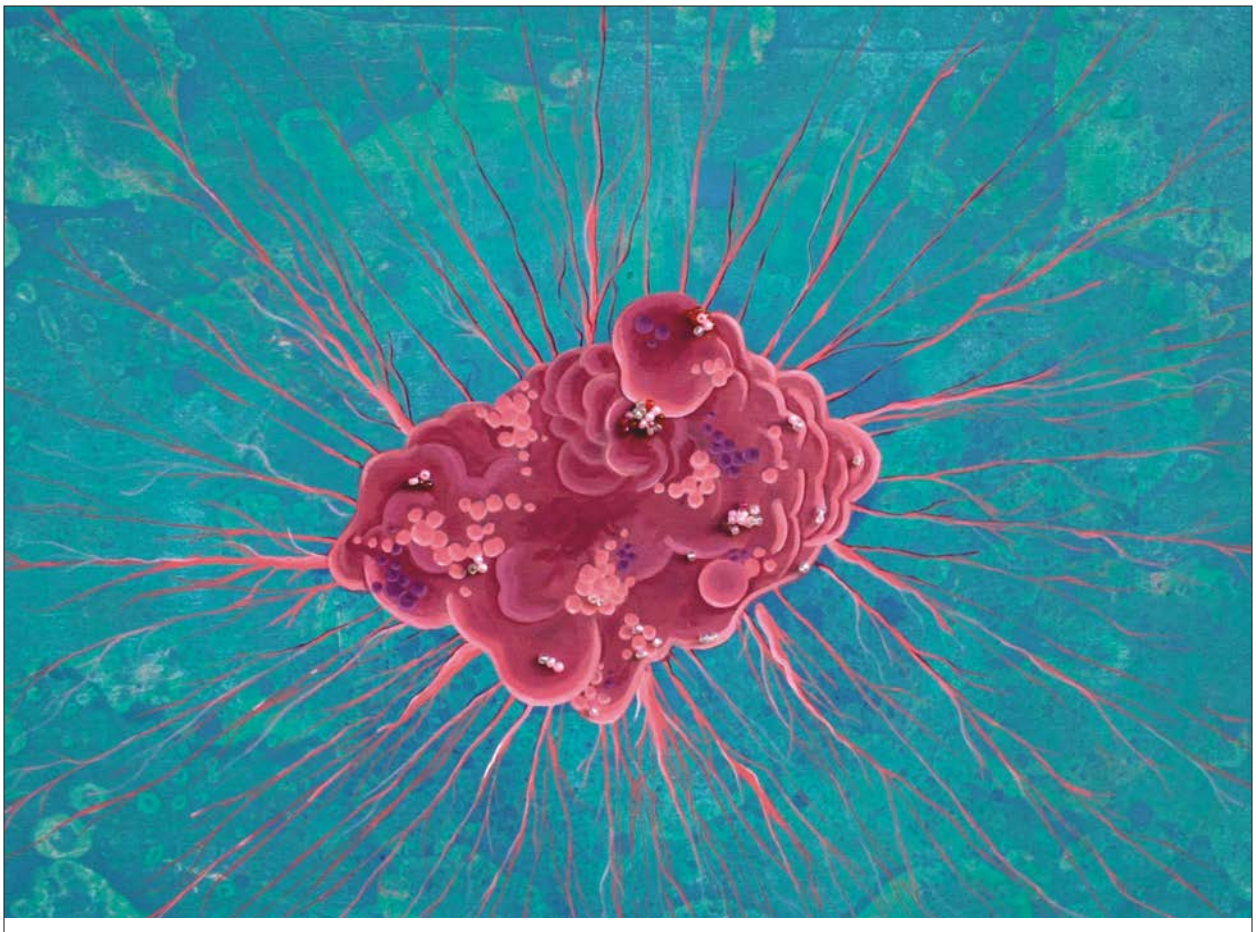

International Journal of Cancer Research

ISSN 1811-9727

Academic Journals Inc.

www.academicjournals.com

International Journal of Cancer Research 3 (1): 1-12, 2007
ISSN 1811-9727

Evaluation of Reference Genes for Normalization of Quantitative Real Time PCR in Non-Small Cell Lung Cancer

[1]Chiaki Endo, [2]Zhifu Sun, [3]Julian R. Molina, [4]John R. Attewell and [2]Ping Yang
[1]Department of Thoracic Surgery, Institute of Development, Aging and Cancer,
Tohoku University, 4-1 Seiryo-Machi, Aoba-ku, Sendai 980-8575, Japan
[2]Department of Health Sciences Research, [3]Department of Medical Oncology,
[4]Department of Laboratory Medicine and Pathology,
Mayo Clinic College of Medicine, 200 First Street SW, Rochester, MN 55905, USA

Abstract: To compare gene expression levels of lung cancer by real time quantitative reverse transcription polymerase chain reaction (qRT-PCR), choosing appropriate control genes for normalization of RNA input is important. We examined whether frequently used control genes are really appropriate. Eighteen lung cancer tissue samples were used in this study. Gene expression levels were measured by qRT-PCR and compared to the data generated by a DNA microarray experiment. The expression levels of four housekeeping genes that are most frequently used as control genes for qRT-PCR published reports in lung cancer were examined. Four test genes (TP53, laminin γ2, TGFα and cathepsin L2) were also examined. These genes had significantly higher expression in the high aggressive disease group. When taken as single genes, the expression levels of these so called control genes (glyceraldehydes-3phosphate dehydrogenase, TATA box binding protein, β actin and β2 microgloblin) were not consistent among the tested samples. However, the geometric mean of the threshold cycle of the TATA box binding protein and β2 microglobulin showed less variation than that of the single control gene. These results indicate that normalization by multiple control genes is more appropriate for comparison of gene expression levels of lung cancer when using qRT-PCR analysis.

