A KNOWLEDGE-BASED EXTENSION OF SAM TO IDENTIFY BIOLOGICAL PATHWAYS ASSOCIATED WITH A PHENOTYPE

I.A. Dinu, T. Mueller, P.F. Halloran, Y. Yasui[†]

University of Alberta, Edmonton, Canada

[†] E-mail: *yutaka.yasui@ualberta.ca*

Significance Analysis of Microarray (SAM) is a useful method designed to find single genes associated with a phenotype; for example, genes that are differentially expressed between cases of a disease and controls. The real biological question often rests on the assessment of biological pathways, or gene sets, associated with a phenotype, however. Single-gene analysis by SAM may no longer be satisfactory for this goal. We propose here a knowledge-based extension of SAM to assess each gene-sets differential expression. SAM ranks the genes according to their t-like statistics. Our method employs a ROC or partial-ROC analysis to quantify the association between gene ranks from SAM and the membership of the genes to the gene set of interest. To measure the significance of this association, we use a permutation-based method, accounting for correlations in the gene expressions. We compare the performance of the proposed method with that of Gene Set Enrichment Analysis of Subramanian et al. (PNAS 2005), by simulation. The method is illustrated on a microarray experiment to identify biological pathways associated with the clinical outcomes of kidney transplantation.