S1-3

Transcriptional Profiling in Neuronal Differentiation of Stem Cells

LEE Yong Sung
Department of Biochemistry, Hanyang University College of Medicine, Seoul, Korea

Neurogenesis is one of the most complex events in embryonic development. However, little information is available regarding the molecular events that occur during neurogenesis. To identify regulatory genes and underlying mechanisms involved in the differentiation of embryonic stem (ES) cells to neurons, gene expression profiles of each developmental stage were analyzed using cDNA microarrays. Out of 10,368 genes studied, 1,633 (16%) known genes were differentially expressed at one or more stage s. At stage 3, during which ES cells differentiate into neural stem cells, modulation of nearly 1,000 genes was obser ved. Most transcription factors (*Otx2, Ptx3, Sox4, 13, 18, Nkx2.2, engrailed, Irx2, BF-2* and *Hes6*), signaling molecules (*Wnt, TGF* and *Shh* family members) and extracellular matrix / adhesion molecules (*collagens, MAPs* and *NCAM*) were up-regulated. However, some genes that may play important roles in maintaining the pluripotency of ES cells (*Kruppel-like factor 2, 4, 5, 9, myeloblast oncogene like2, ZFP 57* and *Esg-1*), were down-regulated.

S1-4

Discovering Significant and Interpretable Patterns from Multifactorial DNA Microarray Experiments Using a Genetic Algorithm and Permutation Test

KIM Ju Han
Seoul National University Medical School, Seoul, Korea

Replication is a cornerstone of scientific research. However, we oftenhave to deal with experimental data, which cannot be replicated easily because of limitations in resource, methodology or knowledge. Factorial designs are advantageous in making it possible to test the separate and combined effects of several variables in a study and to determine whether the independent variables interact significantly. In factorial designs with many factors and/or levels, however, replication is often prohibitively costly. Moreover, biological interpretation of significant higher-order interactions determined by standard statistical methods typically requires a complicated statement. Because we are mainly interested in finding factor-specific effects and their interactions, it is presumed that each observed gene-expression profile is created by a certain combination of underlying factor-specific generative functions with additive noise. A genetic algorithm was created to find, for each gene-expression profile, the nearest factor-specific generative pattern, which is defined as a combination of the underlying factor-specific generative functions. We measured the pattern distance of each gene-expression profile with respect to the nearest factor-specific generative pattern. Permutation testing for the pattern-distance measures successfully determined statistically significant multifactorial expression profiles, yielding straightforward biological interpretations. Genes showing statistically significant drug and/or cancer specific effects were successfully identified in a microarray experiment of gastric cancer cell lines with multifactorial design without replicates. The proposed method may benefit the combined analysis of heterogeneous expression data from the growing public repositories.