Key words: Non-small cell lung cancer, squamous cell carcinoma, RNA, real time PCR, control gene, normalization

Introduction

Researchers have investigated the differences between gene expression profiles of non-small cell lung cancer by stage, histology, or other factors using quantitative real time polymerase chain reaction (qRT-PCR). When comparing gene expression levels between samples, internal control genes are most frequently used to normalize the mRNA fraction. Ideally, the reference mRNA should be expressed at a constant level among all samples across all groups under comparison. Therefore, the housekeeping gene appears to be a good reference, although no single housekeeping gene has been a perfect candidate in all situations (Lossos *et al*., 2003; Beillard *et al*., 2003; Lion, 2001; Lupberger *et al*., 2002) and none has been validated in control experiments.

In the published literature, some genes such as TATA box binding protein, Abelson (ABL1), and protein kinase cGMP-dependent type 1 have been recommended as control genes for qRT-PCR of lymphoid tissues after proper validation (Lossos *et al*., 2003; Beillard *et al*., 2003). To our knowledge, no reports have been found regarding control genes for accurate normalization in lung cancer. In almost all studies of lung cancer, housekeeping genes have been used without proper validation of their

Corresponding Author: Ping Yang, Department of Health Sciences Research, Mayo Clinic, 200 First Street SW, Rochester, Minnesota, 55905, USA

presumed stability of expression. Thus, we evaluated four housekeeping genes frequently used as references in gene expression studies of lung cancer to examine whether or not they are really appropriate for normalization of RNA fraction.

Materials and Methods

Sample Selection

In this study, the same samples were examined using both qRT-PCR and DNA microarray analysis. Samples were separated into two-comparison groups, high vs. low aggressive disease cases as defined in a previous study (Yang *et al*., 2004). Briefly, all analyzed samples came from patients with stage I Squamous Cell Carcinoma (SCC) of the lung that have undergone surgical resection at Mayo Clinic in Rochester, Minnesota. The high aggressive disease group consisted of patients who died within two years; the low aggressive disease group consisted of patients who lived beyond four years. Eligible study subjects were prospectively diagnosed between 1997 and 2000 and followed through 2003. Patients with high and low aggressive disease were one-to-one matched on age, gender, race, tumor histology, size, smoking status and treatment. Eight high aggressive disease samples and ten low aggressive disease samples were included in our analysis. Eight of the ten low aggressive samples were also matched to the samples in the high-aggressive group. Tumors of each patient were snap frozen after resection and stored at -80°C. Demographic data of the 18 patients are shown in Table 1. Matched pairs were from case 1 and 2 to case 15 and 16. Case 17 and 18 were in the low aggressive group and had no pairs, respectively.

Control Gene Selection

We found 25 articles about lung cancer studies published from April 2003 through March 2004, in which target gene expression levels were quantified relative to a housekeeping gene. β actin (ACTB) was used as a control gene in nine articles (Schneider *et al*., 2004; Marchetti *et al*., 2004; Rosell *et al*., 2004; Marchevsky *et al*., 2004; Takahashi *et al*., 2004; Inui *et al*., 2003; Senchenko *et al*., 2003; Harden *et al*., 2003; Inui *et al*., 2003a, b), glyceraldehyde-3-phosphate dehydrogenase (GAPDH) in

Table 1: Patients' characteristics

Case	SEX	Age	Smoking status	*Stage	Vital status	**Survival time (months)	***Group
1	M	63	Former	IA	Alive	72	Low aggressive
2	M	66	Former	IA	Dead	41	High aggressive
3	M	70	Former	IB	Alive	62	Low aggressive
4	M	73	Former	IB	Dead	21	High aggressive
5	M	80	Former	IB	Alive	68	Low aggressive
6	M	76	Former	IB	Dead	21	High aggressive
7	M	78	Former	IB	Alive	60	Low aggressive
8	M	75	Former	IB	Dead	27	High aggressive
9	M	65	Former	IA	Alive	62	Low aggressive
10	M	64	Former	IA	Dead	9	High aggressive
11	F	62	Former	IB	Alive	52	Low aggressive
12	F	66	Former	IB	Dead	23	High aggressive
13	F	62	Former	IA	Alive	54	Low aggressive
14	F	59	Former	IA	Dead	10	High aggressive
15	F	70	Former	IA	Alive	55	Low aggressive
16	F	78	Former	IA	Dead	20	High aggressive
17	F	73	Current	IB	Alive	66	Low aggressive
18	F	76	Current	IB	Alive	72	Low aggressive

*Stage IA = pT1N0M0 and IB = pT2N0M0. **Survival time is calculated from the day of surgery. ***Low aggressive was defined as surviving more than 4 years and high aggressive was defined as surviving less than 2 years. Case 2 relapsed a year after surgery and received chemoradiotherapy and therefore is considered to be in the high aggressive group despite the survival was longer than 2 years. Case 1 and 2 to Case 15 and 16 were one-to-one matched pairs on age, gender, smoking status, tumor size, tumor histology (all were squamous cell carcinomas), tumor stage and treatment (all underwent complete resections)

seven (Whiteside *et al*., 2004; Shi *et al*., 2004; Gu *et al*., 2004; Cho *et al*., 2004; Xue Jun *et al*., 2003; Kawabata *et al*., 2003; Hu *et al*., 2003), TATA box binding protein (TBP) in three (Chen *et al*., 2004; Agathanggelou *et al*., 2003; Kao *et al*., 2003) and β2 microglobulin (B2M) (Falleni *et al*., 2003), β globin (Wu *et al*., 2003), β glucoronidase (Player *et al*., 2003), hypoxanthin phosphoriboxyl transferase (Erovic *et al*., 2003), porphobilinogen deaminase (Steenbergen *et al*., 2004) and type II collagen (Marchevsky *et al*., 2004) in one each. However, justification for the choice of a control gene was not provided in any of the 25 articles. Thus, we selected the four most commonly used housekeeping genes (ACTB, GAPDH, TBP, and B2M) to compare the variation of expression levels among samples and to identify the most suitable gene or gene combinations to be used as normalization.

