Whole blood *ACTB*, *B2M* and *GAPDH* expression reflects activity of inflammatory bowel disease, advancement of colorectal cancer, and correlates with circulating inflammatory and angiogenic factors: Relevance for real-time quantitative PCR

Iwona Bednarz-Misa^{1,A–D,F}, Katarzyna Neubauer^{2,B,C,E,F}, Ewa Zacharska^{3,B,E,F},
Bartosz Kapturkiewicz^{4,B,E,F}, Małgorzata Krzystek-Korpacka^{1,A–D,F}

¹ Department of Medical Biochemistry, Wrocław Medical University, Poland

² Department of Gastroenterology and Hepatology, Wrocław Medical University, Poland

³ Department of Food Hygiene and Consumer Health, Wrocław University of Environmental and Life Sciences, Poland

⁴ First Department of Oncological Surgery of Lower Silesian Oncology Center, Wrocław, Poland

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation;
D – writing the article; E – critical revision of the article; F – final approval of the article

Advances in Clinical and Experimental Medicine, ISSN 1899–5276 (print), ISSN 2451–2680 (online)

Adv Clin Exp Med. 2020;29(5):547–556

Address for correspondence

Iwona Bednarz-Misa

E-mail: iwona.bednarz-misa@umed.wroc.pl

Funding sources

The research was funded by National Science Centre
(grant No. DEC-2011/01/D/NZ5/02835).

Conflict of interest

None declared

Received on November 5, 2019

Reviewed on November 18, 2019

Accepted on March 10, 2020

Published online on May 19, 2020

Cite as

Bednarz-Misa I, Neubauer K, Zacharska E, Kapturkiewicz B,
Krzystek-Korpacka M. Whole blood *ACTB*, *B2M* and *GAPDH*
expression reflects activity of inflammatory bowel
disease, advancement of colorectal cancer, and correlates
with circulating inflammatory and angiogenic factors:
Relevance for real-time quantitative PCR. *Adv Clin Exp Med*.
2020;29(5):547–556. doi:10.17219/acem/118845

DOI

10.17219/acem/118845

Copyright

Copyright by Author(s)

This is an article distributed under the terms of the
Creative Commons Attribution 3.0 Unported (CC BY 3.0)
(<https://creativecommons.org/licenses/by/3.0/>)

Abstract

Background. The effect of bowel inflammation and cancer on the expression of the most prevalent internal controls: *ACTB*, *GAPDH* and *B2M* in whole blood is unknown, although at least *GAPDH* occurred to be tightly regulated and suspected of supporting cancer growth, challenging its suitability as a reference.

Objectives. To evaluate the effect of colorectal cancer (CRC) and active inflammatory bowel disease (IBD) on the stability of *ACTB*, *B2M*, *GAPDH*, *HPRT1*, *SDHA*, and *TBP* leukocyte expression.

Material and methods. Gene expression in controls and CRC and IBD patients (n = 21/18/25) was evaluated in real-time quantitative polymerase chain reaction (RT-qPCR) using NormFinder, geNorm, BestKeeper, and comparative ΔC_t method, and validated by comparison with absolute quantification of interleukin 1 β (IL-1 β) and CCL4.

Results. *HPRT1*, *SDHA* and *TBP* were superior normalizers in CRC and IBD. The highest expression variability was noted in active IBD. *B2M* was significantly lower in CRC but higher in IBD. *GAPDH* was higher in CRC and IBD. *ACTB* and *GAPDH* corresponded with CRC advancement ($p = 0.52$ and $p = 0.53$) and with clinical activity in Crohn's disease ($p = 0.44$ and $p = 0.57$) and ulcerative colitis (*GAPDH*: $p = 0.72$). *ACTB*, *B2M* and *GAPDH* correlated with circulating inflammatory/angiogenic indices, differently in IBD and CRC.

Conclusions. Leukocyte *GAPDH*, *ACTB*, and *B2M* expression is affected by bowel inflammation and cancer, rendering them unsuitable as a reference in CRC and IBD.

Key words: geNorm, NormFinder, BestKeeper, whole blood transcriptome, expression stability

Introduction

Real time quantitative polymerase chain reaction (RT-qPCR) is a powerful tool used to detect even subtle alterations in gene expression in order to unravel pathomechanisms of diseases and to aid research on new biomarkers and/or therapeutic targets. Normalization of target gene expression against reference gene(s), assumed to be uniformly expressed and hence referred to as housekeeping genes (HKG), is a common method of accounting for non-biological variation introduced during sample handling and attributed to differences in amount of loaded templates, transcription efficiencies, contamination with inhibitors, etc. Considering the high sensitivity of the method, the choice of appropriate reference genes is a prerequisite for obtaining valid and reproducible results in studies employing RT-qPCR.¹ However, there is now solid evidence that HKG expression is in fact regulated and might not only differ between cells or tissues but also change in response to endo- or exogenous factors. Particularly, the stability of the most popular reference genes, *GAPDH*, *ACTB* and *B2M*, has recently been challenged, both in cancer disease^{2,3} and in inflammation.^{4,5} Consequently, a need for validating candidate normalizers prior to their application as internal control in order to find these minimally regulated under given experimental conditions has been repeatedly stressed.^{6,7}

Inflammatory bowel disease (IBD) and colorectal cancer (CRC) evoke a response from circulating immune cells, which is why analyzing gene expression patterns in whole blood, owing to its availability, might be advantageous over more relevant but less accessible bowel tissues. Accordingly, profiling gene expression in blood is gaining interest. Recently, it has been proved useful in evaluating various pathological conditions and immune responses, including identifying new biomarkers^{8–10} in IBD¹¹ and CRC.¹² We have previously shown that both bowel inflammation and cancer affect tissue expression of common reference genes.¹³ However, their potential effect in whole blood has not been investigated yet. Hence, we designed our present study to find and validate optimal reference genes for whole blood transcriptome studies involving patients with CRC and IBD, active and non-active, and to examine and compare the effect, if any, bowel cancer and inflammation might have on the expression of popular reference genes in whole blood. We showed that the most popular genes used for whole blood transcriptome analysis are not suitable for CRC and IBD patients due to the high level of variation in their expression. Moreover, we showed that this variation was directional; namely, it reflected the advancement/severity of diseases and the expression of genes associated with inflammatory and angiogenic responses.

