Dedhar, Shoukat [45]

Integrin-linked kinase regulates E-cadherin expression by means of the transcription factor Snail

Shoukat Dedhar

University of British Columbia and British Columbia Cancer Agency, Vancouver, British Columbia V6H3Z6, Canada

The integrin-linked kinase (ILK) is an ankyrin repeat containing serine–threonine protein kinase, which can interact with the cytoplasmic domains of integrin β_1 and β_3 subunits and can be activated by integrins and growth factors. When overexpressed in epithelial cells, ILK inhibits E-cadherin expression, causing an epithelial to mesenchymal transformation. I demonstrate that ILK represses E-cadherin expression by stimulating the expression of Snail, a transcription factor that binds and represses the E-cadherin promoter. Inhibition of ILK in colon cancer cells with a highly selective ILK inhibitor, or dominant-negative ILK, resulted in the inhibition of Snail expression and a concomitant stimulation of expression of E-cadherin. Inhibition of ILK also suppressed β-catenin/TCF transcriptional activity in *APC*$^{-/-}$ colon cancer cells and induced cellular growth arrest as well as growth inhibition of human colon tumors xenografted into SCID mice. Inhibition of ILK may reverse the frequent epithelial-to-mesenchymal transformation associated with tumor invasion and metastasis. ILK is a promising target for the control of cancer progression.

Diehl, Frank [46]

Expression profiling of cervical cancer biopsies using high-resolution dual-label imaging of ^{33}P and ^{3}H on microarrays

Frank Diehl[1], Matthias Nees[2], Heinz Filthuth[3], Magnus von Knebel-Doeberitz[2] & Jörg Hoheisel[1]

[1]*Department of Functional Genome Analysis, German Cancer Research Institute (DKFZ-Heidelberg), Im Neuenheimer Feld 506, 69120 Heidelberg, Germany*
[2]*Department of Surgery, University of Heidelberg, Im Neuenheimer Feld 110, 69120 Heidelberg, Germany*
[3]*Zinsser Analytic, Eschborner Landstrasse 135, 60489 Frankfurt, Germany*

To identify gene clusters that are involved in the development and progression of cervical cancer we hybridized biopsy material to a custom microarray. Since only small amounts of RNA can be isolated from biopsy material we applied a detection system that is more sensitive than fluorescence detection. The approach combines the advantages of competitive hybridization with the high sensitivity of radioactivity on a microarray platform. Instead of using two fluorescent dyes for labeling we co-hybridized ^{33}P- and ^{3}H-labeled probes to our microarray. For the simultaneous detection of the two isotopes we used a new technology that allows the detection of radioactivity in real time with very high resolution (10 μm). The unique features of the instrument are a high detection sensitivity for ^{33}P and ^{3}H and an infinite linear range of detection. To select the genes on the array we made use of public gene expression data generated in the past few years (approximately 2,000 clones). In addition to the genes selected from the literature, we added genes obtained from complementary DNA representational difference analysis experiments (approximately 5,000 clones).

Diehn, Maximilian [47]

From gene expression to the bedside: A high-throughput approach for identifying serum diagnostic markers and immunotherapy targets in human malignancies using DNA microarrays

Maximilian Diehn[1], Michael B. Eisen[2], David Botstein[2] & Patrick O. Brown[1,3]

[1]*Department of Biochemistry, Stanford University School of Medicine, Stanford, California, USA*
[2]*Department of Genetics, Stanford University School of Medicine, Stanford, California, USA*
[3]*Howard Hughes Medical Institute and Stanford University School of Medicine, Stanford, California, USA*

