

MICROARRAYS

Microarray quality in the spotlight again

The formal reports of the work conducted by the MicroArray Quality Control (MAQC) consortium lay the groundwork for microarrays to gain acceptance as a clinical diagnostic and drug development tool.

The reliability of microarrays has fueled controversy in many scientific spheres. Despite some alarming reports on their lack of consistency, several publications by large groups of investigators have also concluded that, with solid planning and procedure harmonization, microarrays can deliver consistent results across platforms and research laboratories. A new set of reports by the MAQC consortium, which appeared in the September issue of *Nature Biotechnology*, is particular in that it directly addresses the possibility of using microarrays in a regulatory and clinical context.

To anyone who has set foot in a QC laboratory, this will be reminiscent of a common squabble between Research and QC departments. Words like 'repeatability' or 'robustness' can have fairly different meanings in the two environments. Investigators trying to discover genes differentially expressed in specific biological situations legitimately have a much more lenient view of the repeatability of their data than clinicians who would use this data to make a diagnostic decision, or a regulatory panel who would try to determine if a preclinical pharmacogenomic profile supports the safety of a new drug.

At the moment the US Food and Drug Administration does not use genomic data to make regulatory decisions, but they encourage sponsors of investigational new drugs to include such data in their application so that regulators can become aware of technical issues they eventually will have to deal with. Unfortunately, such forward-looking measures cannot achieve much more than to demonstrate variability. To start defining a framework for using array data, a systematic analysis of variability causes was necessary. This led the Food and Drug Administration to embark on the MAQC project together with the US Environmental Protection Agency, National Institute for Standards and Technology, major providers of microarray platforms, academic labs and other stakeholders.

Using multiple testing sites for each platform, the group collectively obtained

results for technical replicates of two distinct reference RNA samples (in four different combinations) on six commercially available microarray platforms (Applied Biosystems, Affymetrix, Agilent Technologies, GE Healthcare, Illumina and Eppendorf) and a spotted oligonucleotide array platform from the US National Cancer Institute.

The results are encouraging in that they show good intraplatform consistency from lab to lab, as well as good comparability between microarray platforms and good correlation with other gene-expression assays such as quantitative reverse-transcription PCR. By design, however, this exercise does not provide an absolute measure of the level of consistency between data sets. For example, it used technical replicates of samples that were purportedly very different. It is likely that 'real-life' samples with smaller differences in gene expression will not give the same level of consistency. Nevertheless, the data set characterized the technology platform very well.

Another illustration of the particularity of this study is the consideration of data analysis algorithms. The authors favor a straightforward approach of ranking genes according to their fold-expression change, with a non-stringent *P* value cutoff, a statistical analysis that would most likely not

be suitable in studies geared toward the discovery of gene signatures. This approach, however, performs well when it comes to ensuring the reproducibility of an established signature, like in a diagnostic setting. As Janet Warrington of Affymetrix puts it, "If you are doing discovery work, you want to see everything. But if you are developing a diagnostic signature, the last thing that you want to be doing is discovering new things when you are running that test."

But the regulatory-geared design of the MAQC exercise does not mean that it will not be helpful for research applications. On the contrary, it has been a unique opportunity, for example, to evaluate external RNA controls for overall quality assessment (see **Box 1**). Most importantly, a large data set is now available in combination with the reference samples, which were chosen because their manufacturers have guaranteed availability of the same specific batches for several years. With such tools, performance comparisons will be possible on a much better scale than before.

Veronique Kiermer

RESEARCH PAPERS

MAQC Consortium. The MicroArray Quality Control (MAQC) project shows interplatform reproducibility of gene expression measurements. *Nat. Biotech.* **24** 1151–1161 (2006).

See also other reports in the September issue of *Nature Biotechnology*.

BOX 1 MAQC: A CASE FOR CONTROLS

The MAQC analysis has provided an unprecedented opportunity to evaluate external RNA controls (ERC) for assessment of overall assay performance. Several microarray manufacturers suggest such controls, but the recommended procedure varies between platforms. Some manufacturers recommend spiking a control RNA in the total RNA sample, so that the control will be amplified and labeled with the sample; others suggest the addition of an independently labeled control just before hybridization to the array; others yet favor a combination of both approaches.

In the MAQC study, each platform was tested including the manufacturer's recommended controls providing a unique opportunity to analyze the performance of each type of control. Janet Warrington, who participated in the study, is enthusiastic about the resulting analysis, "This is actually very important data for the whole community to see, to help people understand how you can tease apart failure analysis."

Warrington also points out that this study will be of great value to another consortium, the ERCC, which is working on establishing an extensive set of well-tested ERCs useful across platforms and samples. The MAQC dataset will help make more informed choices for the design of new ERC and ERC-based analyses, undoubtedly contributing to a better tool.