Material and methods

Study population

Sixty-four individuals were enrolled in the current study: 18 CRC patients with histologically confirmed adenocarcinoma, 25 patients with active IBD (13 with Crohn's disease (CD) and 12 with ulcerative colitis (UC)), and 21 controls recruited from patients with adenomas ($n = 4$), colonic diverticulosis ($n = 1$) or IBD in remission ($n = 16$). There were 3 patients with T2N0M0, 1 with T2N × M0, 2 with T3N0M0, 5 with T3N1M0, 2 with T4N0M0, 3 with T4N1M0, 1 with T4N2M1, and 1 with T4N1M1 in cancer group. For the assessment of CD activity, the Crohn's Disease Activity Index (CDAI) combining the evaluation of vital parameters, clinical findings and medical history was applied, and CDAI ≥ 150 was indicative of active disease. For the assessment of UC activity, the Rachmilewitz Index (RI; clinical activity index – CAI), encompassing stool frequency, number of stools with blood, general well-being, abdominal pain/cramp, fever, extraintestinal manifestations, and laboratory tests (erythrocyte sedimentation rate and hemoglobin concentration) was applied and RI ≥ 6 was indicative of active disease.

Female to male ratio in control/CRC/IBD groups was as follows: 7/14, 3/15 and 10/15, $p = 0.257$ and age distribution was as follows: median 46 years (range: 20–78), 62 years (46–77) and 34.8 years (19–60), $p < 0.001$. None of HKG correlated with age.

Ethical considerations

The study protocol was approved by the Medical Ethics Committee of Wroclaw Medical University (approval No. KB-575/2011 from 10 November 2011) and the study was conducted in accordance with the Helsinki Declaration of 1975, as revised in 1983, and informed consent was obtained from all patients.

Analytical methods

Sample selection and RNA isolation, quantification and quality assessment

Whole blood samples (3 mL) were collected into PAXgene Blood RNA Tubes prior to any treatment and stored at -80°C . RNA was isolated and purified with complementary PAXgene Blood RNA Kit (Qiagen, Hilden, Germany) according to manufacturer's instructions, then quantified in duplicates with NanoDrop 2000 (Thermo Fisher Scientific, Waltham, USA) and its purity assessed by calculating ratios of absorbance at 260 nm, 280 nm, and 230 nm. RNA integrity was evaluated with RNA quality indicator (RQI; from 1 – degraded to 10 – intact RNA) using the Experion platform incorporating LabChip microfluidic technology and Experion RNA StdSens analysis kits (BioRad,

Hercules, USA). Only RNA isolates with RQI ≥ 7 were used for RT-qPCR. Possible presence of inhibitors in each RNA isolate was tested by calculating RT-qPCR efficiencies from standard curves prepared by serial dilutions of respective cDNA samples (five-fold dilutions, 6 point-curve, conducted in duplicates using SG qPCR Master Mix from EURx, Gdańsk, Poland). Working dilution of cDNA 1:5 was found to effectively dilute reaction inhibitors and assure near 100% qPCR efficiencies.

Reverse transcription

To assure the same load of RNA template, for all samples 0.25 μg of whole blood RNA per reaction (20 μL), previously quantified in duplicates with NanoDrop 2000 and verified in Experion electrophoresis, was reversely transcribed in C1000 thermocycler (BioRad) using Maxima First Strand cDNA Synthesis Kit for RT-qPCR (Thermo Fisher Scientific) according to the protocol suggested by manufacturer. All samples were accompanied by matching negative transcription (“no-RT”) controls, devoid of reverse transcriptase, subsequently tested to assure lack of contamination with genomic DNA.

Quantitative real-time PCR

Quantitative real-time PCR (qPCRs) were conducted with CFX96 Real-Time PCR system (BioRad) using SsoFast EvaGreen® Supermix (BioRad) and the following cycling conditions: 30 s activation at 95°C, 5 s denaturation at 95°C, annealing/extension for 5 s at 61°C, 40 cycles, followed by melting step (60–95°C with fluorescent reading every 0.5°C). Reaction mixture contained 2 μL of cDNA

(diluted 1:5), 10 μL of $\times 2$ SsoFast EvaGreen® Supermix 1 μL of each 10 nM forward and reverse target-specific primers, and water up to 20 μL . Sequences and specificities (evaluated with RT-qPCR and calculated from a four-fold dilution series with 6 measuring points in triplicates, plotted as Cq vs logarithms of dilution values of the DNA templates) of optimized and validated primers spanning at least 1 intron are presented in Table 1. Primers were synthesized by Genetec Biotech (Hradec Králové, Czech Republic). A mixture of all cDNA samples investigated was used as a template for calculating primers’ specificities. Samples were assessed in 3 technical replicates and accompanied by “no template” control.

Absolute quantification

For absolute quantification, plasmids with *IL-1 β* and *CCL4* inserts were prepared in the following manner: PCR products (10 ng) were cloned into pJET1.2/blunt cloning vector following manufacturer’s protocol (Fermentas UAB, Vilnius, Lithuania) and subsequently used to transform NovaBlue GigaSinglets™ Competent Cells-Novagen (Sigma-Aldrich, St. Louis, USA). Amplified constructs with *IL-1 β* and *CCL4* inserts were isolated using Plasmid DNA Purification Kit (Macherey-Nagel, Düren, Germany) and quantified using NanoDrop 2000. Mean concentrations of *IL-1 β* and *CCL4* plasmids were 4.47 $\mu\text{g}/\mu\text{L}$ and 5.89 $\mu\text{g}/\mu\text{L}$, respectively, and 260/280 and 260/230 ratios were 2.03 and 2.4 for *IL-1 β* and 1.99 and 2.36 for *CCL4*. Ten-fold dilution series was prepared for each plasmid ($E = 102.1\%$, $R^2 = 1.0$ for *IL-1 β* and $E = 100.7\%$ and $R^2 = 0.999$ for *CCL4*) and used for constructing standard curves, in which Cq values were plotted against log of copy

Table 1. Sequences and performance characteristics of primers used in the current study

