

Impact of RNA degradation on gene expression profiles: Assessment of different methods to reliably determine RNA quality

Virginie Copois^a, Frédéric Bibeau^{b,1}, Caroline Bascoul-Mollevi^{d,1}, Nicolas Salvetat^a,
Patrick Chalbos^c, Corinne Bareil^a, Laurent Candeil^a, Caroline Fraslou^e,
Emmanuel Conseiller^e, Virginie Granci^c, Pierre Mazière^a,
Andrew Kramar^d, Marc Ychou^c, Bernard Pau^a,
Pierre Martineau^a, Franck Molina^a, Maguy Del Rio^{a,*}

^a CNRS UMR 5160, CRLC Val d'Aurelle, Montpellier, France

^b Service d'anatomopathologie, CRLC Val d'Aurelle, Montpellier, France

^c Service d'oncologie digestive, CRLC Val d'Aurelle, Montpellier, France

^d Biostatistics Unit, CRLC Val d'Aurelle, Montpellier, France

^e Oncology Therapeutic Department, sanofiaventis, Vitry-sur-Seine, France

Received 30 January 2006; received in revised form 21 July 2006; accepted 27 July 2006

Abstract

DNA microarray technology enables investigators to measure the expression of several 1000 mRNA species simultaneously in a biological specimen. However, the reliability of the microarray technology to detect transcriptional differences representative of the original samples is affected by the quality of the extracted RNA. Thus, it is of critical importance to standardize sample-handling protocols and to perform a quality assessment of RNA preparations. In this report, 59 human tissue samples were used to evaluate the relationships between RNA quality and gene expression. From Affymetrix[®] GeneChip[®] array data analysis of these samples, we compared the performance of the 28S/18S ratio, two computer methods (RIN and degradometer) and our in-house RNA quality scale (RQS) in assessing RNA quality. The optimal RNA reliability threshold was determined for each method using statistical discrimination measures. We showed that RQS, RIN and degradometer have a similar capacity to detect reliable RNA samples whereas the 28S/18S ratio leads to a misleading categorization. Furthermore, we developed a new approach, based on clustering analyses of full chip expression, to control RNA quality after hybridization experiments.

* Corresponding author at: CNRS-UMR 5160, Centre de Recherche en Cancérologie, CRLC Val d'Aurelle-Paul Lamarque, 208 rue des Apothicaires, 34298 Montpellier Cedex 5, France. Tel.: +33 467 613 745; fax: +33 467 613 787/548 610.

E-mail address: mdelrio@valdorel.fnclcc.fr (M. Del Rio).

¹ These authors contributed equally to this work.

The combination of these methods, allowing monitoring of RNA quality prior to and after the hybridization experiments, ensured reliable and reproducible microarray data.

© 2006 Elsevier B.V. All rights reserved.

Keywords: Expression profiles; Gene; Quality; RNA; Scale

1. Introduction

The development of microarray technology has offered the opportunity to analyze the simultaneous expression of thousands of genes in a single experiment (DeRisi et al., 1996). Over the past few years, this powerful technology has been used to explore transcriptional profiles and to obtain molecular expression signatures of the state of activity of diseased cells and patient samples (Xiang et al., 2003). In the field of cancer, microarray analyses may provide information on pathology, progression or resistance to treatment (Pusztai et al., 2003; Chang et al., 2003).

However, the reliability of the microarray technology to detect transcriptional differences representative of original samples is affected by several factors such as array production, RNA extraction, probe labeling, hybridization conditions and image analysis (Schuchhardt et al., 2000). In particular, the quality of data from microarray analysis is strongly related to the quality of the RNA extracted from the tissues which is in turn dependent on the quality of the tissue samples. Multiple studies have demonstrated that the modulation of mRNA stability plays an important role in regulating gene expression (Liebhaber, 1997; Mitchell and Tollervy, 2000; Guhaniyogi and Brewer, 2001). The stability of a given mRNA transcript is controlled by specific interactions between its structural elements, known as *cis*-elements, and trans-acting factors. These interactions are regulated by environmental stimuli such as nutrient levels, cytokines and hormones as well as environmental stresses like hypoxia and tissue injury (Guhaniyogi and Brewer, 2001; Hollams et al., 2002).

Thus, it is essential that the RNA isolated from the tumor tissue specimens be of the highest quality to ensure the reliability of the analysis. Whereas RNA of high quality can be obtained from cell lines, this becomes more difficult when working with tumor biopsies due to the duration of ischemia which is susceptible to influence the integrity of RNA within the tissues

affecting molecular data (Huang et al., 2001; Spruessel et al., 2004).

