Quality control

Doubt is often cast on the reliability of DNA microarrays, but resources are becoming available to help researchers overcome many of the problems inherent in this technology. Michael Eisenstein reports.

It is more than 15 years since DNA microarrays were developed, and in that time they have been adored, attacked and, in an effort to look beyond the hype, appraised. One outcome of this 'soul searching' has been the realization that the flaws inherent to gene-expression arrays are similar to those of other high-throughput platforms. "I don't think microarrays are different from other technologies in that respect, and it's important for people to keep that in mind," says Janet Warrington, vice-president of molecular diagnostics and emerging markets research and development for Affymetrix, based in Santa Clara, California. "I try to point out 'microarray exceptionalism' wherever I find it."

John Quackenbush, a computational biologist at the Harvard School of Public Health, sees several fundamental errors in the way many researchers tackle microarray gene-expression studies. "People tend to go out blindly and do experiments, then go back and try to analyse them and figure out what the question is afterwards. I think that's the first thing you have to avoid," he says. "It's also important to make sure you remove confounding factors from the experiment wherever possible."

Systematically sorting out sources of error can be a daunting process, but a recent series of investigations by people such as Quackenbush into cross-experimental, cross-platform and cross-laboratory variability between array experiments has helped clarify some issues that were preventing comparisons being made between experiments. Several multi-institutional projects are now under way to develop more reliable experimental protocols and controls (see 'Standards and practices'). Meanwhile, many scientists, manufacturers and programmers are working to develop practical tools that could help eliminate unwanted variability from experiments and analyses.

The biggest problems often occur early on. Existing kits for RNA preparation are effective, but many are designed for a 'best-case scenario': large amounts of fresh biological source material. Many researchers are now interested in studying gene expression in a small number of cells, necessitating efficient systems that can work with limited samples.

Affymetrix's tiling and exon arrays: two alternatives for in-depth screening of the human genome.

Various systems have been developed, many of which are based on a linear-amplification procedure known as the Eberwine method. Labelling strategies typically fall into two main categories: direct and indirect. Direct methods, used in systems such as CyScribe, available from GE Healthcare Life Sciences of Little Chalfont, UK, and ChipShot, from Promega of Madison, Wisconsin, typically involve the incorporation of fluorescent-dye-conjugated nucleotides during complementary DNA (cDNA) synthesis.

One favoured indirect-labelling strategy involves incorporating an aminoallyl-modified nucleotide into cDNA or amplified RNA transcripts, and then labelling these with

STANDARDS AND PRACTICES

Leming Shi of the US Food and Drug Administration (FDA) in Rockville, Maryland, is unabashed in his affection for microarrays. But he was disconcerted by the recent publication of several papers challenging the reliability of gene-expression microarray experiments. One article, for example, reported such disagreement that an analysis of 185 genes using three different technologies revealed concordant readings for only four transcripts (P. K. Tan *et al. Nucl. Acid Res.* 31, 5676–5684; 2003).

Subsequently, Shi and his colleagues found that an alternative analytical approach greatly improved cross-platform concordance for these data sets (L. Shi *et al. BMC Bioinformatics* 6 (Suppl. 2), S12; 2005). But the lingering climate of uncertainty, and concerns about the potentially serious implications for the use of microarray data in the FDA drug-approval process, led them to launch the MicroArray Quality Control (MAQC) project. The MAQC brought together research leaders from government, academia and industry to establish tightly controlled 'gold standard' comparisons of microarray systems. They began by identifying commercially available, trustworthy 'standard' RNA samples. But this was just the start. "The MAQC's main goal is to generate a vast reference data set," says Shi. "We have conducted more than 1,000 array hybridizations with these reference samples, plus we're using three alternative technologies and we requested that each system be evaluated at three testing sites." The MAQC recently completed a review of its final data, and will present its findings in a series of articles to be published in *Nature Biotechnology* next month.

Leming Shi hopes to improve the reproducibility and comparability of microarray work.

In the meantime, many in the field are awaiting the outcome of a complementary initative: the External RNA Controls Consortium (ERCC). This evolved from a 2003 meeting headed by the US National Institute of Standards and Technology in Gaithersburg, Maryland. It aims to identify and help make commercially available a collection of reliable RNA 'spike-in' controls, which can be included in any microarray experiment to assess variables such as labelling and hybridization efficiency. Participation has grown rapidly, and ERCC leader Janet Warrington, who is a vice-president at Affymetrix in Santa Clara, California, finds the early progress promising. "A number of organizations that were already using their own controls have donated these — no strings attached — for testing," she says. "So we have a collection of 100 to 150 controls that will be tested across platforms and we have eight sites that have volunteered to carry out the testing."

Both projects have benefited from collaborative environments that have allowed even direct competitors to work together towards a shared goal. "We all share the belief that if we're successful, we'll expand the marketplace for everyone," says Warrington. M.E.