EJSO 2002; 28: 567–568
doi:10.1053/ejso.2002.1333, available online at http://www.idealibrary.com on IDEAL®

EJSO
European Journal of Surgical Oncology

EDITORIAL

The DNA deluge

The past few years have seen a quantum leap in DNA assay technologies.[1] DNA microarray instruments allow the measurement of the presence and relative abundance of many hundreds or thousands of preselected genes or engineered nucleotide sequences from biological and clinicopathological samples. The studies are conducted using computerised, fluorescence based assays on nanoscopic quantities of DNA or RNA on glass slides or nylon membranes. They generate huge volumes of qualitative and semiquantitative data in colourful, graphical format with relatively little technical or intellectual effort. The mass of data challenges the limits of human computation and attention.

The molecular and cell biology journals are now full of advertisements for instruments, assay and calibration kits which transfer the technology from the Nobel prize winning research groups to the clinical laboratories. The clinical journals may now anticipate an exhaustive and an exhausting flood of manuscripts reporting the gene expression profiles of such and such a collection of retrospectively or prospectively collected tissue and tumour samples in clinico-pathological series. A typical manuscript will describe the source of the samples; the source of the assay kit; and a series of molecular biological recipes relating to the techniques of generation of the DNA binding nucleotides on the 'gene chip'; the extraction and hybridisation binding of the many sample genes to the chip; the imaging system, control genes and a verification technique for selected marker genes such as the reverse transcription-polymerase chain reaction (RT-PCR) and Northern hybridisation (for RNA markers); lists of complex polynucleotide CAGT sequences for primers, and analysis software to deconvolute the data.

The results sections of these papers will report the up- or down-regulation of long lists of genes with polysyllabic names and alphanumeric indicators, and will be supported by figures of colour coded grids for many hundreds of genes. The discussion sections will address the significance of the changes in various of the groups of 'headline' genes for the biology and behaviour of the tissue or tumour under study. The more readily understandable analyses will be helped by assay kits which have preselected and pre-organised on the gene chips those groups of genes identified with the cell cycle, with drug resistance and metabolism, with oncogenesis, with apoptosis, invasion, adhesion, cell signalling, growth factors and so on. Such papers will be well written and quite plausible, and will find their way into print with increasing frequency in the general journals.

What are we as surgeons, clinicians, referees and editors to make of this avalanche? Will we face a decade or more of confusion over coffee and bewilderment at bedtime from our favourite journals while the clinical truths are filtered out of the white noise?

We can already make a number of reasonable observations and deductions about microarray studies which should guide us through the digital maze and save the rainforests. The deluge of microarray data will be quite meaningless when offered in isolation, for a number of reasons, and however much techno-molecular virtuosity may appear to have delivered it.

In terms of the patterns of gene expression, comparators are needed with the normal source tissues, otherwise we cannot know the significance of the changes in the expression of individual genes. However, there are pitfalls for the unwary. Even where micro-dissection techniques are used, biological samples are likely to be very heterogenous of cell types, including tumour cell lineages, connective tissue cell lineages, and leukocyte lineages. We cannot be sure that the genetic changes posted relate to the principal tumour cell type. In terms of the technology, there is plenty of scope for error and overinterpretation in highly sophisticated and technically demanding devices, of which the occasional researcher must be very aware.

More importantly, the very complexity of the patterns will ensure that we cannot be certain that changes in the arrays relative to normal tissues describe cause-and-effect relationships. We now believe that many genes and gene groups act in concert in their mutual inter-regulation and in the synthesis of multigenic proteins in a phenomenally complex homeostatic microenvironment.

Most importantly of all, the true effector of gene expression in the three-dimensional world of the cell and tissue is the proteome. The proteome is the total

0748–7983/02/$35.00 © 2002 Published by Elsevier Science Ltd

population of expressed proteins in any one cell which turns over at billions of molecules per second in intricate microenvironments to deliver the stabilities and instabilities which we recognise in cell architecture and physiology. Until we know how DNA microarray expression profiles translate into protein profiles, and how the patterns of gene expression are reflected in fluctuations in the protein population, *ex cathedra* statements on the function of this, that or those genes will be overly speculative guesswork. These revelations are still some years away, as the minds which brought us mass production technologies which delivered the human genome are refocused on the far more demanding technical and conceptual challenges of the human proteome. Until then, we will be challenged to draw any useful conclusions about the gross patterns in the raw gene chip profiles.

We must thus set a number of objectives and guidelines for researchers who would study and report microarray assays in clinical journals, to ensure maximum clarity and minimum obfuscation. Microarray studies should focus on meaningful and carefully thought out biological objectives and be hypothesis driven, rather than machine driven. For example, many of the genes involved in drug metabolism and resistance have now been collected on commercial chips, and these would seem to be a more fruitful approach to clinical oncology material in the applied laboratory than whole genome series. Clinico-pathological researchers would thus do well to focus on macrochips, with a hundred or so gene targets, rather than microchips with thousands. The study of small numbers of knowingly closely related genes and their protein products is likely to be most rewarding.

To communicate clearly and consistently to the 'lay' professional reader, the research community must find a graphical language which distils interminable Northern blots, Southern blots and hybridisations from all other points of the compass, two-dimensional grids with thousands of coloured bars; and we should call for an embargo on the unnecessary publishing in clinical journals of long nucleotide sequences which can remain in the domain of the website and digital archive. The EJSO will promote and advance the best and most relevant work in this field. The onus is upon the research community to adopt and apply these challenging new technologies in a responsible and meaningful way.

D. A. Rew
Consultant Surgeon and Honorary Senior Lecturer Southampton University Hospitals Southampton SO14 0YG, UK

REFERENCE

1. Rew DA. DNA Microarray technology in cancer research. *Eur J Surg Oncol* 2001; **27**: 105–10.

Accepted for publication 22 June 2002