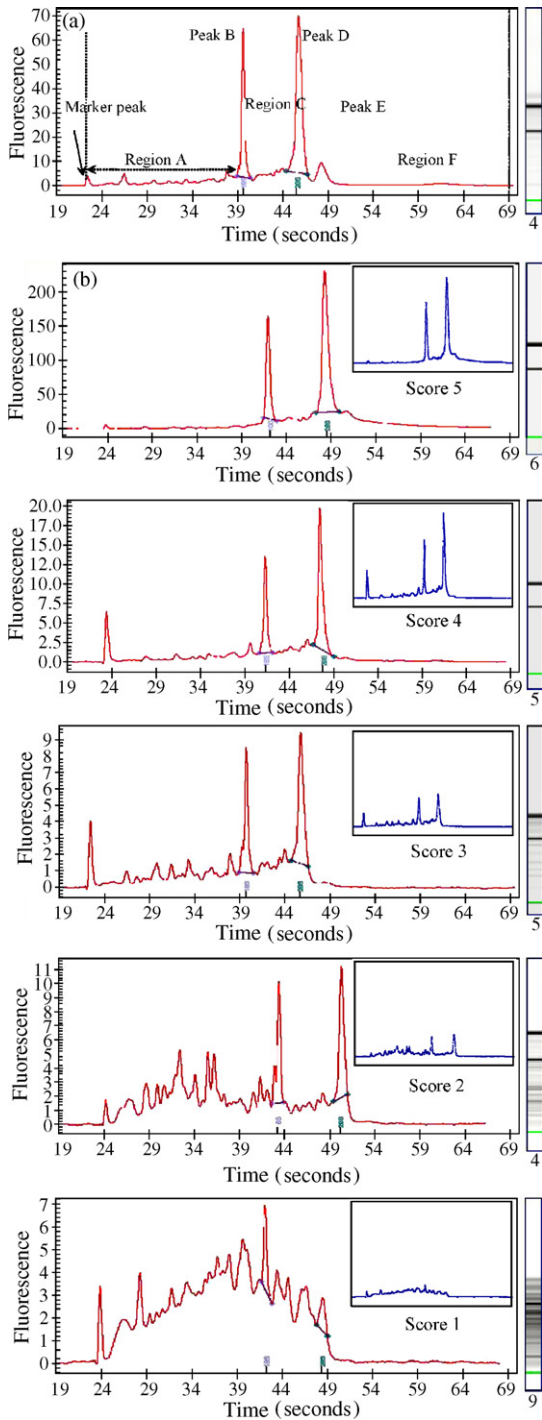
Quality assessment of total RNA is usually determined by quantification on the ethidium bromide gels of 28S and/or 18S ribosomal RNA (Sambrook and Russel, 2001). Recently, two computer methods, the degradometer software (Auer et al., 2003) and the RNA integrity number (RIN) algorithm (Schroeder et al., 2006), have been described.

In this report, we categorized 59 RNA samples from normal and tumor tissues, according to their degree of degradation. We first developed our own in house RNA quality scale (or RQS), and then we verified the consistency of this human evaluation by comparison with the three other methods (28S/18S, RIN and the degradometer). Using data from Affymetrix[®] GeneChip[®] arrays, we defined an RNA reliability threshold to screen RNA samples prior to probe preparation. Finally, we developed a new method to evaluate RNA quality after hybridization experiments. We demonstrated that monitoring RNA quality prior to and after the hybridization experiments is the best way to ensure reliable and reproducible microarray data.

2. Materials and methods

2.1. Tissue sample collection

Tissues were obtained from colorectal cancer patients with synchronous and unresectable liver metastases enrolled in a prospective study at CRLC Val d'Aurelle, aimed at identifying a gene expression profile associated with the response to chemotherapy. We collected 59 samples including 17 primary tumors, 19 normal colon and 23 liver metastases. Immediately after surgical excision of the tumor, a clinical research assistant in charge of sample handling transferred them as quickly as possible to the pathology department. A macroscopic examination was then performed, including localization, measurement of the tumor and assess-



ment of margins. A sample selection for frozen storage was made with precision avoiding necrotic areas and normal tissue, and leaving the resection specimen available for further routine examination. To ensure that a sufficiently high proportion of tumor cells was present, a minimal fragment size of approximately 3 mm^3 was required. Moreover, a touch imprint of the selected fragment was performed and stained with May-Grünwald Giemsa to assess the tumor cellularity. The rest of the specimen was then fixed in 10% formaldehyde solution for 48 h, and a subsequent routine histological examination was made. The ethical committee approved our study design. Patients gave written informed consent.

2.2. RNA preparation

All tissue samples were maintained at -180°C (liquid nitrogen) until RNA extraction and were weighed before homogenization. Then tissue samples were directly disrupted in a lysis buffer using Mixer Mill[®] MM 300 (Qiagen, Valencia, CA). Total RNA was isolated from tissue lysates using the RNeasy[®] mini Kit (Qiagen), and additional DNase digestion was performed on all samples during the extraction process (RNase-Free DNase Set[™] Protocol for DNase treatment on RNeasy[®] Mini spin columns, Qiagen).

Fig. 1. In house RNA quality scale. (a) Description of an electropherogram. Electropherograms were obtained by running RNA 6000 Nano LabChip[®]. Marker peak region A represents low molecular weight RNA; the presence of peaks or smear in this region is due to degradation. Peak B is 18S ribosomal RNA; its height and width vary with the degradation level. Region C represents the region between peaks 18S and 28S, the increase of intensity in this region is related to 28S ribosomal RNA degradation. Peak D is the 28S ribosomal RNA; its height and width vary with the degradation level. Peak E represents precursor nuclear RNA. The presence and the height of this peak are strongly dependent on the nature of the sample. Region F corresponds to high molecular weight RNA. The presence of peaks or smear in this region is often due to genomic DNA contaminations; in this case, the sample must be digested with DNase I and requalified. (b) RNA quality scale (or RQS) for RNA classification. A quality score was given on the basis of the electropherogram evaluation described in Table 1. RNA samples whose quality score was between 5 and 3 were labeled “reliable” for transcriptome analysis; a score of 2 was labeled “doubtful” and a score of 1 was labeled as “unreliable”. To better compare profiles from all scores, insets represent the electropherograms using the same y-axis scale.

2.3. Assessment of RNA quality

After each extraction, a small fraction of the total RNA preparation was taken to determine the quality of the sample and the total RNA yield. Controls were performed by UV spectroscopy to verify RNA concentration and purity. Total RNA profile was analyzed using Agilent RNA 6000 Nano LabChip® kit with the Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA) to determine its quantity and its integrity (Bastard et al., 2002). We classified the total RNA samples using four different scales described below.

