RNA Quality: Defining the Good, the Bad, & the Ugly

by Vivien Marx

Goldilocks, in the course of her serial porridge-tasting, had a pretty precise concept of which porridge was going to be acceptable to her palate. The criteria were her own, based on her own experience, and which may or may not be straightforwardly communicated and easily comparable to the tastes and needs of other porridge afficionados. Moving to the less fairytale-like world of science, RNA quality measurement shares some basic features with the Goldilocks protocol.

There are spectrophotometric methods to assess RNA quality, and one can also run agarose gels. With the introduction of the Agilent 2100 bioanalyzer from Agilent Technologies Inc., Palo Alto, Calif., scientists began looking at the integrity of the ribosomal bands on the mini gel-like image or the ribosomal peak ratio on the electropherogram the instrument delivered. Although the methods have changed, quality assessment has not yet become a quantitative science.

RNA quality is eyeballed, says David Ginzinger, PhD, director of the Genome Analysis Core Facility at the University of California San Francisco. "We are left with a very subjective call: this looks good; this looks kind of good; this looks bad," he says. If data is noisy and error bars too large when looking at samples taken from normal, benign, or metastatic tumors, significant differences can potentially be missed, he says. For a lab with CLIA (Clinical Laboratory Improvement Amendments) certification, strong checks and balances will need to assure that results are bona fide and real because it is samples from patients that will be analyzed to diagnose and treat, he says.

Standardization is perhaps more commonly associated with physics and computer science than biology. "The fact that biological materials tend to be variable—variation is crucial to evolution, after all—make the process of standardization different than in physics, to my mind," says Angela Creager, PhD, a biochemist and historian of science at Princeton University, Princeton, N.J., who has studied the interactions between physics and biology.

In 2002 the Association of Biomolecular Resource Facilities (ABRF) Microarray Research Group (MARG) reported on the need to to separate the methodological variation from the informative variation in the biology itself, when it comes to microarrays.

When Herbert Auer, PhD, director of the microarray unit at the Ohio State University Comprehensive Cancer Center (OSUCCC), Columbus, Ohio, and his colleagues looked at differentially expressed genes in peripheral blood cell samples in normal patients and compared intact and degraded samples independently, they discovered that up to 75% of microarray-based differential gene expression measurements were due solely to the difference in RNA integrity between two samples and were not true biological differences. Besides losing data to noise, there is a risk of false positives, when ". . . we see a difference between patient A and patient B and we believe this difference represents biological differences, when that is just not the case," he says. Such examples illustrate the pressing need for calibration.

Seeking quality RNA

There are comprehensive efforts, for example, the External RNA Controls Consortium (ERCC). There are also more particular proposals, such as new tools to quantitatively ascertain RNA quality. "If we could get a lock on that, it would address an unmet need in the community," says Janet Warrington, PhD, senior director, clinical genomics research and development, Affymetrix Inc, Santa Clara, Calif. "It is very important to understand the integrity of the sample you are hybridizing."
What follows is a look at some practical situations, challenges, and solution proposals pertaining to RNA quality.

The main emphasis at the lab of Federico Monzon, MD, at the University of Pittsburgh department of pathology, at the Medical Center's Shadyside Hospital, is to bring microarray analysis into clinical practice. Many of his samples are composed of prostate tissue used for diagnosis as well as research. "In the hospital environment, the conditions are not ideal for preserving RNA," Monzon says. He and his colleagues cannot just run a sample to see how the research might turn out, he says. Samples are precious, often unique—the result of invasive procedures. One cannot simply go and get more. "We need to find out how a sample is going to behave in analysis," he says.

It's a RIN

Standing at his poster for four hours at the recent American Association of Cancer Research meeting, Odilo Mueller, PhD, assay manager at Agilent Technologies, Palo Alto, Calif., says he spoke to a steady stream of visitors. The topic of interest: RIN or RNA Integrity Number. This month, Agilent Technologies is officially introducing RIN for the company's 2100 bioanalyzer. It will replace the output the machine currently delivers and will become an integrated part of the instrument's software. The beta release of the software is free to download until the next software release, which will be sold, probably later this year.

RIN was developed as a reaction to customer demand for a tool to help characterize facets of the total RNA electropherogram that are telling of sample quality, says Mueller. "We are not trying to create a standard but I think people are standardizing on the bioanalyzer now and will standardize on this tool in the future," he says.

