A Relevant Reference Gene and Normalization for mRNA Real-Time PCR Quantification in Specimens with Distinct Cell Types and Variant Integrity

Kazuyuki SUGAHARA, Akiko UEMURA, Norihiko AKAMATSU, Kazuto TSURUDA, Hiroo HASEGAWA, Katsunori YANAGIHARA, Yasuaki YAMADA, Shimeru KAMIHIRA

Department of Laboratory Medicine, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan

Variability of clinical samples with respect to cell types and quality makes it indispensable to normalize mRNA quantification by real-time reverse transcription-polymerase chain reaction (RT-PCR). The objective of the present study was to elucidate the influence of the difference in RNA integrity and the expression status of control genes commonly used. To compare the expression status of GAPDH, β-actin, PBGD and 18S rRNA in different cell samples, real-time RT-PCR by the LightCycler Technology was applied. The relevance of the above-mentioned normalization by control gene was evaluated through the practical measurement of survivin using normal lymphocytes from 19 healthy donors, adult T-cell leukemia (ATL) cells from 30 patients with ATL and 27 cell lines. The mRNA integrity was found to be tolerable for absolute quantification when the preservation rate of 28S and 18S RNA within total RNA was at least 30%. The expression status of control genes was prominently variable with the expression difference of 4-fold in β-actin, 20-fold in GAPDH, 30-fold in 18S rRNA and 66-fold in PBGD among the different cell types, normal lymphocytes, chronic and acute ATL cells, ATL cell lines, other hematopoietic cell lines and solid tumor cell lines. The survivin mRNA data normalized by such a control gene were influenced by the expression variability of the control genes. These results indicate that the control genes can be used to correct for sample-to-sample variation within the same cell types, but they are not always relevant for normalization in the distinct cell types.

Keywords: mRNA quantification; Normalization; PCR; Survivin

ACTA MEDICA NAGASAKIENSIA 51: 57 - 63, 2006

Introduction

Recently, the evaluation of gene expression levels has been indispensable for understanding molecular pathology in various diseases. For the detection and quantification of mRNA, at present, the real-time reverse transcription-polymerase chain reaction (RT-PCR) is widely and increasingly used because of its high sensitivity, good reproducibility, and dynamic quantification range. For successful mRNA quantification using real-time RT-PCR, there remains the problem of bias and variance to be resolved, including sample-to-sample variations, variation in RNA integrity, RT efficiency differences and cDNA sample loading variation. In particular, careful normalization with the use of internal control is essential to compare mRNA levels among specimens with different cell type. A major problem for real-time RT-PCR quantification is the degradation level of isolated RNA (integrity), the accurate loading dose of cDNA, and data normalization among variant samples. Accurate quantification and quality assessment of the starting RNA sample are important for absolute quantification methods that normalize target mRNA expression levels against total RNA dose, e.g. copies per total RNA. For quantification in real-time PCR, two strategies can generally be performed, i.e. absolute and relative. This absolute quantification relates the PCR signal to input copy number using a calibration, while relative quantification measures the relative changes in mRNA expression level. The reliability of the absolute quantification, therefore, mostly depends upon the condition of identical amplification efficiencies of both the native target and the calibration curve. As compared to absolute quantification, relative quantification is easier and simpler to carry out because it does not necessitate calibration curve; the expression levels by relative quantification are represented by the ratio of a target gene versus a housekeeping gene (HKG) or reference gene, and it is adequate to compare changes in the gene expression levels within the intra-case samples. Survivin is a member of 8 inhibitors of apoptotic proteins detected

Address correspondence: Shimeru Kamihira, M.D., Department of Laboratory Medicine, Nagasaki University Graduate School of Biomedical Sciences, 1-7-1 Sakamoto, Nagasaki 852-8501 JAPAN
TEL: +81-(0)95-849-7407, FAX: +81-(0)95-849-7422, E-mail: kamihira@nagasaki-u.ac.jp