We developed an in-house RNA quality scale (RQS). This is a visual method which is essentially based on the RNA electropherogram profiles (Fig. 1) obtained from the Agilent 2100 Bioanalyzer using the RNA 6000 Nano Labchip kit. The main features considered for RNA quality evaluation are described in Table 1. The first feature is the size and the shape of the 18S and 28S rRNA peaks (Fig. 1a, peaks B and D). The second feature is the stability of the baseline of the RNA electropherogram (Fig. 1a, region A). In some cases, when classification of an RNA sample is uncertain, we take also in consideration the decrease of the 28S/18S ratio, and the appearance of additional peaks and elevation of the baseline in the region between ribosome peaks (Fig. 1a, region C). We classified RNA into different groups, representing the level of integrity of total RNA on a scale from 1 to 5. We considered that only the samples scored from 3 to 5 (or slightly degraded

to intact RNA) are reliable samples for transcriptome analysis.

The 28S/18S ratio is considered as the gold standard for evaluation of RNA integrity (Sambrook and Russel, 2001). It is commonly accepted that intact RNA has an rRNA band ratio ≥ 1.8 .

The degradometer software (Version 1.41) provides quantitative data as regards to the integrity and concentration of eukaryotic total RNA (Auer et al., 2003). The software calculates a degradation factor (DegFact) which represents the percentage of degradation product against the 18S peak. The higher the degradation factor, the more the sample is considered to be degraded. The samples are divided into four groups as a function of their RNA degradation: DegFact > 24 (strong degradation), DegFact > 16 (severe degradation) and DegFact > 8 (degradation can be detected) and DegFact < 8 (intact RNA). For more details see <http://www.dnaarrays.org/downloads.php>.

The RIN software uses an algorithm that has been developed to extract information about RNA sample integrity from a Bioanalyzer electrophoretic trace obtained with the Eucaryote Total RNA Nano assay. The RIN algorithm is based on a combination of different features including the total RNA ratio (the fraction of the area in the region of 18S and 28S compared to the total area under the curve), the height of the 28S peak, the fast area ratio (the fraction of the area in the fast region compared to the total area under the curve) and the marker height (Schroeder et al., 2006). Ten categories were defined

Table 1
In house quality scale (RQS) description

	Criterion no. 1		Criterion no. 2
Electropherogram regions	Peak B: ribosomal 18S RNA peak	Peak D: ribosomal 28S RNA peak	Region A: between marker peak and 18S peak
Score 5	Slim and well-resolved peak	Slim and well-resolved peak	Flat baseline
Score 4	Well resolved peak	Slight diminution of peak intensity that may become broader	Unstable baseline (when concentration >50 ng/ μ l)
Score 3	Slight diminution of peak intensity	Diminution of peak intensity that may become broader	Elevation of unstable baseline
Score 2	Peak starts to collapse	Strong diminution of peak intensity with a variable elevation of baseline	Elevation of baseline with appearance of peaks which intensities attempt to reach ribosomal peaks
Score 1	Peak starts to become indiscernible from degradation products	Strong diminution (or disappearance) of the peak with a high elevation of baseline	Smear: numerous peaks of various size and high intensity

ranging from 1 (total degraded RNA) to 10 (total intact RNA).

2.4. Complex probe preparation and hybridization on AffymetrixTM GeneChip[®] arrays

Synthesis of the first strand cDNA was performed by a T7-linked oligo-dT primer, followed by second strand synthesis. Labeled cRNA probes were then generated by reverse transcription followed by in vitro transcription, incorporating biotin labeling, as part of the standard Affymetrix protocol (<http://www.affymetrix.com/support/technical/manuals.affx>). Probes were synthesized from 59 total RNA samples of various quality and were hybridized on the Human Genome GeneChip[®] arrays U133 set A, that contains almost 22,000 probes (Affymetrix, Santa Clara, CA), according to the manufacturer's instructions. After hybridization, the probes were scanned using a laser scanner, and signal intensity for each transcript and detection call (present, absent, or marginal) were determined using MAS 5.0 Software (Affymetrix). Inter-array normalization was performed using a set of internal standard genes (normalization set) leading to the determination of a scaling factor.

2.5. Statistical analysis

To examine whether the variables RIN, 28S/18S ratio and DegFact were different between the groups defined by RQS, we used the Kruskal–Wallis test which is a non-parametric alternative to one-way ANOVA.

The medians of RIN, 28S/18S ratio and degradometer were compared between the reliable and unreliable groups using the Wilcoxon non-parametric test. *p*-Values less than 0.05 were considered statistically significant.

The proportion of cases classified as reliable by the RQS were compared between the reliable and unreliable groups using Pearson's chi-square test.

The area under the ROC curves is used to study the accuracy of markers and to judge their capacity in the discrimination between two well-defined populations (Kramar et al., 2001). A marker can take on a wide range of values. An optimal cut-off can be searched for and summarized through the ROC curve, which can be also used to compare the performance of the marker relative to others. By calculating the sensitivity and the specificity for each value of the cut-off, we traced

the ROC curve, which provides a visual description of the trade-off between false-positive rates and true-positive rates for all possible cut-off values. Youden's index, defined as the sum of sensitivity and specificity minus one, was used in the choice of the optimal cut-off, if the "weight" of sensitivity and the specificity can be considered equal. The most commonly used global index of diagnostic accuracy is the area under the ROC curve (AUC). This index is particularly useful to compare several ROC curves, and an AUC of 1 represents a perfect test; an AUC of 0.5 represents a worthless test.