Symbol	Gene name; function of encoded protein	Accession No.	Primer sequence 5'→3' (forward/reverse)	Amp. size	E [%]
<i>ACTB</i> ¹	actin, β ; cytoskeletal structural protein	NM_001101.3	F: caccattggcaatgagcgggt R: aggtctttgaggatgtccacgt	135 bp	100.9
<i>B2M</i> ¹	β -2-microglobulin; β -chain of MHC class I molecules	NM_004048.2	F: ccaactgaaaagatgagtagtcct R: ccaatccaaatgcggcatctca	126 bp	104.3
<i>GAPDH</i> ¹	glyceraldehyde-3-phosphate dehydrogenase; glycolytic enzyme	NM_002046.4	F: gtctcctctgacttcaacagcg R: accaccctgttgctgtagccaa	131 bp	105.4
<i>HPRT1</i>	hypoxanthine phosphoribosyl-transferase; purine metabolism	NM_000194.2	F: tgacactggcaaaacaatgca R: ggtcctttcaccagcaagct	94 bp	103.2
<i>SDHA</i>	succinate dehydrogenase subunit A; subunit of respiratory chain complex	NM_004168.2	F: agaggcacggaaggagtcac R: caccacatctgtctcatcagtagg	267 bp	94.8
<i>TBP</i>	TATA-box-binding protein; general transcription factor	NM_003194.4	F: tataatcccaagcgggttgctg R: ctggctcataactactaaattgtg	283 bp	109.7
<i>IL-1β</i> ¹	interleukin (IL)-1 β	NM_000576.2	F: ccacagaccttcaggagaatg R: gtgcagttcagtgatcgtagcag	131 bp	100.1
<i>CCL4</i> ¹	macrophage inflammatory protein (MIP)-1 β	NM_002984.2	F: ggtcatcacgtactcctggac R: gcttcctcgcaactttgtgtag	140 bp	103.5

Amp – amplicon; E – efficiency; ¹ primer sequences were as proposed by Origene (www.origene.com). Remaining primers were designed using Beacon Designer Probe/Primer Design Software (BioRad) as previously described (manuscript submitted). Forward and reverse primer sequences are denoted by “F” and “R”, respectively.

number. The qPCR were run in triplicates using SsoFast EvaGreen® Supermix (BioRad) and conditions described in the above section.

Circulating cytokines and growth factors

For the purpose of correlation analysis, data on circulating cytokines and growth factors: IL1- β , IL-4, IL-6, IL-8, IL-12, G-CSF, GM-CSF, FGF2, MCP-1 (CCL2), macrophage inflammatory protein (MIP)-1 α (CCL3), PDGF-BB, and tumor necrosis factor α (TNF- α), were retrieved from our earlier study and were available for 26 patients (16 with CRC and 10 with IBD).¹⁴ Cytokines/growth factors were measured by means of flow cytometry-based method incorporating Luminex xMAP® technology, which utilizes magnetic microspheres conjugated with monoclonal antibodies. Measurements were conducted on BioPlex 200 platform with HRF (BioRad) using custom-made multiplexes validated by the manufacturer (BioRad).

Data on C-reactive protein (CRP) levels at the time of blood sample collection for current study was retrieved from patients' medical records.

Statistical analysis

Technical replicates were averaged prior to any analyses and differences in primer efficiencies (Table 1) were taken into account by calculating efficiency (E)-corrected Cq values (CqE) using the following formula: $CqE = \log(E^{\Delta Cq})$ base 2, where $E = 10^{(-1/\text{slope of standard curve})}$.

Expression stability was evaluated using the 4 most popular algorithms: geNorm utility in qbasePLUS v. 2.4 software (Biogazelle BE, Ghent, Belgium),¹⁵ calculating stability M value described as the average pair-wise variation of a specific gene as compared with other candidate genes, NormFinder software v. 0.953 (available as MS Excel Add-in at www.mdl.dk/publicationsnormfinder.htm),¹⁶ calculating intra- and inter-group variability combined into stability value, BestKeeper,¹⁷ using pair-wise correlation analysis of candidate genes yielding BestKeeper index, and comparative ΔC_t method,¹⁸ comparing relative expression

of pairs of genes within each sample and calculating mean standard deviation (SD) for each gene. RefFinder comprehensive tool (available at www.leonxie.com) was used to create the final ranking integrating the results obtained with geNorm, NormFinder, BestKeeper, and comparative ΔC_t method by assigning an appropriate weight to an individual gene, calculating the geometric mean of their weights and re-ranking the candidate genes accordingly. The smaller geometric mean, the more stably expressed reference gene.

Number of genes sufficient as normalizers was determined with geNorm algorithm. Remaining analyses were conducted using MedCalc Statistical Software v. 12.7.7 (MedCalc Software bvba, Ostend, Belgium) on absolute or normalized relative quantities (NRQ) calculated with qbasePLUS. Normalized relative quantities is a Cq of a given sample referred to the average Cq across all samples for specific gene, with differences in PCR amplification efficiencies taken into account, and normalized against geometric mean Cq of a set of genes selected as reference. If not otherwise stated, data was presented as geometric means with 95% confidence interval (95% CI). Between-group differences were analyzed using one-way analysis of variance (ANOVA) with post-hoc analysis with the Tukey–Kramer test (multigroup comparisons) or using t-test for independent samples with Welch correction when necessary (two-group comparisons). Correlation analysis was conducted using Pearson or Spearman rank correlation tests. Data distribution was tested using χ^2 test and the homogeneity of variances was tested using Levene's test. Frequency analysis was conducted using χ^2 test. All calculated probabilities were two-tailed and p-values ≤ 0.05 were considered statistically significant.

