Analysis of Microarray Time Course Data

Tanya Logvinenko[1], **David Schoenfeld**[2],**Douglas Hayden**[3]

Keywords: Microarray Data, Time Course Data, Analysis of Variance, Two-Sample Comparison, Cluster Analysis, Fold Change.

1 Introduction

With the development of microarray technology, it became possible to study biological processes over time. For example, of interest can be changes in gene expression profile over time. In case of a disease, discovering which genes are induced or repressed at different stages can help in the understanding of the disease process, as well as in the development of the treatment procedures. We will address a problem of analyzing time-course microarray data. The data that will be used is part of the Inflammation and the Host Response to Trauma collaborative project (`www.gleugrant.org`). It was obtained from eight healthy human volunteers who were randomized into two groups of four individuals each. One of the groups was exposed to *in-vivo* LPS (lipo-polysacharide stimulation), whereas the other group was administered a placebo treatment. Six microarrays per individual were obtained over the course of one day. We will study the efficiency of different methods of time-course microarray data analysis applied to this data set.

2 Methods

A variety of methods for analysing time-course microarray data were developed recently. To detect differentially expressed over time genes, one can apply

[1]Division of Biostatistics, Massachusetts General Hospital, 50 Staniford St, Suite 560, Boston, MA 02114. E-mail: `tlogvinenko@partners.org`

[2]Division of Biostatistics, Massachusetts General Hospital, 50 Staniford St, Suite 560, Boston, MA 02114. E-mail: `dschoenfeld@partners.org`

[3]Division of Biostatistics, Massachusetts General Hospital, 50 Staniford St, Suite 560, Boston, MA 02114. E-mail: `doug@gcrc.mgh.harvard.edu`

standard statistical models like analysis of variance [3] to expression profiles of each genes, or examine fold changes of gene expressions over time. A statistical technique, *SAM*, for discovering genes differentially expressed between different samples was developed [5]. Modified to account for the higher degree of similarity between observations obtained from the same subject, rather than from different subjects at the same time point, it can be applied to time-course microarray data. A whole range of clustering procedures can be used for discovering genes with the same expression profiles over time [1, 2, 4]. In cases when differential expression of genes is caused by a known factor (such as a disease or a trauma), a control sample can be examined. In such situations, we propose directly comparing distributions of the gene expressions from the two groups. We will examine a few techniques and models applied to time-course microarray data, and present their advantages and disadvantages.

References

[1] Bar-Joseph, Z., Gerber, G., Gifford, D., Jaakkola, T., and Simon, I. (2003). A new approach to analyzing gene expression time series data. *Proc. Natl. Acad. Sci. USA*, 100:10146–10151.

[2] Eisen, M., Spellman, P., Brown, P., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA*, 95:14863–14868.

[3] Kerr, M., Martin, M., and Churchill, G. (2000). Analysis of variance for gene expression microarray data. *J. Comput. Biol.*, 7:819–837.

[4] Tseng, G. and Wong, W. (2003). Tight clustering: A resampling-based approach for identifying stable and tight patterns in data.

[5] Tusher, V., Tibshirani, R., and Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. USA*, 98:5116–5121.