2.6. Quality assessment of microarray data: "dispersion tree" approach

The dispersion of global chip expression was evaluated using a very simple original method we called "dispersion tree". After normalization, each chip experiment is represented by an expression vector of dimension *n* (*n* corresponds to the number of genes on the chip). The Euclidian distance between these vectors (representing all chip experiments and their gene intensities) was calculated. The resulting distance matrix was used to perform a clustering of all the experiments. The clustering (Kitsh algorithm) and the resulting unrooted "dispersion tree" graphical representation (drawtree algorithm) were performed with the package PHYLIP v3.6 (Felsenstein, 2005). Given the large number of genes on the chip all the experiment distances should be of the same order, leading to a low dispersion graphical representation. Any experiments clearly dispersed compared to the other experiments should be considered suspicious or different from the others (either an experimental quality problem or very different biological conditions).

3. Results

3.1. Assessment of RNA quality

The quality of each RNA sample extracted from 59 normal (32%) and tumor (68%) tissues was first determined by the RNA quality scale (RQS), a visual RNA quality evaluation established in our laboratory. Among the 59 RNA samples analyzed, 46 (78%) were scored from 3 to 5 and, thus, considered as reliable samples.

The RNA quality of these samples was also assessed in parallel using 28S/18S rRNA ratio provided by bio-sizing software and two recent computer methods, the degradometer software (Auer et al., 2003) and the RIN (RNA integrity number) algorithm (Schroeder et al., 2006). Three RNA samples could not be scored by RIN algorithm due to their low RNA concentration, and two other RNA samples presented a “BLACK” degradation alert by degradometer software, indicating that they could not be reliably interpreted. On the remaining 54 RNA samples, we showed that only RIN and degradometer factor (DegFact) were negatively correlated ($\rho = -0.92$) for RNA categorization (Fig. 2). Using a linear regression model, we estimated the relationship between the RIN and DegFact variables with a linear regression line fit to the data. In other words, RNA samples scored between 8 and 16 by DegFact would be scored between 7 and 8.3 by RIN. We then evaluated whether our visual interpretation was consistent with computer analytic methods. The first group corresponding to an RQS score of 1 gave a mean RIN number of 2.70 and a mean DegFact of 39.58, indicating unambiguously that these RNA samples were degraded. For the groups with a 3–5 RQS score, the mean RIN number was >8 (8.16) and the mean DegFact was <8 (7.39), corresponding to the expected scores for good RNA quality. The groups with an RQS score of

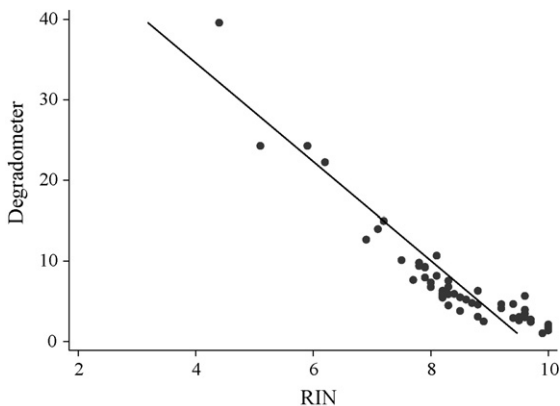


Fig. 2. For the RIN and degradometer methods, assessment of the RNA quality score obtained for each of 54 human tissue samples was plotted. A linear regression attempts to explain this relationship with a straight line fit to the data. The linear equation was $\text{RIN} = -162.982 \times \text{DegFact} + 9.634924$. The residuals were normally distributed, and the coefficient of determination was good ($r^2 = 0.85$).

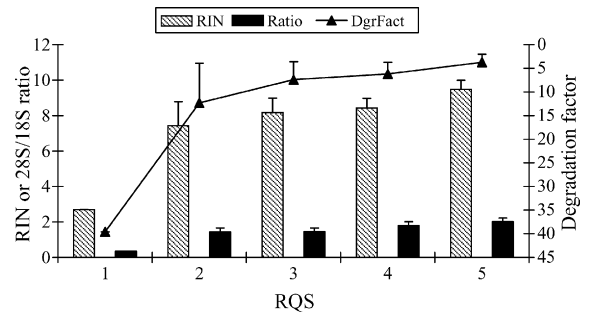


Fig. 3. RNA sample categorization. RNA quality of human tissue samples was evaluated by RIN, degradometer and 28S/18S ratio and compared with the RQS groups (x-axis). Histograms represent the mean values of RIN and 28S/18S ratio (left y-axis), and mean values of DegFact are indicated by a triangle plot (right y-axis).

2 appeared as the borderline groups with an important variability in the results: mean RIN = 7.43 ± 1.35 and mean DegFact = 12.34 ± 8.43 (Fig. 3). In addition, groups scored 2 or 3 by RQS were not discriminated by the 28S/18S ratio; mean values being similar for both groups. Using the Kruskal–Wallis test, we showed that the variables RIN ($p = 0.0001$), 28S/18S ratio ($p = 0.0001$) and DegFact ($p = 0.0003$) were significantly different between the groups defined by the RQS.

These results showed a good concordance between the RQS and the two computer methods, thus emphasizing the difficulty in determining the threshold for the reliability of RNA samples for transcriptome analyses.

3.2. RNA quality and gene expression profiles

To explore the relationships between the four RNA quality categorization and the gene expression profiles, we collected Affymetrix® GeneChip® array data after the hybridization of the 59 samples on HG-U133 gene chip array A.