Results

Finding optimal reference genes for studying whole blood transcriptome from IBD and CRC patients

The expression stability of 6 common reference genes, namely *ACTB*, *B2M*, *GAPDH*, *HPRT1*, *SDHA*, and *TBP*, was

Table 2. The expression stability of selected reference genes as calculated using different algorithms: NormFinder, geNorm, BestKeeper, and comparative ΔC_t method and integrated using RefFinder

NormFinder		geNorm		BestKeeper		Comparative ΔC_t		Comprehensive ranking (RefFinder)	
gene	S ¹	gene	S ¹	gene	S ¹	gene	S ¹	gene	S ¹
<i>HPRT1</i>	0.281	<i>TBP</i>	0.332	<i>ACTB</i>	0.462	<i>TBP</i>	0.561	<i>TBP</i>	1.682
<i>TBP</i>	0.296	<i>SDHA</i>	0.34	<i>GAPDH</i>	0.489	<i>SDHA</i>	0.585	<i>SDHA</i>	2.06
<i>SDHA</i>	0.361	<i>HPRT1</i>	0.385	<i>SDHA</i>	0.515	<i>HPRT1</i>	0.59	<i>HPRT1</i>	2.59
<i>ACTB</i>	0.496	<i>ACTB</i>	0.502	<i>TBP</i>	0.528	<i>ACTB</i>	0.667	<i>ACTB</i>	2.828
<i>GAPDH</i>	0.499	<i>GAPDH</i>	0.545	<i>HPRT1</i>	0.623	<i>GAPDH</i>	0.68	<i>GAPDH</i>	3.976
<i>B2M</i>	0.829	<i>B2M</i>	0.661	<i>B2M</i>	0.907	<i>B2M</i>	0.912	<i>B2M</i>	6

¹stability value calculated using different algorithms and interpreted as follows: the lower S, the more stable the expression of a given gene. Genes listed in descending order.

Fig. 1. Expression variability of selected reference genes between controls and patients with CRC or active IBD. Bars present intra-group variability in each of studied groups calculated using NormFinder

CRC – colorectal cancer; IBD – inflammatory bowel disease.

Fig. 2. Determination of optimal number of reference genes to be used as normalizers in studies on whole blood transcriptome in IBD and CRC

Optimal number was determined using geNorm algorithm based on pairwise variation analysis. Bars represent the magnitude of the change in normalization factor following the inclusion of an additional reference gene. GeNorm V values exceeding 0.15 are indicative of a significant effect, pointing at the necessity to include the added gene in a panel of normalizers.

evaluated using the 4 most popular algorithms: NormFinder, geNorm, BestKeeper, and comparative ΔC_t method, and integrated using RefFinder in a form of comprehensive ranking (Table 2). Although the exact order differed, most approaches yielded concordant results with *SDHA*, *TBP* and *HPRT1* expression found more stable than that of *GAPDH* or *ACTB*. Regardless of the evaluation method, *B2M* was uniformly found to be the least stable gene in the sample set investigated (Table 2). It also displayed the highest variability within each of the examined cohort, while *GAPDH* and *ACTB* expression varied particularly in IBD group (Fig. 1).

GeNorm algorithm was employed to evaluate the optimal number of reference genes to be used as normalizers.

For studies on whole blood transcriptome in mixed cohorts of CRC and IBD patients or for studies on IBD patients, respective V value dropped below 0.15 (threshold) when normalization factors based on the 3 or 4 most stable reference genes were compared, indicating that normalization of target gene expression should be based on 3 most stable reference genes (Fig. 2), that is, *SDHA*, *TBP* and *HPRT1*. For studies on CRC patients, normalization against *SDHA* and *TBP* was found sufficient.

Validation of selected normalizers

The performance of reference genes was tested on 2 target genes, *CCL4* and *IL-1 β* , known to be differently expressed in CRC and IBD. Results obtained using relative quantification method in which target gene expression was normalized either against geometric mean of *SDHA*, *TBP* and *HPRT1* or against individual popular normalizers – *ACTB*, *B2M* or *GAPDH*, were compared with these yielded by absolute quantification method with transcript copy number.

CCL4 was significantly overexpressed in CRC as compared to IBD and insignificantly downregulated in IBD as compared to controls (Fig. 3A). While results obtained using relative quantification with *SDHA*, *TBP* and *HPRT1* as normalizers were concordant with those obtained using absolute quantification, normalization against *B2M* and *GAPDH* substantially overestimated between-group differences in *CCL4* expression.

IL-1 β was significantly underexpressed in CRC as compared to IBD, while the slight upregulation in IBD as compared to controls was not significant (Fig. 3B). Again, the results obtained with the absolute method and using the proposed panel of reference genes were concordant, while those obtained using either *B2M* or *GAPDH* underestimated CRC-IBD difference and wrongly implied the downregulation of *IL-1 β* in IBD as compared to controls. For both target genes, of the most popular reference genes, normalization against *ACTB* yielded the closest results to those obtained with absolute quantification or by using a panel of normalizers.

Effect of inflammation and cancer on *GAPDH*, *ACTB* and *B2M* expression in whole blood

Using the panel of normalizers validated here, we compared the effect, if any, of bowel inflammation and cancer on the whole blood expression of *ACTB*, *B2M* and *GAPDH*. As depicted in Table 3, both *B2M* and *GAPDH* expressions differed significantly between groups. *GAPDH* expression was regulated in active IBD as compared to controls by 1.7-fold (95% CI = 1.3–2.1; $p < 0.001$). It was also upregulated as compared to CRC – by 1.8-fold (1.3–2.5; $p < 0.001$). *B2M* expression in active IBD was upregulated by 1.5-fold (1.1–2; $p = 0.009$) and by 1.5-fold (1.1–1.9;

Fig. 3. Expression of target genes in CRC and IBD as compared to controls quantified using absolute and relative quantification methods with various reference genes as normalizers: (a) *CCL4*; (b) *IL-1β*

Light grey bars represent geometric means of normalized relative quantities (scale on left Y axis) calculated for colorectal cancer patients (denoted as CRC_) or inflammatory bowel disease patients (IBD_) using geometric mean of *SDHA*, *TBP* and *HPRT1* expression as reference (_S/T/H) or individual expression of *B2M* (_B), *ACTB* (_A), or *GAPDH* (_G) as reference. Dark grey bars represent geometric means of absolute quantities (scale on right Y axis; denoted as "_abs") obtained from standard curves based on serial dilutions of plasmid DNA with *CCL4* and *IL-1β* inserts. Expression ratios, CRC-to-IBD and controls-to-IBD (for *CCL4* in part A) or IBD-to-CRC and IBD-to-controls (for *IL-1β* in part B), are given above the bars. Statistical significance ($p < 0.05$) is denoted with asterisks.