Test Gene Selection

Gene-expression levels of four test genes, TP53, laminin γ2 (LAMC2), TGFα (TGFA), and cathespin L2 (CTSL2) were examined by qRT-PCR with normalization to the control genes. Test genes were randomly selected from a list of genes with significantly higher expression on the high aggressive disease group as determined by data generated from a DNA microarray experiment (Yang *et al*., 2004).

RNA Extraction

Thirty mm^3 of each tissue were sectioned at 20 or 35 μm thickness, collected in buffer RLT (Qiagen, Valencia, CA) supplemented with β-mercaptoethanol and homogenized using PT 1200C rotor/stator (Kinematica AG, Luzern, Switzerland). Total RNA was isolated using the RNeasy kit (Qiagen, Valencia, CA) following the manufacturer's specifications. Quality and quantity of RNA samples were controlled by spectrophotometry and the Agilent 2100 Bioanalyzer. Hematoxylin-Eosin stained frozen tissue sections were made to confirm that more than 90% of the tested tissue sample consisted of tumor cells (Yang *et al*., 2004).

Real Time PCR

Total RNA was treated with DNase I (Invitrogen Life Technologies, Inc., Carlsbad, CA). Oligo-dT primers and SuperScript II Reverse Transcriptase (Invitrogen Life Technologies, Inc.) were used in reverse transcription, following the manufacturer's specifications. The qRT-PCR was performed using a template from the reverse transcription reaction by the ABI Prism 7700 Sequence Detection System (Applied Biosystems, Foster City, CA) with SYBR Green I Mastermix or TaqMan Universal Mastermix (Applied Biosystems). Each PCR reaction mixture contained 1 μL of template (50 ng RNA input) in a final volume of 25 μL with final concentrations of 1×SYBR Green I or 1×TaqMan Mastermix. Thermal cycling conditions comprised an initial incubation at 50°C for two min, DNA polymerase activation at 95°C for 10 min and 40 cycles of denaturation at 95°C for 15 sec and annealing at 60°C for one min. Each measurement was performed in duplicate and the threshold cycle (Ct) was determined.

Sequences of primers and dual labeled probes were as follows:
GAPDH (107 bp), forward: 5'-CATCCATGACAACTTTGGTATCGT-3', reverse: 5'-CCATCACGCCACAGTTTCC-3', TaqMan probe: 5'-AAGGACTCATGACCACAGTCCATGCCA-3'; TBP (110 bp), forward: 5'-ACCAGGTGATGCCCTTCTGT-3', reverse: 5'-GCAGCACGGTATGAGCAACT-3'; ACTB (130 bp), forward: 5'-TCCTCTCCCAAGTCCACACA-3', reverse: 5'-GGCACGAAGGCTCATCATT-3'; B2M (99 bp), forward: 5'-AGGAGGGCTGGCAACTTAGA-3', reverse: 5'-GCTTTGAGTGCAAGAGATTGAAGA-3'; TP53 (139 bp), forward: 5'-GCTGGCATTTGCACCTACCT-3', reverse: 5'- GGGAGAGGGAGAGGGAACA-3'; LAMC2

(98 bp), forward: 5'-TCGGCAGGTTCCCTTACATT-', reverse: 5'-TTGGCCCACAGACCTTTACC-3'; TGFA (91 bp), forward: 5'- CCTGTGTTTCCTTGCCCTTT-3', reverse: 5'- CCAGGTGTGTGGCAGCTTAC-3'; CTSL2 (117 bp), forward: 5'-GCCATTCGTCCTTCCAGTTC-3', reverse: 5'- TCGAATTTGCTCCTTCAAAGC-3'.

Microarray

The same samples were used for the microarray analysis as described elsewhere (Yang *et al*., 2004). Briefly, hybridization, washes, and scanning were performed following the manufacturer's protocols (Affymetrix, Santa Clara, CA). Two-GeneChip®-array sets from Affymetrix (HG-U133) were used, containing 22,283 (chip A) and 22,645 (chip B) probe sets derived from approximately 33,000 well-substantiated human genes. Reproducibility of the chip-set was examined by repeated tests on the same sample. The correlation coefficient is 0.99 for the two duplicate chip sets (from two different manufacturer lots) tested on the same sample three weeks apart.

The microarray data analysis has been reported previously (Yang *et al*., 2004). In brief, Affymetrix Microarray Analysis Suite, version 5 (MAS5), was used to process scanned chip images. An expression value (signal) and a detection call for each probe set were calculated according to the One-Step Tukey's Biweight Estimate and detection algorithm (Affymetrix, GeneChip Expression Analysis--Data Analysis Fundamentals, Part No. 701190 Rev. 2, 2002). Each chip was globally scaled with a target value of 1,500.

Data Analysis

For quantitative analysis of qRT-PCR, the amplification process was monitored and each amplification curve consisted of at least three distinct phases (initial lag phase, exponential phase, and plateau phase). The template copy number could be estimated with greater precision from the number of cycles needed for the signal to reach an arbitrary threshold. The threshold must intersect the signal curve in its exponential phase, in which the signal elevation correlates with product accumulation. The threshold cycle (Ct) is the intersection point. The Delta Ct is the difference in threshold cycles between the target and reference. Normalized amount of target (X) is $X = K (1+E)^{-\text{delta Ct}}$, where K is a constant and E is the efficiency of amplification. The value of the delta Ct provides a simple tool to estimate relative changes (Wilhelm *et al*., 2003).

The coefficient of variation (CV) and the standard deviation (SD) were used as statistical measures of the variation of a set of data from its mean. The CV is the ratio of the SD to the average. When the SD changes with the average, the CV is more useful to summarize the variation as long as the average is greater than zero. Correlation coefficients were determined by rank correlation using a nonparametric Spearman's test since the number of samples was relatively small and the relationship between the Ct values and microarray signals were non-linear. The correlation was considered significant when the p-value was less than 0.05.