Received April 13, 2006; Accepted June 9, 2006
usually only in fetal tissues. Survivin is now one of notable genes used as a universal biomarker for diagnostics and therapeutics, because it is one of the top four transcripts among 3.5 million human transcriptomes uniformly upregulated in cancer tissues, but not in normal tissues. All of these facts contribute to survivin being a possible candidate as a universal biomarker in clinical oncology. Our laboratory also pays attention to the clinical relevance and pathological significance of survivin for leukemias including adult T-cell leukemia (ATL), in which the disease is shown to have up-regulation of survivin and should be discriminated from healthy carriers with HTLV-1. Our preliminary studies have shown the potential of survivin for screening the disease using real-time quantitative RT-PCR. However, the positive rate in leukemias of the survivin mRNA density was somewhat discrepant according to the usage of either absolute or relative quantification. This suggested the importance of the choice of control gene. We therefore conducted the present study to investigate which reference control genes may be suitable for data normalization when we use quantitative real-time RT-PCR by LightCycler Technology.

Materials and Methods

Specimens

Peripheral blood mononuclear cells (PBMCs) in 19 healthy donors, leukemia cells in 30 ATL and 27 malignant cell lines were used for the present study. ATL patients were diagnosed by routine diagnostic procedures including morphology, histochemistry, surface immunophenotyping and monoclonal integration of HTLV-1 provirus into the genomes using HTLV-1 whole-length and pX probes by Southern blot analysis. ATL consisted of 5 chronic subtypes and 25 acute subtypes. We analyzed a panel of the following 27 cell lines: 7 ATL-derived cell lines (ST1, KOB, KK-1, OMT, RS04, SHI and NAK established in our laboratory), 6 other hematopoietic (HTP)-derived cell lines (K562, HL60 and THP-1 purchased from ATCC, Manassas, VA; SP49 and SU-DHL-1 purchased from DSMZ, Brauschweig, Germany; CAL-1 established in our laboratory), 13 lung cancer-derived cell lines (H460, H1299, H526, PC-6, H69, H82, H441, H358, H23, A549 and H520 purchased from ATCC, Manassas, VA; SW1573 and Lu134-A purchased from RIKEN BioResource, Tokyo, Japan) and 1 prostate cancer-derived cell line (LNCaP purchased from ATCC, Manassas, VA).

Cells and mRNA preparations

PBMCs were separated with Ficoll-density centrifugation. Separated cells were divided into two fractions: one part for immediately processing RNA extraction and the other part for cryopreserving in liquid nitrogen. Total RNA was extracted from cells with the ISOGEN kit (NipponGene, Toyama, Japan) and was treated with DNase (Message Clean TM kit: GeneHunter, Nashville, TN) to remove contaminated DNA. Complementary DNA (cDNA) was synthesized from 1 μg RNA by incubation with ThermoScript RT (15 units/μL) and 2.5 μM oligo (dT) 20 primer (THERMO SCRIPT RT-PCR System: Invitrogen, Carlsbad, CA). cDNA synthesis was performed according to the manufacturer’s instructions. cDNA was stored at -80°C until real-time quantitative PCR was performed.

The integrity of isolated RNAs was assessed according to electrophoretic patterns of 28S and 18S rRNA by a denaturing agarose (1.5%) gel system containing 0.66 M formaldehyde. Figure 1 shows electrophoretic features by the denaturing gel system on total RNA samples; lanes 1 to 4 correspond to total RNAs of the 28S and 18S preservation rate of 31% or more, 30-21%, 20-11%, and 10% or less, respectively. The standard controls with different integrity were artificially produced by degrading-treatment to add the ISOGEN into cell specimens at different time (0, 5, 10 and 60 min). The quality of total RNA was assessed by the preservation rate (%) of 28S/8S rRNA relative to total RNA.

![Figure 1](image)

**Figure 1.** Assessment of the integrity of total RNA according to electrophoretic patterns with denaturing agarose gel electrophoresis. The preservation rate of 28S and 18S rRNAs for total tRNAs is 31% or more, 30-21%, 20-11% and 10% or less in lanes 1 to 4, respectively.

Real-time quantitative PCR

mRNA quantification for survivin and housekeeping genes (Table 1) of glyceraldehyde-3-phosphate dehydrogenase (GAPDH), β-actin, porphobilogen deaminase (PBGD), and 18S rRNA was performed with real-time RT-PCR using a LightCycler thermal cycle instrument (Roche Applied Science, Mannheim, Germany). All reactions were run in duplicates. The primers and hybrid probes used for survivin were as follows:

- forward primer: 5'-TTCTCAAGGCCACCGCATC-3';
- reverse primer: 5'-ACAGAAGGAAAGGCACACC-3';
- probe-sensor: 5'-ATCCACTGCCCCACTGAAACGGA-3' FITC;
- probe-anchor: LC Red640 5'-CCAGACTGTTCGCCAGTGTGTTCTTCT-3' P.

