Der, Sandy

Characterization of interferon-regulated pathways using oligonucleotide microarrays

Sandy D. Der, Mike J. de Veer, Aimin Zhou, Robert H. Silverman & Bryan R.G. Williams

Department of Cancer Biology, Lerner Research Institute, Cleveland Clinic Foundation, Cleveland, USA

Type I (α, β) and type II (γ) interferons (IFN) possess pleiotropic functions, including antiviral, antiproliferative and immunomodulatory activities, which are likely mediated by the induction of unique subsets of IFN-stimulated genes (ISG) that are differentially regulated by IFN subtypes and in different cell types. Oligonucleotide arrays (Affymetrix) were used to investigate differential gene expression in the human fibrosarcoma cell line, HT1080, or primary mouse embryonic fibroblasts in response to IFNs or virus infection. In all samples, the detection of known IFN-regulated genes helped to validate the experimental design and conditions. Combined analysis of these data sets, each with several hundreds of differentially expressed genes, has led to the idenfication of groups of IFN-regulated genes with similar functions, previously unknown IFN-regulated pathways and potential to crosstalk with other signalling pathways. For example, ISGs with similar functions included apoptosis regulators, chemokines, genes involved with antigen processing and presentation, GTPases, heat shock proteins and members of the ubiquitin system. Signalling molecules identified as potentially novel ISGs included known kinases, phosphatases and signalling adapter molecules. The functional significance of each gene relating to IFN functions requires further study on an individual basis, but the continuous refinement of a global view for IFN-regulated pathways may identify additional yet unknown areas of IFN action and enable better design of therapeutic strategies based on IFN-regulated mechanisms.

Diehn, Max

Large-scale identification of secreted and membrane-associated gene products using DNA microarrays

Max Diehn

Medical Scientist Training Program & Department of Biochemistry, Stanford University School of Medicine, Stanford, California 94305, USA

We describe a rapid and efficient method for identifying secreted and transmembrane gene products. From yeast and human cell cultures, mRNA species bound to rough endoplasmic reticulum (rER)-associated polysomes were separated from free and cytoplasmic polysome-associated mRNAs by sedimentation equilibrium or sedimentation velocity. The representation of mRNA species in each population was quantified by hybridization to high-density DNA microarrays. Transcripts known to encode transmembrane or secreted proteins were significantly enriched in the rER-associated fractions, whereas those known to encode cytoplasmic proteins were enriched in the fractions containing free and cytoplasmic polysome-associated mRNAs. Based on our data, we are able to predict with reasonable confidence that over 275 human and 300 yeast genes encode previously unrecognized transmembrane, ER/golgi-resident or secreted proteins. The combination of fractionation of nucleic acids on the basis of physical properties and analysis on DNA microarrays is a general method that can be applied to a diverse set of biological questions.

Dougherty, Edward

Nonlinear stochastic determination for gene expressions via cDNA microarrays

Edward R. Dougherty[1], Seungchan Kim[1], Michael L. Bittner[2], Yidong Chen[2], Krishnamoorthy Sivakumar[3], Paul Meltzer[2] & Jeffrey M. Trent[2]

[1]Department of Electrical Engineering, Texas A&M University, College Station, Texas, USA
[2]Cancer Genetics Laboratory, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland, USA
[3]Department of Electrical Engineering, Washington State University, Pullman, Washington, USA

A basic problem in multivariate systems is finding predictive relations between the measurable variables in the system. The extent to which a target variable can be predicted via measurement of a set of predictor variables reveals the level of interaction between the variables. We want to quantify the degree to which the target variable is statistically determined by the predictor variables, where statistical determination does not imply physical causality. In signal processing systems, stochastic determination provides a window into the overall control structure. In the simple case of two variables possessing a bivariate normal distribution, the best predictive relation is linear and the correlation coefficient provides the desired coefficient of determination. The problem can still be framed in terms of correlation for multivariate linear systems (albeit, not in such a straightforward form), but not for nonlinear systems.

We examine quantification of nonlinear multivariate stochastic determination among gene expression levels using cDNA microarrays. The method allows incorporation of knowledge of other relevant conditions, such as the application of particular stimuli or the presence of inactivating gene mutations, as predictive elements affecting the target expression level. The approach is general and can be applied to any class of nonlinear predictor functions. For our talk, prediction is based on a ternary perceptron, to which is input one of three values for each predictor gene: +1 [up-regulated], -1 [down-regulated] or 0 [invariant]. External conditions are quantified as +1 [present] or 0 [not present]. Our reason for choosing a perceptron is twofold: first, it is intuitive; second, for n predictor variables, it has only $n + 1$ parameters to estimate and therefore requires much less data than more general nonlinear predictors.

Owing to the large number of genes and a severe limitation on experimental replication when using microarrays, any individual predictor derived from the data lacks statistical significance relative to the population. Since our interest is measuring degrees of determination, we can derive the best predictors from the sample data and then estimate the determination coefficients for various predictor sets from the sample data. These estimates will be biased high, but can be used as sample coefficients to relatively quantify the degrees to which various expression-level sets stochastically determine a target expression level. A more accurate, more computationally costly way to proceed is to employ a bootstrap method to estimate the coefficients of determination. The bias will on average be low, but we can again obtain relative quantification.

The entire procedure is supported by software that has a number of facilities to simplify data analysis, and which provides graphics for visualising experimental data and sequential increases in determination as additional expression levels are used for prediction.