Results

Specificity of the PCR amplification for each primer pair was determined using agarose gel electrophoresis and confirmed to give a single amplified band of a corresponding size. The standard curves for GAPDH ($y = -3.471x+19.94$, $r^2 = 0.955$), TBP ($y = -3.636x+25.935$, $r^2 = 0.993$), ACTB ($y = -3.239x+16.422$, $r^2 = 0.953$), and B2M ($y = -3.405x+15.772$, $r^2 = 0.989$) were almost parallel, which indicates the efficiencies of the amplifications are approximately equal among the four control genes (Hamalainen *et al*., 2001).

The CV of each control gene was 0.090 for GAPDH, 0.048 for TBP, 0.065 for ACTB, and 0.062 for B2M and the CV values of the four genes were small, with the smallest being for TBP (Fig. 1).

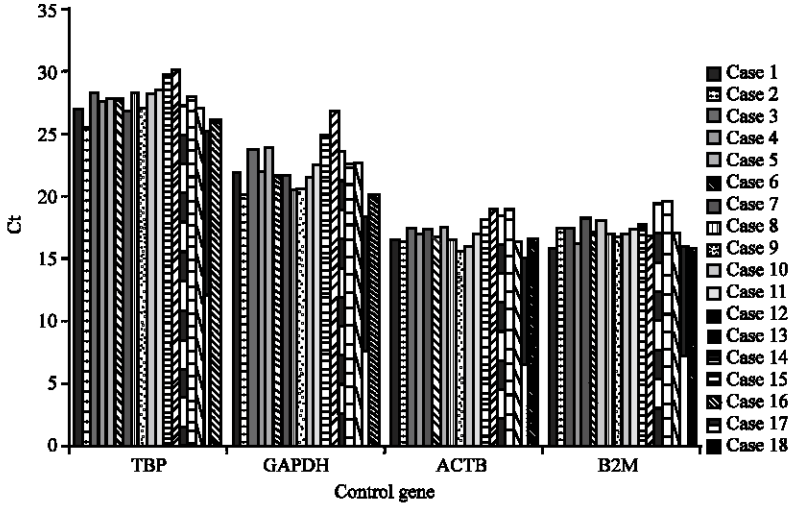

Fig. 1: Ct values of 18 cases with four control genes (TBP, GAPDH, ACTB, and B2M). Each bar represents each sample. Ct: threshold cycle, SD: standard deviation, CV: coefficient of variation

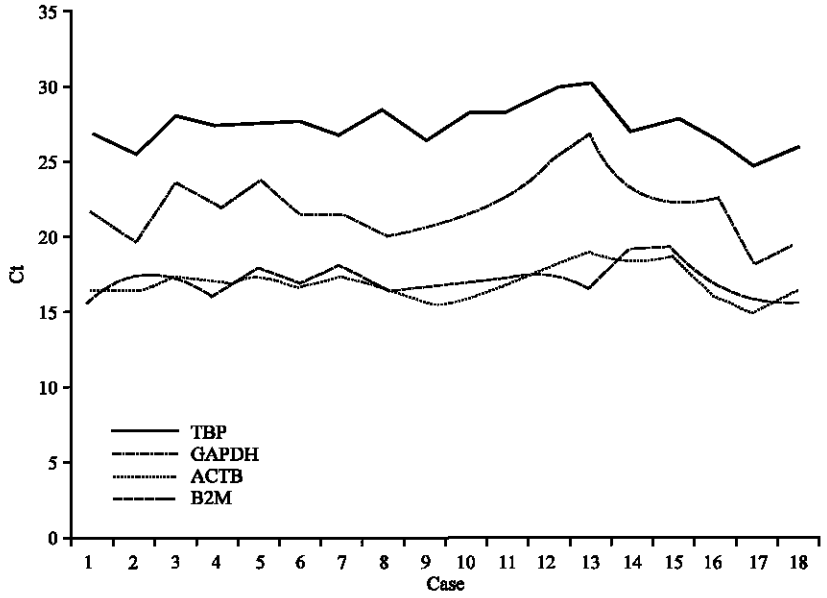

Fig. 2: Ct values of four control genes (TBP, GAPDH, ACTB, and B2M) with 18 cases. Each line represents each control gene. Three lines (TBP, ACTB, and B2M) were nearly parallel and flat and changed within 5 Ct. Ct: threshold cycle

Figure 2 shows variations in control gene expression. Each line represents the data from one control gene. Connecting lines of TBP, ACTB, and B2M are nearly parallel and flat within a 5 Ct difference. This indicates that variations in Ct values could be reflected as inevitable differences like variations in the initial amount of RNA and the efficiency of the cDNA synthesis and the expression levels of these genes (TBP, ACTB and B2M) are relatively constant through all samples. If there are substantial differences of gene expression levels among comparison groups, the connecting lines may not run

Table 2: Spearman's correlation coefficient between the microarray data and *delta Ct values of four test genes normalized by four control genes

	Test gene			
Control gene	TP53	TGFA	LAMC2	CTSL2
GAPDH	-0.52**	-0.43	-0.14	-0.50**
TBP	-0.61**	-0.57**	-0.12	-0.56**
ACTB	-0.62**	-0.48**	-0.08	-0.57**
B2M	-0.62**	-0.52**	-0.15	-0.58**
Geometric mean of TBP and B2M	-0.64**	-0.54**	-0.16	-0.57**

*The Delta Ct was the Ct value of the test gene minus that of the control gene. **$p<0.05$ by Spearman's test. A lower Delta Ct value means a higher gene expression level and a higher value of microarray data means a higher gene expression level. Accordingly, the negative correlation means the same direction of gene expression change. The Delta Ct for TP53 normalized by any control gene showed a statistically significant correlation to the microarray data. The Delta Ct for TGFAα normalized by any control gene but GAPDH and that for CTSL2 normalized by any control gene also showed significant correlation. However, the Delta Ct for LAMC2 normalized by any control gene showed no significant correlation