Table 3. Relationship between the whole blood expression of *ACTB*, *B2M* and *GAPDH*, and the disease characteristics

Characteristics	n	<i>ACTB</i>	<i>B2M</i>	<i>GAPDH</i>
Cohort:		$p = 0.127$	$p = 0.001$	$p < 0.001$
controls	21	0.87 (0.78–0.98)	0.89 (0.73–1.1) ^{1,2}	0.79 (0.69–0.91) ^{1,2}
CRC	18	1.01 (0.85–1.21)	0.75 (0.6–0.94) ^{2,3}	0.9 (0.76–1.05) ^{2,3}
active IBD	25	1.11 (0.9–1.36)	1.35 (1.07–1.69) ^{1,3}	1.32 (1.09–1.6) ^{1,3}
CRC stage		$\rho = 0.52, p = 0.029$	$\rho = 0.02, p = 0.951$	$\rho = 0.53, p = 0.025$
Local progression (T)		$\rho = 0.16, p = 0.520$	$\rho = 0, p = 0.965$	$\rho = 0, p = 0.972$
Lymph node status (N):		$p = 0.084$	$p = 0.958$	$p = 0.051$
N0	7	0.89 (0.7–1.07)	0.77 (0.51–1.17)	0.79 (0.65–0.92)
N1/2	10	1.2 (0.91–1.49)	0.78 (0.56–1.09)	1.05 (0.8–1.3)
Distant metastases (M):		$p = 0.242$	$p = 0.561$	$p = 0.097$
M0	16	1.04 (0.85–1.22)	0.74 (0.58–0.94)	0.9 (0.75–1.05) ^{2,3}
M1	2	1.36	1.36	1.27
IBD phenotype:		$p = 0.347$	$p = 0.339$	$p = 0.755$
CD	14	1.22 (1.01–1.46)	1.22 (0.85–1.8)	1.28 (1.08–1.53)
UC	11	0.99 (0.63–1.54)	1.52 (1.19–1.95)	1.37 (0.9–2.07)
IBD activity ⁴ :				
CD: CDAI	20	$\rho = 0.44, p = 0.052$	$\rho = 0.36, p = 0.118$	$\rho = 0.57, p = 0.009$
UC: RI	12	$\rho = 0.34, p = 0.274$	$\rho = 0, p = 1$	$\rho = 0.72, p = 0.009$

¹significantly different from CRC; ²significantly different from active IBD; ³significantly different from controls; ⁴correlation with clinical activity was calculated for all patients for whom it was available, also these with clinically inactive disease (CDAI < 150 or RI < 6). Data presented as means of normalized relative quantities (against geometric mean of *SDHA*, *TBP* and *HPRT1*) with 95% confidence intervals (95% CI) or as Spearman rank correlation coefficients (ρ). CD – Crohn's disease; UC – ulcerative colitis; CDAI – Crohn's disease activity index; RI – Rachmilewitz index.

$p = 0.004$) as compared to controls and CRC, respectively. While *GAPDH* in CRC was significantly higher than in controls, the expression of *B2M* was significantly lower. Between-group differences in *ACTB* expression did not reach statistical significance when analyzed with ANOVA and Tukey–Kramer post-hoc test, but when the controls and IBD patients were compared using t-test for independent samples, the difference was significant ($p = 0.043$).

In CRC, *ACTB* and *GAPDH* expressions increased along with the disease advancement, being associated with the metastatic potential rather than local progression. In IBD, gene expression was not affected by the disease phenotype (CD vs UC), but *ACTB* positively correlated with clinical activity of CD and *GAPDH* also with the activity of UC.

Table 4. Association between the whole blood expression of *ACTB*, *B2M* and *GAPDH*, and the systemic levels of inflammatory markers and growth factors

Marker	<i>ACTB</i>	<i>B2M</i>	<i>GAPDH</i>
CRP	$r = 0.51, p = 0.012^1$	–	$r = 0.63, p = 0.001^1$
IL-1 β	$r = 0.64, p = 0.010^2$	–	$r = 0.74, p < 0.0001$
IL-4	–	–	$p = 0.62, p = 0.053^1$
IL-6	$r = 0.54, p = 0.032^2$	–	$r = 0.63, p = 0.009^2$
IL-8	$r = 0.49, p = 0.053^2$	–	$r = 0.64, p = 0.046^1$
IL-12	–	–	$r = 0.71, p = 0.023^1$
FGF2	$r = 0.65, p = 0.007^2$	$r = -0.49, p = 0.055^2$ $r = 0.56, p = 0.094^1$	$r = 0.53, p = 0.035^2$
G-CSF	$r = 0.66, p = 0.005^2$	$r = -0.56, p = 0.025^2$	$r = 0.81, p < 0.0001$
GM-CSF	$r = 0.65, p = 0.006^2$	–	$r = 0.43, p = 0.096^2$
MIP-1 α	$r = 0.63, p = 0.009^2$	$r = 0.65, p = 0.044^1$	$r = 0.54, p = 0.032^2$
TNF- α	$r = 0.57, p = 0.020^2$	$r = -0.65, p = 0.006^2$	$r = 0.79, p < 0.001$

¹correlation observed in IBD patients; ²correlation observed in CRC patients. Data on circulating cytokines and growth factors was available for 26 patients (16 with CRC and 10 with IBD). Data presented as Pearson correlation (r) or Spearman rank correlation (p) coefficients.

GAPDH, *ACTB* and *B2M* correlation with inflammatory and angiogenic indices

To further explore the association between popular reference genes and inflammation or tumor angiogenic potential, we investigated whether their expression was related to the levels of *IL-1 β* and *CCL4* transcripts in whole blood and whether there was a correlation with circulating mediators of inflammation and angiogenesis.

Concerning whole blood, *ACTB* positively correlated with *CCL4*, exclusively in CRC patients ($r = 0.57, p = 0.014$), and with *IL-1 β* in active IBD ($r = 0.41, p = 0.42$). It also tended to correlate with *IL-1 β* in CRC ($r = 0.45, p = 0.061$). *B2M* correlated with *CCL4* and *IL-1 β* neither in CRC nor in IBD patients. *GAPDH*, in turn, correlated exclusively with *CCL4* expression: positively in CRC ($r = 0.58, p = 0.013$) but negatively in IBD ($r = -0.45, p = 0.023$).

