Meta-analysis of Japanese Toxicogenomics data: differences between in vivo and in vitro models

Carlos A. O. de Biagi Júnior[1, 2*], Richa Batra[5, 6], Jan Baumbach[3, 4], José L. Rybarczyk Filho[1, 2]

[1]Institute of Biosciences of Botucatu, Univ.Estadual Paulista, [2]Institute of Biotechnology, Institute of Biosciences of Botucatu, Univ.Estadual Paulista, [3]Department. of Mathematics and Computer Science, University of Southern Denmark, Odense, Denmark, [4]Computational Systems Biology group, Max Planck Institute for Informatics, Saarbrücken, Germany, [5]Department of Dermatology and Allergy, Technical University of Munich, Munich, Germany, [6]Institute of Computational Biology, Helmholtz Zentrum Munich, Munich, Germany

**cbiagijr@ibb.unesp.br*

Toxicogenomics is an emerging field to decipher effects of a drug at the molecular level in model systems. One of the main questions is if we can replace the in vivo study by in vitro study. To answer this question we used the data generated by the Japanese Toxicogenomics Project for the Rattus norvegicus liver, with in vivo and in vitro experiments, and Homo sapiens, with only in vitro experiments, treated with 131 drugs (approved by FDA) in different dosages and treatment durations, recorded in a total of 20000 microarray chips. We perform a comparative analysis of the in vivo and in vitro models at modular level using modular map [SEGAL, Eran et al. A module map showing conditional activity of expression modules in cancer. Nature genetics, v. 36, n. 10, p. 1090-1098, 2004.], through a package in R we develop for this methodology, such that each module is a cluster of genes with common gene signature. The analysis made with the comparison between the in vitro Homo sapiens microarray data and REACTOME gene sets, whose data were normalized with the "MAS5" method, yielded a total of 315 clusters. When multiple clusters have similar signatures, we can extract a module from these clusters. This module reflects more clearly the genes that participate in a specific biological process, since it consists of genes whose expression matches the signature of the cluster. Various processes and functions were identified among the clusters obtained previously, including: cell cycle, signal transduction, immune system, metabolism, protein metabolism, gene expression, transport, etc. As a result we obtained a global map showing modules that are induced or repressed in different conditions. The next step of our analysis is to link the modules with clinical conditions.