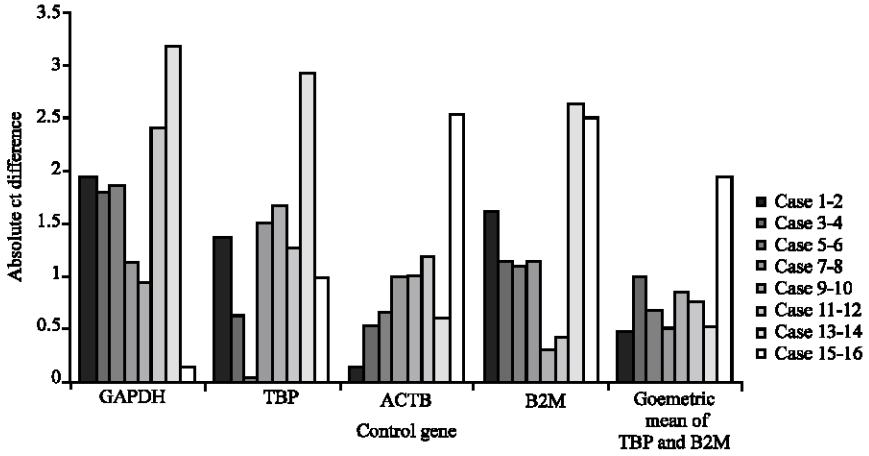

Fig. 3: Difference of Ct values for four control genes (GAPDH, TBP, ACTB, and B2M) between the matched pairs. Mean difference was less than 2 Ct for all control genes. ACTB showed the smallest mean and SD, both of which were less than 1 Ct. ACTB showed the smallest 95% CI, as well. Ct: threshold cycle, SD: standard deviation, 95% CI: 95% confidential interval

parallel. For example, if one gene expressed higher and one gene expressed lower in the high aggressive group and yet another gene expressed constant in both groups, the connected lines of the three genes would not have run parallel or probably have crossed each other.

As our samples were one-to-one matched except for one variable, survival time, control genes are supposedly expressed at the same level between the pairs unless they are associated with survival. Figure 3 shows the difference of gene expression levels (i.e., the difference of the Ct value) between the matched pairs. Since ideally the Ct difference between the pairs is equal or close to zero, the best control gene should have the smaller mean of the Ct difference with a smaller variation. The gene expression differences between pairs had the smallest mean, SD and 95% confidential interval (CI) when normalized by ACTB.

The correlation between qRT-PCR data and microarray data was examined. In general, the microarray data is normalized by multiple control genes and its normalization is thought to be reliable, so the normalized Ct values (i.e., Delta Ct values) should be comparable to the microarray data. Table 2 shows that Delta Ct values of the all test genes except LAMC2 were well correlated with the microarray data.

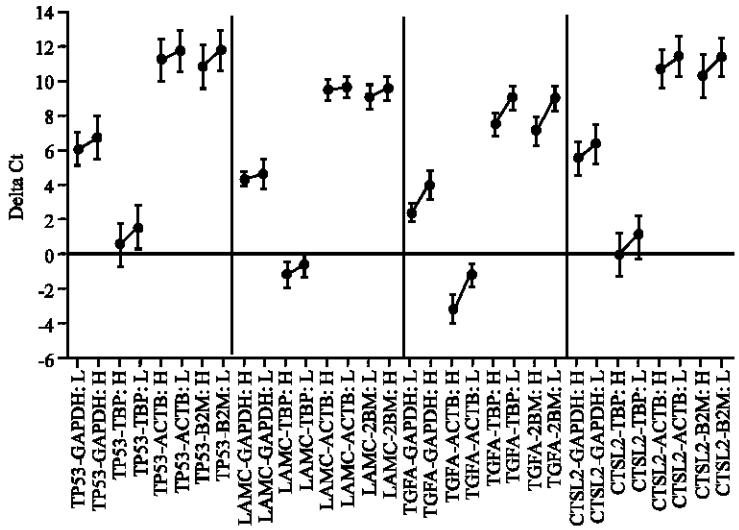

Fig. 4: Delta Ct values of four test genes in the low and high aggressive groups normalized by four control genes. The error bar represents the mean±standard deviation. All test genes but LAMC2 showed lower Delta Ct values in the high aggressive group. LAMC2 normalized by GAPDH and ACTB showed nearly equal values between the two groups. H: High aggressive group, L: Low aggressive group

Since the four test gene expression levels were significantly higher for the high aggressive group than the low aggressive group in the microarray data, the Delta Ct of the four genes in the high aggressive group is bound to show a smaller value than that of the low aggressive group if the normalization is conducted appropriately. Our results showed this was true regardless of which control gene was used for normalization (Fig. 4). However, LAMC2 normalized by GAPDH or ACTB showed that the values of the Delta Ct between both groups were almost equal, emphasizing the importance of appropriately selecting control genes in order to detect a subtle difference of expression level.

Discussion

Accurate normalization of gene-expression levels in qRT-PCR analysis is very important for reliable results, especially when the relationship between the clinical findings and the subtle gene-expression level difference is studied. Several systematic studies of normalization procedures have been reported (Tricarico *et al*., 2002; Vandesompele *et al*., 2002; Hamalainen *et al*., 2001; Liu *et al*., 2002) and extensive analyses in the cancer field have been conducted to evaluate different control genes and to select a suitable control gene for the normalization of expression data (Lossos *et al*., 2003; Beillard *et al*., 2003; Lupberger *et al*., 2002). In contrast, evaluation of control genes has never been investigated for lung cancer. As a standard approach, experiments use only one control gene for sample normalization. This is usually done because of convenience, availability in the lab or because it was done in a similar experiment but, without any validation for selection. No recent published reports were found for control genes in lung cancer, in which the reason for selection of the control gene was stated. In this study, four housekeeping genes were selected as control genes because they have been most frequently used for normalization of qRT-PCR data in lung cancer reports published from April 2003 to March 2004. We compared the Ct values of the control genes with each other and also with their microarray data.

