

Measurable impact of RNA quality on gene expression results from quantitative PCR

Joëlle Vermeulen¹, Katleen De Preter¹, Steve Lefever¹, Justine Nuytens¹, Fanny De Vloed¹, Stefaan Derveaux^{1,2}, Jan Hellemans^{1,2}, Frank Speleman¹ and Jo Vandesompele^{1,2,*}

¹Centre for Medical Genetics, Ghent University Hospital, De Pintelaan 185, B-9000 Ghent, Belgium and

²Biogazelle, Technologiepark 3, B-9052 Ghent, Belgium

Received August 13, 2010; Revised January 24, 2011; Accepted January 25, 2011

ABSTRACT

Compromised RNA quality is suggested to lead to unreliable results in gene expression studies. Therefore, assessment of RNA integrity and purity is deemed essential prior to including samples in the analytical pipeline. This may be of particular importance when diagnostic, prognostic or therapeutic conclusions depend on such analyses. In this study, the comparative value of six RNA quality parameters was determined using a large panel of 740 primary tumour samples for which real-time quantitative PCR gene expression results were available. The tested parameters comprise of microfluidic capillary electrophoresis based 18S/28S rRNA ratio and RNA Quality Index value, *HPRT1* 5'–3' difference in quantification cycle (Cq) and *HPRT1* 3' Cq value based on a 5'/3' ratio mRNA integrity assay, the Cq value of expressed Alu repeat sequences and a normalization factor based on the mean expression level of four reference genes. Upon establishment of an innovative analytical framework to assess impact of RNA quality, we observed a measurable impact of RNA quality on the variation of the reference genes, on the significance of differential expression of prognostic marker genes between two cancer patient risk groups, and on risk classification performance using a multigene signature. This study forms the basis for further rational assessment of reverse transcription quantitative PCR based results in relation to RNA quality.

INTRODUCTION

Gene expression quantification plays a central role in a wide variety of studies, including biomedical research with clinical relevance. Among the various methods available for gene expression analysis, the reverse transcription quantitative polymerase chain reaction (RT-qPCR) is the most rapid, sensitive, accurate and precise method and its use in clinical diagnostic procedures is presently growing exponentially (1–5).

While there is conflicting literature data, it is often suggested that RNA integrity and purity are important in order to obtain reliable results (6–9). RNA degradation can occur due to inadequate sample handling, prolonged storage, suboptimal storage conditions or inter-laboratory shipment of samples (10,11). RNA may be degraded through exposure to heat or UV, or cleavage by RNase enzymes. In addition, the presence of inhibiting components such as urea, salts, phenol, heparin or other agents used during sampling or RNA extraction may also compromise with results (12). It would seem, therefore, that a rigorous assessment of RNA integrity and purity is essential before using RNA samples in downstream applications, especially if diagnostic, therapeutic or prognostic conclusions will be drawn. Unfortunately, proper RNA quality control is lacking in a substantial number of studies (4). While it is recently listed as a required element in the Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) guidelines (13), there remains a great need to explore in detail the implications of RNA quality on the final results.

Various methods have been proposed for the assessment of RNA integrity, most often through measurement of the size of the ribosomal subunit RNA molecules. Importantly though, in RT-qPCR analyses, messenger

*To whom correspondence should be addressed. Tel: +32 479 353563; Fax: +32 9 332 6549; Email: joke.vandesompele@ugent.be

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

RNA is the target and not the ribosomal RNA transcripts. Therefore, it might be more appropriate to directly assess the quality of the mRNA fraction. In addition, PCR-based methods assessing RNA quality might be more relevant given the fact that the targets are also measured using the same technology when doing RT-qPCR.

In this study, two qPCR-based assays using minute amounts of RNA for investigation of mRNA integrity or purity were first validated in RNA samples from cultured neuroblastoma cells and subsequently applied along with microfluidic-based capillary electrophoresis on a large panel of RNA samples extracted from neuroblastoma tumours recently used in a qPCR-based prognostic multigene signature validation study (14).

MATERIAL AND METHODS

Sample preparation

Total RNA was extracted from 6 neuroblastoma cell lines and 740 fresh frozen neuroblastoma tumour biopsies according to three different methods in collaborating laboratories as described in Vermeulen *et al.* (14). Concentration of each RNA sample was measured using the Nanodrop 1000 Spectrophotometer (Thermo Scientific).

Reverse transcription of RNA and cDNA synthesis was carried out using the iScript Select cDNA Synthesis Kit (Bio-Rad). In brief, 10 ng of each total RNA sample was reverse transcribed using anchored oligo-dT primers and iScript Select reverse transcriptase according to the manufacturer's instructions and subsequently diluted with nuclease-free water (Sigma) to 0.5 ng/ μ l cDNA (total RNA equivalents) and stored at -20°C .

Prior to the cDNA synthesis, a DNase treatment was performed using the RQ1 RNase-free DNase according to the manufacturer's instructions (Promega).

A sample pre-amplification method was applied starting from 20 ng total RNA from the neuroblastoma tumour samples yielding sufficient cDNA to measure more than 1000 target genes (WT-Ovation, NuGEN). Briefly, this linear and isothermal pre-amplification method starts with randomly primed whole transcriptome double-stranded cDNA synthesis followed by SPIA-based DNA pre-amplification to generate single-stranded cDNA pre-amplification product (15–17).

Assessment of RNA purity and integrity

SPUD assay. To assess the purity of the RNA samples, a quantitative PCR-based assay was used to detect PCR (enzyme) inhibitors (9). A potato nucleic acid sequence of 101 bp, lacking homology with any other known human sequence, was amplified with specific primers by real-time qPCR using an iQ5 real-time PCR detection system (Bio-Rad) (Supplementary Data). PCR amplification mixtures (15 μ l) contained SYBR Green I Master Mix buffer (7.5 μ l) (Eurogentec), 0.15 μ l fluorescein (1 μM), 2 μ l SPUD template (5000 molecules/ μ l), 0.75 μ l matching forward and reverse primer (5 μM), 2.85 μ l nuclease-free water and 1 μ l of either RNA from the sample to be tested

(10 ng/ μ l), inhibitory agent (positive control) or nuclease-free water (negative control to determine the reference Cq value). The cycling conditions comprised 10 min polymerase activation at 95°C and 40 cycles of 15 s at 95°C and of 60 s at 60°C , followed by a dissociation run from 60°C to 95°C for melting curve analysis. Subsequently, Cq values were compared to the Cq value of the negative control assay and a difference in Cq or delta-Cq (ΔCq) > 1 was used as a cut-off designating the presence of inhibitors of the SPUD assay.

The 5'/3' ratio mRNA integrity assay. This assay aims at measurement of the integrity of a reference gene mRNA that is considered to be representative of the integrity of all mRNAs in a given RNA sample (12). The principle of this qPCR-based assay is based on the fact that anchored oligo-dT primed reverse transcription proceeds from the 3' poly-A tail to the 5' start, being interrupted if mRNA is fragmented due to degradation. Two qPCR assays targeting either the 3' end or the 5' start of the low abundant reference gene *HPRT1* were developed and validated using our *in silico* analysis pipeline (18). The assays had an efficiency of 88.6% (± 0.4 SEM) and 94.9% (± 1.0 SEM), respectively, based on 6-point, 4-fold dilution series (Supplementary Data). Real-time qPCR was performed on oligo-dT primed cDNA in a 384-well plate instrument (LC480, Roche). Real-time qPCR amplifications were performed in 7.5 μ l containing 3.75 μ l 2 \times SYBR Green I master mix (Roche), 0.375 μ l forward and reverse primer (5 μM each), 1 μ l nuclease-free water and 2 μ l cDNA (1 ng total RNA equivalents). The cycling conditions were comprised of 3 min polymerase activation at 95°C and 55 cycles of 15 s at 95°C and 30 s at 60°C , followed by a dissociation curve analysis from 60°C to 95°C .

