were within fourfold in each tissue. Taking these results as a starting point, we studied the expression of a new set of candidate genes in naive and differentiating T cells. Among the selected 7 genes, we discovered 3 genes with the desired stability ($\leq$ 2-fold change) in their expression throughout the differentiation process. In other words, the expression of these genes was relatively stable when freshly isolated naive cells and samples from the *in vitro* activated T cells were compared.

In addition to the requirement of a stable expression for a reference gene, the overall accuracy of the detection is also an essential determinant of quantification. Moreover, as several large-scale microarray studies have already indicated, the amplitude of the vast majority of changes at the transcriptional level is small ($\leq$2- to 4-fold) (20–31). This further emphasizes the importance of appropriate methods with quantitative capacity. The real-time RT-PCR detection, which occurs in the linear range of the PCR, leads to very accurate and reproducible measurements of the target molecules (5). Additionally, the dynamic range of at least 5 to 6 orders of magnitude ensures a very broad quantification potential. In several recent papers, the accuracy and reproducibility of the real-time RT-PCR detection has been validated (32–37). In all studies, both intraassay and interassay variability of real-time detection has been generally low (4–24%). This level of reproducibility has also been confirmed in our study and repeatedly observed in our other real-time detection measurements with similar principles of detection design.

Only one paper was identified comparing quantification with real-time RT-PCR (TaqMan) by multiple methods including the use of an absolute standard curve, relative standard curve, or a comparative ($\Delta C_T$) method (35). That study found the performance of all three to be equivalent. Accuracy of the real-time detection together with near 100% efficiency of PCR for both a target and the reference gene has several outcomes that make the comparative real-time RT-PCR method even more compelling. Namely, when the $\Delta C_T$ value truly reflects the target transcript concentration of the

**TABLE 2**

Calculated Error in Real-Time RT-PCR with Varying Sample Concentrations

| Fold dilution | PCR efficiency of 100% (theoretical $C_T$) | PCR efficiency of 110% (calculated $C_T$) | Fold error ($2^{\Delta CT}$) |
|---|---|---|---|
| Undiluted | 25.00 | 25.00 | None |
| 10 | 28.63 | 28.30 | 1.3 |
| $10^2$ | 32.26 | 31.60 | 1.6 |
| $10^3$ | 35.86 | 34.90 | 1.9 |

sample, consequently, in repeated measurements, it is not necessary to occupy every run with a standard curve. Also in case the overall cDNA synthesis is not as robust as desired, the comparative quantification method and the use of $\Delta C_T$ values would bypass the problem, assuming that both the target and the reference gene in the sample during synthesis are equally well transcribed. For comparative quantitative RT-PCR, factors like the purity and the integrity of the RNA sample must be considered. However, in our case the samples used for our comparative expression profiling represent an extensively purified subpopulation of hematopoietic cells and are probably less prone to such variation.

Transcription of the classical housekeeping genes appears to be subject to regulatory factors that have not been recognized. The cells in our experimental culture conditions are exposed to cytokines immediately after isolation and additionally induced to proliferate. We found that selecting an endogenous reference gene among the classical housekeeping genes can result in faulty interpretations of the research data. Nonetheless, based on extensive screening, it was possible to find a set of genes whose expression remained stable throughout the differentiation. Without validation, the quantitative interpretation of expression profiling information can be misleading or ambiguous. All approaches generate candidate lists that include false positives and false negatives. Using multiple quantitative methods in carefully controlled studies aids in filtering out the candidates that may be the result of experimental artifacts.

Although it is clear that the results presented here cannot directly be extrapolated to other cell types, we believe that the study illustrates the benefits of large-scale technologies when searching for new endogenous reference genes for comparative quantification. The combination of microarray analysis and real-time RNA detection methods provides a powerful approach for quantifying gene activity.

## ACKNOWLEDGMENTS

We thank Marju Niskala and Paula Suominen for technical assistance and Ilana Saarikko for mathematical advice.