Although GAPDH varied more than 5 Ct, 32 fold in Fig. 1, similar variation has also been found in other publications evaluating control genes. Ct values of B2M varied more than 5 Ct among non-Hodgkin's lymphomas (Lossos *et al*., 2003) and those of TBP varied more than 5 Ct among mononuclear cells (Beillard *et al*., 2003). In comparison with the variation of the Ct value (Fig. 1-3), TBP, B2M, and ACTB were considered to be more constant than GAPDH. However, there is no single control gene which was constantly expressed in all analyses, although no control gene showed the outlier. GAPDH was not the best choice as a control in any of the analyses reported here.

Moreover, the Delta Cts of the three test genes that were normalized with the four control genes were correlated to the microarray data (Table 2 and Fig. 4). Even when a test gene expression level showed a subtle difference like LAMC2 (the mean Delta Ct difference was less than one Ct), TBP, and B2M showed the same tendency of expression level as in the microarray data (Fig. 4). Its tendency, however, showed no statistical significance, which was due to the small number of samples and the small difference of the mean expression level between the two groups. Only a 2.8 Ct difference can be significantly detected with nine samples in each group if it is assumed that the SD is 2 and the type I error and power are 5 and 80%, respectively. More than 64 samples in each group are needed to detect a 1 Ct difference with 5% type I error and 80% power. Therefore, a set of TBP and B2M should be considered as appropriate control genes for normalization of the qRT-PCR data for lung cancer studies.

Beillard *et al*. (2003) reported that the selected control genes should fulfill the absence of pseudogenes and high or medium expression. Since the Ct values of three of the control genes reported here (TBP, B2M, and ACTB) were between 15 and 30, these gene-expression levels can be considered not to be very high or low. TBP and B2M have no pseudogenes (Beillard *et al*., 2003; Vandesompele *et al*., 2002). Ideally, the control gene expression level would not change over conditions. The gene expression levels of some frequently used control genes have been reported to change under certain conditions. Lossos *et al*. (2002) reported that alterations in cell morphology associated with tumor progression or transformation were related to changes in ACTB expression. Aviles *et al*. (1992) reported how B2M expression varied in different lymphomas. GAPDH expression has been reported to exhibit marked variability among normal colon epithelium (Bustin, 2000), prostate cancer (Ripple *et al*., 1995), during the cell cycle (Mansur *et al*., 1993), and to correlate to VEGF165/189 expression level (Tricarico *et al*., 2002). Hypoxia is a common condition in solid tumors and it is known to stimulate GAPDH expression as well as VEGF expression (Graven *et al*., 1994; Zhong *et al*., 1999). As smoking and COPD, which are risk factors of lung cancer, are related to hypoxia, the relationship between hypoxia and lung cancer may be more significant than between hypoxia and other solid tumors. Therefore, it is important that normalization using GAPDH for lung cancer studies is carefully considered. Moreover, Vandesompele *et al*. (2002) reported the possible errors related to the common practice of using only one housekeeping gene for normalization and recommended normalization by the geometric mean of multiple housekeeping genes. In the present study, the CV of the geometric mean of TBP and B2M was 0.043, of TBP and ACTB was 0.058, of B2M and ACTB was 0.058, and of TBP, ACTB and B2M was 0.048, with the CV of the geometric mean of TBP and B2M being the smallest value (Table 3). In Table 2, the geometric mean of TBP and B2M showed good correlation to the microarray data. Moreover, in Fig. 3, the geometric mean of TBP and B2M showed the smallest mean of the Ct difference with the smallest variation. Consequently, normalization of the real time PCR data by the geometric mean of TBP and B2M could be most appropriate.

Table 3: Coefficient of variation (CV) for geometric mean of Ct values Of TBP, ACTB and B2M

	TBP	ACTB	B2M	TBP&B2M	TBP&ACTB	B2M&ACTB	TBP&ACTB&B2M
CV	0.048	0.066	0.062	0.043	0.058	0.058	0.048

Ct: Threshold cycle, CV: Coefficient of variation, The geometric mean of Ct values of TBP and B2M showed the least CV

Since we examined only four housekeeping genes, examination of other housekeeping genes might show more constant expression level. Some investigators have utilized microarray data for choosing control genes. Wikman *et al*. (2002) chose phopholipase 2A as a reference due to its constant expression at a moderate level in the cDNA array data for lung adenocarcinomas. However, they did not confirm its stability using real time PCR. Hamalainene *et al*. (2001) selected control genes based on the results of an oligonucleotide microarray for T helper cells and evaluated them using real time PCR.

We examined the validity of the four control genes used frequently for normalization of qRT-PCR data in lung cancer. Our results show that among the four housekeeping genes examined, there is no single control gene that expresses consistently in all analyses. If the microarray data for the same samples can be utilized as in the present study, we may select control genes of qRT-PCR according to the microarray data. But such a situation must be rare in the qRT-PCR study for lung cancer. Based on the data presented in this study, it is suggested that normalization using multiple control genes (geometric mean of TBP and B2M) is more appropriate than a single control gene.

This study has been the first report on the evaluation of control genes for qRT-PCR study in lung cancer, although a relatively small number of samples were used. In leukemia studies, multi-center trials have been conducted in order to evaluate candidate control genes using real time PCR (Beillard *et al*., 2003). Such trials should be planned for lung cancer studies in the near future.

Acknowledgments

The authors would like to thank Ms. Susan Ernst for her technical assistance with the manuscript and Drs. Farhad Kosari and Marie Kube for their helpful advice on laboratory experiments. This research was partially supported by: NIH Grant R01 CA 84354 (Yang), Mayo Foundation Funds, and Tohoku University, Department of Thoracic Surgery, Institute of Development, Aging and Cancer, Japan.

References

Agathanggelou, A., I. Bieche, J. Ahmed-Choudhury, B. Nicke and R. Dammann *et al*., 2003. Identification of novel gene expression targets for the Ras association domain family 1 (RASSF1A) tumor suppressor gene in non-small cell lung cancer and neuroblastoma. Cancer Res., 63: 5344-5351.

