Am J Physiol Gastrointest Liver Physiol 290: G1067–G1074, 2006.
First published January 6, 2006; doi:10.1152/ajpgi.00234.2005.

Innovative Methodology

American Journal of Physiology – Gastrointestinal and Liver Physiology

Normalizing genes for quantitative RT-PCR in differentiating human intestinal epithelial cells and adenocarcinomas of the colon

Anders Bondo Dydensborg, Elizabeth Herring, Joëlle Auclair, Eric Tremblay, and Jean-Francois Beaulieu

Canadian Institutes of Health Research Group in Functional Development and Physiopathology of the Digestive Tract, Département d'Anatomie et de Biologie Cellulaire, Faculté de Médecine, Université de Sherbrooke, Sherbrooke, Québec, Canada

Submitted 23 May 2005; accepted in final form 26 December 2005

Dydensborg, Anders Bondo, Elizabeth Herring, Joëlle Auclair, Eric Tremblay, and Jean-Francois Beaulieu. Normalizing genes for quantitative RT-PCR in differentiating human intestinal epithelial cells and adenocarcinomas of the colon. *Am J Physiol Gastrointest Liver Physiol* 290: G1067–G1074, 2006. First published January 6, 2006; doi:10.1152/ajpgi.00234.2005.—As for other mRNA measurement methods, quantitative RT-PCR results need to be normalized relative to stably expressed genes. Widely used normalizing genes include β-actin and glyceraldehyde-3-phosphate dehydrogenase. It has, however, become clear that these and other normalizing genes can display modulated patterns of expression across tissue types and during complex cellular processes such as cell differentiation and cancer progression. Our objective was to set the basis for identifying normalizing genes that displayed stable expression during enterocytic differentiation and between healthy tissue and adenocarcinomas of the human colon. We thus identified novel potential normalizing genes using previously generated cDNA microarray data and examined the alterations of expression of two of these genes as well as seven commonly used normalizing genes during the enterocytic differentiation process and between matched pairs of resection margins and primary carcinomas of the human colon using real-time RT-PCR. We found that ribosomal phosphoprotein P0 was particularly stable in all intestinal epithelial cell extracts, thereby representing a particularly robust housekeeping reference gene for the assessment of gene expression during the human enterocytic differentiation process. On the other hand, β-2-microglobulin generated the best score as a normalizing gene for comparing human colon primary carcinomas with their corresponding normal mucosa of the resection margin, although others were found to represent acceptable alternatives. In conclusion, we identified and characterized specific normalizing genes that should significantly improve quantitative mRNA studies related to both the differentiation process of the human intestinal epithelium and adenocarcinomas of the human colon. This approach should also be useful to validate normalizing genes in other intestinal contexts.

mRNA analysis; intestinal cell differentiation; colon cancer cells

REAL-TIME QUANTITATIVE RT-PCR (qRT-PCR) is rapidly becoming the method of choice for gene expression assessments because it requires relatively little biological material and is efficient. However, there are substantial drawbacks to the use of this technique for accurate quantification of transcript levels. Indeed, the efficiency of the RT reaction is very sensitive to the quality of the input mRNA (15), whereas the PCR step greatly amplifies the transcript and thus intrinsic differences in the quality of the input cDNA are correspondingly greatly amplified (8). The common way to deal with these difficulties is to express the transcript level of the gene of interest normalized to another gene commonly termed the housekeeping or normalizing gene. This method is based on the assumption that the expression level of the normalizing gene does not change from sample to sample.

The differentiation of a tissue (33) or the formation of carcinomas (7) are complex cellular processes that contribute to the difficulty of assessing gene expression changes using RT-PCR. Indeed, inherent difficulties in identifying a "gold standard" normalizing gene for cDNA input in RT-PCR analysis experiments have become apparent from reports of variation of expression across tissue types (26), developmental stages (9), and cancer types (1, 11). One proposed alternative for addressing this problem is to use a weighted average of several normalizing genes (24, 32). However, this option suffers from the drawback that it only normalizes a gene of interest against the average variation of many normalizing genes. While the ubiquitous normalizing gene may not exist, another more promising strategy is to identify a single normalizing gene for a single tissue type (27), pathology (12), or experimental design (5, 16, 18, 28).

The aim of the present study was to identify and validate such potential normalizing genes to be used for studies of proliferating and differentiating human intestinal epithelial cells as well as human colon adenocarcinomas. In the present study, we identified novel candidate normalizing genes using cDNA microarray data from human intestinal epithelial cell models reflecting different stages of enterocytic differentiation and selected two of these for validation along with seven commonly used normalizing genes. We compared the variation of expression of these genes by validating their stability during enterocytic differentiation using real-time qRT-PCR on stringently purified and quantified RNA samples. We found that the ribosomal phosphoprotein P0 (RPLP0) gene (10) is the most suited for normalization purposes during the enterocytic differentiation process. Among the nine chosen candidate genes further tested for variation of expression between adenocarcinomas and their corresponding histologically normal resection margins, we found that the β-2-microgobulin (B2M) gene is best for normalizing gene expression, although other genes were found to represent acceptable alternatives.

MATERIALS AND METHODS

cDNA Microarrays

Tissues. Specimens of the ileum from five fetuses ranging from 16 to 20 wk of age (postfertilization) were obtained after legal therapeu-

Address for reprint requests and other correspondence: J.-F. Beaulieu, Département d'Anatomie et de Biologie Cellulaire, Faculté de Médecine, Université de Sherbrooke, Sherbrooke, Québec, Canada J1H 5N4 (e-mail: Jean-Francois.Beaulieu@usherbrooke.ca).