We also examined possible associations between the expression of popular reference genes in whole blood and the levels of circulating inflammatory and angiogenic indices and hematopoietic factors: CRP, IL-1 β , IL-4, IL-6, IL-8, IL-12, G-CSF, GM-CSF, FGF2, MCP-1 (*CCL2*), MIP-1 α (*CCL3*), PDGF-BB, TNF- α , and VEGF-A. As shown in Table 4, *ACTB* and *GAPDH* both correlated with a marker of systemic inflammation, CRP, yet exclusively in IBD patients. *ACTB* displayed a number of positive correlations with inflammatory indices, which, however, were significant solely in CRC patients. *GAPDH*, in turn, correlated with IL-6, FGF2, GM-CSF, and MIP-1 α exclusively in CRC, with IL-4, IL-8, and IL-12 solely in IBD and with IL-1 β , G-CSF, and TNF- α in both groups of patients. *B2M* expression was associated only with FGF2 and MIP-1 α in IBD, and with G-CSF, FGF2 and TNF- α in CRC. While the observed associations of *B2M* were positive in IBD, they were negative in CRC patients. None of the examined popular reference genes correlated with circulating MCP-1, VEGF-A or PDGF-BB.

Discussion

A body of evidence has been gathered showing that there is no universal reference gene(s) that would be suitable for all tissues or pathophysiological conditions and that using invalidated normalizers, the expression of which is regulated under experimental conditions, might compromise the study leading to erroneous conclusions.^{1,6} There are no previous reports validating reference genes for whole blood analysis (encompassing a mixture of RNA from numerous blood cell types) in a mixed cohort including patients with inflammatory and cancer disease. Thus, we started our research from selection and validation of potential reference genes. As shown by the literature survey, *GAPDH*, *ACTB* and *B2M* are the most popular reference genes, frequently used singlehandedly. Of these, *GAPDH* was rather poorly rated in blood-derived RNA studies, displaying high variability in neonatal whole blood with varying degrees of hypoxia and acidosis,¹⁹ whole blood of patients with multiple sclerosis,²⁰ in neutrophils from healthy individuals,²¹ and in platelets²² or peripheral blood mononuclear cells²³ from patients with a history of cardiac inflammatory conditions. It also fluctuated during the activation of T lymphocytes²⁴ or LPS-stimulation of monocytes.²⁵ Under some of these conditions, *ACTB*^{19,21,22,26} or *B2M*^{20,21,25} have been expressed stably, but similarly to *GAPDH*, *ACTB* expression has been altered upon the activation of lymphocytes^{24,27} or monocytes.²⁵ In reticulocytes, in turn, *GAPDH* and not *B2M* have been found to be a better reference gene.¹⁸ Of other evaluated reference genes, *HPRT1*,^{18,21,26} *TBP*²⁵ and *SDHA*¹⁸ have been repeatedly ranked well and, as such, they have been included in the current study.

There are several software programs available, which, using various algorithms, are dedicated to aiding in the evaluation of expression stability. We used the 4 most popular ones: NormFinder,¹⁶ geNorm utility in qBasePLUS,¹⁵ Best-Keeper,¹⁷ and comparative ΔCt method,¹⁸ and employed

RefFinder to summarize the results. Except for BestKeeper, all algorithms yielded concordant results with *HPRT1*, *SDHA* and *TBP*, outperforming *ACTB*, *B2M* and *GAPDH*. *B2M* was uniformly found the least stable of evaluated reference genes. All genes displayed the highest level of variability in IBD. Consequently, in studies involving patients with active disease normalization against 3 genes is optimal, while 2 – *TBP* and *SDHA* – are sufficient for CRC patients. To validate selected reference genes as normalizers, we analyzed expression of 2 target genes – *IL-1β* and *CCL4* – using both absolute and relative quantification, the latter employing either a set of stable reference genes (*HPRT1/SDHA/TBP*) or *ACTB*, *B2M* and *GAPDH* individually. Only normalization against a set of reference genes yielded results of relative *IL-1β* and *CCL4* expression analysis comparable to these obtained with absolute quantification of copy number. Normalization against *GAPDH* or *B2M* substantially overestimated between-group differences in *CCL4* and underestimated differences in *IL-1β* expression, rendering these genes not suitable for whole blood expression analysis in IBD and CRC patients. However, the results obtained with *ACTB* as a reference did not differ much, although normalization against *ACTB* overestimated *CCL4* downregulation in active IBD, falsely signifying the difference.

Random fluctuations in reference genes are likely to mask subtle alterations in expression of target genes but directional changes in genes used for normalization might result in invalid conclusions.⁶ It is evident now that, contrary to previous assumptions, *GAPDH* expression is tightly regulated at transcriptional and posttranslational levels.^{28,29} Consistently, several studies have demonstrated that *GAPDH* might be upregulated in response to hypoxia, a common feature of cancer and inflammation.³⁰ Moreover, it has been implied that *GAPDH* overexpression may constitute a growth advantage for tumor cells.³¹ However, the effect of inflammation and cancer on whole blood expression of *GAPDH* as well as that of *ACTB* or *B2M* has not been evaluated yet. Hence, having validated set of stably expressed reference genes, we examined whether variability in *GAPDH*, *B2M* and *ACTB* in whole blood of CRC and IBD patients was random or specifically affected by the examined conditions.

Complementing the findings on *GAPDH* overexpression in colonic tumors at protein level,³² we found *GAPDH* expression to be upregulated in whole blood of CRC patients. The elevation was even more profound in IBD patients. Moreover, we showed that *GAPDH* upregulation was not random but corresponded with clinical activity of both CD and UC in IBD patients as well as with cancer advancement in CRC patients, in whom it tended to be associated with gaining metastatic potential rather than with local progression of the tumor. Furthermore, it positively and strongly correlated with a number of circulating inflammatory indices, either exclusively in IBD patients (CRP, IL-8, IL-12) or CRC patients (IL-6, FGF2, MIP-1α) or both

(IL-1β, G-CSF, TNF-α), most of which display pro-angiogenic properties as well. As all these cytokines are inter-related, it is likely that some of *GAPDH* associations are mediated by others. Unfortunately, the number of observations is not sufficient to allow for a multivariate analysis, impeding identification of variables independently associated with whole blood *GAPDH* expression.