Aviles, A., G. Zepeda, J.C. Diaz-Maqueo, L. Rodriguez, R. Guzman, E.L. Garcia and A. Talavera, 1992. Beta 2 microglobulin level as an indicator of prognosis in diffuse large cell lymphoma. Leuk. Lymphoma, 7: 135-138.

Beillard, E., N. Pallisgaard, V.H. van der Velden, W. Bi, R. Dee and E.J. van der Schoot *et al*., 2003. Evaluation of candidate control genes for diagnosis and residual disease detection in leukemic patients using 'real-time' quantitative reverse-transcriptase polymerase chain reaction (RQ-PCR) - a Europe against cancer program. Leukemia, 17: 2474-2486.

Bustin, S.A., 2000. Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. J. Mol. Endocrinol., 25: 169-193.

Chen, H.W., S.L. Yu, J.J. Chen, H.N. Li and Y.C. Lin *et al*., 2004. Anti-invasive gene expression profile of curcumin in lung adenocarcinoma based on a high throughput microarray analysis. Mol. Pharmacol., 65: 99-110.

Cho, N.H., K.P. Hong, S.H. Hong, S. Kang, K.Y. Chung and S.H. Cho, 2004. MMP expression profiling in recurred stage IB lung cancer. Oncogene, 23: 845-851.

Erovic, B.M., J. Pammer, D. Hollemann, M. Woegerbauer, S. Geleff, M.B. Fischer, M. Burian, F. Frommlet and C. Neuchrist, 2003. Motility-related protein-1/CD9 expression in head and neck squamous cell carcinoma. Head Neck, 25: 848-857.

Falleni, M., C. Pellegrini, A. Marchetti, B. Oprandi and F. Buttitta, 2003. Survivin gene expression in early-stage non-small cell lung cancer. J. Pathol., 200: 620-626.

Graven, K.K., R.F. Troxler, H. Kornfeld, M.V. Panchenko and H.W. Farber, 1994. Regulation of endothelial cell glyceraldehyde-3-phosphate dehydrogenase expression by hypoxia. J. Biol. Chem., 269: 24446-24453.

Gu, J., L. Zhang, S.G. Swisher, J. Liu, J.A. Roth and B. Fang, 2004. Induction of p53-regulated genes in lung cancer cells: Implications of the mechanism for adenoviral p53-mediated apoptosis. Oncogene, 23: 1300-1307.

Hamalainen, H.K., J.C. Tubman, S. Vikman, T. Kyrola, E. Ylikoski, J.A. Warrington and R. Lahesmaa, 2001. Identification and validation of endogenous reference genes for expression profiling of T helper cell differentiation by quantitative real-time RT-PCR. Anal. Biochem., 299: 63-70.

Harden, S.V., Y. Tokumaru, W.H. Westra, S. Goodman, S.A. Ahrendt, S.C. Yang and D. Sidransky, 2003. Gene promoter hypermethylation in tumors and lymph nodes of stage I lung cancer patients. Clin. Cancer Res., 9: 1370-1375.

Hu, Y., G. Cherton-Horvat, V. Dragowska, S. Baird, R.G. Korneluk, J.P. Durkin, L.D. Mayer and E.C. LaCasse, 2003. Antisense oligonucleotides targeting XIAP induce apoptosis and enhance chemotherapeutic activity against human lung cancer cells *in vitro* and *in vivo*. Clin. Cancer Res., 9: 2826-2836.

Inui, N., K. Kitagawa, S. Miwa, T. Hattori, K. Chida, H. Nakamura and M. Kitagawa, 2003a. High expression of Cks1 in human non-small cell lung carcinomas. Biochem. Biophys. Res. Commun., 303: 978-984.

Inui, N., S. Sasaki, T. Suda, K. Chida and H. Nakamura, 2003b. The loss of retinoic acid receptor alpha, beta and alcohol dehydrogenase3 expression in non-small cell lung cancer. Respirology, 8: 302-309.

Kao, Y.R., J.Y. Shih, W.C. Wen, Y.P. Ko, B.M. Chen, Y.L. Chan, Y.W. Chu, P.C. Yang, C.W. Wu and S.R. Roffler, 2003. Tumor-associated antigen L6 and the invasion of human lung cancer cells. Clin. Cancer Res., 9: 2807-2816.

Kawabata, S., M. Oka, H. Soda, K. Shiozawa and K. Nakatomi *et al*., 2003. Expression and functional analyses of breast cancer resistance protein in lung cancer. Clin. Cancer Res., 9: 3052-3057.

Lion, T., 2001. Current recommendations for positive controls in RT-PCR assays. Leukemia, 15: 1033-1037.

Liu, W. and D.A. Saint, 2002. A new quantitative method of real time reverse transcription polymerase chain reaction assay based on simulation of polymerase chain reaction kinetics. Anal. Biochem., 302: 52-59.

Lossos, I.S., A.A. Alizadeh, M. Diehn, R. Warnke and Y. Thorstenson, 2002. Transformation of follicular lymphoma to diffuse large-cell lymphoma: Alternative patterns with increased or decreased expression of c-myc and its regulated genes. Proc. Natl. Acad. Sci. USA., 99: 8886-8891.

Lossos, I.S., D.K. Czerwinski, M.A. Wechser and R. Levy, 2003. Optimization of quantitative real-time RT-PCR parameters for the study of lymphoid malignancies. Leukemia, 17: 789-795.

Lupberger, J., K.A. Kreuzer, G. Baskaynak, U.R. Peters, P. le Coutre and C.A. Schmidt, 2002. Quantitative analysis of beta-actin, beta-2-microglobulin and porphobilinogen deaminase mRNA and their comparison as control transcripts for RT-PCR. Mol. Cell Probes, 16: 25-30.

Mansur, N.R., K. Meyer-Siegler, J.C. Wurzer and M.A. Sirover, 1993. Cell cycle regulation of the glyceraldehyde-3-phosphate dehydrogenase/uracil DNA glycosylase gene in normal human cells. Nucleic Acids Res., 21: 993-998.

