Analysis of Gene Expression and Drug Activity Data by Knowledge-based Association Mining

D. Berrar[1], W. Dubitzky[1], M. Granzow[2], R. Eils[1]

[1]Intelligent Bioinformatics Systems, German Cancer Research Center; [2]phase-it, Intelligent Solutions AG

It is a major goal of the research field of pharmacogenomics to identify genes that play a crucial role in the effect pathway of a certain drug as they might be a proper target for an advanced therapeutic strategy. To obtain data that can be screened for reaching this goal it is necessary to (1) retrieve gene expression activity data and (2) to gain drug activity measurements from e.g. certain tumor tissues. To gather this information in a controlled manner, the in vitro analysis of tumor cell can be an example. The data sets examined in this study is described by Scherf et al. [1] and comprise data from 1,376 gene expression profiles of 60 cell lines (NCI60) and the growth inhibition power of 1,400 chemical compounds on the same cell lines. Growth inhibition is measured with a sulphorhodamine B assay that, briefly, measures the concentration necessary to inhibit growth by 50% in comparison to untreated controls (GI_{50}). The cell lines includes 9 categories, i.e. cell cultures from colorectal, renal, ovarian, breast, prostate, lung, central nervous system, leukaemia and melanoma cancers. The identification of gene/drug relationships relevant for each of the cell line categories might highlight mechanisms of gene-drug interaction that provide (1) information about new targets for new drugs in order to improve therapy and (2) distinguish specific cancer types from each other.