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

 0193-1857/06 $8.00 Copyright © 2006 the American Physiological Society

Downloaded from ajpgi.physiology.org on October 25, 2006

tic abortion. Only specimens obtained rapidly (within 60 min) were used. The project was in accordance with the protocol approved by the Institutional Human Research Review Committee at the Université de Sherbrooke for the use of human material. Fetal small intestinal epithelia were separated from the mesenchyme using Matrisperse (Collaborative Biomedical Products, Becton-Dickenson Labware; Mississauga, ON, Canada), and fraction purity was confirmed by RT-PCR as described previously (21, 23).

Cell culture. Two epithelial cell lines, which together with the samples obtained from tissues recapitulate the crypt-villus axis of the fetal intestine (20), were used in this study. The HIEC-6 cell line was generated from the normal fetal human intestine (22). These cells express a number of crypt cell markers but no villus cell markers, appear to be unable to differentiate, and are thus considered to be intestinal stemlike cells (20, 22). HIEC cells were used at *passages 14–17* and grown as described (22). The Caco-2/15 cell line, a stable clone of the parent Caco-2 cell line (25), has been characterized elsewhere (4, 31). Although of cancerous origin, these cells are unique in that upon confluence, they spontaneously undergo a gradual villus-like enterocytic differentiation process, similar to that observed in the epithelium of the intact fetal small and large intestine (19, 20). Cells between *passages 14* and *22*, used for the generation of cDNA microarrays, were cultured in plastic dishes as previously described (31) and analyzed at 0–25 days postconfluence.

Three other human cell lines were used to construct a reference sample for the cDNA microarrays: *1*) the HT-29 cell line, a cell line derived from a carcinoma of the colon (3); *2*) the A549 cell line, a pulmonary adenocarcinoma cell line (30) obtained from Dr. Jacques Bérard; and *3*) the SKOV3 cell line, a cell line derived from an ovarian carcinoma obtained from Dr. Claudine Rancourt (17). These cells were grown as previously described (3, 17, 30).

qRT-PCR

Tissues, RNA purification, quantification, and RT of samples used for real-time qRT-PCR. Total RNA from our RNA bank was originally purified using TriPure (Roche; Indianapolis, IN) and stored at −80°C. The integrity of triplicate total RNA samples of proliferatively active HIEC-6 cells, differentiated fetal epithelial cells (see *Tissues*), and Caco-2/15 cells at several differentiation stages [subconfluent (SC), 0 days postconfluent (PC), 2–5 days PC, 10–15 days PC, and 25–35 days PC] was verified by visual analysis of 18S and 28S bands on an ethidium bromide-stained 1.5% agarose gel. The same procedure was followed for the seven patient-matched pairs of resection margins (RM) and primary adenocarcinoma samples. These samples were provided by the Cooperative Human Tissue Network, which is funded by the National Cancer Institute, and consisted of one well-differentiated, five moderately differentiated, and one poorly differentiated colon adenocarcinoma. It is noteworthy that in contrast to the cell lines and villus cells representing pure epithelial cell preparations, the colonic tissues used herein were heterogeneous, containing both epithelial and stromal cells. Because the RT reaction is extremely sensitive to variations in salt concentration and contamination by carry-over reagents from previous steps (15), the samples were repurified using RNeasy columns (Qiagen; Mississauga, ON, Canada) to obtain equivalent salt concentrations and purity of the samples. The RiboGreen RNA Quantitation Kit (Molecular Probes; Hornby, ON, Canada) was used according to the manufacturer's instructions to accurately quantify the samples. Fluorescence emission was read by an Mx3000P (Stratagene; La Jolla, CA) functioning as a plate reader. First-strand cDNA synthesis using Omniscript (Qiagen) was performed on 2 μg total RNA using a combination of oligo(dT)$_{12–18}$ and random hexamers (Amersham Pharmacia; Baie d'Urfé, QC, Canada) as primers. To ensure an RT performance of equal quality, all samples were reverse transcribed simultaneously using a single mastermix.

Real-time PCR. All reactions were performed in an Mx3000P (Stratagene) starting with 10 min of *Taq* activation at 95°C, followed by 40 cycles of melting (95°C, 30 s), primer annealing at the temperature appropriate for each primer (55–60°C, 45 s), and extension (72°C, 45 s) ending with a melting curve analysis to validate the specificity of the PCR products. Fluorescence data were acquired after each annealing step. Amplification efficiencies ranged from 93% to 104%. The genes investigated were 5-methyltetrahydrofolate-homocystine methyltransferase (MTR; Accession No. NM_000254), RNA polymerase II (DNA directed) polypeptide A (POLR2A; Accession No. NM_000937), B2M (Accession No. NM_004048), β-actin (ACTB; Accession No. NM_001101), GAPD (Accession No. NM_002046), ribosomal protein S14 (RPS14; Accession No. NM_005617), mannosidase-α, class 1B, member 1 (MAN1B1; Accession No. NM_016219), RPLP0 (Accession No. BC019014), hypoxanthine phosphoribosyltransferase 1 (HPRT1; Accession No. NM_000194), sucrase-isomaltase (α-glucosidase) (SI; Accession No. NM_001041), and c-Myc (MYC; Accession No. NM_002467). Table 1 summarizes the primers used in this study. Primers were generated using the primer formation software Primer3 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) with attention given to avoiding primer-dimer formation by stringent use of the maximum 3′ self-complimentarity function of the Primer3 program. The absence of primer-dimers was verified postamplification by melting curve analysis.