Four reference samples were tested in all runs and used as inter-run calibrators. The *HPRT1* 3' and 5' Cq values were determined and the difference in Cq value between both assays was calculated and defined as the 5'–3' ΔCq . In principle, the more degraded the RNA sample, the higher the 5'–3' ΔCq . As the 5' Cq was below detection level for various samples (hence no 5'–3' ΔCq could be calculated), the 3' Cq value in itself was also evaluated as an alternative RNA quality parameter.

Microfluidic capillary electrophoresis. About 1 ng of each total RNA isolate was analysed on a High Sensitivity Chip (Experion, software version 3.0, Bio-Rad) in order to determine a 18S/28S rRNA ratio and an RNA quality index (RQI) (according to the manufacturer's instructions).

Alu expression. Alu repeat sequences are the most abundant repeats in the human genome (~ 1 million copies). By random integration during evolution, these repeats got embedded in the 3'-UTR of thousands of coding genes and can thus be used as a reference for the amount of mRNA (19). An RT-qPCR assay was designed and validated in-house (forward primer: CATGGTGAAACCCCGTCTCTA, reverse primer: GCCTCAGCCTCCGAGTAG) and RT-qPCR was performed in a 384-well

plate instrument (LC480, Roche) using the same conditions as described above.

Normalization factor. A normalization factor was calculated based on the arithmetic mean Cq value of four stably expressed reference genes (*HMBS*, *HPRT1*, *SDHA*, *UBC*) (20). In principle, the more degraded the RNA sample, the higher the normalization factor is.

Gene expression analysis of prognostic marker genes

Gene expression analysis was performed according to a procedure described elsewhere (14). In brief, a qPCR assay was designed for 59 prognostic genes and 5 reference sequences (*HMBS*, *HPRT1*, *SDHA*, *UBC* and Alu) and validated using our *in silico* analysis pipeline (18) (Supplementary Data). Real-time qPCR was performed in a 384-well plate instrument (LC480, Roche) and data analysed using qbase^{PLUS} 1.4 [http://www.qbaseplus.com,(21)].

Statistical analysis

Assessment of impact of RNA quality on qPCR results was performed with six individual quality parameters, including 18S/28S rRNA ratio, RQI, *HPRT1* 5'–3' dCq, *HPRT1* 3' Cq value, Alu Cq value and a normalization factor based on the arithmetic mean Cq value of four reference genes (*HMBS*, *HPRT1*, *SDHA*, *UBC*).

The gene-specific variation for each reference gene was calculated as the standard deviation of the log₂ transformed expression of the reference gene under study normalized with the exponentiated arithmetic mean Cq value of the three other reference genes. The Mann–Whitney test was used to measure the significance of differential expression of a marker gene between the HR and non-HR subgroups. The HR subgroup comprised neuroblastoma patients >12 months at diagnosis with INSS Stage 4 tumours (irrespective of MYCN status) or with INSS Stages 2 and 3 tumours with MYCN amplification and patients younger than 12 months with INSS Stages 2–4 tumours with MYCN amplification. The non-HR subgroup comprised of all other patients. Reference gene variation and significance of differential expressions were measured in sample groups with increasing numbers of either best quality samples or worst quality samples starting with 50 samples up to the total number of samples analysed. This resulted in two curves, one for each quality group, denoting the difference of the reference gene expression variation or significance of differential expression between the two RNA quality groups in function of increasing group size. The mean difference between the best quality and worst quality sample groups was calculated as the area between the curves divided by the numbers of group comparisons (i.e. the total number of intervals across which the reference gene expression variation was calculated) and constitutes a measure for the impact of the RNA quality parameter under investigation on the results (reference gene expression variation or significance of differential expression).

A null distribution interval was determined upon 100 random permutations and calculation of a 90%

distribution interval (1.64*standard deviation). Values outside this distribution are significantly different ($P < 0.05$, one-sided).

A multigene expression signature was built using 30 training samples using either the Prediction Analysis of Microarrays (PAM) method (22) or the correlation signature method (23) and tested on the remaining samples. The R-language for statistical computing (version 2.6.2) was used to train and test the prognostic signature and to evaluate its performance by receiver operating characteristic (ROC) area under the curve (AUC) analyses using the Bioconductor MCR estimate (24) and the ROC packages, respectively.

Multivariate logistic regression analysis was performed using SPSS (version 17). Therefore the six RNA quality parameter values were first re-scaled between 0 (best RNA quality) and 1 (worst RNA quality) to make them comparable.

Data availability

The experimental data are available as Supplementary Data in RDML (Real-time PCR Data Markup Language) format, a structured and universal data standard for exchanging quantitative PCR (qPCR) data (25) (http://www.rdml.org) according to the MIQE guidelines (13). The primer sequences and target genes of each assay are available in the rdml files together with their corresponding RTPrimerDB ID (18) (http://www.rtpriimerdb.org).

RESULTS

Validation of the methods for RNA purity and integrity assessment

SPUD assay. A dilution series consisting of six 10-fold serial dilution points, starting from 1 000 000 molecules down to 10 molecules of SPUD template was created using 10 ng/μl yeast tRNA as carrier. Heparin (0.4 U/ml) was added as inhibitory agent to the individual wells. Figure 1 shows that addition of heparin results in PCR (enzyme) inhibition leading to higher Cq values compared to the reference Cq value of the negative control (Figure 1a) and to a difference in Cq or dCq > 1 for each dilution point (Figure 1b) (Supplementary Data). This agent thus constitutes an useful positive control for PCR inhibition of an assay containing 10 000 SPUD molecules.

The 5'/3' ratio mRNA integrity assay. Six RNA samples from cultured neuroblastoma cells were artificially degraded by heat exposure (10 or 20 min at 80°C) and subsequently subjected to 5'/3' ratio mRNA integrity and microfluidic electrophoresis analyses. A clear increase in 5'–3' dCq was noticed upon exposure of the originally high quality RNA samples to an elevated temperature (Figure 2a) (Supplementary Data). Total RNA electropherograms were in accordance with the above described results showing a progressive reduction in size of the 18S and 28S peaks and an elevation of the base line

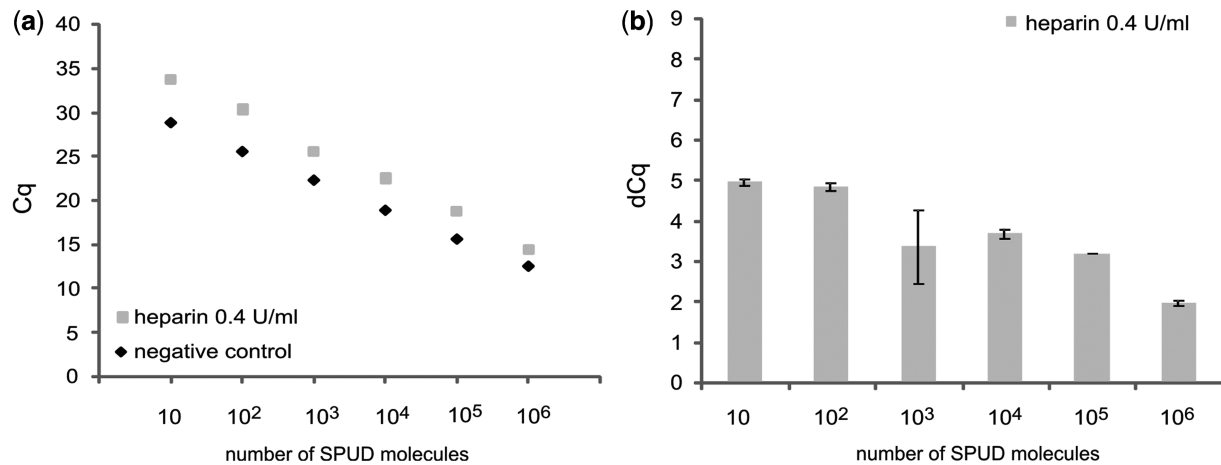


Figure 1. qPCR-based SPUD assay for the detection of PCR (enzyme) inhibitors. (a) Dilution series consisting of six 10-fold serial dilution points, starting from 1 000 000 molecules down to 10 molecules of SPUD template. Individual wells were supplemented with water (negative control; four replicates for each dilution point) to determine the reference Cq value shown as black diamonds, or with an inhibitory agent (heparin, two replicates for each dilution point; grey squares) leading to an increase in Cq values. (b) Difference in Cq (dCq) (y-axis) between the Cq values of the control assays performed in the presence of water and the Cq values of the assays containing the inhibitory agent. The dCq values >1 indicate the presence of PCR (enzyme) inhibitors.