3.2.1. Housekeeping gene expression

To compare the four scales, we had to determine an independent cut-off threshold that separates reliable and unreliable RNA samples. We chose the 3'/5' ratios for the glyceraldehyde-3-phosphatase (GAPDH) housekeeping gene. It is known that the 3'/5' ratio gives an indication of the integrity of the starting RNA (Affymetrix guidelines). Moreover, a positive correlation between the DegFact and the 3'/5' ratio of

Table 2
Description of RNA quality assessment methods on GAPDH 3'/5' ratio defined-groups

	RNA sample quality		<i>p</i>
	Unreliable	Reliable	
RIN; median [range]	7.5 [4.4; 9.4]	8.6 [6.9; 10]	0.001 ^a
28S/18S ratio; median [range]	1.44 [1.0; 1.7]	1.8 [1.1; 2.4]	0.001 ^a
RQS; number of samples (%)			
1	1 (9.1%)	0(0%)	
2	8 (72.7%)	2(4.7%)	
3	2 (18.2%)	10(23.3%)	<0.001 ^b
4	0(0%)	14(32.5%)	
5	0(0%)	17(39.5%)	
Degradometer; median [range]	10.11 [2.9; 39.6]	5.21 [1; 14]	0.002 ^a

^a Wilcoxon non-parametric test.

^b Pearson's chi-square test.

two housekeeping genes has been demonstrated (Auer et al., 2003), and this ratio has also been used as a measure of RNA quality (Croner et al., 2004). Thus, we considered that an increase in the 3'/5' ratio for the GAPDH gene is a relevant indicator of RNA degradation. We defined the 3'/5' ratio threshold for GAPDH as the mean 3'/5' ratio + S.D., which was 1.25. RNA samples whose 3'/5' ratio > 1.25 were considered as unreliable samples; and samples whose 3'/5' ratio < 1.25 were considered as reliable ones. As shown in Table 2, the two groups were found to be significantly different, whatever the scale used. We performed ROC analyses to evaluate the ability of the different quality assessment methods to discriminate unreliable RNA samples from reliable ones. As presented in Table 3, all areas under the ROC curve for the GAPDH data were high (ranging from 0.814 to 0.969), which enabled us to conclude that the four methods were able to discriminate reliable RNA samples from unreliable ones. With the highest AUC and the narrow 95% confidence inter-

vals (0.89–0.994), our RQS seemed to be the most discriminant scale. Moreover, for each quality assessment method, we determined the optimal threshold to separate both RNA populations, using Youden's index (Fig. 4 and Table 3). These results showed that a reliable RNA sample needs to be scored ≥ 3 for RQS, >1.63 for 28S/18S ratio, >7.8 for RIN and <8.7 for DegFact. Similar results were obtained with the second housekeeping gene, beta actin (ACTB), present on the HG-U133 gene chip array A (Supplementary data online). Using these optimal thresholds, we calculated the specificity (probability of detecting reliable RNA samples) and the sensitivity (probability of detecting unreliable samples) for the four RNA quality scales. We obtained a high specificity for RQS, RIN and DegFact (95.3%, 93% and 86%, respectively) while 28S/18S ratio specificity was low (69.7%). Inversely, the higher sensitivity was obtained for 28S/18S ratio (90.9%) followed by RQS (81.8%), DegFact (63.4%) and RIN (54.6%).

3.2.2. Global expression profiles

The distribution of global gene expression was represented by a "dispersion tree". Each sample was associated with its different RNA quality assessment score (Fig. 5). We observed that sample dispersion was homogeneous except for those experiments covered by the gray cluster and corresponding to samples 10, 11, 12, 27, 28, 29, 37, 40 and 41. The clustering observed from duplicated experiments from identical RNA samples (samples 6, 37 and 57) showed that experimental variation was not responsible for the specific dispersion observed for samples in the gray cluster. The quality score from the different RNA quality assessment methods clearly showed that highly dispersed experiments are associated with low RNA quality scores. These results demonstrate the impact of RNA quality on the distribution of global gene expression.

All the gray cluster-related samples were systematically categorized by RQS as unreliable samples, whereas the 28S/18S ratio indicated that sample 10 was a reliable sample. The RIN and the degradometer showed similar evaluations as RQS, except for samples 27 and 40. In those cases where the score of RQS was different from RIN or DegFact scores (samples 13, 27, 36, 40, 47, 49, 51 or 55), it was difficult to determine which score was correct. As these samples were

Table 3
Generalized AUC and scale thresholds determined by ROC analysis

	Generalized AUC	95% confidence interval	Optimal threshold
RQS	0.969	[0.89; 0.994]	≥ 3
RIN	0.828	[0.664; 0.927]	7.8
28S/18S ratio	0.819	[0.596; 0.94]	1.63
Degradometer	0.814	[0.594; 0.936]	8.7