As previously reported, reactions were performed in a 20 μL volume with 0.5 μM primers for survivin, 4.5 mM MgCl2, and 0.25 μM hybridization probes using the LightCycler FastStart DNA Master Hybridization kit (Roche Applied Science, Mannheim, Germany). Each 20 μL reaction contained 1 μL template cDNA, corresponding to 50 ng total RNA. The reaction conditions were 95°C for 10 min for activation of the Taq polymerase and then 40 cycles of 10
sec each at 95°C (denaturation) followed by 5 sec at 63°C (annealing) and 13 sec at 72°C (extension). A standard curve was generated using a 10-fold dilution series of 5 different plasmid concentrations, driving from a clone in pGEM-T Easy Vector (Promega KK, Tokyo, Japan) inserted with the 48 to 494 region of the survivin sequence. The quantification data were analyzed with LightCycler software.

Similarly, by modifying slightly the technique for real-time quantitative PCR of survivin, GAPDH, β-actin, PBGD mRNAs and 18S rRNA were quantified using LightCycler™-Primer Set (Roche Diagnostic GmbH, Mannheim, Germany) of Human GAPDH (#110304), Human β-actin (#270504), Human PBGD (#281003) and 18s RNA (#061103), and the LightCycler FastStart DNA Master Hybridization kit, respectively, according to the manufacturer's instructions.

The absolute copy amount of an unknown cDNA sample was calculated from the standard cDNA curves and was reported as copies/50 ng total RNA. Under the assumption that each type of cell of samples harbors HKGs transcripts of the same expression levels as normal PBMCs, each sample was also normalized relative to the control gene expression level of normal PBMCs, on the basis of its internal control gene content, according to the following formula: normalized data of \( \frac{N_{\text{survivin}}}{N_{\text{GAPDH}}} = \text{constant} \times \frac{\text{sample survivin copies}}{\text{sample control gene copies}} \). The constant is the mean number of copies of control gene in normal PBMCs.

### Results

**Assessment of the integrity of total RNA and mRNA**

For the overall quality of RNA preparation, denaturing agarose gel electrophoresis containing 7 mol urea was performed. A denaturing gel system is suggested because most RNA (rRNA) forms extensive secondary structure via intra-molecular base pairing, and this prevents the RNA from migrating according to its size, mainly of 28S, 18S and 5S. We first evaluated the relationship between the overall quality of isolated total RNA and the quality of mRNA. The former was assessed on the basis of the preservation rate of 28S/18S by denaturing gel electrophoresis patterns, and the latter was assessed from the expression level of GAPDH. As shown in Figure 2, when the preservation rate of 28S/18S rRNA was about 30% or more, the GAPDH expression levels were almost constant with a variation of 7.9×10^3 to 3.9×10^3/total RNA, while when the preservation rate was 30% or less, the expression levels gradually declined along with the reduction of the preservation rate. This suggests that samples preserving 30% or more of 28S and 18S rRNA are tolerable for absolute quantification of mRNAs.

**Correction for degradation variance of RNA by normalization**

In clinical practice, samples of poor quality are sometimes unavoidable. We therefore examined the necessity for normalizing the absolute quantification data into the relative quantification data using GAPDH control. The integrity of RNA was divided into 4 groups according to the 28S/18S preservation rate: grade 1 (G-1): 31% or over; G-2: 30-21%; G-3: 20-11%; and G-4: 10% or less. Both the absolute expression levels of survivin and GAPDH mRNAs monotonously declined with degradation of total RNA integrity from G-1 to G-4 (Figure 3 A), while the relative quantification levels of survivin mRNA relative to GAPDH were nearly stable for different RNA
integrity except for G-4 where they largely varied (Figure 3 B).