Statistical Analysis

All statistical analyses were performed using the statistical software Graph Pad Prism 3.02 (Graph Pad Software; San Diego, CA). For linear regression analysis, we applied the following numeric denotation to the different cellular samples representing different

Table 1. *Primers used in this study*

Gene Symbol	Sense Primer	Antisense Primer	Accession No.
RPLP0	5′-GCAATGTTGCCAGTGTCTG-3′	5′-GCCTTGACCTTTTCAGCAA-3′	BC019014
GAPD	5′-TCTCCTCTGACTTCAACAGCGAC-3′	5′-CCCTGTTGCTGTAGCCAAATTC-3′	NM_002046
HPRT1	5′-TGACACTGGCAAAACAATGCA-3′	5′-GGTCCTTTTCACCAGCAAGCT-3′	NM_000194
B2M	5′-GTGCTCGCGCTACTCTCTC-3′	5′-GTCAACTTCAATGTCGGAT-3′	NM_004048
POLR2A	5′-ATCTCTCCTGCCATGACACC-3′	5′-AGACCAGGCAGGGGAGTAAC-3′	NM_000937
RPS14	5′-GGCAGACCGAGATGAATCCTCA-3′	5′-CAGGTCCAGGGGTCTTGGTCC-3′	NM_005617
MAN1B1	5′-ACCGTGGAGAGCCTGTTCTA-3′	5′-GTTTGGGTCATCGGAGAAGA-3′	NM_016219
ACTB	5′-CCTCGCCTTTGCCGATCC-3′	5′-GGATCTTCATGAGGTAGTCAGTC-3′	NM_001101
MTR	5′-TGTGGAGACTCGCAGACATC-3′	5′-CCTCAACCTGATCCTTGGAA-3′	NM_000254
SI	5′-GAGGACACTGGCTTGGAGAC-3′	5′-ATCCAGCGGGTACAGAGATG-3′	NM_001041
MYC	5′-CCTACCCTCTCAACGACAGC-3′	5′-CTCTGACCTTTTGCCAGGAG-3′	NM_002467

RPLP0, ribosomal phosphoprotein P0; HPRT1, hypoxanthine phosphoribosyltransferase 1; B2M, β-2-microglobin; POLR2A, RNA polymerase (DNA directed) polypeptide A; RPS14, ribosomal protein S14; MAN1B1, mannosidase-α, class 1B, member 1; ACTB, β-actin; MTR, 5-methyltetrahydrotolate-homocystine methyltransferase; SI, sucrase-isomaltase; MYC, c-Myc.

American Journal of Physiology – Gastrointestinal and Liver Physiology

Downloaded from ajpgi.physiology.org on October 25, 2006

stages of differentiation: HIEC: 1, Caco-2/15 SC: 2, Caco-2/15 0 PC: 3, Caco-2/15 2–5 PC: 4, Caco-2/15 10–15PC: 5, Caco-2/15 25–35 PC: 6, and villus epithelial: 7. These numeric values were then plotted against the average cycle threshold (C_t) values obtained for the individual gene at the corresponding differentiation stage. Goodness of fit analysis (r^2) was performed on the obtained linear regression lines to evaluate the strength of the regression and to assess the usability of the individual genes as normalizing genes during differentiation. ANOVA was performed as a one-way analysis of variance without post hoc testing, and paired *t*-tests were performed as two-tailed tests. A *P* value below <0.05 was considered significant in all analyses. The Normfinder applet (2) was courtesy of Dr. Ørntoft (Aarhus University Hospital; Aarhus, Denmark). Briefly, this applet describes expression values and performs both a separate analysis of the expression values in sample subgroups as well as an estimation of the inter- and intragroup variations providing a "stability factor." The geNorm applet (32) was obtained from the Gene Quantification homepage (www.gene-quantification.org). Briefly, this applet performs an analysis of expression stability leading to a stability index value. Furthermore, the applet provides a normalization factor based on the geometric average of all input normalization genes.

RESULTS

Differentiation of Enterocytes

Samples represent different stages of differentiation. We initially verified that our stringently purified and reverse transcribed RNA samples truly represented differentiating cells by analyzing the expression level of the brush-border enzyme SI, a well-known and accepted marker of differentiation. Bescause our samples were stringently purified, quantified, and reverse transcribed, we compared the obtained C_t values directly. The C_t value represents the number of thermal cycles needed to reach a certain threshold value of fluorescence and is inversely correlated to the amount of input cDNA. Under perfect conditions, a ΔC_t of 1 equals a doubling of input material. The HIEC-6 and SC Caco-2/15 samples did not yield detectable amplicons and were thus attributed the nomination of nondetectable (ND). As expected, Fig. 1*A* demonstrates a dramatic increase in transcript levels as a function of differentiation. The decline shown in the C_t value corresponds to an ~325-fold increase in gene expression from newly confluent Caco-2/15 cells to Matrisperse-extracted villus epithelial cells, demonstrating that the samples indeed represent intestinal epithelial cells at different stages of differentiation.

