Poster abstracts

Dressman, Marlene

Signing on to expression profiles

Marlene M. Dressman[1], Eric Holland[2], Michael H. Polymeropoulos[1] & Christian N. Lavedan[1]

[1]Novartis Pharmaceutical Corp., Pharmacogenetics, Gaithersburg, Maryland, USA
[2]MD Anderson Cancer Center, Houston, Texas, USA

Unique expression patterns of a large number of genes can now define cell specific signatures under a given context. Cell specific signatures will be invaluable to understanding differentiation, cell type specificity and the cellular biology of environmental responses. Expression signatures could greatly enhance drug development by increasing efficacy and decreasing toxicity of therapeutic compounds. We have used (Affymetrix) oligonucleotide microarrays to survey and compare the expression profiles of 7000 human genes from cells of various origins, including lymphocytes, umbilical vein endothelial cells (HUVEC), a breast adenocarcinoma cell line (MCF7) and a glioblastoma cell line (U373MG). Environmental context such as time in culture, state of differentiation, media composition, CO_2 content and culture confluency may alter cell signatures. Chemical compound manipulation of cell lines provides one mechanism to dissect function at the gene expression level. For example, extended treatment of endothelial cells with vascular endothelial growth factor (VEGF) results in greater than two-fold changes in expression level of approximately 100 genes, while extended treatment of glioma cell lines with platelet derived growth factor (PDGF) results in greater that two-fold changes in 39 genes. Analysing expression data from multiple cell lines can also be used to identify similarities of expression profiles for all cell lines. We have observed that an average cell type expresses 2335 out of 7000 genes assayed (approxiamately 33 %). Sixty-four percent of the 2335 genes are expressed in endothelial, U373MG, and MCF7 cells, with a total of 630 expressed at similar levels in all three cell types despite variation in growth media. Additional cell types are being analyzed to refine the universal and cell-specific signatures.

Dudoit, Sandrine

Statistical methods for the characterisation of tumour types using cDNA microarray data

Sandrine Dudoit

Statistics, University of California at Berkeley, Berkeley, California, USA

We explore an ensemble of multivariate statistical methods for the analysis of gene expression data from cDNA microarray experiments. The statistical questions we investigate are motivated by the experimental program carried out in the laboratories of professors Brown and Botstein at Stanford University to characterise the molecular variations among cancers of the breast, prostate, liver and brain, based on transcript abundance for tens of thousands of genes in several hundred tumour samples. The aims of the statistical analysis are to assist the biologists in (i) developing a classification or "taxonomy" of tumours based on gene expression data, and (ii) identifying a subset of "marker" genes that characterise the different tumour types defined in (i).

The statistical analysis for addressing aims (i) and (ii) involves a synthesis of approaches from the fields of cluster analysis and discriminant analysis, also known in the pattern recognition literature as unsupervised and supervised learning, respectively. Unsupervised methods (e.g. projection methods, partitioning methods such as k-medoids, hierarchical clustering and self-organising maps) are investigated to identify possible tumour types, some already recognised (e.g. tumour site of origin) others (most?) not. A concomitant question is the develop-

ment of statistics for comparing different clusterings. Supervised methods (e.g. linear discriminant analysis, nearest neighbour methods, neural networks and tree-structured classifiers) are used to examine in greater detail the definitions of tumour types from unsupervised learning methods and identify a subset of marker genes which provide a more compact description of these tumour types. Our analysis focuses on the CART (Classification and Regression Tree) method and its extensions to perform variable (marker gene) selection and develop a classifier for tumour types. We apply re-sampling methods such as bagging and boosting to improve the accuracy of the classifier and the variable selection process.

Duffield, Giles

Application of microarray technology to identify novel components of the mammalian circadian clock

Giles E. Duffield[1], Jill A. Wahleithner[1], Jonathan D. Best[1], William J. Schwartz[2], Jennifer J. Loros[1] & Jay C. Dunlap[1]

[1]Department of Biochemistry, Dartmouth Medical School, Hanover, New Hampshire 03755, USA
[2]Department of Neurology, University of Massachusetts Medical School, Worcester, Massachusetts 01655, USA

Circadian rhythms are integral to eukaryotic life, structuring temporal patterns of physiology and behavior. These cycles are not a passive response to the external environment, but are driven by an internal biological clock with an endogenous period approximating 24 hours. This biological clock and its constituent outputs are represented by regular cycles in gene expression that continue in the absence of any environmental stimuli. The suprachiasmatic nuclei of the hypothalamus (SCN) are the site of a central circadian clock in the mammal. The aim of our current study is to identify and characterise genes that are components of the central oscillator or that are elements of the input or output (clock controlled genes) pathways of the mammalian circadian clock. Time-of-day specific cDNA libraries have been constructed from mRNA extracted from the SCN (0.8 mm^3 punched tissue sections) and the cerebella of 2-month old male C57BL/6 mice maintained on a 12:12 light:dark cycle. Animals were sacrificed in three different groups: 0.2-0.7 h after lights on (zeitgeber time [ZT] 0.2-0.7, morning), ZT5.5-6 (midday) and ZT11-11.5 (evening, 0.5-1 h before lights off) (ZT0 = lights on). These timepoints and light-dark cycle (entrainment) conditions were chosen since (1) known clock genes and clock-regulated genes peak in expression at various time points after the dark/light transitions, and (2) the amplitudes of rhythmicity are greatest under entrainment conditions. By collecting tissue under entrainment conditions we have been able to use light-induced immediate early genes such as *c-fos* as internal controls to validate the method. For each collection, brain tissue from 90-120 animals was extracted and frozen rapidly (approximately 1.5 min between sacrifice and freezing of each tissue punch), yielding amounts of poly A+ RNA sufficient to construct representative libraries containing even rare transcripts.

Data will be presented describing differential gene expression in the mammalian brain based on location (e.g. SCN versus cerebellum) and time-of-day (e.g. midday versus evening). We anticipate that microarray screening will identify novel clock genes, whose expression changes as a function of time of day, and thus in turn will reveal components of the central oscillator, or components of the input or output pathways to the clock.