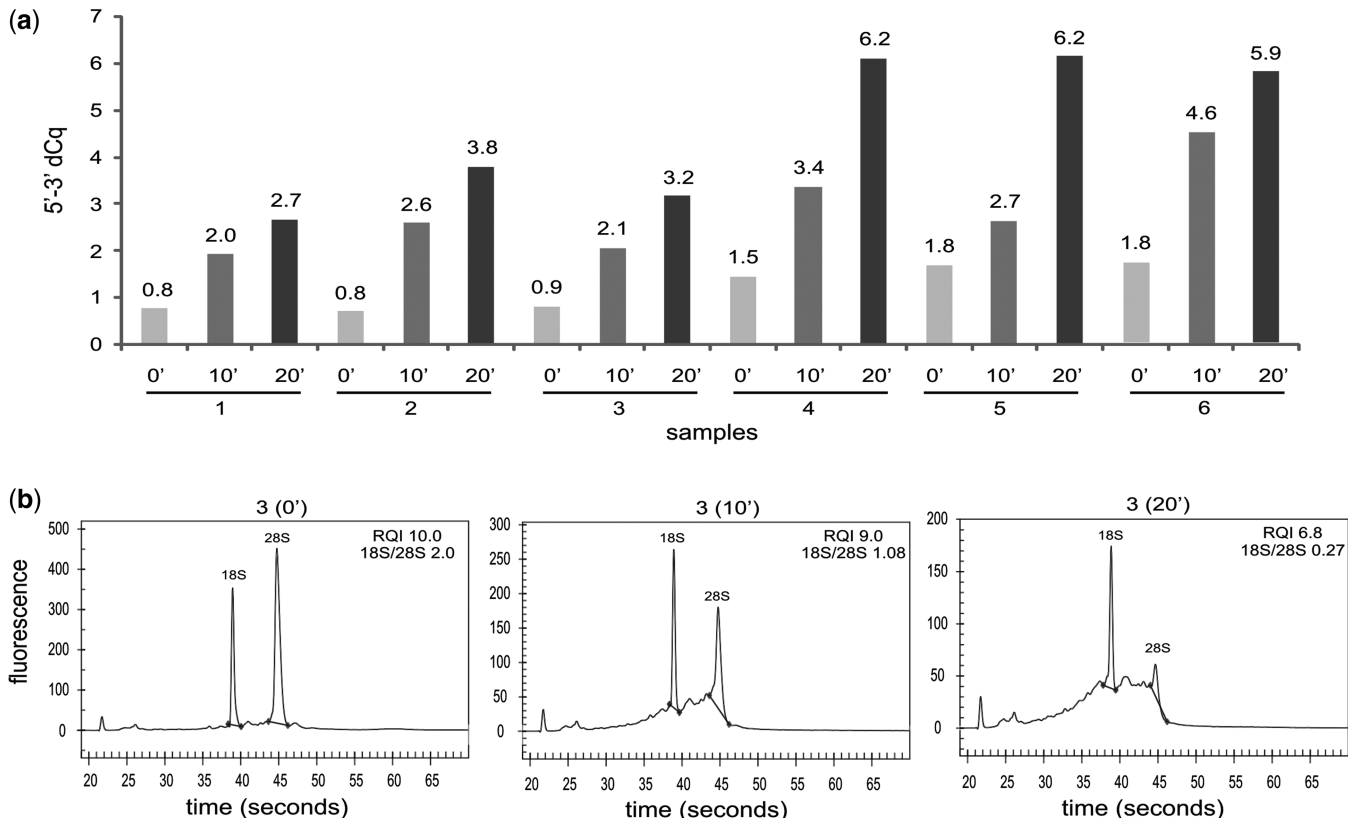


Figure 2. (a) qPCR based 5'/3' ratio mRNA integrity assay for 6 RNA samples from cultured neuroblastoma cells artificially degraded by heat exposure (at 80°C). The x-axis represents the samples (intact, 10 min heat exposure, 20 min heat exposure). The longer the heat exposure, the more degraded the sample, the higher the difference in Cq (5'-3' dCq). (b) Microfluidic capillary electrophoresis for one representative sample. The longer the heat exposure, the more degraded the sample, the lower the RNA Quality Index (RQI).

resulting in a decrease of the RQI as shown in Figure 2b for one representative RNA sample. Assay characteristics, RT and PCR variability, and suboptimal RNA quality contribute to 5'-3' dCq values deviating from zero in the six untreated samples.

Assessment of RNA purity and integrity in a large series of tumour samples

The SPUD assay was used for the detection of inhibitors in a large panel of RNA samples extracted from fresh frozen neuroblastoma biopsies in 11 different laboratories

using 5 different protocols. Out of 740 samples, all but 8 (1.1%, confirmed in a separate run) were within 1 cycle of the negative control and thus considered to be free of inhibitors affecting the SPUD assay. Based on this, we extrapolate that the applied extraction methods generally provide RNA free of PCR inhibitors.

In order to evaluate the RNA integrity of the neuroblastoma samples, six different quality parameters were determined and compared, including 18S/28S rRNA ratio (S), RQI (R), *HPRT1* 5'–3' difference in Cq or dCq (D), *HPRT1* 3' Cq value (C), the Cq value of expressed Alu repeat sequences (A) (measure for the overall mRNA content) and a normalization factor based on the arithmetic mean Cq value of four stably expressed reference genes (N) (*HMBS*, *HPRT1*, *SDHA*, *UBC*) (Supplementary Table S1). S, R, D and C are measured on unamplified RNA; A and N on pre-amplified cDNA. The frequency distribution and cumulative frequency of the different parameters are displayed in Figure 3 and Table 1. All parameters were significantly correlated to each other (Figure 4, $P < 0.001$). Not unexpectedly, the highest correlation was found between Alu Cq (A) and the normalization factor (N) ($r^2 = 0.68$).

Influence of RNA integrity on single gene level

Variation of reference genes. In principle, normalized reference gene expression levels should be similar in all samples, hence displaying a low variation across all samples. Here, we wanted to determine if RNA quality has an impact on the remaining variation of normalized reference gene expression. To this purpose, the gene-specific variation of each reference gene (standard variation) was calculated for sample subgroups with increasing numbers of either best quality samples or worst quality samples starting with 50 samples up to all 615 samples for which expression and clinical data were available. All RNA quality parameters were investigated. RNA quality has a clear influence on the noise of the reference genes as shown in Figure 5 for one representative reference gene (*UBC*) and RNA quality parameter (*HPRT1* 5'–3' dCq). The variation in the 50 most compromised RNA samples is >2-fold in the 50 most intact samples (3.5 versus 1.5). A significant difference in reference gene variation between best quality and worst quality samples was observed for 18 out of 24 combinations of 4 reference genes with 6 RNA quality parameters (Table 2). On average, the 5'–3' dCq and the normalization factor had the largest impact on reference gene variation when comparing best quality with worst quality samples.

We also tested the effect of different combinations of the six RNA quality parameters but could not find significantly larger effects for specific combinations (data not shown).