Analysis of the slope of the regression lines. We next used previously generated cDNA microarray data (E. Tremblay, J. Auclair, and J.-F. Beaulieu, unpublished data) generated with probes from a variety of cell lines representing different stages of human epithelial cell differentiation to identify housekeeping genes with low variation as a function of differentiation. Two such genes were identified: MAN1B1 and MTR. We then analyzed these genes in conjunction with seven other widely used normalizing genes such as GAPD and ACTB. To validate the usability of the nine candidate genes as normalizing genes during differentiation, we performed quantitative PCR on the stringently purified and reverse transcribed samples representing different levels of differentiation. To evaluate the variation of gene expression during differentiation, we plotted the average C_t value of the individual genes against the corresponding differentiation stage and performed linear regression analysis on the computer-generated line of best fit representing the individual gene. A straight line can be described as $y = mx + b$, where *y* denotes the ordinate value, *x* represents the abscissa value, *m* is the slope of the line, and *b* identifies the *y*-intercept of the line. Because our aim was to identify the gene with the least variation of expression as a function of differentiation, plotting gene expression as a function of differentiation should yield a regression line with a slope (or *m*) as close to zero as possible. Figure 1*B* shows the C_t values of the nine candidate genes during the course of differentiation, whereas Table 2 summarizes the statistical analysis of the same genes. MAN1B1, GAPD, and ACTB displayed positive regression with differentiation ($m = 0.24$, 0.21, and 0.28, respectively), whereas MTR, POLR2A, and RPS14 showed negative regression with differentiation ($m = -0.22$, -0.27, and -0.11, respectively). On the other hand, RPLP0, B2M, and HPRT1 displayed slope values near zero ($m = -0.02$, -0.03, and 0.06, respectively; Fig. 1*B* and Table 2).

Analysis of goodness of fit of linear regression. r^2 Is obtained by computing the relationship between an identified regression line with a hypothetical horizontal line through the mean of all data. An r^2 value of zero means that the best way of describing the data is a horizontal line through the mean of all the data points (i.e., an *m* value of 0). With our data, an r^2 value near zero therefore demonstrated low variation of gene expression as a function of differentiation. With the use of this parameter (low r^2 value) to choose the best normalizing genes, RPLP0, B2M, and HPRT1 again revealed quite stable patterns of expression during differentiation, with r^2 values of 0.01 for all three (Table 2).

Analysis of intra- and intergroup variance. m And r^2 values are obtained using the full data set and are thus an "average" of the entire data set. Thus a certain tolerance of outliers that can "disappear" in the full data set is a precondition of these methods. To evaluate if any such outliers were present at any stage(s) of differentiation, we performed one-way ANOVA on the data from the nine candidate genes and found that five of the genes had statistically significantly ($P < 0.05$) different levels of expression from one differentiation stage to another. The four genes that did not have statistically relevant alterations of expression from differentiation stage to differentiation stage were RPLP0, B2M, RPS14, and MTR (Table 2).

We next evaluated the SD of the C_t values of the nine candidate genes. The four genes with the lowest SDs were RPLP0, B2M, RPS14, and ACTB (0.43, 0.58, 0.61, and 0.61, respectively), whereas POLR2A, MTR, and HPRT1 were the only genes scoring SD values higher than 1.000 (Table 2), representing an input cDNA variation of more than twofold (24). These three genes all showed high variation in the nondifferentiated stages, whereas the differentiated stages displayed considerably less variation (Fig. 1*B*). To exclude the possibility of experimental error as the cause of the high variation, we repeated the PCR run, resulting in analogous results (data not shown). Finally, we analyzed the data sets using the Excel applets Normfinder and geNorm, which take into account both the intra- and the inter-group variation when calculating the stability factor of a given set of data (2, 32). This approach identified RPLP0 and RPS14 as the two genes with the least variability (Table 2). Overall, we thus found that RPLP0 was the gene that most frequently was identified as having the least variation using all the different statistical approaches.

American Journal of Physiology – Gastrointestinal and Liver Physiology

Downloaded from ajpgi.physiology.org on October 25, 2006

American Journal of Physiology – Gastrointestinal and Liver Physiology

Fig. 1. *A*: samples represent different levels of differentiation. Threshold cycle (C_t) values of sucrase-isomaltase obtained using the different samples are shown. Samples were HIEC cells [subconfluent (SC), newly confluent (0PC), 2–5, 10–15, and 25–35 days postconfluence (PC)] and villus epithelium from the ileum. ND, not detected. Shaded area, 95% confidence boundaries. The decrease in C_t value represents an ~325-fold increase in gene expression. $P < 0.001$ by one-way ANOVA. Outcome variables are means ± SE. *B*: variation of expression levels of the nine candidate genes during differentiation. The C_t values obtained for each gene at the different differentiation stages are shown. The genes investigated were 5-methyltetrahydrofolate-homocystine methyltransferase (MTR), RNA polymerase II (DNA directed) polypeptide A (POLR2A), β-2-microglobin (B2M), β-actin (ACTB), GAPD, ribosomal protein S14 (RPS14), mannosidase-α, class 1B, member 1 (MAN1B1), ribosomal phosphoprotein P0 (RPLP0), and hypoxanthine phosphoribosyltransferase 1 (HPRT1). Shaded area, 95% confidence boundaries. Outcome variables are means ± SE.

Analysis of normalizing gene modulation in differentiating Caco-2/15 cells only. To further characterize the data sets for enterocyte differentiation, we investigated the same parameters as above on the nine candidate genes for the Caco-2/15 cell line only. The results are summarized in Table 2. POLR2, HPRT1, RPS14, and MTR displayed a strong tendency toward modulation during the course of Caco-2/15 differentiation ($m = -0.82, -0.35, -0.31$, and -0.52, respectively), whereas GAPD, MAN1B1, and, to a certain extent, ACTB showed less modulation ($m = -0.04$, 0.08, and 0.12, respectively). RPLP0 retained a low m value (-0.07), whereas B2M showed less stable expression during Caco-2/15 differentiation (-0.13; Table 2) than with all three cell lines (-0.03; Table 2). Likewise, it was MAN1B1, GAPD, and RPLP0 that scored the lowest values for the r^2 test (0.05, 0.01, and 0.04, respectively) with Caco-2/15 cells only. Using ANOVA, we also obtained statistically significant differences in expression between differentiation stages for only POLR2A and HPRT1 (0.01 and 0.03, respectively). The SD data were not significantly modulated by the exclusion of the data from the HIEC and villus cells except for HPRT1 (0.76 vs. 1.47). The three genes with the lowest SDs were RPLP0, MAN1B1, and ACTB (0.43, 0.4, and 0.41, respectively). Finally, analysis of stability values showed RPLP0, RPS14, and MAN1B1 to be the three genes with the best scores (0.31, 0.28, and 0.29, respectively).