As the miniaturization trend in this lab technology has evolved from gel electrophoresis to capillary electrophoresis to chip electrophoresis, which the bioanalyzer performs, its output too is evolving. Currently, RNA samples are usually gauged according to ribosomal peak heights on the electropherogram, in particular the ratio between distinctive 18S and 28S ribosomal peaks. "A ratio close to 2 is commonly considered good. [A ratio less than] 1 is cause for caution," Mueller says. Monzon says that with all the extraction methods his lab has tried, he hardly ever gets a ratio of 2, usually values are between 1.5 to 1.8. As time went on they grew more confident about their data, but, as many researchers report, found an appropriate quality assessment lacking.

Although older Agilent brochures state that RNA quality can "easily" be determined by visually inspecting the electropherogram, Mueller admits that the ribosomal ratio is not as informative as one might like. Intact samples can have a wide range of ribosomal ratios and the ratio does not take the entire spectrum of RNA degradation into account, he says. "This number doesn't tell you a lot. There is a certain relation between that number and the integrity, but it is a weak relation," he says.

RIN is set up to allow users faster quality control, Mueller says. "Having this number, you don't have to interpret electropherograms anymore," he says. "If a 9 is shipped and arrives a 6, something happened in transit." With this kind of measurement, Agilent believes that comparison between labs and within a lab should become easier. When publishing results of a microarray experiment, scientists could publish their RIN, too, Mueller says. For example, ". . . you may find that whenever you have a 7, the microarray works; when you have a 5 or lower, it doesn't." The RIN offers repeatability and quality control to determine whether you use that sample or discard it, saving another researcher time, he says.

The new software opens all previous bioanalyzer files, can reanalyze the old validation, and can correlate with the RIN. Scientists will then be able to determine the cut-off threshold for the usefulness of samples. The algorithm at work behind the scenes takes the electropherogram output and calculates the RIN between 1 and 10, with 1 describing degraded RNA and 10 corresponding to intact RNA. The software itself was developed as a collaboration between Agilent, Quantiom Bioinformatics GmbH, Weingarten/Baden, Germany, a company specializing in neural networks, and a computer scientist completing his master's thesis and seeking to put his programming skills to work on a biologically relevant problem. Mueller jumped at the chance to give him one. "That is how we came to employ neural networks, which allowed us to choose relevant parameters, weigh them," he says. The adaptive learning process involved training the algorithm on 1,200 eukaryotic total RNA samples, comprising human, rat, and mouse samples. "We developed it for and tested it with these types of samples," he says. They tested it for the nano not the pico assay and have not tested it extensively for prokaryotic samples. The samples were

provided by the German Cancer Research Center (DKFZ) in Heidelberg, Germany, an institution with which Agilent and Mueller have collaborated in the past.

The algorithm itself is proprietary, says Mueller. "We look at the entire electrophoretic trace to determine the integrity status of a sample. The main part we look at is the presence or absence of degradation products, which can be seen between the two ribosomal bands and below the 18S band."

Weighing in on these new options

"I am very happy that Agilent has developed this new software," says Nadia Novoradovskaya, MD, PhD, staff scientist at Stratagene Inc., La Jolla, Calif., "We have just been waiting for this." Scientists will begin trying it out and their experience will reveal how useful it is, she says. "I think the RIN number will offer better measurement of RNA integrity than the current ratio," she says.
Because Stratagene produces large volumes of RNA, so-called universal reference RNA, which is used to standardize, for example, microarray experiments, they need quality controls for their RNA. The company uses the bioanalyzer to accomplish this task for each individual RNA produced as well as when they blend RNA into the human reference or mouse reference RNA. Up until now, they decided internally not to use RNA less than a 1.5 for the 18S/28S ratio, which is a less than perfect number in her view, since these figures are tissue-specific. "We set up our own internal standard to control the quality of the universal reference RNA . . .," she says. "We are going to begin using the new software and from our experience we will determine a number that we will prefer to use," she says. The company ships RNA precipitated in ethanol and on dry ice, which they feel is pretty stable. The RIN will permit quality control to make sure the shipment has the same quality as when it left the company.
At UCSF's Comprehensive Cancer Center, discussions are underway to more closely connect research and clinic by creating a clinical diagnostics lab. David Ginzinger, PhD, director of the research-focused Genome Analysis Core Facility is involved in this process. He is also a RIN software beta tester.
"I think it is going to be very helpful for a number of studies, " he says. He and his colleagues are using RIN to stratify their samples, for example as a screening technique before spending any money on analysis, whether it is microarrays or reverse transcription PCR (RT-PCR). There will not be a universal RIN for all RNA or all experiments, he says, but probably values specific to tissues and/or protocols.
Implications for RIN will vary by technique, he says. "If a sample is pretty degraded, we probably won't bother doing a microarray experiment on it, but it will be fine to do RT-PCR on it," Ginzinger says. Thus, RIN scoring might indicate which technique is most promising for a particular sample. "I think there is a lot of value to be added to our analysis by having a RIN scoring system," he says.
In the operating room, time is of the essence, for both patient and tissue sample, with patients, of course, taking priority. Pancreatic cancer research, involving work on a disease which few survive, might well benefit from more gene expression study with microarrays. As Ginzinger says, enzyme-rich tissues such as in the pancreas self-degrade within minutes and so does the RNA quality. With an RNA quality scoring system, researchers might be able to clarify a protocol for working with patient samples. "This would give us an answer [for physicians, such as] 'If you can't get it to us in 2.5 hours, then we don't want the sample,' " Ginzinger says.
Monzon says he has been requesting this kind of software from Agilent for a while, has been testing the RIN software, and is finding it "quite useful." He has verified previous experiments with it. "We reran all those samples on the RIN software and got the data; it was very nice to see confirmed impressions," he says. He has also looked at other software to evaluate electropherograms, compared them and is tallying his results for an upcoming publication.