Tumor-specific membrane-associated and secreted proteins represent a clinically important class of proteins because they have the potential to serve as serum diagnostic markers or immunotherapy targets. We have recently shown that, in addition to allowing gene expression profiling of human tumors, DNA microarrays can be used to identify new membrane-associated and secreted proteins. This can be accomplished by separating messenger RNA species bound to membrane-associated polysomes from other mRNAs and then quantifying the distribution of transcripts in the two fractions by hybridization to DNA microarrays[1]. Using this technique we have now identified approximately 4,000 human gene products that are likely to encode previously unrecognized secreted or membrane proteins. By combining this data set with gene expression profiling data sets from a variety of human malignancies (including those of the breast, liver, kidney, prostate, ovaries, brain and immune system) we have further identified a subset of these 4,000 genes that are candidates for the development of new serum diagnostics and immunotherapies. Many of these markers show more robust tumor-specific gene expression profiles than those currently used at the bedside. We present a paradigm for assigning priority to gene expression data for human malignancies and selecting candidates for downstream analysis that have the highest likelihood of achieving clinical success.

1. Diehn M. *et al. Nature Genet.* **25**, 58–62 (2000).

Divine, Kevin K. [48]

Inactivation of p16 by methylation in human lung adenocarcinoma

Kevin K. Divine[1], Frank D. Gilland[2], Christine A. Stidley[3], Therese J. Bocklage[3], Ann Schwartz[4], Dennis L. Cook[1] & Steven A. Belinsky[1]

[1]*Lovelace Respiratory Research Institute, Albuquerque, New Mexico, USA*
[2]*University of Southern California, USA*
[3]*University of New Mexico, USA*
[4]*Karmanos Cancer Institute, USA*

Adenocarcinoma (AdC) of the central and peripheral lung may arise through distinct molecular mechanisms resulting from differences in exposure of these lung compartments to carcinogen. Inactivation of the p16 gene by aberrant promoter hypermethylation, an epigenetic mechanism, has been implicated in AdC development associated with tobacco. We investigated the frequency of p16 methylation in AdC from three different populations: smokers, former uranium miners who

smoked, and those who had never smoked (NS). We also examined the effect of central versus peripheral tumor location on gene inactivation. We found that p16 was methylated in 53% (36/68), 59% (16/27), and 49% (27/55) of peripheral tumors from smokers, former uranium miners, and those in the NS group. Central AdCs seem to have a higher frequency of p16 methylation (21/30; 70%) than peripheral tumors. This higher frequency of p16 methylation in central AdCs parallels our findings for squamous carcinoma, a predominantly central lesion. These studies also indicate that radon exposure does not interact synergistically with tobacco to target this gene for inactivation, but they do implicate p16 as a major pathway for AdC development in the NS group. The interaction of p19 with p16 methylation will also be discussed.

Dougherty, Edward [49]

Clustering algorithms: can anything be concluded?

Edward Dougherty[1], Junior Barrera[2], Marcel Brun[2], Seungchan Kim[1], Roberto Marcondes[2], Yidong Chen[3], Michael Bittner[3] & Jeffrey Trent[3]

[1]*Department of Electrical Engineering, Texas A&M University, College Station, Texas, USA*
[2]*Center for Bioinformatics, University of São Paulo, São Paulo, Brazil*
[3]*Cancer Genetics Laboratory, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland, USA*

