【S-1】

Toxicogenomics -A phenotype-independent approach-

Jun Kanno, M.D., Ph.D.
NIHS, Japan

The whole genome sequences, mapped for humans and rodents, and technical capability of monitoring whole genome expression in a high throughput fashion enable us to perform the "whole genome profiling". The major characteristics of this profiling from the toxicological point of view are that the overt phenotypes are not the essential factors for the construction of toxicity database/informatics. In other words, the amount of currently available phenotypic endpoints can be much smaller than what the "whole genome profiling" might generate. This "gene first, phenotype second" approach named as "Phenotype-Independent Toxicology" in combination with the currently available toxicology based on "phenotype first, gene second" strategy will compile the future predictive toxicology.

To build up such large-scale database, multiple experiments must be conducted and data collected over a certain period of time. Commonly, DNA microarray data are normalized primarily against relative mRNA concentration of certain housekeeping gene(s) or total mRNA quantity, or any other empirical baselines that are considered to be universal among samples. Although such relative methods are satisfactory for detection of genes that are changing drastically among test groups, small changes of many genes are not well analyzed. It is also noted that, the version-ups and/or changes in the make of microarrays, which may take place in a timeframe of database construction, cannot be effectively handled by such normalization methods. To overcome these problems, we have developed a method that will absolutize the gene expression values in a "per one cell" basis. Once absolutized, data from different microarrays and studies can be compared without further normalization. This method is made primarily for Affymetrix GeneChips, but can be expanded to other platforms and RT-PCR, as long as they fulfill the demands of this system.

This "per one cell" absolutization will enable us to express data in a linear scale from zero. This method will facilitate the analysis of low expression genes including those of knocked out genes in gene knock out mice. Unsupervised clustering is much easier for

this type of data than for ratio data to the concurrent control. This method will utilize almost all gene data that a microarray generates, and will contribute to predictive toxicology by the development of a large-scale database/ınformatics of higher precision.