

## **Multivariate High Dimensional Visualization and Analysis of Microarray Data Incorporating Simultaneous Spatial and Temporal Components**

### **Introduction**

There are currently several scientific visualization tools available for scientists that allow the researcher to model, simulate, and visualize complex biological/biomedical systems data. The functionality and features of these software tools vary depending on the hierarchy level (cellular, molecular, etc.) of the biological/biomedical system to be explored as well as the type of data being collected. This research effort aims to develop a tool that will augment the current toolset for the visualization and analysis of complex microarray datasets that have multiple variables of interest. *Trypanosoma cruzi* (*T. cruzi*) parasite microarray data will be used as the prototype development data; however, the resulting tool could be used to visualize datasets that contain more than 5-dimensions or variables of interest and may include time as well.

In an effort to make a new and novel contribution to the field of microarray visualization and analysis a comprehensive review of past and present methods for visualizing such data will be initiated as follows.

### **Summer Research Goals**

The goals for this summer are: (1) to complete a comprehensive literature review of current software tools used in the analysis of high dimensional data. The focus of this research effort will be on software used in visualizing microarray data, which is inherently high-dimensional. It is the aim of this project to create a generalized tool that can be used and applied to any multivariate high-dimensional dataset that includes, but is not limited to microarray data, (2) to develop a needs assessment survey to be used in determining which currently available visualization tools are being used by Life and Biomedical/Health Scientists in the analysis of microarray data, and most importantly ascertain an understanding of what, if any, unmet visualization needs exist among those scientists in order to develop tools that could potentially aid in the analysis of their microarray data; and (3) to submit a request for approval to execute the survey to Virginia Commonwealth University (VCU) IRB board.

### **Survey Design**

The survey will be composed of questions designed around extracting the knowledge of the target audience's use and familiarity with existing microarray analysis software in current use. The desired outcome of the survey is to obtain information that can be used to determine what visualization needs are not being met by current software and to subsequently address those needs in the development of the aforementioned visualization tool.

The target survey audience will include life scientists on the campuses of Virginia Commonwealth University and The University of Alabama at Birmingham (UAB).

Specifically, the survey will be distributed to VCU's Departments of Biology and Microbiology as well as the Massey Cancer Center and distributed to similar departments within the UAB Community, including the UAB Comprehensive Cancer Center. This is a use assessment survey so no personal data will be collected.

### **Results**

The response and feedback from the survey will help to determine what new visualization features and functions are desired and should be implemented in new software tools in order to meet the current and future needs of life and biomedical/health researchers.