Expression performance of control genes in different cell types

Figure 4 shows the respective absolute expression levels of GAPDH, β-actin, 18S rRNA, and PBGD in the 6 different cell groups of normal PBMCs, chronic ATL cells, acute ATL cells, ATL cell lines, HPT cell lines, and solid cell lines. As shown in Figure 4, the median of control gene mRNA expression levels among the 6 cell types was variable with the difference of 20-fold in GAPDH (Figure 4 A), 4-fold in β-actin (Figure 4 B), 30-fold in 18S rRNA (Figure 4 C) and 66-fold in PBGD (Figure 4 D). In particular, the expression levels
of GAPDH and PBGD showed a tendency to increase in order of acute ATL cells, ATL cell lines, HPT cell lines, and solid cell lines, whereas β-actin was relatively uniformly expressed among the 6 different cell groups. The expression of 18S rRNA was observed to be apparently down-regulated in the malignant cell line groups compared to the in vivo cell group.

Absolute and relative quantification of survivin expression

The absolute quantification of survivin mRNA produced by our real-time RT-PCR system varied from 1.65×10^5 to 1.75×10^6 copies/total RNA (Figure 5). Table 2 presents the relative quantification of survivin mRNA normalized to 4 respective control genes for 6 cell types. The expression patterns differed largely among cell types depending on the expression performance of the control gene used. For example, acute ATL cells, as compared to normal PBMCs, were estimated to harbor mRNA expression by 9-fold and 19-fold using GAPDH and 18s rRNA, respectively.

Table 2 provides the positive rate for survivin in each control gene with reference to a cut off value of the mean+2×(standard deviation) in normal PBMCs; the positive rates for survivin were similar among the absolute and the 4 relative normalized data.

Table 2. Expression levels of survivin mRNA presented as absolute quantification and as normalized to control genes

<table>
<thead>
<tr>
<th>Group</th>
<th>Sample size</th>
<th>Absolute quantification</th>
<th>Control genes used for normalization</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>GAPDH</td>
</tr>
<tr>
<td>Normal donors</td>
<td>0.39±0.24</td>
<td>0.24±0.17</td>
<td>β-actin</td>
</tr>
<tr>
<td>ATL chronic</td>
<td>2.60±3.91</td>
<td>0.77±0.52</td>
<td>18S rRNA</td>
</tr>
<tr>
<td>ATL acute</td>
<td>10.0±10.0</td>
<td>2.20±2.01</td>
<td>PBGD</td>
</tr>
<tr>
<td>ATL cell lines</td>
<td>7.0±51.2</td>
<td>4.86±4.30</td>
<td></td>
</tr>
<tr>
<td>HTP cell lines</td>
<td>55.8±35.2</td>
<td>6.24±7.40</td>
<td></td>
</tr>
<tr>
<td>SOLID cell lines</td>
<td>105.0±55.3</td>
<td>40.1±34.2</td>
<td></td>
</tr>
</tbody>
</table>

*Number of copies per 50 ng total RNA.
*Data are presented by multiplying 1000.
*Mean±standard deviation.

Discussion

Four methods are available to quantify mRNA: northern blotting, RNase protection assay, RT-PCR and cDNA array. Among them,
RT-PCR is the most common and convenient for quantifying the copy number of target genes. In general, for measuring copy number using real-time RT-PCR, there are two quantification strategies; absolute and relative. Absolute quantification relates the PCR signal to input copy number using a calibration curve, while relative quantification determines the changes in steady-state mRNA levels of a gene across multiple samples and expresses it relative to the levels of an internal control RNA. At present, if accurate in calibration curve and total RNA dose, it is generally acceptable to report absolute quantification by copy number per total RNA. However, absolute quantification does not necessarily properly reflect the true RNA state of the cells, because there remain problems such as the integrity of total RNA, accuracy calibration curve, and the quality of the denominator for the template. This requires appropriate normalization for each specimen. In general, mRNA expression level of a gene is represented relatively to that of a control gene. If one uses high-quality RNA with accurate RNA loading and precise calibration curve, absolute quantification is considered accurate. In the present study, we have used blood or cultured specimens with lymphoid cells to quantify the mRNA levels of survivin and commonly used internal HKGs and rRNA to investigate their availability in clinical diagnostic assays. We therefore first investigated how we could guarantee the quality of mRNA within total RNA. Since the total RNA consists of about 80% of rRNA, 15% of tRNA and 5% of mRNA, the integrity of rRNA is considered to be surrogate for guarantee of total RNA including the quality of mRNA. The present study confirmed this as follows: when the preservation rate of 28S/18S within total RNA was 30% or more, mRNA of GAPDH seemed to be expressed at a constant level, indicating the closeness of integrity of 28S/18S rRNA to the quality of mRNA. The absolute quantification data obtained from samples of 30% or more 28S/18S in RNA integrity are therefore significant and are guaranteed. However, in clinical practice, the practical samples sometimes possess several defects such as poor quality, different sample origins and distinct cell types. To resolve these problems in specimens of poor quality, we examined whether one can normalize absolute values into relative data. Our results show that, if the cell types are same, the relative data normalized to GAPDH is useful, as an alternative to absolute quantification, for evaluating the mRNA profile even for RNAs of poor quality. The relative data even for RNAs of poor quality were thus shown to be relevant for assessing target gene expression levels under the condition that the origins and types of the cells are same.