Clustering algorithms are used in various applications, including the clustering of genes by expression profiles. Clustering algorithms find clusters, and these are often visually satisfying. However, the worth of a clustering algorithm must be judged by the degree to which the clusters it produces agree with the true classes. For instance, when clustering discrete time series from expression data, genes whose expression profiles show similar behavior are clustered, indicating possible co-regulation. The ability of the clustering algorithms to identify truly co-regulated genes is limited by both the uniqueness of the transcription pattern and the extent of variance of the biological processes. Independently regulated processes that produce similar transcription profiles can be mistakenly clustered as a single class if the variance in the observed patterns is of the same magnitude as the difference in the patterns. We consider ways to test the precision of clustering for given sets of transcription patterns and given levels of pattern variance, analyzing this issue for various clustering algorithms, including K-means, fuzzy C-means, hierarchical correlation-based clustering and Euclidean-distance clustering. The analysis applies to clustering in general; however, we focus on expression-based time series. We postulate a model having gene classes whose expression profiles exhibit congruency. A profile is modeled as a time-expression template defining the congruency class, plus noise. Clusters are estimates of the congruency classes. We measure clustering precision by the expected number of misclassifications. Precision depends on several factors: class separation, experimental variability and the number of sample replications. Very poor results are obtained for a single replicate, but there can be vast improvement with as few as three replicates. Moreover, performance varies substantially among clustering algorithms. Application of the inference analysis to real data requires estimation of model parameters from the data. Once these parameters have been estimated, the inference analysis can be applied to predict the expected numbers of errors on the basis of various algorithms and the number of replications. Intuitively the analysis gives the number of errors one would expect given the data. Since the analysis requires means and variances for the congruency classes, we apply a clustering algorithm to the raw data to form congruency classes with which to seed the algorithm. Clustering algorithms are applied to the seed classes.

Duffy, Carol [50]

Oligonucleotide array analysis of telomerase activation by the human papillomavirus type 16 E6 protein

Carol Duffy & Aloysius Klingelhutz

Department of Microbiology, University of Iowa, College of Medicine, Iowa City, Iowa 52242, USA

Human papillomavirus type 16 (HPV-16) is strongly associated with the development of cervical cancer. The E6 and E7 proteins from HPV-16 possess several functions that contribute to oncogenesis. HPV-16 E7 is known to bind and inactivate the retinoblastoma tumor suppressor protein. Expression of E7 alone extends the life span of newborn foreskin epithelial cells about twofold, whereas expression of E6 and E7 together can result in cellular immortalization without a significant crisis. HPV-16 E6 functions in the activation of telomerase and the targeting of p53 for ubiquitin-mediated degradation. E6 is known to alter the expression of several cellular genes by both p53-dependent and -independent mechanisms. Several cellular transcription factors (such as c-Myc, Sp1 and WT1) have been reported to be involved in telomerase regulation in either a cell-specific or -nonspecific manner. We propose that HPV-16 E6 activates telomerase through the upregulation of a cellular telomerase activator or the downregulation of a cellular telomerase repressor. To identify cellular genes involved in the activation of telomerase by E6 we have performed an oligonucleotide array analysis of messenger RNA from cells expressing E7 and either wild-type or mutant E6. The E6 mutants we are studying differentially activate telomerse and target p53. Array analysis of these mutants will help us identify changes in gene expression specifically associated with telomerase activation.

Duggan, David J. [51]

Molecular profiling of prostate cancer progression

David J. Duggan[1], Jun Luo[2], Jurga Sauvageot[2], Michael Bittner[1], Yidong Chen[1], Jeff Trent[1] & William B. Isaacs[2]

[1]*Cancer Genetics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland, USA*
[2]*Department of Urology, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA*

Prostate cancer is the most commonly diagnosed malignancy and the second most common cause of cancer death in North American males. It is largely unknown why some tumors progress to aggressive, potentially life-threatening disease whereas others remain latent for decades. Although cancer classification based on gene expression has been successful for several types of malignancies, expression analysis in prostate cancer is complicated by the heterogeneity of the tumor (even within a single specimen) and by contaminating normal cells. We have developed an approach that combines trimming and sectioning of frozen samples with associated histopathological review to isolate and verify percent tumor and grade within a given specimen. Messenger RNA from more than 40 normal, benign prostatic hyperplasia, and low- and high-grade prostate tumor samples was isolated and compared with a single reference to generate over 300,000 data points. Gene expression profiles of all prostate samples were correlated with Gleason score, clinical stage and other phenotypic and pathologic descriptors. Using a series of statistical tools, including multidimensional scaling, we have recognized unique clusters of samples that seem to correlate with tumor aggressiveness. Statistical analy-