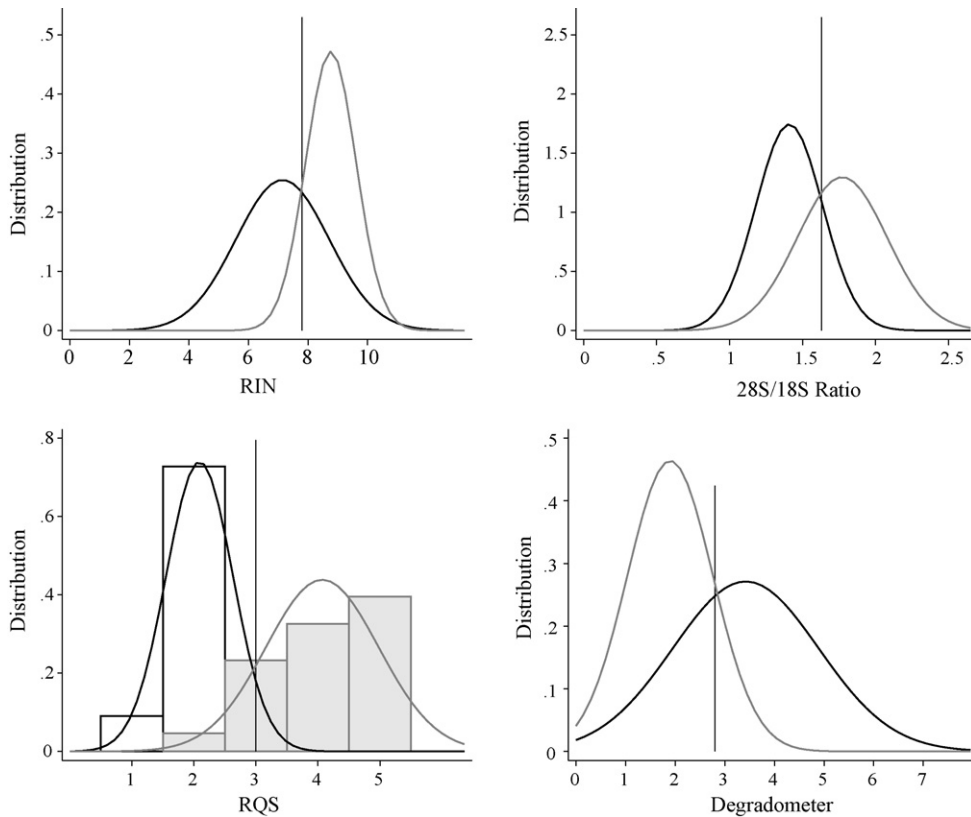


Fig. 4. Distribution of the reliable (black line) and unreliable (gray line) RNA sample population as a function of the four RNA quality assessment methods. The degradometer distribution is represented after a Box–Cox transformation with parameter equal to 0.24 to normalize the data. Vertical lines represents optimal threshold in discriminating both populations.

borderline samples, with RQS = 2 or RIN or DegFact scores near the threshold, dispersion was the only reliability criterion considered. Also, samples 27 and 40 were classified as unreliable, whereas samples 13, 36, 47, 49, 51 and 55 were classified as reliable ones. Sample 56 should be considered as an unreliable sample, even though it is not included in the gray cluster. This sample appeared dispersed and was evaluated, without any ambiguity, as highly degraded (RQS = 1; RIN = 2.4). The 28S/18S ratio gave a contradictory indication for 10 samples in comparison with the three other methods. This was probably due to its lack of specificity.

Thus, the combination of the “dispersion tree” and the RQS method for evaluation of RNA quality was found to be a very stringent procedure in identifying spurious samples that should be treated with care for further microarray data analyses.

4. Discussion

The limiting factor for obtaining meaningful gene expression data is the quality of the initial RNA preparation. Obtaining high quality RNA from human samples requires a succession of processes and a multidisciplinary organization. Indeed, it involves the participation of surgeons, pathologists, clinical research assistants and molecular biologists. We developed a standardized tissue acquisition/processing method aimed at reducing the time (less than 30 min) between tissue excision and storage in liquid nitrogen while allowing a pathologist examination of the tumor sample. Huang et al. (2001) demonstrated that the majority of RNA alterations considered as experimentally significant occur after 20 min of ischemic time. The degradation of RNA, which is not correlated with ischemic time, may then be explained by other factors as for

