

Poster presentation

Open Access

Robust detection method for differential expression studies

Shesh N Rai*^{1,2} and Stanley Pounds³

Address: ¹Biostatistics Shared Facility, James Graham Brown Cancer Center, Louisville, KY 40292, USA, ²Department of Bioinformatics and Biostatistics, University of Louisville, Louisville, KY 40292, USA and ³St. Jude Children's Research Hospital, Memphis, TN 38105, USA

Email: Shesh N Rai* - shesh.raai@louisville.edu

* Corresponding author

from UT-ORNL-KBRIN Bioinformatics Summit 2008
Cadiz, KY, USA. 28–30 March 2008

Published: 8 July 2008

BMC Bioinformatics 2008, **9**(Suppl 7):P19 doi:10.1186/1471-2105-9-S7-P19

This abstract is available from: <http://www.biomedcentral.com/1471-2105/9/S7/P19>

© 2008 Rai and Pounds; licensee BioMed Central Ltd.

Background

Differential gene expression analysis typically applies a statistical hypothesis testing procedure to the (log-transformed) data for each gene and then adjusts for multiple testing. However, it is not always clear whether a parametric or non-parametric test of hypothesis is preferred for a specific application. In a less complex application, a statistical test or graphical assessment of the fit of the assumed parametric model (e.g. the normal distribution) could be used to make this determination. However, in the analysis of microarray data it is unclear how to assess assumption adequacy for each gene and use it to guide subsequent analysis. Therefore, we introduce the principle of empirical Bayes assumption adequacy averaging and use it to develop a robust method to perform differential expression analysis. The proposed method combines results from the t-test and nonparametric rank-sum test according to the results of assessing the assumption of normality using the Shapiro-Wilk procedure. Through this averaging process, the proposed method is able to rely more heavily on the statistical test that the data suggests is superior for each individual gene. Subsequently, we observed that the proposed method has desirable power and robustness properties in a series of traditional and bootstrap-based simulation studies. Additionally, the proposed method showed greater concordance in gene selection across two studies of gene expression in acute myeloid leukemia than did the t-test or rank-sum test. Finally, we note that the principle of empirical Bayes assumption adequacy averaging can be used to develop robust procedures for a wide variety of experiments.