

ArrayBank™, a community microarray database and knowledgebase with integrated analysis tools.

John Kokinis*†, David Jones†, Gary Lindstrom*, Ron Lundstrom†

The Huntsman Cancer Institute†, University of Utah, Salt Lake City, UT 84112
The Department of Computer Science*, University of Utah, Salt Lake City, UT 84112
john.kokinis@hci.utah.edu

ABSTRACT

Currently there exist three notable deficiencies in the area of microarray data analysis: First, there are no widely used and publicly available repositories of expression data. Second, the few existing tools used to analyze these data are generally limited to clustering genes of like expression and do not correlate gene expression with known or putative biologic knowledge. Third, there are few methods for effectively sharing, and/or reusing, data and/or putative conclusions associated with those data. In August of 1999 we embarked upon a research effort aimed at addressing these issues. Our approach to this large task is to combine a relational database of raw microarray data augmented with descriptive data, a knowledgebase of gene expression profiles annotated with putative functional information, and an open-ended set of distributed and easy-to-use analysis tools that make maximum use of the relational database model. Particular attention is focused on providing the ability to effectively discover and compare patterns of gene expression between a priori unrelated data. This poster describes our progress to date: the implementation of a community relational database, initial analysis tools that utilize the relational data model, and the storage of annotated gene expression motifs. We also mention our approach for correlating a priori unrelated expression data by building a knowledgebase of annotated expression profile motifs and providing a tool to search the database for "similar" patterns of expression.