Our observations are consistent with the phenomena of tumor metabolism reprogramming, which occurs to satisfy increased energetic demands of proliferating cancer cells under diminished oxygen availability. One of its manifestations is the acceleration of glycolysis upon combined actions of mutations in c-myc proto-oncogene and the up-regulation of hypoxia-inducible factor (HIF)-1α, the binding sites which can be found in *GAPDH* promoter.²⁹ It also contains a binding site for c-jun/AP-1, transcription factor involved in inflammatory responses³³ as well as tumor angiogenesis.³⁴ Accordingly, new non-glycolytic functions have recently been attributed to *GAPDH*, e.g., in cell death regulation. Although the exact role it may play remains unclear, as *GAPDH* has been demonstrated to both induce cancer cell senescence³⁵ and contribute to cancer progression²⁸ in colon adenocarcinoma cell lines, *GAPDH* has been demonstrated to display pro-survival activities.³⁶

B2M was uniformly selected as the least stable reference gene in the whole cohort, independently from statistical method of evaluation. It also displayed the highest degree of variability within each study group. Similarly to our previous observations in colorectal tissue (manuscript submitted), whole blood *B2M* was significantly down-regulated in cancer but upregulated in active IBD. Accordingly, Bianchini et al.³⁷ identified *B2M* as the most downregulated gene from among 19.2K genes screened and suggested that the phenomenon might be related to the escape mechanism from NK-mediated lysis developed by tumor cells. Shrout et al.³ showed that *B2M* down-regulation in colorectal tumors was a strong prognostic indicator of lymph node metastases. In the investigated sample set, whole blood *B2M* did not correspond with cancer advancement nor was it correlated with clinical scores of IBD activity. Nevertheless, consistently with its observed underexpression in CRC and overexpression in IBD, *B2M* expression negatively correlated with FGF2, G-CSF and TNF-α in the former and positively with FGF2 and MIP-1α in the latter.

Expression of *ACTB* in whole blood was the least variable of the 3 most popular reference genes, and normalization against *ACTB* was even found superior using BestKeeper. However, although the differences between groups did not reach statistical significance when analyzed with ANOVA, *ACTB* clearly tended to be higher in active IBD than in controls (the difference was significant when t-test was used). Moreover, it positively correlated with CRP levels and clinical severity score in CD, rendering it unsuitable as a reference in IBD. In CRC, in turn, it tended to be higher in patients with lymph node metastases and

positively correlated with inflammatory and angiogenic factors: IL-1 β , IL-6, FGF2, G-CSF, GM-CSF, MIP-1 α , and TNF- α .

Taken together, our results demonstrate that none of the most popular reference genes, namely, *GAPDH*, *ACTB* or *B2M*, are suitable for transcriptional studies on whole blood from CRC and/or IBD patients. Firstly, their expression differs between conditions, showing that bowel inflammation and cancer affect not only their tissue,¹³ but also whole blood expression. Even more importantly, however, the observed variability is not random, as the alterations in their expression levels correspond, to a varying degree, with CRC advancement and the clinical activity of IBD (*GAPDH* and *ACTB*) as well as correlate with the levels of circulating mediators of inflammation and angiogenesis. The clinical relevance of this observation needs to be elucidated; however, regardless of whether the *ACTB*, *B2M*, and *GAPDH* regulation is unintended or plays an active role in inflammatory responses and/or cancer, they definitely does not constitute good internal controls. The set of normalizers proposed here instead, that is, *TBP*, *SDHA* and *HPRT1*, have recently been successfully used by our group for the analysis of whole blood transcriptome of IBD patients.^{38,39}

Conclusions

The expression of frequently used normalizers for whole blood transcriptomic analysis, that is, *GAPDH*, *ACTB*, and *B2M*, is directionally affected by bowel inflammation and cancer, rendering them unsuitable as references in CRC and IBD. The expression of *HPRT1*, *SDHA* and *TBP* was stable across CRC and IBD patients, allowing for their recommendation as normalizers in studies involving both groups of patients.

ORCID iDs

Iwona Bednarz-Misa <https://orcid.org/0000-0001-7244-2017>
 Katarzyna Neubauer <https://orcid.org/0000-0003-3650-9311>
 Ewa Zacharska <https://orcid.org/0000-0002-1840-9731>
 Bartosz Kapturkiewicz <https://orcid.org/0000-0002-2894-4834>
 Małgorzata Krzystek-Korpacka <https://orcid.org/0000-0002-2753-8092>