Differential expression of prognostic marker genes. In the second step, we studied the impact of RNA quality on the significance of differential expression of marker genes between two risk groups of cancer patients (HR versus non-HR) using the six RNA quality parameters in all

samples for which clinical and expression data were available ($n = 615$). The significance of differential expression of each marker gene ($n = 59$) between the HR and non-HR group was measured using the same procedure of growing sample subgroups of opposite RNA quality as used for assessment of reference gene variation. Instead of calculating the variation, the negative 10 log of the P -value of the Mann–Whitney test was used. The results clearly show an influence of RNA quality on single gene differential expression for a substantial number of genes (Figure 6 and Supplementary Table S2). For 27 genes, significance of differential expression is higher for high-quality samples for at least one RNA quality parameter (as exemplified for 5'/3' ratio for target *SLC25A5* in Figure 6a, mean difference in significance is 8.29); 8 genes seem not to be sensitive to RNA quality (as exemplified for 5'/3' ratio assay for target *PLAT* in Figure 6b, mean difference in significance is 0.13); for 19 genes, significance of differential expression seems to be better in low-quality samples for at least one RNA quality parameter (as exemplified for 5'/3' ratio assay for target *CAMTA1* in Figure 6c, mean difference in significance is –4.95); and for 5 genes there are conflicting data. S, N and A parameters seem to result in more genes where better quality RNA leads to more significant results. R, D and C parameters seem to result in as many or more genes where there is an opposite effect.

In practice, gene expression studies are often based on small sample sizes (e.g. 15 low risk patients versus 15 high risk patients). Therefore, we also evaluated the impact of RNA quality parameters on the differential expression of genes within such an experimental set-up. For this analysis, we performed 100 000 random sampling of 15 patients from the HR group and 15 patients from the non-HR group. Within each of the sampling, we performed a Mann–Whitney test for differential expression analysis and plotted the –10 log P -values as boxplots, grouped according to the number of samples belonging to the 75% percentile of the best quality. This approach was applied on the representative genes depicted in Figure 6a–c for quality parameter D (cut-off for good quality, 75% = 3.54). The same conclusions can be obtained, i.e. for *SLC25A5*, the differential expression is more significant when testing on a higher proportion of high-quality samples (Figure 6d), while for *PLAT*, there is no bias (Figure 6e) and for *CAMTA1*, a reduction in significance can be appreciated (Figure 6f).

Influence of RNA integrity on multigene level

After establishment of a clear link between RNA quality and single gene results, we now studied the impact of RNA quality on risk classification performance using a multigene signature in the samples for which clinical data were available (30 training samples and 585 test samples) (Table 3). We used two different classification algorithms PAM (22) and correlation signature (23) and determined the sensitivity, specificity, positive predictive value, negative predictive value and overall accuracy as ROC–AUC analysis of either a 6-gene or a 59-gene

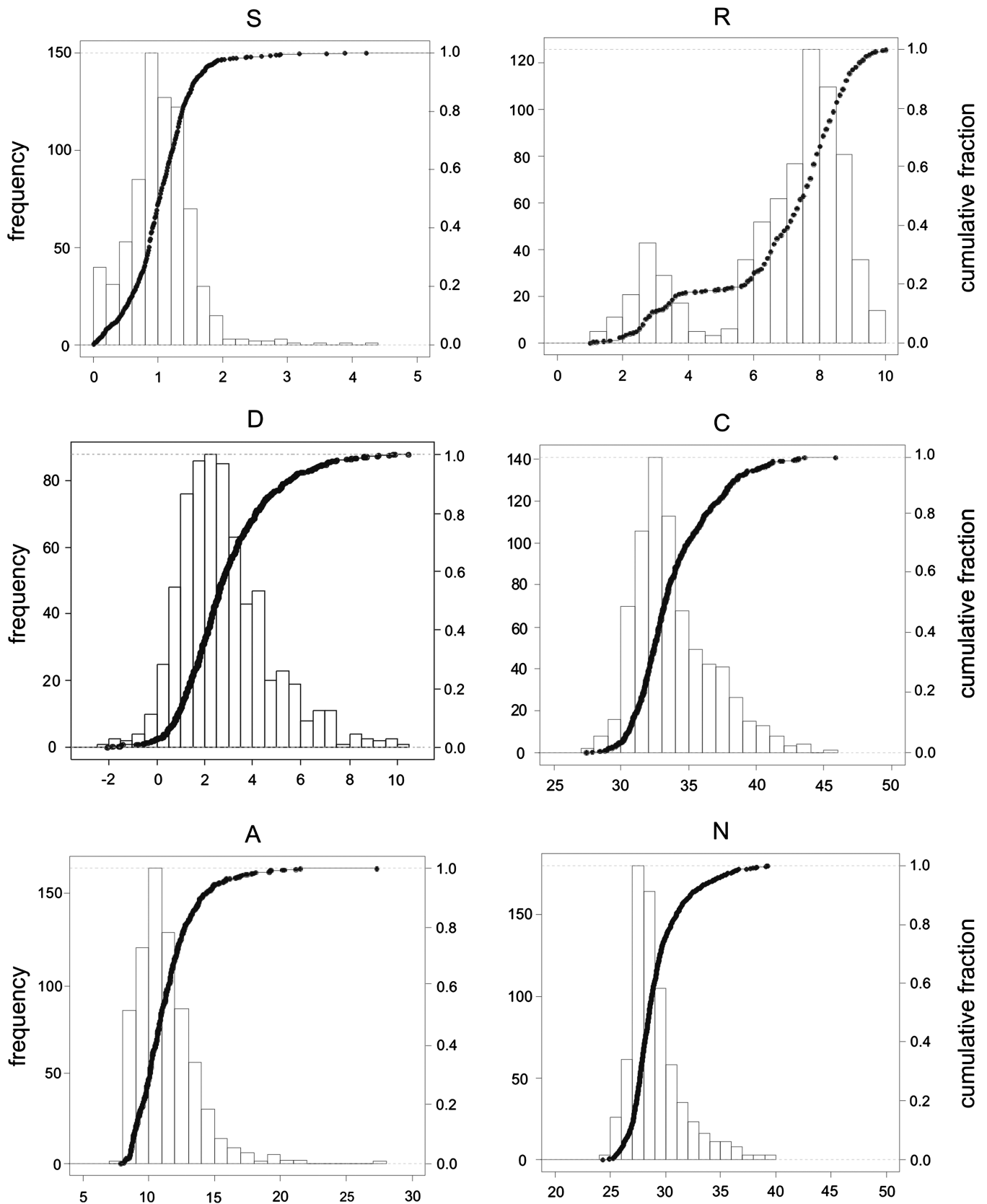


Figure 3. Frequency distribution (left axis) and cumulative frequency (right axis) of six RNA quality parameters measured in 740 neuroblastoma tumour samples. S, 18S/28S rRNA ratio determined by microfluidic capillary electrophoresis; R, RNA Quality Index determined by microfluidic capillary electrophoresis; D, *HPRT1* 5'-3' dCq (difference in quantification cycle value); C, *HPRT1* 3' Cq value; A, Alu Cq value; N, normalization factor based on the arithmetic mean Cq value of four reference genes (*HMBS*, *HPRT1*, *SDHA*, *UBC*).

Table 1. Frequencies of six RNA quality parameters measured in 740 neuroblastoma tumour samples

n = 740	Median (range)	Mean	Not available (%)
S	1.0 (0.0 to 4.2)	1	0 (0)
R	7.6 (1 to 10)	6.8	6 (0.8)
D	2.5 (-2.1 to 10.5)	2.8	53 (7.2)
C	33.2 (27.4 to 45.8)	33.8	14 (1.9)
A	10.9 (7.8 to 27.3)	11.3	30 (4.1)
N	28.5 (24.3 to 39.2)	29.1	30 (4.1)

S, 18S/28S rRNA ratio determined by microfluidic capillary electrophoresis; R, RNA Quality Index determined by microfluidic capillary electrophoresis; D, *HPRT1* 5'-3' dCq; C, *HPRT1* 3' Cq value; A, Alu Cq value; N, normalization factor based on the arithmetic mean Cq value of four reference genes (*HMBS*, *HPRT1*, *SDHA*, *UBC*).

classifier. RNA quality of the tumour samples from survivors was comparable with those from non-survivors (data not shown). The impact of RNA quality was determined using the same procedure as described in the previous paragraph. The 5'-3' dCq had the largest overall impact on classification performance when comparing best quality and worst quality samples. The correlation signature method appears less sensitive to RNA quality than the PAM method and the 59-gene classifier is less sensitive to RNA quality than the 6-gene classifier when the correlation signature method is used.