Downloaded from ajpgi.physiology.org on October 25, 2006

American Journal of Physiology – Gastrointestinal and Liver Physiology

Table 2. *Regression and statistical data of the nine candidate genes analyzed during enterocytic differentiation*

Gene Symbol	Slope	r^2	*P* Value (ANOVA)	SD	Stability Value: Using the Normfinder applet	Stability Value: Using the geNorm applet
All three cell lines						
RPLP0	−0.02±0.05	0.01	0.19	0.43	0.42	0.50
GAPD	0.21±0.10	0.20	0.004	0.71	0.59	ND
HPRT1	−0.06±0.16	0.01	<0.0001	1.21	0.81	0.94
B2M	−0.03±0.09	0.01	0.48	0.58	0.76	0.80
POLR2A	−0.27±0.20	0.09	0.01	1.25	0.76	1.16
RPS14	−0.11±0.08	0.09	0.06	0.61	0.37	0.50
MAN1B1	0.24±0.08	0.32	0.001	0.66	0.47	0.74
ACTB	0.28±0.06	0.53	< 0.0001	0.61	0.52	0.72
MTR	−0.22±0.19	0.06	0.62	1.23	0.84	1.27
Caco-2/15 only						
RPLP0	−0.07±0.09	0.04	0.50	0.43	0.31	0.44
GAPD	−0.04±0.12	0.01	0.53	0.50	0.38	ND
HPRT1	−0.35±0.20	0.19	0.03	0.76	0.57	0.64
B2M	−0.13±0.12	0.08	0.66	0.53	0.51	0.49
POLR2A	−0.82±0.19	0.58	0.01	1.05	0.51	0.80
RPS14	−0.31±0.1	0.36	0.06	0.53	0.28	0.53
MAN1B1	0.08±0.10	0.05	0.30	0.40	0.29	0.44
ACTB	0.12±0.09	0.11	0.54	0.41	0.45	0.45
MTR	−0.52±0.26	0.24	0.23	1.08	0.61	0.91

Regression and statistical data were analyzed as described in Fig. 1*B*. Slope was determined as the *m* value of the equation of a straight line, $y = mx + b$, when the cycle threshold (C_t) value is plotted against the differentiation stage (Fig. 1*B*). r^2 is the goodness of fit of the determined regression line of each gene. *P* values were determined using the null hypothesis that the average C_t value of one given differentiation stage does not differ from another. SD values were determined as SDs of C_t values across all differentiation stages. Stability values were determined using the norm finder (2) or geNorm (32) applets; low values represent little variation. ND, not defined. The geNorm applet uses an elimination-based algorithm to calculate the stability value. Thus the least stable normalizing gene is eliminated first (no stability value is given), and no difference is made between the 2 best genes.

Healthy Tissue Versus Adenocarcinomas

Cancer samples were first tested for the expression of the oncogene MYC in the tumor samples compared with the corresponding RMs. As expected (13), we found a significant upregulation of the expression of MYC in the cancer samples compared with the RMs (Fig. 2*A*). The ΔC_t value of 1.69 between the RMs and adenocarcinoma samples represented a 3.22-fold ($2^{1.69}$, considering 100% PCR efficiency) upregulation ($P < 0.01$).

Normalizing gene expression in adenocarcinomas versus corresponding RMs. The expression levels of the nine putative normalizing genes were evaluated in the tumors and their corresponding RMs (Fig. 2*A*). It was found that the expression was not significantly altered between the two stages for B2M, MAN1B1, POLR2A, ACTB, HPRT1, and MTR, whereas the expression did vary in the case of RPLP0, RPS14, and GAPD. Bidirectional regulation of gene expression of the same magnitude in individual patient-matched pairs of tissue can lead to a crude average gene modulation that appears unaffected even though there is substantial alteration of gene expression. We therefore transformed the individual ΔC_t values from the patient-paired samples to absolute (positive) numbers to analyze the average gene modulation between the RMs and adenocarcinomas for the nine candidate genes. As shown in Fig. 2*B*, B2M scored the lowest overall absolute ΔC_t value at 0.32. Three other genes stood out with low differences in expression between the tumors and RMs: ACTB, MAN1B1, and MTR (ΔC_t: 0.43, 0.49, and 0.48, respectively). In particular, POLR2A, which displayed hardly any modulation when the crude average was observed (Fig. 2*A*), showed a high modulation of gene expression using absolute values (Fig. 2*B*).

Expression of genes of interest: averages of normalizing genes versus selected normalizing genes. The expression levels of genes of interest for both intestinal cell differentiation (SI) and colon adenocarcinomas (MYC) were analyzed relative to the geometric average (23, 31) of either all nine analyzed normalizing genes or the best three or four normalizing genes as determined previously (see above) versus the single best gene for each experimental setting. As shown in Table 3, all three approaches led to similar conclusions, although some differences in the levels of induction were observed.