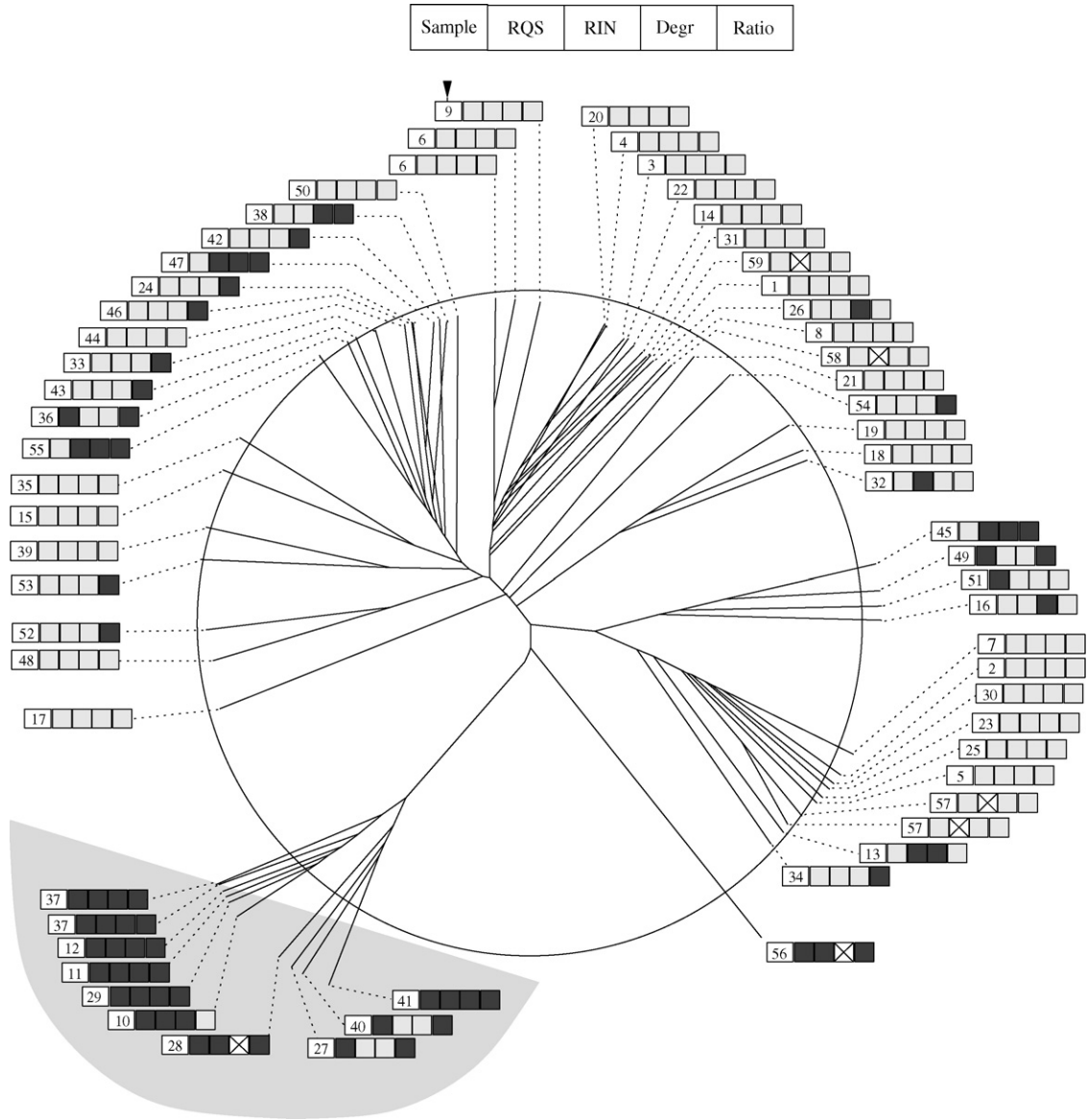


Fig. 5. “Dispersion tree” approach. “Dispersion tree” represents the distribution of global expression of Affymetrix® GeneChip® array. Each sample, represented by a number, is associated with its RNA quality evaluation obtained by the four different RNA quality assessment methods, as indicated at the top of the figure. In accordance with the established scale thresholds, the reliable samples are represented in gray and the unreliable samples in black; X indicates the samples not evaluated by RIN or degradometer scale. The gray area indicates the dispersed samples.

example, the warm ischemia time (Almeida et al., 2004).

Nowadays, there is no standardized procedure for quality assessment of total RNA. RNA quality is usually determined by quantification on ethidium bromide gels of 28S and/or 18S ribosomal RNA (Sambrook

and Russel, 2001). However, this evaluation is not precise enough to detect subtle changes in RNA integrity. Thus, to assess the quality of RNA we took advantage of the lab-on-a-chip technology, which is based on miniaturization of RNA capillary electrophoresis and which is more sensitive (Copois et al., 2003). Using

this technology, it is possible to visualize a wide spectrum of total RNA components on an electropherogram profile. Based on these electropherogram profiles, we elaborated our own RNA quality scale called RQS to monitor RNA integrity. This scale allows categorization of RNA preparations based on their reliability for transcriptome analysis. More recently, two computer methods for characterization of RNA degradation (RIN and degradometer) have been described (Auer et al., 2003; Schroeder et al., 2006). These methods allowed us to categorize 95% (RIN) and 97% (degradometer) of the RNA samples tested. From GAPDH expression level measured on the Affymetrix® GeneChip® array HG-U133A, we used statistical measurements of discrimination to compare the RQS, the latter two computer methods and the 28S/18S ratio. We found that the all RNA quality assessment methods had a capacity to discriminate samples, and we determined, for each method, the optimal RNA reliability threshold. The differences between these methods resided in their capacity to detect reliable samples (or their specificity). Because of its low specificity, the discrimination provided by the 28S/18S ratio could lead to the exclusion of reliable samples prior to transcriptome experiments. As previously indicated from reverse transcription real-time PCR data (Miller et al., 2004). The 28S/18S ratio can be a misleading indicator of the state of the mRNA. Inversely, RQS, RIN and DegFact were more valuable methods and showed an equivalent capacity to evaluate the reliability of samples with regards to transcriptome analysis. The two computer methods were well correlated and easy to use. RQS was a more subjective method that needs experimented operators, but this measurement allowed us to categorize all the RNA samples, even those with a low RNA concentration. Nevertheless, it is important to keep in mind that these methods are based on the limits of the sensitivity of the Bioanalyzer to detect the RNA samples. If the sample concentration is below the threshold required for detection, these methods cannot be used.

As categorization remains tricky for some borderline samples, we developed a new approach to evaluate RNA quality after hybridization experiments. This method is based on clustering analyses of full chip expression and represented by a “dispersion tree”. We clearly showed that the dispersion tree approach allowed us to underline some hybridization experi-

ments that were dispersed compared with the overall population. These experiments were always associated with unreliable RNA quality scores when using the RQS leading to the exclusion of these samples from data analyses. Furthermore, this second approach to assess RNA quality enabled us to make a decision on borderline samples. Altogether, we demonstrated that the combination of the “dispersion tree” and RQS, RIN or DegFact for evaluation of RNA quality was a very stringent approach to identify unreliable samples or experiments. Such a qualification of reliability of microarray experiments is of key importance for further sophisticated data mining analysis techniques. Our approach could contribute to reducing the background noise of microarray result analyses by pointing out unreliable samples.

Finally, an increasing number of researchers are now convinced that the success of microarray experiments depends on RNA quality assessment and that it is necessary to establish defined quality control criteria for sample quality to distinguish between analytical and biological variability. If monitoring RNA quality prior to and after hybridization experiments seems to be the best way to ensure reliable and reproducible microarray data, the key to the successful procurement of high quality RNA in a hospital environment is the establishment of a multidisciplinary collaboration from the surgeons to the molecular biologists.