Multivariate logistic forward conditional regression testing all RNA quality parameters in a model indicated that the 5'-3' dCq and normalization factor quality

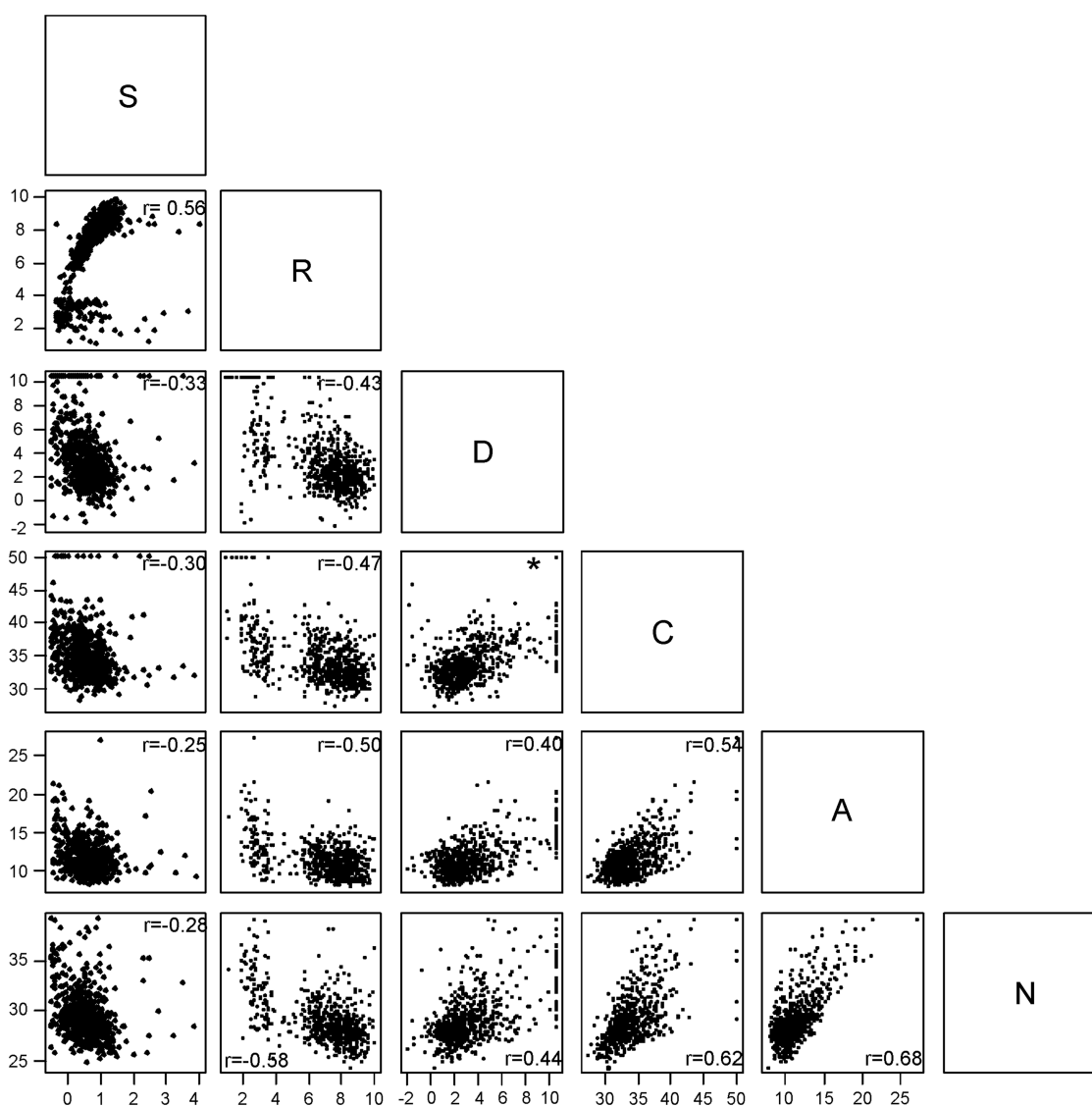


Figure 4. Correlation scatterplots between six RNA quality parameters measured on 740 neuroblastoma tumour samples indicate a significant correlation (Pearson, $P < 0.001$). Missing values for D and C due to severely degraded RNA were replaced by the value from the sample with the worst quality according to the parameter under study (these imputed values were not taken into account for the correlation analysis). Asterisks indicate correlation between 3' Cq and 5'-3' dCq values was not calculated because 3' Cq is included in the calculation of 5'-3' dCq (not independent parameters). S, 18S/28S rRNA ratio determined by microfluidic capillary electrophoresis; R, RNA Quality Index determined by microfluidic capillary electrophoresis; D, *HPRT1* 5'-3' dCq (difference in quantification cycle value); C, *HPRT1* 3' Cq value; A, Alu Cq value; N, normalization factor based on the arithmetic mean Cq value of four reference genes (*HMBS*, *HPRT1*, *SDHA*, *UBC*).

parameters were the only significant independent parameters to have an impact on outcome prediction using the 59-gene PAM signature [odds ratios 4.65 (95% CI: 1.19–18.18; and 17.86 (95% CI: 4.17–76.92), respectively]. These parameters were also found to be independently statistically significant together with other parameters using the 6-gene PAM signature and the 59-gene correlation signature, and the normalization factor was the only significant independent parameters using the 6-gene correlation signature (Supplementary Table S3).

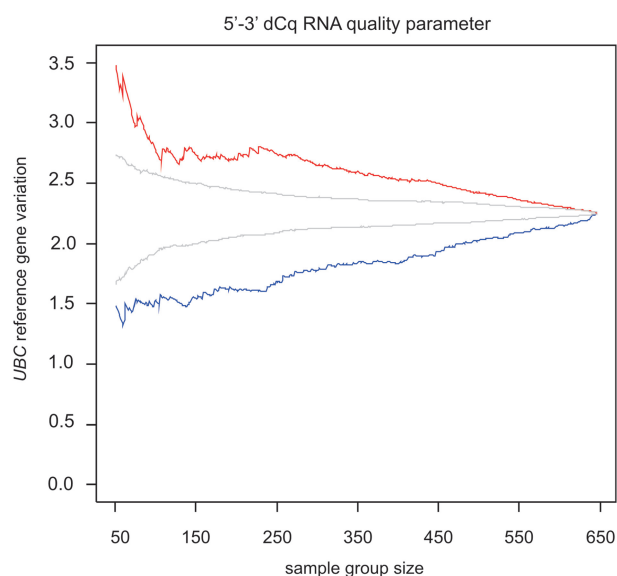


Figure 5. Variation of reference gene *UBC* calculated as the standard deviation of the log transformed gene expression levels after normalization with the arithmetic mean Cq value of three other reference genes (*HMBS*, *SDHA*, *HPRT1*). Variation is measured in increasing numbers of both best quality samples (blue curve) and worst quality samples (red curves) starting with 50 samples up to all 615 samples (quality here defined by 5'-3' dCq value). The null distribution was determined upon 100 random permutations (grey curves denote 90% distribution range). Similar results were obtained for the other reference genes and RNA quality parameters.