DISCUSSION

In this work, we provided the basis for identifying and characterizing genes that can serve as normalizing genes for qRT-PCR analyses related to the differentiation process of the human intestinal epithelium and adenocarcinomas of the human colon. Using this approach, we found the RPLP0 gene to be particularly stable in all intestinal epithelial cell extracts, thereby representing a robust housekeeping reference for assessing gene expression during the human enterocytic differentiation process. On the other hand, B2M was found to be the gene best suited for normalizing input mRNA quantity and quality in studies related to pathological changes in extracts of intestinal mucosa thus containing both epithelial and nonepithelial cells, as demonstrated by comparing matched pairs of primary carcinomas and RMs of the human colon. Although these housekeeping genes may not be optimal for other settings (such as distinct cell lines, pathologies, and/or species), the approach described herein can serve as a paradigm to identify valid housekeeping genes in an intestinal context.

Downloaded from ajpgi.physiology.org on October 25, 2006

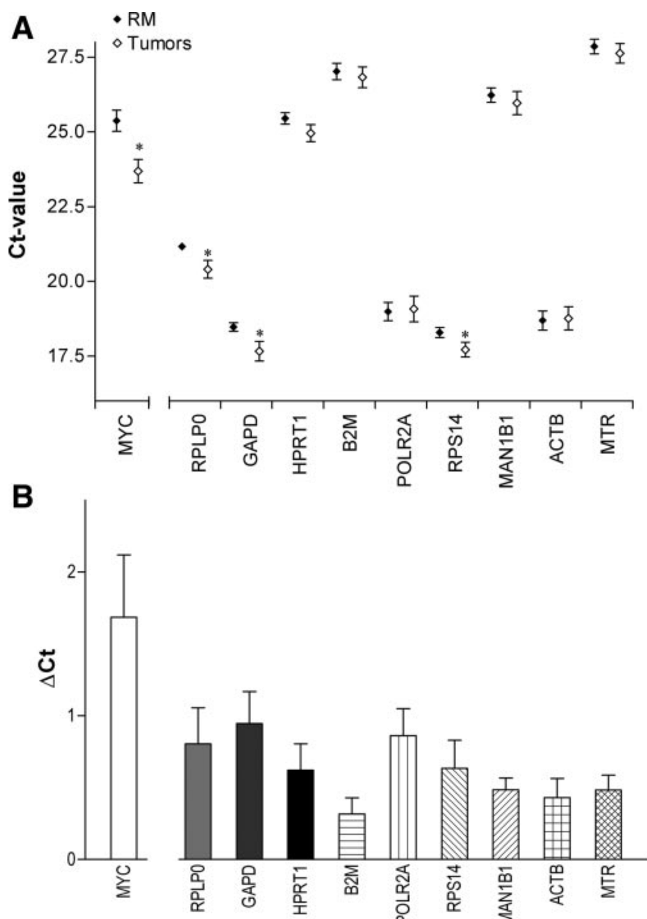

Fig. 2. Variation in gene expression between primary tumors and corresponding resection margins (RMs). *A*: average C_t value obtained for primary tumors and histologically normal RMs for the nine candidate genes and the colon cancer marker c-Myc (MYC). $*P < 0.05$ by paired *t*-test. Outcome variables are means ± SE. *B*: average absolute modulation of gene expression (ΔC_t) between tumors and RMs. Outcome variables are means ± SE.

We employed several different statistical analyses to analyze the levels of gene expression in the intestinal cell samples representing distinct differentiation stages. Distinct statistical analyses were performed to consider complementary parameters and information relative to each gene. Individually, this allowed us to address specific parameters such as trend of expression, internal variation, and relative variation as well as inter- and intragroup variation. Collectively, these analyses led to the identification of optimal housekeeping genes. Indeed, some of the genes, such as POLR2A, MTR, and HPRT1, were rapidly discarded based on their high SD values, which could be ascribed to biological rather than input cDNA variation because it was not seen with the other genes. In fact, considering all the various analyses performed, only three genes stood up to our assessment: RPLP0, RPS14, and B2M. Both RPLP0 and B2M had near-zero slope values; however, B2M was, in our opinion, difficult to accept as the "gold standard" for normalization during differentiation of enterocytes due to the relative high variation observed in the undifferentiated samples, which was further reflected by the poor stability value obtained using the Normfinder and geNorm applets (2, 32). These analytic tools scored RPS14 and RPLP0 as the most stable normalizing genes, although RPS14 displayed a tendency toward a negative slope. Therefore, taking into consideration all the parameters, RPLP0 was selected on the basis that it clearly displayed the best overall stability. Although other experimental conditions may require further validation for the identification of housekeeping genes, it is noteworthy that these analyses were performed across two very different cell lines (HIEC-6 and Caco-2/15) as well as in primary tissue (epithelial villus fractions of the small intestine) strengthening the potency of RPLP0 as the ideal normalizing gene for mRNA quantification in human intestinal epithelial cells. Interestingly, performing the same analyses on Caco-2/15 cells only led to the identification of additional genes, such as GAPD and MAN1B1, that appeared to perform nearly as well as RPLP0. This finding of stable expression in a larger subset of candidate genes in a single cell line is consistent with a more homogeneous cellular background. This further underscores both the remarkably stable expression of RPLP0 across all the cell lines as well as the necessity to carefully validate potential normalizing genes under particular experimental settings.