The problem of relevant normalization against a control gene remains to be elucidated for different sample origins or cell types. In general, the ideal control gene is thought to express its transcripts at a constant level in all cell types and tissues, as well as in cancer cells. Although β-actin expression in the present study appeared to be mostly constant in different cell types, recent reports with the use of BestKeeper tool and GeNorm tool have pointed out that β-actin is not always suitable as universal control gene. Another report, however, supports the usefulness of β-actin because of a small variation in expression with only 2- to 3-fold discrepancy in K562 and leukocytes. Of the GAPDH expression, normal PBMCs and chronic ATL cells showed the performance similar to that of β-actin, while acute ATL cells and malignant cell lines were highly expressed than chronic ATL cells, indicating up-regulation of GAPDH during the cell-cycle or in proliferating cells. Indeed, the study by Mansur displayed a 19-fold increase in the amount of GAPDH mRNAs in relation to DNA synthesis. The up-regulation in cancers has also already been reported in hepatomas, malignant cell lines and prostate cancer. 18S rRNA and PBGD are strongly variable among groups, especially among the in vivo cell groups and cell line groups. rRNA is considered useful because its expression levels are less likely to vary under conditions that affect the expression of mRNAs. Our findings, however, showed that its expression levels are greatly different between normal and malignant cells, indicating that the relative dose of rRNAs is less in malignant cells despite the increase of the total RNA dose. This may overestimate the expression levels of target gene mRNA. PBGD contrasted with 18S rRNA regarding the mRNA expression levels in distinct cells. Taken together, all of these findings suggest that each gene tested in the present study has its merits and drawbacks, at least, for data normalization. In contrast to our results, there is a report based on the study using serum-stimulated fibroblasts, that 18S rRNA is suitable as an internal control gene. This different conclusion may have resulted from the characteristics of cells used, serum-stimulated fibroblasts and fresh or cultured ATL cells. Moreover, Radonic et al. have described that the ranking of the control genes for usefulness is dependent on the assessment strategy such as the BestKeeper. The data of absolute quantification presented in Table 2 seem reasonably comparable to our previous semi-quantitative data. However, according to the absolute quantification strategy in real-time RT-PCR, these data are not always guaranteed to be correct, because total RNAs and components of rRNA, tRNA and mRNA are not necessarily constant on a per cell basis, especially in normal or malignant cells. It suggests the necessity of normalization to reference genes. Tricarico et al. have reported findings, similar to our results presented in Table 3, that VEGF mRNA quantification normalized to total RNA is correlated to its protein levels, and have pointed out that normalization to HKGs or 18S RNA is inappropriate. In particular, the expression of 18S rRNA and PBGD was markedly discrepant between in vitro cell lines and freshly isolated in vivo cells. Such performance of control gene likely affects the outcome of the normalized data. Indeed, in contrast to our high expression data of survivin in all of lung cancer cell lines, only 36.8% in lung cancer in vivo tissues was reported to be positive with the usage of 18S rRNA. Thus, although nobody knows yet which results are true, change of the outcome by normalized data is likely due to, at least in part, different performance of the control gene.

In conclusion, normalization is essential, but it sometimes leads to inadequate estimation due to the characteristics of expression performance in the control genes used. Thus, at present, absolute quantification under appropriate conditions is useful to some extent for mRNA expression assessment in distinct cell types.
Acknowledgement

This study was supported in part by Grant-in-Aid Scientific Research (No.13557228/No.15659136) from the Japan Society for the Promotion of Science.

References