Table 2. Mean difference in normalized reference gene standard variation between best quality and worst quality samples based on six RNA quality parameters in 615 neuroblastoma tumour samples

RNA quality parameter	<i>HMBS</i>	<i>HPRT1</i>	<i>SDHA</i>	<i>UBC</i>	Mean
S	0.12	0.19	0.06	0.41 ^a	0.12
R	0.25	0.42 ^a	0.18	0.32 ^a	0.37
D	0.49 ^a	0.66 ^a	0.72 ^a	0.81 ^{a,b}	0.62
C	0.58 ^a	0.66 ^a	0.47 ^a	0.78 ^a	0.62
A	0.06	0.60 ^a	0.59 ^a	0.51 ^a	0.56
N	0.60 ^a	0.75 ^a	0.72 ^a	0.81 ^a	0.72
90% null distribution interval ^c	(-0.27 to 0.28)	(-0.24 to 0.26)	(-0.31 to 0.33)	(-0.24 to 0.23)	

^aSignificant difference ($P < 0.05$) (smaller in permuted samples than in ranked samples based on RNA quality).

^bMean difference in reference gene *UBC* variation according to 5'-3' dCq (example shown in Figure 5).

^cBased on 100 random permutations. Grey boxes: the value with the largest mean difference (quality parameter with biggest impact on variation) for each reference gene.

S, 18S/28S rRNA ratio determined by microfluidic capillary electrophoresis; R, RNA Quality Index determined by microfluidic capillary electrophoresis; D, *HPRT1* 5'-3' dCq (difference in quantification cycle value); C, *HPRT1* 3' Cq value; A, Alu Cq value; N, normalization factor based on the arithmetic mean Cq value of three reference genes (i.e. excluding reference gene under evaluation).

DISCUSSION

An often underestimated critical issue underlying the reliability of gene expression results is the quality of the RNA samples. To assess the impact of RNA quality on the results, a useful quality parameter is needed as well as a measurable outcome. In this study, we have examined the quality of an unprecedentedly large series of 740 clinical RNA samples using six RNA quality parameters. We did not aim to identify the best parameter measuring RNA quality. Rather, we developed and applied an analytical framework using novel methods for evaluation of RNA quality in relation to qPCR results. Undisputedly, we demonstrated a measurable influence of RNA quality on the gene expression results.

A significant—albeit imperfect—positive correlation was found between all RNA quality parameters; each parameter appears to have a different appreciation of RNA quality. While the Alu repeat sequence expression level and the normalization factor based on four reference genes were determined on randomly primed pre-amplified material, *HPRT1* 5'-3' dCq and *HPRT1* 3' Cq value were measured on cDNA obtained from anchored oligo-dT priming of original RNA. Therefore, a possible explanation for the lower correlation between these parameters is the use of a different RT priming strategy, resulting in successful pre-amplification of partially compromised RNA samples in case of random priming (23). This might also explain why some samples classified as bad quality based on RQI or 18S/28S rRNA ratio (measured on total RNA) turn out to be better quality samples based on the other methods (measured on cDNA).

Upon careful interpretation of the impact of RNA quality on the results using the novel methods, we clearly observed an effect on the variation of reference gene expression, on the significance of differential expression of prognostic marker genes and on the classification performance using a multigene signature. In contrast to some reports in the literature (8,26), the results obtained from this study indicate that the process of normalization does not completely resolve the effect of compromised RNA quality on the final results. A substantial impact

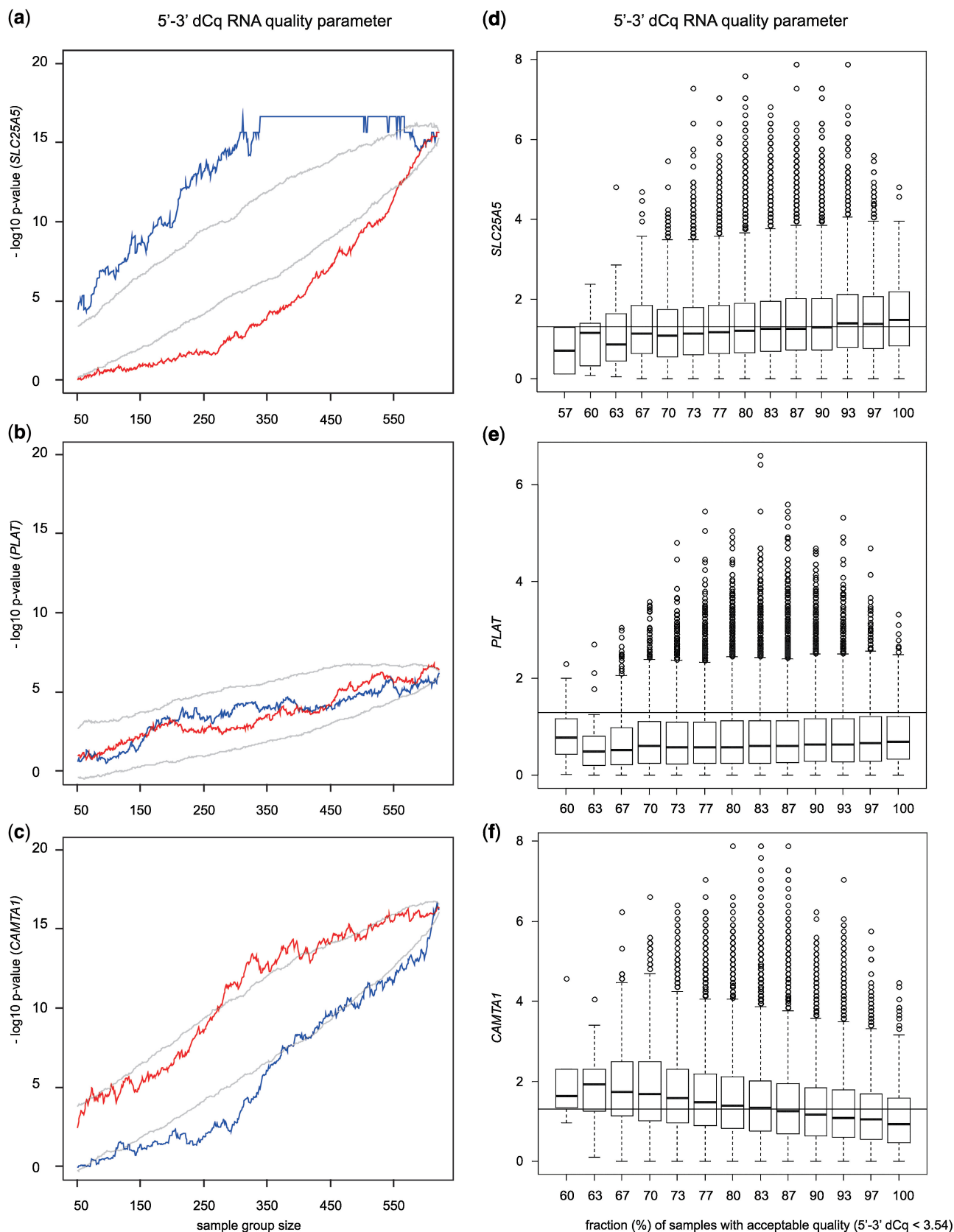


Figure 6. The effect of RNA quality on the significance of differential expression of a marker gene between tumours from two risk groups of neuroblastoma patients (HR versus non-HR) (Mann–Whitney test) was investigated using two different approaches. (i) The negative 10log of the *P*-value is calculated for increasing numbers of both best quality samples (blue curve) and worst quality samples (red curves) starting with 50 samples up to all 615 samples (a–c). The null distribution was determined upon 100 random permutations (grey curves denote 90% distribution range).