For the identification of normalizing genes to be used for studies related to adenocarcinomas in the colon, we mainly considered the average modulation obtained using the absolute ΔC_t values rather than the crude average, which does not always reflect individual pair variation. With this approach, we found that the classic normalizing gene B2M performed slightly better than others such as ACTB, MTR, and MAN1B1, whereas other classic normalizing genes such as GAPD and POLR2A appeared to significantly vary between control and tumor tissue. Minute alterations in B2M expression as a consequence of tumor development is in accordance with reported results from the colon (6) and the kidney (14). The strength of these studies, including ours, is the use of matched primary tumors and corresponding histologically normal RMs, enabling an accurate assessment of the difference in gene expression levels between healthy and adenocarcinoma tissue. Our finding that ACTB is little affected by tumor formation is also in agreement with previous studies (6, 14). Therefore, based on

Table 3. *Modulation of gene expression using a distinct approach of normalization*

Experimental Setting (Gene of Interest)	Geometric Average: All normalizing genes	Geometric Average: Selected normalizing genes	Identified Normalizing Gene
Cell differentiation (SI)	765	405 (B2M; RPS14,and RPLP0)	599 (RPLP0)
Colon adenocarcinomas (MYC)	1.67	1.97 (B2M; ACTB, MAN1B1,and MTR)	2.96 (B2M)

Modulation of gene expression was determined as a consequence of normalizing using a geometric average of all the candidate genes versus selected candidate genes versus using the identified optimal genes. Cell differentiation is shown as the average fold upregulation in the villus epithelial fraction compared with newly confluent Caco-2/15 cells. Colon adenocarcinomas is shown as the average fold upregulation in cancer samples compared with normal resection margins.

American Journal of Physiology – Gastrointestinal and Liver Physiology

Downloaded from ajpgi.physiology.org on October 25, 2006

the current analyses, it appears that B2M represents the best normalizing gene for studies related to gene expression in adenocarcinomas of the human colon, although ACTB, MTR, and MAN1B1 represent acceptable alternatives. It cannot, however, be excluded that these normalizing genes, as well as others not tested herein, might also be acceptable controls for distinct subtypes of colonic malignancy.

In conclusion, we herein identified RPLP0 as a very well suited, if not ideal, normalizing gene for mRNA quantification in human intestinal epithelial cells at various stages of differentiation, whereas B2M was confirmed to be the best housekeeping gene for studying gene expression in human colon primary carcinomas. Interestingly, a comparison of the expression levels of genes of interest for intestinal cell differentiation (SI) and colon carcinomas (MYC) using these single optimal housekeeping genes versus an average of several normalizing genes (23, 31) led to similar results. The identification of different optimal normalizing genes for studies of differentiation versus adenocarcinoma formation is not surprising considering the fact that the differentiation studies were performed on pure epithelial cell cultures and extracts, whereas the mRNA for cancers and resections margins were extracted from mucosa. In support of this, differences in gene expression profiles between primary cancers and the surrounding stroma have been recently demonstrated (29).

ACKNOWLEDGMENTS

We thank Drs. Jean Gosselin, Michel Lessard (Département de Pathologie, Centre Hospitalier Universitaire de Sherbrooke), and Jacques Poisson (Département de Chirurgie, Centre Hospitalier Universitaire de Sherbrooke) for the excellent cooperation in providing matched colon cancer specimens and pathological expertise. We also acknowledge the collaboration of the Cooperative Human Tissue Network, funded by the National Cancer Institute, in obtaining additional sets of matched colon cancer specimens.

GRANTS

This work was supported by Canadian Institutes of Heath Research Grants MOP 62914 and MOP 57727 (to J. F. Beaulieu). J. F.Beaulieu is the recipient of a Canadian Research Chair in Intestinal Physiopathology; A. B. Dydensborg is supported by a studentship from the Danish Research Agency (642-02-0055).