Other programmers, other software

A number of researchers interviewed for this article credit Herbert Auer as having significantly contributed to raising awareness on how RNA degradation affects microarray analysis. [see, for H. Auer, *et al. Nature Genetics* , vol. 35, pp. 292-293 (December 2003)].

Auer points out that the gold standard for RNA integrity is the 3'/5'ratio of the microarray analysis on, for example, Affymetrix chips. "Whenever there is degradation, the signal from the 3null ends is much higher than from 5null," Auer says. However, it is not practical to do the entire experiment first, to figure out the 3null/5null ratio. To begin to quantitate RNA degradation and avoid degradation bias in gene expression experiments, a few years ago he and physicist colleague Karl Kornacker, PhD, began developing software, called Degradometer software, and began offering it to the community in December 2003. It is available as a free download at www.dnaarrays.org . The software was not developed with Agilent, but rather the two were developed in parallel. Auer and Kornacker, in their attempt to obtain more quantitative information from the electropherogram, hadn't published about it before December 2003.
"The 18S and 28S peak: there are some rumors that the ratio of the peaks should be informative for RNA integrity, but there is no scientific background for this," he says. Customs, it appears, can quickly become habits. For Degradometer, Auer chose to look at the complete chromatogram, as representing three classes of RNA: 28S, 18S ribosomal RNA (rRNA) as well as the so-called small RNAs such as 5S rRNA and transfer RNAs (tRNAs). As RNA degrades, the 18S and 28S peaks decrease and degradation peaks occur in the molecular weight range between the small RNAs and the 18S signal, he says. "Molecular biology knows that these products do not exist *in vivo* usually," he says. "This is an artifact from preparation." The software obtains the ratio between these degradation peaks and the 18S peak, and multiplies that by 100 to deliver a degradation factor. The software also normalizes the data to account for the fact that not every run contains the same amount of RNA. Auer also tried to avoid the problems he sees in Agilent's current software, in the way peaks are analyzed in that the baseline of the triangle (the bottom of the inside of the peak) is generated automatically, and in his view, often erroneously.
Reactions from the academic as well as corporate community have been good, Auer says, judging from requests he has received from the United States, Japan, and Europe. He, too, has beta-tested Agilent's software. "They have to validate to an accepted [authority] to show their RIN is reliable," he says. "If they have a nice correlation between RIN and 3null/5null ratio, then it is a great tool." Both tools are a definite improvement over last year, he says, when nothing was available for establishing RNA integrity. Mueller says Agilent's software is "rugged compared to other software available" and is based on a very large data set.

Comparative impressions

Virginie Copois is a research engineer at Centre National de Recherche Scientifique (CNRS) UMR-5160 at the Cancer Research Center, (Centre de Recherche et de Lutte Contre le Cancer or CRLCC) Val D'Aurelle in Montpellier, France, both a clinic and a research institution. CNRS has special units (Unitnull Mixte de Recherche or UMR) founded by the government and a company, in this case it is Bio-Rad Laboratories Inc., Hercules, Calif. In an effort begun by Marc Ychou, MD, Bernard Pau, PhD, Frnulldnullric Bibeau, MD, and Aventis Pharmaceuticals, they are studying chemotherapeutic drug resistance particularly in colon cancer, using gene expression profiles of patients. Monzon, Ginzinger, Auer and Copois also report that surgeons in their clinics, who put the patient's well-being first, increasingly understand the requirements of their research. "

The CNRS researchers compared four methods: RIN; 28S/18S ratio; their "in-lab scale" as well as the %Dgr/18S of the Degradometer for their capacity to discriminate poor and acceptable RNA quality for transcriptome analysis. The conclusion thus far: RIN is a very "accurate, useful tool." "We conclude that the best way to discriminate samples was our scale, RIN and %Dgr/18S all with the same level of discrimination," she says. "We confirm that the discrimination made with 28S/18S ratio was not significant enough." Further statistical analysis with Affymetrix chips confirmed these impressions for the group and they are currently working to publish their results.

