

Rigorous assessment of gene set enrichment tests

Haroon Naeem, Ralf Zimmer, Robert Küffner

*haroon.naeem@campus.lmu.de, ralf.zimmer@bio.ifi.lmu.de, robert.kueffner@bio.ifi.lmu.de
Department of Informatics, Ludwig-Maximilians Universität, Amalienstr. 17, 80333 München, Germany*

Abstract

Background: Several enrichment tests are available to determine the differential expression of gene sets. Such tests were originally proposed for analyzing gene sets associated with biological processes. The objective evaluation of tests on real measurements has not been possible as it is difficult to decide *a priori*, which processes will be affected in given experiments.

Results: We propose a rigorous assessment of enrichment tests based on regulators such as transcription factors (TFs) and microRNAs (miRNAs). Gene sets are defined based on the known targets of given regulators. In contrast to processes, TFs and miRNAs are amenable to direct manipulations, e.g. regulator over-expression or deletion experiments. We assessed the ability of 12 different tests to predict the manipulations from expression measurements in *E. coli*, *S. cerevisiae* and human. We also analyzed how performance depends on the quality and comprehensiveness of the known regulator targets via an additional permutation approach. Overall, we find that ANOVA and Wilcoxon's test consistently performed better than for instance Kolmogorov-Smirnov and hypergeometric tests. Combining all evaluated tests into an unweighted consensus is a particularly robust choice for the analysis of novel scenarios.

Conclusion: We present a first large study to rigorously assess and compare the performance of gene set enrichment tests. Our results provide a guide for the selection of existing tests as well as a basis for the development and assessment of novel tests.