(continued)

Table 3. Impact of RNA quality on risk classification performance in 585 neuroblastoma tumour samples

	SENS (%)	SPEC (%)	AUC (%)	PPV (%)	NPV (%)		SENS (%)	SPEC (%)	AUC (%)	PPV (%)	NPV (%)
COR 59						COR 6					
S	-4.2	-6.4	-5.3	-9.6	-1.3	S	-2.4	-1.6	-2.0	-3.2	-0.5
R	-8.4	0.1	-4.1	-6.7	-1.3	R	-8.0	6.5	-0.7	-1.8	-0.5
D	6.9 ^a	9.7	8.3	-3.4	5.3	D	5.0	16.2	10.6	-1.9	7.2
C	0.8	7.8	4.3	0.3	2.5	C	-5.5	14.3	4.4	1.3	2.3
A	1.6	-3.1	-0.8	0.9	-0.5	A	-13.8	12.3	-0.7	7.2	-3.9
N	-2.5	-5.0	-3.8	0.6	-2.3	N	1.3	-2.9	-0.8	4.4	-1.5
PAM 59						PAM 6					
S	-5.4	1.8	-1.8	-1.1	-0.7	S	-9.7	-1.3	-5.5	-6.2	-2.2
R	-8.0	7.1	-0.4	1.8	0.0	R	-13.9	5.9	-4.0	-2.7	-1.8
D	0.6	15.2	7.9	5.1	5.0	D	-0.7	14.6	6.9	0.4	5.0
C	-13.9	12.6	-0.6	2.4	0.0	C	-14.7	12.1	-1.3	-0.4	-0.3
A	-2.8	8.7	3.0	11.6	-0.8	A	-6.9	4.8	-1.1	5.5	-2.4
N	-5.7	-0.2	-3.0	4.5	-3.0	N	-9.0	-1.6	-5.3	1.8	-4.0
COR 59 versus PAM 59 ^b	0.0762	<0.01	0.478	<0.01	0.610	PAM 6 versus COR 6 ^b	0.0927	0.234	<0.01	0.158	0.0651
COR 59 versus COR 6 ^b	0.351	<0.05	<0.05	<0.01	0.876	PAM 6 versus PAM 59 ^b	<0.01	<0.05	<0.01	<0.01	<0.05

Values represent the mean difference in performance between two groups of divergent RNA quality.

^aWorst RNA quality samples show 6.9% lower sensitivity compared to best quality samples. Grey boxes: the value with the largest difference (quality parameter with biggest impact on classification performance) for each performance parameter.

^bComparison of the values of all RNA quality parameters for a specific performance between two classification methods (*t*-test). COR/PAM: correlation signature/prediction analysis of microarray method using 59 or 6 genes. SENS, sensitivity; SPEC, specificity; AUC, overall accuracy as receiver operating characteristic area under the curve analysis; PPV, negative predictive value; NPV, positive predictive value; S, 18S/28S rRNA ratio determined by microfluidic capillary electrophoresis; R, RNA Quality Index determined by microfluidic capillary electrophoresis; D, *HPRT1* 5'-3' dCq (difference in quantification cycle value); C, *HPRT1* 3' Cq value; A, Alu Cq value; N, normalization factor based on the arithmetic mean Cq value of four reference genes (*HMBS*, *HPRT1*, *SDHA*, *UBC*).

of RNA quality on the standard deviation of each of the four reference genes when normalized by the three other reference genes was observed. This is in accordance with our previous report indicating that reference gene expression stability is influenced by RNA quality and that genes display varying sensitivity to RNA degradation (7). As lower RNA quality generally results in higher Cq values, it is important to note that the observed increase in variation is not simply due to sampling noise that occurs when the number of input molecules are low. Indeed, the measured Cq values of the reference genes are found in a range well below values at which sampling noise is expected; furthermore, the same observations are made when RNA quality metrics are used that are not depending on the input amount (such as 18S/28S rRNA ratio, RQI and 5'-3' dCq).

RNA quality has also a noticeable influence on the significance of differential expression of individual marker genes between two divergent risk groups of cancer patients. Some genes appear to be sensitive to RNA quality, while others are not. Surprisingly, for a few genes, the results seem better when RNA was of lower quality. This puzzling observation is in accordance to findings within the EU FP7 Spedia project on

standardization and improvement of pre-analytical procedures for *in vitro* diagnostics (M. Kubista, Personal communication). Upon extensive correlation analyses using different parameters, including qPCR assay and gene expression specific characteristics such as amplicon length, transcript length, distance to 3'-end, assay amplification efficiency, mean Cq value and magnitude of differential expression between the two risk groups, no clear explanation was found as to why some genes were more sensitive to RNA quality than others (data not shown).

Our data further show that the performance of gene expression based classification in function of RNA integrity is influenced by the number of genes included in the classifier and by the nature of the applied classification algorithm. The correlation signature algorithm seems to be the least sensitive to RNA quality and an expression signature built with a larger number of genes results in more robust classification.

Overall, the 5'-3' dCq and normalization factor quality parameters appear to have the largest influence on the qPCR expression results obtained on fresh frozen biopsies. As such, they 'appear to constitute the most' useful parameters to qualify RNA samples. The advantage of using a 5'/3' ratio assay or the normalization factor to

Figure 6. Continued

(ii) The impact of RNA quality parameters on the differential expression of genes between 100 000 random samplings of 15 patients belonging to two risk groups (HR versus non-HR) (Mann-Whitney test) (d-f). Each boxplot represents differential expression between HR and non-HR groups according to the fraction of samples with best quality (dCq 5'-3' < 3.54). Representative gene (*SLC25A5*) for which differential expression is more significant in best quality samples (a and d), representative gene (*PLAT*) for which expression results are not sensitive to RNA quality (b and e), representative gene (*CAMTA1*) for which differential expression is more significant in worst quality samples (c and f).

assess the integrity of an RNA sample before its use in a gene expression study is that it specifically ‘addresses’ the integrity of a messenger RNA molecule. This is not the case for other methods such as microfluidic electrophoresis that predominately inspect the ribosomal RNA profile to infer RNA quality. Undoubtedly, such methods provide an indication of total RNA quality but are not necessarily most appropriate to predict the integrity of mRNA transcripts that form the actual template in RT-qPCR analyses. Of note, the 5′–3′ dCq value in itself is not only depending on the RNA quality, but may also be influenced by reverse transcriptase efficiency and assay performance (qPCR efficiency). The last two factors do not need to be taken into account if the RNA quality parameter is used to rank the samples in the same study according to RNA quality. However, if the aim is to establish a quality cut-off value, these factors should be considered. The evaluation of qPCR assays targeting the 3′-end and the 5′ start of other reference gene is needed in order to confirm our results and establish 5′–3′ dCq as a valuable RNA quality parameter.

Clearly, further studies are warranted to establish a cut-off value for inclusion of a given sample in a gene expression study. We propose that pilot experiments are initiated that include positive and negative control samples in order to establish a study-specific cut-off. It is expected that this value will depend on the observed expression difference, the target abundance, the intra-group expression variability, the sensitivity to degradation of the target, the gene expression measurement method (RT-qPCR versus microarray versus massively parallel sequencing), and the nature of the samples (e.g. fresh frozen versus formalin fixed paraffin embedded). The cut-off value will also depend on the purpose of the study; a more stringent assessment of RNA quality is probably needed when a therapeutic decision is required for an individual patient compared to drawing statistical conclusions for a group of samples. Nonetheless, the inability to diagnose or assess prognosis of patients due to inferior RNA quality is unacceptable. Therefore, efforts should be made to overcome this problem and to increase the percentage of eligible cases for gene expression profiling in a clinical setting. For instance, laboratories should be trained to use standard operating procedures for the extraction and storage of high quality RNA. Furthermore, random primed reverse transcription could allow samples with some degree of RNA fragmentation to become eligible (27). The number of useable samples can also be increased if gene expression profiling is performed immediately after sampling of the tumour (which is clearly a compromising factor in retrospective studies) as it is known that storage of RNA samples might lead to degradation (10,11).

