

ANALYSIS OF [PSI+] PRION AFFECTED YEAST MICROARRAY DATA USING VARIOUS MICROARRAY SPOT FINDING AND STATISTICAL ANALYSIS PROGRAMS

*John Brothers II*¹, *Amanda Watkins*², *Yolande Tra*³, and *Irene Evans**⁴

¹Bioinformatics, Department of Biological Sciences, Rochester Institute of Technology, Rochester, NY 14623, jfb4497@rit.edu

²Biotechnology, Department of Biological Sciences, Rochester Institute of Technology, Rochester, NY 14623, fizzywater313@hotmail.com

³Department of Mathematics and Statistics, Rochester Institute of Technology, Rochester, NY 14623, yvtsma@rit.edu

⁴Department of Biological Sciences, Rochester Institute of Technology, Rochester, NY 14623, imesbi@rit.edu

Microarray technology has reshaped biological research by giving researchers the ability to analyze a vast amount of data containing the measurements of the transcription output of every single gene in a species' genome. The focus of our laboratory research is to study the effects of the [PSI+] prion on the gene expression of *Saccharomyces cerevisiae* (Baker's yeast) by comparing the [PSI+] prion affected yeast with a [psi-] control strain of the yeast. In order to successfully solve the question of which genes are affected by the [PSI+] yeast prion, quality control laboratory techniques were used in order to produce replication between various microarray grids. The laboratory process of preparing microarrays can be unwieldy and a great deal of non-biologically influenced variability can influence the data output of the microarray. Over 6400 gene spots on the arrays must be gridded and analyzed, and various microarray programs can be used such as MagicTool, Scanalyze, Cluster, R, and GeneSpring. Each of these tools has their own plethora of benefits and limitations such as ease of use and accuracy of data mining. Originally, MagicTool was the program of choice for both mining the data from the microarrays, and for subsequent statistical analysis of those microarrays. However, Scanalyze was found to contain a more accurate method of spot finding and data mining. The programming environment R and the programs Cluster and GeneSpring were found to contain superior statistical packages and functions than the ones found in MagicTool. As we have acquired more and more replicates from the laboratory, using these more efficient and more statistically sound programs will guarantee the accuracy of our final differentially expressed gene list. Confirmation of these differentially expressed genes will be done using such techniques as Real Time RT-PCR. This will facilitate discovering what effects the [PSI+] prion is having on *S. cerevisiae*.