Statistical Analysis Using DNA Microarray in the Expression Level of Genes in Prostata Cancer Patients

H. M. Ruivo[1], F. M. Ramos[1] , E. M. R. Reis[2]
[1]Laboratory for Computing and Applied Mathematics - LAC
Brazilian National Institute for Space Research - INPE
C. Postal 515 – 12245-970 – São José dos Campos - SP
[2]Biochemistry Departament - Chemistry Institute - IQ
University of São Paulo – USP
05508-900 – São Paulo - SP
BRAZIL

E-mail: [1]{heloisa.ruivo, fernando}@lac.inpe.br; [2] emreis@iq.usp.br

Keywords: microarray, statistical analysis, cluster

Nowadays, data analysis has a prominent role in the life and natural sciences (Amaratunga & Cabrera, 2004). In molecular biology, for instance, Microarrays (MA) have revolutionized research and generated considerable excitement. MA are experiments that simultaneously provide information about expression levels of thousands of genes, and are, consequently, finding wide use in biomedical research (Simon *et al*, 2002). MA contain several hundred thousand data points arranged in the form of a table, with rows corresponding to genes and columns to samples (Domany, 2003).

Through the analysis (Simon *et al*, 2003) of these data it is possible to group genes according to the existence of similarities among their expression profiles in the several analyzed conditions. This approach has been used to diagnose and to classify tumors in relevant subsets, as well as to identify molecular markers that allow to predict the clinical evolution of a cancer or to suggest a chemotherapy strategy for new patients diagnosed with cancer.

The objective of this work is to apply an integrated free software package for DNA microarray data, called BRB ArrayTools (Simon & Lam, 2006), to the analysis of prostate tumor samples. The analysis results are compared with those published in (Reis *et al*, 2004). Our main finding is a robust set of 52 genes that may be used to classify prostate tumors in terms of their clinical evolution.

REFERENCES

Reis, E. M. R. et al. (2004), *Antisense intronic non-coding RNA levels correlate to the degree of tumor differentiation in prostate cancer*, Oncogene, 23, 6684-6692.

Domany, E. (2003*), Cluster Analysis of Gene Expression Data*, Journal of Statistical Physics, 110, 3–6.

Simon, R. and Radmacher, M. D. and Dobbin, K. (2002), *Design of Studies Using DNA Microarrays*, Genetic Epidemiology, 23, 21-36.

Simon, R. and Lam, A. P. (2006), *Biometric Research Branch (BRB) Version 3.4 – User's Manual* - National Institute Cancer.

Amaratunga, D. and Cabrera, J. (2004), *Exploration and Analysis of DNA Microarray and Protein Array Data*, Wiley-Interscience, New Jersey.

Simon, R. M. and Kom, E. L. and McShane, L. M. and Radmacher, M. D. and Wright, G. W. and Zhao, Y. (2003), *Statistics for Biology and Healyh*, Springer-Verlag, New York.