Morrison, Norman

Robust normalization of microarray data over multiple experiments

Norman Morrison[1], Magnus Rattray[1], Martin Brutsche[2], Stephen G. Oliver[3], Andrew Hayes[3], Nianshu Zhang[3], Chris Penkett[4], Jacqui Lockey[4], Sudha Rao[4], Ian Hayes[4], Ray Jupp[4] & Andy Brass[1]

[1]Department of Biochemistry, 2.205 Stopford Building, University of Manchester, Oxford Rd, Manchester M13 9PT, UK
[2]North West Lung Research Centre, South Manchester University Hospital, Wythenshaw, Southmoor Rd, Manchester M23 9LT, UK
[3]Department of Biomolecular Sciences, UMIST, PO Box 88, Sackville St, Manchester M60 1QD, UK
[4]Rhône-Poulenc Rorer Limited, JA3-3, Rainham Road South, Dagenham, Essex RM10 7XS, UK

Microarrays have set the stage for an explosion of large-scale expression data, driven by a diversity of genome sequencing projects. The technology has already demonstrated its applications in analysis of model systems, such as the response of mammalian fibroblasts to serum and sporulation in yeast. The comparison of data between multiple experiments run as a time series or under different conditions is not a trivial task. Although the analysis is challenging, it has the potential to answer some of the most interesting questions regarding information mining on gene expression patterns or function. To address these questions we have investigated standardization methods over multiple expression analysis experiments covering systems from high-density microarrays (~40,000 individual gene transcripts) to membrane applications (~500 individual gene transcripts). By making the assumption that global changes in gene activity are negligible, we show that normalization over the entire set of gene expression values in a given profile (provided that profile is not biased for examination of a particular system) provides a more statistically robust method than using housekeeping gene expression values. We also show that there is no significant reason for normalizing with a reduced subset of genes over a given range of expression. We have compared expression data derived from two different technological systems (glass slide and filter based); both of these systems have an intra-experimental distribution close to log-normal. We therefore normalize by mapping logged expression values within each experiment to a standard distribution with zero mean and unit variance. This transformation can be seen to effectively reduce to a minimum intra- and extra-experimental variances when analysing replicate experiment data. These methods are currently being applied in the statistical analysis of differential expression among patient groups and in the analysis of model organisms subject to certain conditions.

Mousses, Spyro

Identification of genes involved in hormone-independent prostate cancer by cDNA microarrays, followed by in vivo analysis of selected genes using tissue microarray analysis

S. Mousses, L. Bubendorf, J. Kononen, M. Bittner, Y. Chen, M. Kolmer, A. Elkahloun, P. Koivisto, T. Pretlow, P. Schraml, G. Sauter & O.-P. Kallioniemi

CGB/NHGRI/NIH, Bethesda, Maryland 20892, USA
Case Western University, Cleveland, Ohio, USA
University of Tampere, Finland
University of Basel, Basel, Switzerland

Although prostate cancer initially responds and regresses in response to androgen-depletion therapy, most human prostate cancers will re-grow as an androgen-independent tumour. The goal of our study was to apply functional genomics to identify gene expression changes involved in this process. Two high-throughout technologies, cDNA microarrays and tissue micorarrays, were applied to explore the molecular mechanisms underlying hormone-refractory prostate cancer.

A cDNA microarray consisting of 6,048 transcripts was constructed (including 4,032 known genes, and 2,016 ESTs) and applied to the identification of differentially regulated genes between a panel of hormone-sensitive human prostate cancer xenografts (CWR22), and their hormone-refractory derivatives (CWR22R). A tumour tissue microarray[1] with 269 clinical specimens (27 benign prostates, as well as 50 incidental, 138 clinically localized and 54 hormone-refractory recurrent cancers) was constructed to investigate whether candidate genes discovered by the cDNA microarray are also involved *in vivo* in human prostate cancer progression. For example, upregulation of one of such gene, insulin growth factor binding protein 2 (IGFBP2), was significantly (P<0.0001) associated with hormone-refractory tumours using the tissue array analysis. In conclusion, the combination of cDNA and tissue microarray analysis is a powerful strategy to identify novel cancer associated genes, and to rapidly explore their role in clinical prostate cancer.

1. Kononen, J. *et al. Nature Med.* **4**, 844-847 (1998).

Müller, Uwe

Quantitation of sequence copy-number changes in genomic DNA through GenoSensor-based comparative genomic hybridization

U.R. Müller, Y.P. Bao, D. Che, N. Lermer, W.R. Li, J. She, T. Ruffalo, A. Prokhorova, D. Lane & S. Seelig

Vysis Inc., Downers Grove, Illinois, USA

Comparative genomic hybridization has been adapted to the microarray format with the GenoSensor™ system, which consists of a unique chromium-coated chip surface, multi-colour fluorescent hybridization chemistry and a wide-field, non-scanning, CCD-based imaging system. In genomic assays differentially labelled test (green) and reference (red) DNA are co-hybridized to the chip in the presence of Cot-1 DNA (to suppress repeat sequences). Custom software determines the green/red ratios for each pixel under each target spot to deduce sequence gains or losses. The combination of these features provides for high sensitivity and a linear dose response over several orders of magnitude in model systems. For appli-