## REFERENCES

1. Ferre, F. (1992) Quantitative or semi-quantitative PCR: Reality versus myth. *PCR Methods Appl.* **2,** 1–9.
2. Freeman, W. M., Walker, S. J., and Vrana, K. E. (1999) Quantitative RT-PCR: Pitfalls and potential. *Biotechniques.* **26,** 112–122, 124–125.
3. Higuchi, R., Dollinger, G., Walsh, P. S., and Griffith, R. (1992) Simultaneous amplification and detection of specific DNA sequences. *Biotechnology (N. Y.)* **10,** 413–417.
4. Higuchi, R., Fockler, C., Dollinger, G., and Watson, R. (1993) Kinetic PCR analysis: Real-time monitoring of DNA amplification reactions. *Biotechnology (N. Y.)* **11,** 1026–1030.
5. Heid, C. A., Stevens, J., Livak, K. J., and Williams, P. M. (1996) Real time quantitative PCR. *Genome Res.* **6,** 986–994.
6. Gibson, U. E., Heid, C. A., and Williams, P. M. (1996) A novel method for real time quantitative RT-PCR. *Genome Res.* **6,** 995–1001.
7. Gilliland, G., Perrin, S., Blanchard, K., and Bunn, H. F. (1990) Analysis of cytokine mRNA and DNA: Detection and quantitation by competitive polymerase chain reaction. *Proc. Natl. Acad. Sci. USA* **87,** 2725–2729.
8. de Leeuw, W. J., Slagboom, P. E., and Vijg, J. (1989) Quantitative comparison of mRNA levels in mammalian tissues: 28S ribosomal RNA level as an accurate internal control. *Nucleic Acids Res.* **17,** 10137–10138.
9. Bhatia, P., Taylor, W. R., Greenberg, A. H., and Wright, J. A. (1994) Comparison of glyceraldehyde-3-phosphate dehydrogenase and 28S-ribosomal RNA gene expression as RNA loading controls for northern blot analysis of cell lines of varying malignant potential. *Anal. Biochem.* **216,** 223–226.
10. Shimokawa, T., Kato, M., Ezaki, O., and Hashimoto, S. (1998) Transcriptional regulation of muscle-specific genes during myoblast differentiation. *Biochem. Biophys. Res. Commun.* **246,** 287–292.
11. Warrington, J. A., Nair, A., Mahadevappa, M., and Tsyganskaya, M. (2000) Comparison of human adult and fetal expression and identification of 535 housekeeping/maintenance genes. *Physiol. Genomics.* **2,** 143–147.
12. Sornasse, T., Larenas, P. V., Davis, K. A., de Vries, J. E., and Yssel, H. (1996) Differentiation and stability of T helper 1 and 2 cells derived from naive human neonatal CD4+ T cells, analyzed at the single-cell level. *J. Exp. Med.* **184,** 473–483.
13. Yang, L. P., Byun, D. G., Demeure, C. E., Vezzio, N., and Delespesse, G. (1995) Default development of cloned human naive CD4 T cells into interleukin-4- and interleukin-5-producing effector cells. *Eur. J. Immunol.* **25,** 3517–3520.
14. Hamalainen, H., Meissner, S., and Lahesmaa, R. (2000) Signaling lymphocytic activation molecule (SLAM) is differentially expressed in human Th1 and Th2 cells. *J. Immunol. Methods.* **242,** 9–19.
15. Mironov, A. A., Fickett, J. W., and Gelfand, M. S. (1999) Frequent alternative splicing of human genes. *Genome Res.* **9,** 1288–1293.
16. Steen, A. M., Luthman, H., Hellgren, D., and Lambert, B. (1990) Levels of hypoxanthine phosphoribosyltransferase RNA in human cells. *Exp. Cell Res.* **186,** 236–244.
17. Neumann, H., Schmidt, H., Cavalie, A., Jenne, D., and Wekerle, H. (1997) Major histocompatibility complex (MHC) class I gene expression in single neurons of the central nervous system: Differential regulation by interferon (IFN)-gamma and tumor necrosis factor (TNF)-alpha. *J. Exp. Med.* **185,** 305–316.
18. Duggan, D. J., Bittner, M., Chen, Y., Meltzer, P., and Trent, J. M. (1999) Expression profiling using cDNA microarrays. *Nature Genet.* **21,** 10–14.
19. Lockhart, D. J., and Winzeler, E. A. (2000) Genomics, gene expression and DNA arrays. *Nature* **405,** 827–836.
20. Zhu, H., Cong, J. P., Mamtora, G., Gingeras, T., and Shenk, T. (1998) Cellular gene expression altered by human cytomegalovirus: Global monitoring with oligonucleotide arrays. *Proc. Natl. Acad. Sci. USA* **95,** 14470–14475.
21. Der, S. D., Zhou, A., Williams, B. R., and Silverman, R. H. (1998) Identification of genes differentially regulated by interferon alpha, beta, or gamma using oligonucleotide arrays. *Proc. Natl. Acad. Sci. USA* **95,** 15623–15628.
22. Amundson, S. A., Bittner, M., Chen, Y., Trent, J., Meltzer, P., and Fornace, A. J., Jr. (1999) Fluorescent cDNA microarray

hybridization reveals complexity and heterogeneity of cellular genotoxic stress responses. *Oncogene* **18,** 3666–3672.