References

- Bustin SA, Murphy J. RNA biomarkers in colorectal cancer. *Methods*. 2013;59(1):116–125.
- Rubie C, Kempf K, Hans J, et al. Housekeeping gene variability in normal and cancerous colorectal, pancreatic, esophageal, gastric and hepatic tissues. *Mol Cell Probes*. 2005;19(2):101–109.
- Shrout J, Yousefzadeh M, Dodd A, et al. Beta(2)microglobulin mRNA expression levels are prognostic for lymph node metastasis in colorectal cancer patients. *Br J Cancer*. 2008;98(12):1999–2005.
- Glare EM, Divjak M, Bailey MJ. β -Actin and GAPDH housekeeping gene expression in asthmatic airways is variable and not suitable for normalising mRNA levels. *Thorax*. 2002;57(9):765–770.
- Congiu M, Slavin JL, Desmond PV. Expression of common housekeeping genes is affected by disease in human hepatitis C virus-infected liver. *Liver Int*. 2011;31(3):386–390.
- Dheda K, Huggett JF, Bustin SA, Johnson MA, Rook G, Zumla A. Validation of housekeeping genes for normalizing RNA expression in real-time PCR. *Biotechniques*. 2004;37(1):112–119.
- Radonić A, Thulke S, Mackay IM, Landt O, Siebert W, Nitsche A. Guideline to reference gene selection for quantitative real-time PCR. *Biochem Biophys Res Commun*. 2004;313(4):856–862.
- Vartanian K, Slottke R, Johnstone T, et al. Gene expression profiling of whole blood: Comparison of target preparation methods for accurate and reproducible microarray analysis. *BMC Genomics*. 2009;10:2.
- Ghosh S, Dent R, Harper ME, Gorman SA, Stuart JS, McPherson R. Gene expression profiling in whole blood identifies distinct biological pathways associated with obesity. *BMC Med Genomics*. 2010;3:56.
- Batliwalla FM, Baechler EC, Xiao X, et al. Peripheral blood gene expression profiling in rheumatoid arthritis. *Genes Immun*. 2005;6(5):388–397.
- Burczynski ME, Peterson RL, Twine NC, et al. Molecular classification of Crohn's disease and ulcerative colitis patients using transcriptional profiles in peripheral blood mononuclear cells. *J Mol Diag*. 2006;8(1):51–61.
- DePrimo SE, Wong LM, Khatry DB, et al. Expression profiling of blood samples from an SU5416 Phase III metastatic colorectal cancer clinical trial: A novel strategy for biomarker identification. *BMC Cancer*. 2003;3:3.
- Krzystek-Korpacka M, Diakowska D, Bania J, Gamian A. Expression stability of common housekeeping genes is differently affected by bowel inflammation and cancer: Implications for finding suitable normalizers for inflammatory bowel disease studies. *Inflamm Bowel Dis*. 2014;20(7):1147–1156.
- Krzystek-Korpacka M, Diakowska D, Kapturkiewicz B, Bębenek M, Gamian A. Profiles of circulating inflammatory cytokines in colorectal cancer (CRC), high cancer risk conditions, and health are distinct. *Cancer Lett*. 2013;337(1):107–114.
- Vandesompele J, De Preter K, Pattyn F, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol*. 2002;3(7):RESEARCH0034.
- Andersen CL, Jensen JL, 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*. 2004;64(15):5245–5250.
- Pfaffl MW, Tichopad A, Prgomet C, Neuvians TP. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper-Excel-based tool using pair-wise correlations. *Biotechnol Lett*. 2004;26(6):509–515.
- Silver N, Best S, Jiang J, Thein SL. Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR. *BMC Mol Biol*. 2006;7:33.
- Maron JL, Arya MA, Seefeld KJ, et al. pH but not hypoxia affects neonatal gene expression: Relevance for housekeeping gene selection. *Matern Fetal Neonatal Med*. 2008;21(7):443–447.
- Stamova BS, Apperson M, Walker WL, et al. Identification and validation of suitable endogenous reference genes for gene expression studies in human peripheral blood. *BMC Med Genomics*. 2009;2:49.
- Zhang X, Ding L, Sandford AJ. Selection of reference genes for gene expression studies in human neutrophils by real-time PCR. *BMC Mol Biol*. 2005;6:4.
- Zsóri KS, Muszbek L, Csiki Z, Shemirani AH. Validation of reference genes for the determination of platelet transcript level in healthy individuals and in patients with the history of myocardial infarction. *Int J Mol Sci*. 2013;14(2):3456–3466.
- Dabek J, Wilczok J, Kulach A, Gasior Z. Altered transcriptional activity of gene encoding GAPDH in peripheral blood mononuclear cells from patients with cardiac syndrome X – an important part in pathology of microvascular angina? *Arch Med Sci*. 2010;6(5):709–712.
- Bas A, Forsberg G, Hammarström S, Hammarström ML. Utility of the housekeeping genes 18S rRNA, beta-actin and glyceraldehyde-3-phosphate-dehydrogenase for normalization in real-time quantitative reverse transcriptase-polymerase chain reaction analysis of gene expression in human T lymphocytes. *Scand J Immunol*. 2004;59(6):566–573.
- Piehler AP, Grimholt RM, Ovstebø R, Berg JP. Gene expression results in lipopolysaccharide-stimulated monocytes depend significantly on the choice of reference genes. *BMC Immunol*. 2010;11:21.
- Spinsanti G, Zannolli R, Panti C, et al. Quantitative real-time PCR detection of TRPV1-4 gene expression in human leukocytes from healthy and hyposensitive subjects. *Mol Pain*. 2008;4:51.

27. Røge R, Thorsen J, Tørring C, Ozbay A, Møller BK, Carstens J. Commonly used reference genes are actively regulated in in vitro stimulated lymphocytes. *Scand J Immunol.* 2007;65(2):202–209.
28. Colell A, Green DR, Ricci JE. Novel roles for GAPDH in cell death and carcinogenesis. *Cell Death Differ.* 2009;16(12):1573–1581.
29. Krasnov GS, Dmitriev AA, Snezhkina AV, Kudryavtseva AV. Deregulation of glycolysis in cancer: Glyceraldehyde-3-phosphate dehydrogenase as a therapeutic target. *Expert Opin Ther Targets.* 2013;17(6):681–693.
30. Eltzschig HK, Carmeliet P. Hypoxia and inflammation. *N Engl J Med.* 2011;364(7):656–665.
31. Guo C, Liu S, Sun MZ. Novel insight into the role of GAPDH playing in tumor. *Clin Transl Oncol.* 2013;15(3):167–172.
32. Cuezva JM, Krajewska M, de Heredia ML, et al. The bioenergetic signature of cancer: A marker of tumor progression. *Cancer Res.* 2002;62(22):6674–6681.
33. Schonhaler HB, Guinea-Viniegra J, Wagner EF. Targeting inflammation by modulating the Jun/AP-1 pathway. *Ann Rheum Dis.* 2011;70(Suppl 1):i109–112.
34. Yin Y, Wang S, Sun Y, et al. JNK/AP-1 pathway is involved in tumor necrosis factor- α induced expression of vascular endothelial growth factor in MCF7 cells. *Biomed Pharmacother.* 2009;63(6):429–435.
35. Nicholls C, Pinto AR, Li H, et al. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) induces cancer cell senescence by interacting with telomerase RNA component. *Proc Natl Acad Sci U S A.* 2012;109(33):13308–13313.
36. Azam S, Jouvett N, Jilani A, et al. Human glyceraldehyde-3-phosphate dehydrogenase plays a direct role in reactivating oxidized forms of the DNA repair enzyme APE1. *J Biol Chem.* 2008;283(45):30632–30641.
37. Bianchini M, Levy E, Zucchini C, et al. Comparative study of gene expression by cDNA microarray in human colorectal cancer tissues and normal mucosa. *Int J Oncol.* 2006;29(1):83–94.
38. Neubauer K, Misa IB, Diakowska D, Kapturkiewicz B, Gamian A, Krzystek-Korpacka M. Nampt/PBEF/visfatin upregulation in colorectal tumors, mirrored in normal tissue and whole blood of colorectal cancer patients, is associated with metastasis, hypoxia, IL1 β , and anemia. *Biomed Res Int.* 2015;2015:523930.
39. Neubauer K, Bednarz-Misa I, Walecka-Zacharska E, et al. Oversecretion and overexpression of nicotinamide phosphoribosyltransferase/pre-b colony-enhancing factor/visfatin in inflammatory bowel disease reflects the disease activity, severity of inflammatory response and hypoxia. *Int J Mol Sci.* 2019;20(1):166.