Marchetti, A., F. Barassi, C. Martella, A. Chella, S. Salvatore, A. Castrataro, F. Mucilli, R. Sacco and F. Buttitta, 2004. Down regulation of high in normal-1 (HIN-1) is a frequent event in stage I non-small cell lung cancer and correlates with poor clinical outcome. Clin. Cancer Res., 10: 1338-1343.

Marchevsky, A.M., J.A. Tsou and I.A. Laird-Offringa, 2004. Classification of individual lung cancer cell lines based on DNA methylation markers: Use of linear discriminant analysis and artificial neural networks. J. Mol. Diagn., 6: 28-36.

Player, A., J. Gillespie, T. Fujii, J. Fukuoka, T. Dracheva, D. Meerzaman, K.M. Hong, J. Curran, G. Attoh, W. Travis and J. Jen, 2003. Identification of TDE2 gene and its expression in non-small cell lung cancer. Int. J. Cancer, 107: 238-243.

Ripple, M.O. and G. Wilding, 1995. Alteration of glyceraldehyde-3-phosphate dehydrogenase activity and messenger RNA content by androgen in human prostate carcinoma cells. Cancer Res., 55: 4234-4236.

Rosell, R., K.D. Danenberg, V. Alberola, G. Bepler, J.J. Sanchez, C. Camps, M. Provencio, D. Isla, M. Taron, P. Diz and A. Artal, 2004. Ribonucleotide reductase messenger RNA expression and survival in gemcitabine/cisplatin-treated advanced non-small cell lung cancer patients. Clin. Cancer Res., 10: 1318-1325.

Schneider, S., J. Yochim, J. Brabender, K. Uchida, K.D. Danenberg, R. Metzger, P.M. Schneider, D. Salonga, A.H. Holscher and P.V. Danenberg, 2004. Osteopontin but not osteonectin messenger RNA expression is a prognostic marker in curatively resected non-small cell lung cancer. Clin. Cancer Res., 10: 1588-1596.

Senchenko, V., J. Liu, E. Braga, N. Mazurenko, W. Loginov, Y. Seryogin, I. Bazov, A. Protopopov, F.L. Kisseljov, V. Kashuba, M.I. Lerman, G. Klein and E.R. Zabarovsky, 2003. Deletion mapping using quantitative real-time PCR identifies two distinct 3p21.3 regions affected in most cervical carcinomas. Oncogene, 22: 2984-2992.

Shi, W., H. Chen, J. Sun, C. Chen, J. Zhao, Y.L. Wang, K.D. Anderson and D. Warburton, 2004. Overexpression of Smurf1 negatively regulates mouse embryonic lung branching morphogenesis by specifically reducing Smad1 and Smad5 proteins. Am. J. Physiol. Lung Cell Mol. Physiol., 286: L293-300.

Steenbergen, R.D., D. Kramer, B.J. Braakhuis, P.L. Stern, R.H. Verheijen, C.J. Meijer and P.J. Snijders, 2004. TSLC1 gene silencing in cervical cancer cell lines and cervical neoplasia. J. Natl. Cancer Inst., 96: 294-305.

Takahashi, Y., J. Hamada, K. Murakawa, M. Takada, M. Tada, I. Nogami, N. Hayashi, S. Nakamori, M. Monden, M. Miyamoto, H. Katoh and T. Moriuchi, 2004. Expression profiles of 39 HOX genes in normal human adult organs and anaplastic thyroid cancer cell lines by quantitative real-time RT-PCR system. Exp. Cell Res., 293: 144-153.

Tricarico, C., P. Pinzani, S. Bianchi, M. Paglierani, V. Distante, M. Pazzagli, S.A. Bustin and C. Orlando, 2002. Quantitative real-time reverse transcription polymerase chain reaction: Normalization to rRNA or single housekeeping genes is inappropriate for human tissue biopsies. Anal. Biochem., 309: 293-300.

Vandesompele, J., De K. Preter, F. Pattyn, B. Poppe, N. Van Roy, A. De Paepe and F. Speleman, 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome. Biol., 3: RESEARCH0034.

Whiteside, M.A., D.T. Chen, R.A. Desmond, S.A. Abdulkadir and G.L. Johanning, 2004. A novel time-course cDNA microarray analysis method identifies genes associated with the development of cisplatin resistance. Oncogene, 23: 744-752.

Wikman, H., E. Kettunen, J.K. Seppanen, A. Karjalainen, J. Hollmen, S. Anttila and S. Knuutila, 2002. Identification of differentially expressed genes in pulmonary adenocarcinoma by using cDNA array. Oncogene, 21: 5804-5813.

Wilhelm, J. and A. Pingoud, 2003. Real-time polymerase chain reaction. Chembiochem., 4: 1120-1128.

Wu, Y., B. Han, H. Sheng, M. Lin, P.A. Moore, J. Zhang and J. Wu, 2003. Clinical significance of detecting elevated serum DcR3/TR6/M68 in malignant tumor patients. Int. J. Cancer, 105: 724-732.

Xue Jun, H., A. Gemma, Y. Hosoya, K. Matsuda, M. Nara, Y. Hosomi, T. Okano, F. Kurimoto, M. Seike, K. Takenaka, A. Yoshimura, M. Toyota and S. Kudoh, 2003. Reduced transcription of the RB2/p130 gene in human lung cancer. Mol. Carcinog., 38: 124-129.

Yang, P., Z. Sun, M. Aubry, F. Kosari, W. Bamlet, C. Endo, R. Molina and G. Vasmatzis, 2004. Study design considerations in clinical outcome research of lung cancer using microarray analysis. Lung Cancer, 46: 215-226.

Zhong, H. and J.W. Simons, 1999. Direct comparison of GAPDH, beta-actin, cyclophilin and 28S rRNA as internal standards for quantifying RNA levels under hypoxia. Biochem. Biophys. Res. Commun., 259: 523-526.