REFERENCES

1. **Aerts JL, Gonzales MI, and Topalian SL.** Selection of appropriate control genes to assess expression of tumor antigens using real-time RT-PCR. *Biotechniques* 36: 84–91, 2004.
2. **Andersen CL, Jensen JL, and Ørntoft TF.** Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res* 64: 5245–5250, 2004.
3. **Basora N, Desloges N, Chang Q, Bouatrouss Y, Gosselin J, Poisson J, Sheppard D, and Beaulieu JF.** Expression of the alpha9beta1 integrin in human colonic epithelial cells: resurgence of the fetal phenotype in a subset of colon cancers and adenocarcinoma cell lines. *Int J Cancer* 75: 738–743, 1998.
4. **Beaulieu JF and Quaroni A.** Clonal analysis of sucrase-isomaltase expression in the human colon adenocarcinoma Caco-2 cells. *Biochem J* 280: 599–608, 1991.
5. **Biederman J, Yee J, and Cortes P.** Validation of internal control genes for gene expression analysis in diabetic glomerulosclerosis. *Kidney Int* 66: 2308–2314, 2004.
6. **Blanquicett C, Johnson MR, Heslin M, and Diasio RB.** Housekeeping gene variability in normal and carcinomatous colorectal and liver tissues: applications in pharmacogenomic gene expression studies. *Anal Biochem* 303: 209–214, 2002.
7. **Boland CR.** Neoplasia of the gastrointestinal tract. In: *Textbook of Gastroenterology* (2nd ed.), edited by Yamada T. Philadelphia, PA: Lippincott, 1995, p. 578–595.
8. **Bustin SA.** Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. *J Mol Endocrinol* 29: 23–39, 2002.
9. **Calvo EL, Boucher C, Coulombe Z, and Morisset J.** Pancreatic GAPDH gene expression during ontogeny and acute pancreatitis induced by caerulein. *Biochem Biophys Res Commun* 235: 636–640, 1997.
10. **de Cremoux P, Bieche I, Tran-Perennou C, Vignaud S, Boudou E, Asselain B, Lidereau R, Magdelenat H, Becette V, Sigal-Zafrani B, and Spyratos F.** Inter-laboratory quality control for hormone-dependent gene expression in human breast tumors using real-time reverse transcription-polymerase chain reaction. *Endocr Relat Cancer* 11: 489–495, 2004.
11. **de Kok JB, Roelofs RW, Giesendorf BA, Pennings JL, Waas ET, Feuth T, Swinkels DW, and Span PN.** Normalization of gene expression measurements in tumor tissues: comparison of 13 endogenous control genes. *Lab Invest* 85: 154–159, 2005.
12. **Dheda K, Huggett JF, Bustin SA, Johnson MA, Rook G, and Zumla A.** Validation of housekeeping genes for normalizing RNA expression in real-time PCR. *Biotechniques* 37: 112–114, 116, and 118–119, 2004.
13. **Erisman MD, Rothberg PG, Diehl RE, Morse CC, Spandorfer JM, and Astrin SM.** Deregulation of c-myc gene expression in human colon carcinoma is not accompanied by amplification or rearrangement of the gene. *Mol Cell Biol* 5: 1969–1976, 1985.
14. **Ferguson RE, Carroll HP, Harris A, Maher ER, Selby PJ, and Banks RE.** Housekeeping proteins: a preliminary study illustrating some limitations as useful references in protein expression studies. *Proteomics* 5: 566–571, 2005.
15. **Freeman WM, Walker SJ, and Vrana KE.** Quantitative RT-PCR: pitfalls and potential. *Biotechniques* 26: 112–125, 1999.
16. **Hamalainen HK, Tubman JC, Vikman S, Kyrölä T, Ylikoski E, Warrington JA, and Lahesmaa R.** 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, 2001.
17. **Lane D, Cartier A, L'Esperance S, Cote M, Rancourt C, and Piche A.** Differential induction of apoptosis by tumor necrosis factor-related apoptosis-inducing ligand in human ovarian carcinoma cells. *Gynecol Oncol* 93: 594–604, 2004.
18. **Malarstig A, Tenno T, Jossan S, Aberg M, and Siegbahn A.** A quantitative real-time PCR method for tissue factor mRNA. *Thromb Res* 112: 175–183, 2003.
19. **Ménard D and Beaulieu JF.** Human intestinal brush border membrane hydrolases. In: *Membrane Physiopathology*, edited by Bkaily G. Norwell: Kluwer Academic, 1994, p. 319–341.
20. **Pageot L, Perreault N, Basora N, Francoeur C, Magny P, and Beaulieu JF.** Human cell models to study small intestinal functions: recapitulation of the crypt-villus axis. *Microsc Res Tech* 49: 394–406, 2000.
21. **Perreault N and Beaulieu JF.** Primary cultures of fully differentiated and pure human intestinal epithelial cells. *Exp Cell Res* 245: 34–42, 1998.
22. **Perreault N and Beaulieu JF.** Use of the dissociating enzyme thermolysine to generate viable human normal intestinal epithelial cell cultures. *Exp Cell Res* 224: 354–364, 1996.
23. **Perreault N, Herring-Gillam FE, Desloges N, Belanger I, Pageot LP, and Beaulieu JF.** Epithelial vs mesenchymal contribution to the extracellular matrix in the human intestine. *Biochem Biophys Res Commun* 248: 121–126, 1998.
24. **Pfaffl MW, Tichopad A, Prgomet C, and Neuvians TP.** Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper–Excel-based tool using pair-wise correlations. *Biotechnol Lett* 26: 509–515, 2004.
25. **Pinto M, Robine-Leon S, Appay MD, Kedinger M, Triadou N, Dussaulx E, Lacroix B, Simon-Assmann P, Haffen K, Fogh J, and Zweibaum A.** Enterocyte-like differentiation and polarization of the human colon cacinoma cell line Caco-2 in culture. *Biol Cell* 47: 323–330, 1983.
26. **Radonic A, Thulke S, Mackay IM, Landt O, Siegert W, and Nitsche A.** Guideline to reference gene selection for quantitative real-time PCR. *Biochem Biophys Res Commun* 313: 856–862, 2004.
27. **Schmid H, Cohen CD, Henger A, Irrgang S, Schlöndorff D, and Kretzler M.** Validation of endogenous controls for gene expression analysis in microdissected human renal biopsies. *Kidney Int* 64: 356–360, 2003.
28. **Schmittgen TD and Zakrajsek BA.** Effect of experimental treatment on housekeeping gene expression: validation by real-time, quantitative RT-PCR. *J Biochem Biophys Methods* 46: 69–81, 2000.
29. **Sugiyama Y, Farrow B, Murillo C, Li J, Watanabe H, Sugiyama K, and Evers BM.** Analysis of differential gene expression patterns in colon cancer and cancer stroma using microdissected tissues. *Gastroenterology* 128: 480–486, 2005.

American Journal of Physiology – Gastrointestinal and Liver Physiology

Downloaded from ajpgi.physiology.org on October 25, 2006

30. **Tani T, Lehto VP, and Virtanen I.** Expression of laminins 1 and 10 in carcinoma cells and comparison of their roles in cell adhesion. *Exp Cell Res* 248: 115–121, 1999.
31. **Vachon PH and Beaulieu JF.** Transient mosaic patterns of morphological and functional differentiation in the Caco-2 cell line. *Gastroenterology* 103: 414–423, 1992.
32. **Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, and Speleman F.** Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 3: RESEARCH0034, 2002.
33. **Walters JR.** Recent findings in the cell and molecular biology of the small intestine. *Curr Opin Gastroenterol* 21: 135–140, 2005.

American Journal of Physiology – Gastrointestinal and Liver Physiology

Downloaded from ajpgi.physiology.org on October 25, 2006