

# Genomic Approach to identify Gene Signature for Outcome Prediction of Lung Cancer Patients

Pan-Chyr Yang, M.D., Ph.D.

Department of Internal Medicine, National Taiwan University Hospital

DNA microarray analysis is a powerful tool that can high throughput probe the expression of thousands of genes simultaneously. This technique has been widely adopted by the research community to study a variety of biologic processes and disease mechanisms. Microarray is very useful to profile the differential gene expression and identify genes with potential relevance to a wide range of biologic processes, such as the progression of cancer, the causes of specific diseases such as bronchial asthma. Recently, the DNA microarrays are extensively applied to subclassify diseases and identify specific biomarkers for early detection and prediction of clinical outcomes. To facilitate the microarray data analysis, we recently constructed a comprehensive web server, the composite regulatory signature database (CRSD), that integrates six well-known and large-scale databases, including the human UniGene, mature microRNAs, putative promoter, TRANSFAC, pathway, and Gene Ontology databases, and can be applied in investigating microarray gene expression data, complex regulatory behaviors involving the gene expression, microRNA regulatory and TF regulatory signatures. The CRSD and the online tutorial are publicly available at <http://biochip.nchu.edu.tw/crsd1/>. Here we demonstrate several examples of the applications of DNA microarray and pathway analysis to study the pathogenic mechanisms in respiratory diseases.

To identify gene signature that can predict survival and metastasis in lung cancer patients, we examined the gene expressions in surgical specimens of NSCLC patients using microarray and real-time RT-PCR and correlated the results with survival. We used the risk score and decision-tree methods to develop a gene-expression predictive model to predict the clinical outcome of NSCLC. Sixteen genes that correlated with patient survival were identified using microarray and risk score analysis. We selected 5-gene and developed a risk predictive model based on RT-PCR and decision-tree approach. The 5-gene signature is an independent predictor of cancer recurrence and overall survival of NSCLC patients. We validated the model in an independent cohort of NSCLC patients and also in an independent set of published microarray data.