In her experience, the Degradometer software has been a somewhat less convenient tool to use. She also says she found a limitation of RIN software in the analysis of low concentrated samples, those with less than 50 ng/uL. "This should be possible with the combination of RIN and Pico kit," Copois says. Agilent offers both an RNA Pico LabChip kit for total RNA assays of 200 to -5,000 pg/uL and the Nano LabChip kit for total RNA assays of 5 to 500 ng/uL). Pointing to a recent

publication the researchers have completed, she says, "The Pico kit coupled with RIN should be indispensable when working with LCM [laser capture microdissection] samples.

In his comparative experiments, Federico Monzon found both tools to be useful. With the Degradometer, he has run into problems with an area of the electropherogram of concern to him. "We have a lot of tumor samples that undergo apoptosis, and apoptosis has a very specific pattern of degradation that was not being addressed by the Degradometer," he says. It is based on the low end of the size spread, he says. "Anything between the 18S and 28S peaks, which is where apoptosis comes into play, is not taken into account." Monzon says.

What now has commenced is discussion over the utility of these new tools. Agilent's Mueller says that "feedback is welcome and appreciated" from scientists who have downloaded RIN software and tried it.

In Monzon's view, a fundamental issue remains, independent of which RNA integrity measurement researchers choose is whether degradation bias is reproducible. Scientists need to figure out whether degradation effects are random or hitting specific genes.

"What we are going to have to do as a community is look at how predictive that RNA quality assessment is of success in a microarray experiment," says John Quackenbush, PhD, investigator at The Institute for Genomic Research (TIGR), Rockville, Md. If it turns out to be predictive in this sense, the standard will be incredibly valuable, he says.

Quackenbush believes in the importance of open standards. "What I mean by 'open' is one that is clearly defined, that is not some proprietary algorithm," he says. The calculation is defined such that it is clear how the quality measurement is reached. Researchers can possibly make slight variations on it to suit their needs or develop one built on the calculation. This facet brings to his mind the call for standardzing electropherograms in DNA sequencing. "What eventually arose was Phred [Phil's Read EDitor] quality scores from Phil Green and his group [at the University of Washington, Seattle]," he says. While they may not be the optimal quality scores, says Quackenbush, they are based on much real data and became an objective measure for which the calculation method was disclosed. The Phred scores grew to be accepted as a way to characterize sequence quality and enabled exchange across sequencing methods.

Defining RNA quality more homogenously rather than with the current in-house lab standards at labs around the world appeals, it seems, to many researchers. It is like any other experiment or instrument, says Affymetrix's Warrington, it follows the "garbage in, garbage out" principle. "It doesn't really matter what platform you are on, you want to make sure the starting material is good," she says. "It is like cooking; if you want good food, you start with good ingredients."

Sample Acquisition	**RNA Extraction**	**Assay Method**	**Data Acquisition**	**Data Analysisand Preparation**
• Tissue	• RNA purity	• Microarray cDNA oligo	• Reporter label used	• Feature extraction
• Donor	• RNA integrity		• Labeling method	• Standard curve
• Collection method	• RNA stability		• Image scanner	• Normalization
• Tissue processing	• RNA storage	• QRT-PCR	• Detector	• Q-PCR reaction efficiency
• Sample storage		• Other		
• Sample age				

Alert	Degradation factor (%Deg/18S)
	0-8
Yellow	8-16
Orange	16-24
Red	>24
Black	18S and/or 28S peaks not reliably detected

SIDEBAR: What is RIN?

by Vivien Marx

The acronym stands for RNA Integrity Number (RIN) and is now being introduced by Agilent Technologies for its 2100 bioanalyzer.

It is new software developed with neural networks and adaptive learning trained on 1,200 eukaryote total RNA samples (human, mouse, and rat). The algorithm analyzes the entire electropherogram, offering a measure of RNA quality on a scale of 1-10.

The new software will open all previous bioanalyzer files, can reanalyze the old validation, correlate them with the RIN. The software is downloadable for free as a beta release and is purchasable with the next version scheduled to come out later this year. See www.agilent.com/chem/RIN