23. Loftus, S. K., Chen, Y., Gooden, G., Ryan, J. F., Birznieks, G., Hilliard, M., Baxevanis, A. D., Bittner, M., Meltzer, P., Trent, J., and Pavan, W. (1999) Informatic selection of a neural crest-melanocyte cDNA set for microarray analysis. *Proc. Natl. Acad. Sci. USA* **96,** 9277–9280.
24. Khan, J., Bittner, M. L., Saal, L. H., Teichmann, U., Azorsa, D. O., Gooden, G. C., Pavan, W. J., Trent, J. M., and Meltzer, P. S. (1999) cDNA microarrays detect activation of a myogenic transcription program by the PAX3–FKHR fusion oncogene. *Proc. Natl. Acad. Sci. USA* **96,** 13264–13269.
25. ter Linde, J. J., Liang, H., Davis, R. W., Steensma, H. Y., van Dijken, J. P., and Pronk, J. T. (1999) Genome-wide transcriptional analysis of aerobic and anaerobic chemostat cultures of *Saccharomyces cerevisiae. J. Bacteriol.* **181,** 7409–7413.
26. Geiss, G. K., Bumgarner, R. E., An, M. C., Agy, M. B., van't Wout, A. B., Hammersmark, E., Carter, V. S., Upchurch, D., Mullins, J. I., and Katze, M. G. (2000) Large-scale monitoring of host cell gene expression during HIV-1 infection using cDNA microarrays. *Virology* **266,** 8–16.
27. Chang, Y. E., and Laimins, L. A. (2000) Microarray analysis identifies interferon-inducible genes and Stat-1 as major transcriptional targets of human papillomavirus type 31. *J. Virol.* **74,** 4174–4182.
28. Kaminski, N., Allard, J. D., Pittet, J. F., Zuo, F., Griffiths, M. J., Morris, D., Huang, X., Sheppard, D., and Heller, R. A. (2000) Global analysis of gene expression in pulmonary fibrosis reveals distinct programs regulating lung inflammation and fibrosis. *Proc. Natl. Acad. Sci. USA* **97,** 1778–1783.
29. Travers, K. J., Patil, C. K., Wodicka, L., Lockhart, D. J., Weissman, J. S., and Walter, P. (2000) Functional and genomic analyses reveal an essential coordination between the unfolded protein response and ER-associated degradation. *Cell* **101,** 249–258.
30. Webb, G. C., Akbar, M. S., Zhao, C., and Steiner, D. F. (2000) Expression profiling of pancreatic beta cells: Glucose regulation of secretory and metabolic pathway genes. *Proc. Natl. Acad. Sci. USA* **97,** 5773–5778.
31. Muta, H., Boise, L. H., Fang, L., and Podack, E. R. (2000) CD30 signals integrate expression of cytotoxic effector molecules, lymphocyte trafficking signals, and signals for proliferation and apoptosis. *J. Immunol.* **165,** 5105–5111.
32. Wingo, S. T., Ringel, M. D., Anderson, J. S., Patel, A. D., Lukes, Y. D., Djuh, Y. Y., Solomon, B., Nicholson, D., Balducci-Silano, P. L., Levine, M. A., Francis, G. L., and Tuttle, R. M. (1999) Quantitative reverse transcription-PCR measurement of thyroglobulin mRNA in peripheral blood of healthy subjects. *Clin. Chem.* **45,** 785–789.
33. Bieche, I., Onody, P., Laurendeau, I., Olivi, M., Vidaud, D., Lidereau, R., and Vidaud, M. (1999) Real-time reverse transcription-PCR assay for future management of ERBB2-based clinical applications. *Clin. Chem.* **45,** 1148–1156.
34. Raggi, C. C., Bagnoni, M. L., Tonini, G. P., Maggi, M., Vona, G., Pinzani, P., Mazzocco, K., De Bernardi, B., Pazzagli, M., and Orlando, C. (1999) Real-time quantitative PCR for the measurement of MYCN amplification in human neuroblastoma with the TaqMan detection system. *Clin. Chem.* **45,** 1918–1924.
35. Johnson, M. R., Wang, K., Smith, J. B., Heslin, M. J., and Diasio, R. B. (2000) Quantitation of dihydropyrimidine dehydrogenase expression by real-time reverse transcription polymerase chain reaction. *Anal. Biochem.* **278,** 175–184.
36. Schmittgen, T. D., Zakrajsek, B. A., Mills, A. G., Gorn, V., Singer, M. J., and Reed, M. W. (2000) Quantitative reverse transcription-polymerase chain reaction to study mRNA decay: Comparison of endpoint and real-time methods. *Anal. Biochem.* **285,** 194–204.
37. Rajeevan, M. S., Vernon, S. D., Taysavang, N., and Unger, E. R. (2001) Validation of array-based gene expression profiles by real-time (kinetic) RT-PCR. *J. Mol. Diagn.* **3,** 26–31.