The results from this study demonstrate that monitoring RNA quality is of critical importance to obtain meaningful and reliable gene expression data and to ensure reproducibility of the results. This study confirms the need of proper RNA integrity control and proposes a framework to assess the value of an RNA quality parameter to measure the impact on the gene expression results.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Els De Smet and Nurten Yigit for their excellent technical assistance. We also would like to acknowledge Arnaud Remy, Mary Grace Brubacher, Isabelle De Roose and their team (Bio-Rad) for support with the implementation and use of the microfluidic capillary electrophoresis (Experion). We are indebted to all members of the International Society of Paediatric Oncology, European Neuroblastoma Group (SIOPEN), the Gesellschaft fuer Paediatrische Onkologie und Haematologie (GPOH-Germany), and the Children’s Oncology Group (COG-United States) for providing tumour samples and clinical history of patients.

FUNDING

Belgian Foundation Against Cancer (SCIE2006-25); the Children Cancer Fund Ghent; the Foundation Fournier Majoie pour l’Innovation; the Belgian Society of Paediatric Haematology and Oncology, the Belgian Kid’s Fund (to J.V.M.); the Foundation pour la recherche Nuovo-Soldati (to J.V.M.); the Fund for Scientific Research Flanders (to K.D.P. and J.H.L.); the Institute for the Promotion of Innovation by Science and Technology in Flanders (to S.D.R.); the Fund for Scientific Research Flanders (G.0198.08); the Ghent University Research Fund [Bijzonder Onderzoeksfonds (BOF) to J.V.S. and S.L.F.]; the European Community under the FP6 [project: STREP: EET-pipeline (037260)]; Belgian program of Interuniversity Poles of Attraction, initiated by the Belgian State, Prime Minister’s Office, Science Policy Programming. Funding for open access charge: VLK (Flemish League Against Cancer).

Conflict of interest statement. None declared.

REFERENCES

- Weis, J.H., Tan, S.S., Martin, B.K. and Wittwer, C.T. (1992) Detection of rare mRNAs via quantitative RT-PCR. *Trends Genet.*, **8**, 263–264.
- Bustin, S.A. (2000) Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *J. Mol. Endocrinol.*, **25**, 169–193.
- Ginzinger, D.G. (2002) Gene quantification using real-time quantitative PCR: an emerging technology hits the mainstream. *Exp. Hematol.*, **30**, 503–512.
- Bustin, S.A. (2008) Real-time quantitative PCR – opportunities and pitfalls. *Eur. Pharm. Rev.*, **4**, 18–23.
- Murphy, J. and Bustin, S.A. (2009) Reliability of real-time reverse-transcription PCR in clinical diagnostics: gold standard or substandard? *Expert Rev. Mol. Diagn.*, **9**, 187–197.
- Imbeaud, S., Graudens, E., Boulanger, V., Barlet, X., Zaborski, P., Eveno, E., Mueller, O., Schroeder, A. and Auffray, C. (2005) Towards standardization of RNA quality assessment using user-independent classifiers of microcapillary electrophoresis traces. *Nucleic Acids Res.*, **33**, e56.
- Perez-Novo, C.A., Claeys, C., Speleman, F., Van Cauwenberge, P., Bachert, C. and Vandesompele, J. (2005) Impact of RNA quality on reference gene expression stability. *Biotechniques*, **39**, 5254, 56.

8. Fleige, S. and Pfaffl, M.W. (2006) RNA integrity and the effect on the real-time qRT-PCR performance. *Mol. Aspects Med.*, **27**, 126–139.
9. Nolan, T., Hands, R.E., Ogunkolade, W. and Bustin, S.A. (2006) SPUD: a quantitative PCR assay for the detection of inhibitors in nucleic acid preparations. *Anal. Biochem.*, **351**, 308–310.
10. Holland, N.T., Smith, M.T., Eskenazi, B. and Bastaki, M. (2003) Biological sample collection and processing for molecular epidemiological studies. *Mutat. Res.*, **543**, 217–234.
11. Schoor, O., Weinschenk, T., Hennenlotter, J., Corvin, S., Stenzl, A., Rammensee, H.G. and Stevanovic, S. (2003) Moderate degradation does not preclude microarray analysis of small amounts of RNA. *Biotechniques*, **35**, 1192–1196, 1198–1201.
12. Nolan, T., Hands, R.E. and Bustin, S.A. (2006) Quantification of mRNA using real-time RT-PCR. *Nat. Protoc.*, **1**, 1559–1582.
13. Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L. *et al.* (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem.*, **55**, 611–622.
14. Vermeulen, J., De Preter, K., Naranjo, A., Vercruyse, L., Van Roy, N., Hellemans, J., Swerts, K., Bravo, S., Scaruffi, P., Tonini, G.P. *et al.* (2009) Predicting outcomes for children with neuroblastoma using a multigene-expression signature: a retrospective SIOPEN/COG/GPOH study. *Lancet Oncol.*, **10**, 663–671.
15. Dafforn, A., Chen, P., Deng, G., Herrler, M., Iglehart, D., Koritala, S., Lato, S., Pillarisetty, S., Purohit, R., Wang, M. *et al.* (2004) Linear mRNA amplification from as little as 5 ng total RNA for global gene expression analysis. *Biotechniques*, **37**, 854–857.
16. Kurn, N., Chen, P., Heath, J.D., Kopf-Sill, A., Stephens, K.M. and Wang, S. (2005) Novel isothermal, linear nucleic acid amplification systems for highly multiplexed applications. *Clin. Chem.*, **51**, 1973–1981.
17. Vermeulen, J., Derveaux, S., Lefever, S., De Smet, E., De Preter, K., Yigit, N., De Paepe, A., Pattyn, F., Speleman, F. and Vandesompele, J. (2009) RNA pre-amplification enables large-scale RT-qPCR gene-expression studies on limiting sample amounts. *BMC Res. Notes*, **2**, 235.
18. Lefever, S., Vandesompele, J., Speleman, F. and Pattyn, F. (2009) RTPrimerDB: the portal for real-time PCR primers and probes. *Nucleic Acids Res.*, **37**, D942–945.
19. Vandesompele, J. (2005) *2nd International qPCR Symposium & Application Workshop*. Center of Life Science in Freising Weihenstephan, Technische Universität München, Germany.
20. Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A. and Speleman, F. (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.*, **3**, Research0034.11.
21. Hellemans, J., Mortier, G., De Paepe, A., Speleman, F. and Vandesompele, J. (2007) qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol.*, **8**, R19.
22. Tibshirani, R., Hastie, T., Narasimhan, B. and Chu, G. (2002) Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proc. Natl Acad. Sci. USA*, **99**, 6567–6572.
23. Liu, R., Wang, X., Chen, G.Y., Dalerba, P., Gurney, A., Hoey, T., Sherlock, G., Lewicki, J., Shedden, K. and Clarke, M.F. (2007) The prognostic role of a gene signature from tumorigenic breast-cancer cells. *N. Engl. J. Med.*, **356**, 217–226.
24. Ruschhaupt, M., Huber, W., Poustka, A. and Mansmann, U. (2004) A compendium to ensure computational reproducibility in high-dimensional classification tasks. *Stat. Appl. Genet. Mol. Biol.*, **3**, Article37.
25. Lefever, S., Hellemans, J., Pattyn, F., Przybylski, D.R., Taylor, C., Geurts, R., Untergasser, A. and Vandesompele, J. (2009) RDML: structured language and reporting guidelines for real-time quantitative PCR data. *Nucleic Acids Res.*, **37**, 2065–2069.
26. Antonov, J., Goldstein, D.R., Oberli, A., Baltzer, A., Pirota, M., Fleischmann, A., Altermatt, H.J. and Jaggi, R. (2005) Reliable gene expression measurements from degraded RNA by quantitative real-time PCR depend on short amplicons and a proper normalization. *Lab Invest.*, **85**, 1040–1050.
27. Stangegaard, M., Dufva, I.H. and Dufva, M. (2006) Reverse transcription using random pentadecamer primers increases yield and quality of resulting cDNA. *Biotechniques*, **40**, 649–657.