Acknowledgements

This work was supported by funds from Sanofi-Aventis and from the “Centre National de la Recherche Scientifique”. The authors would like to thank Dr. S.L. Salhi for presubmission editorial assistance, Dr. O. Müller (Agilent Technologies) for valuable advice and the surgical team from the CRLC Val d’Aurelle for strong participation in this work.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.jbiotec.2006.07.032](https://doi.org/10.1016/j.jbiotec.2006.07.032).

References

- Almeida, A., Paul Thiery, J., Magdelenat, H., Radvanyi, F., 2004. Gene expression analysis by real-time reverse transcription polymerase chain reaction: influence of tissue handling. *Anal. Biochem.* 328, 101–108.
- Auer, H., Lyianarachchi, S., Newsom, D., Klisovic, M.I., Marcucci, G., Kornacker, K., Marcucci, U., 2003. Chipping away at the chip bias: RNA degradation in microarray analysis. *Nat. Genet.* 35, 292–293.
- Bastard, J.P., Chambert, S., Ceppa, F., Coude, M., Grapez, E., Loric, S., Muzeau, F., Spyrtos, F., Poirier, K., Copois, V., et al., 2002. RNA isolation and purification methods. *Ann. Biol. Clin. (Paris)* 60, 513–523.
- Chang, J.C., Wooten, E.C., Tsimelzon, A., Hilsenbeck, S.G., Gutierrez, M.C., Elledge, R., Mohsin, S., Osborne, C.K., Chamness, G.C., Allred, D.C., et al., 2003. Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast cancer. *Lancet* 362, 362–369.
- Copois, V., Bret, C., Bibeau, F., Brouillet, J.P., Del Rio, M., Berthe, M.L., Maudelonde, T., Boulle, N., 2003. Assessment of RNA quality extracted from laser-captured tissues using miniaturized capillary electrophoresis. *Lab. Invest.* 83, 599–602.
- Croner, R.S., Guenther, K., Foertsch, T., Siebenhaar, R., Brueckl, W.M., Stremmel, C., Hlubek, F., Hohenberger, W., Reingruber, B., 2004. Tissue preparation for gene expression profiling of colorectal carcinoma: three alternatives to laser microdissection with preamplification. *J. Lab. Clin. Med.* 143, 344–351.
- DeRisi, J., Penland, L., Brown, P.O., Bittner, M.L., Meltzer, P.S., Ray, M., Chen, Y., Su, Y.A., Trent, J.M., 1996. Use of a cDNA microarray to analyse gene expression patterns in human cancer. *Nat. Genet.* 14, 457–460.
- Felsenstein, J., 2005. PHYLIP (Phylogeny Inference Package) Version 3.6. Distributed by the Author. Department of Genome Sciences, University of Washington, Seattle.
- Guhaniyogi, J., Brewer, G., 2001. Regulation of mRNA stability in mammalian cells. *Gene* 265, 11–23.
- Hollams, E.M., Giles, K.M., Thomson, A.M., Leedman, P.J., 2002. mRNA stability and the control of gene expression: implications for human disease. *Neurochem. Res.* 27, 957–980.
- Huang, J., Qi, R., Quackenbush, J., Dauway, E., Lazaridis, E., Yeatman, T., 2001. Effects of ischemia on gene expression. *J. Surg. Res.* 99, 222–227.
- Kramar, A., Faraggi, D., Fortune, A., Reiser, B., 2001. mROC: a computer program for combining tumour markers in predicting disease states. *Comput. Methods Prog. Biomed.* 66, 199–207.
- Liebhaber, S.A., 1997. mRNA stability and the control of gene expression. *Nucleic Acids Symp. Ser.* 29–32.
- Miller, C.L., Diglisic, S., Leister, F., Webster, M., Yolken, R.H., 2004. Evaluating RNA status for RT-PCR in extracts of postmortem human brain tissue. *Biotechniques* 36, 628–633.
- Mitchell, P., Tollervy, D., 2000. mRNA stability in eukaryotes. *Curr. Opin. Genet. Dev.* 10, 193–198.
- Pusztai, L., Ayers, M., Stec, J., Hortobagyi, G.N., 2003. Clinical application of cDNA microarrays in oncology. *Oncologist* 8, 252–258.
- Sambrook, J., Russel, D.W., 2001. *Molecular Cloning: A Laboratory Manual*, 3rd ed. Cold Spring Harbord Laboratory Press, Cold Spring Harbord, NY.
- Schroeder, A., Mueller, O., Stocker, S., Salowsky, R., Leiber, M., Gassmann, M., Lightfoot, S., Menzel, W., Granzow, M., Ragg, T., 2006. The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol. Biol.* 7, 3.
- Schuchhardt, J., Beule, D., Malik, A., Wolski, E., Eickhoff, H., Lehrach, H., Herzel, H., 2000. Normalization strategies for cDNA microarrays. *Nucleic Acids Res.* 28, E47.
- Spruessel, A., Steimann, G., Jung, M., Lee, S.A., Carr, T., Fentz, A.K., Spangenberg, J., Zornig, C., Juhl, H.H., David, K.A., 2004. Tissue ischemia time affects gene and protein expression patterns within minutes following surgical tumor excision. *Biotechniques* 36, 1030–1037.
- Xiang, Z., Yang, Y., Ma, X., Ding, W., 2003. Microarray expression profiling: analysis and applications. *Curr. Opin. Drug Discov. Dev.* 6, 384–395.