BioNet: an R-Package for the functional analysis of biological networks

Daniela Beisser1, Gunnar W. Klau2, Thomas Dandekar1, Tobias Müller1,* and Marcus T. Dittrich1,*

1Department of Bioinformatics, Biocenter, University of Würzburg, Am Hubland, 97074 Würzburg, Germany and 2Life Sciences group, CWI, Science Park 123, 1098 XG Amsterdam, The Netherlands

ABSTRACT

Motivation: Increasing quantity and quality of data in transcriptomics and interactomics create the need for integrative approaches to network analysis. Here, we present a comprehensive R-package for the functional analysis of biological networks including an exact and a heuristic approach to identify functional modules.

Results: The BioNet package provides an extensive framework for integrated network analysis in R. This includes the statistics for the integration of transcriptomic and functional data with biological networks, the scoring of nodes as well as methods for network search and visualization.

Availability: The BioNet package and a tutorial are available from http://bio.net

Contact: marcus.dittrich@biozentrum.uni-wuerzburg.de; tobias.mueller@biozentrum.uni-wuerzburg.de

Received on December 14, 2009; revised on February 5, 2010; accepted on February 22, 2010

1 INTRODUCTION

Integrated analysis of microarray data in the context of biological networks such as protein–protein interaction (PPI) networks has become a major technique in systems biology. The primary objective is the identification of functional modules (significantly differentially expressed subnetworks) within large networks. This can be achieved by computing a score for each node in the network reflecting its functional relevance. Subsequently, a network search algorithm is required to find the highest scoring subgraph. In fact, this problem has been proven to be NP-hard (Ideker et al., 2002). Various heuristic approaches have been proposed, most of them inspired by the seminal work from Ideker et al. (2002) that used a simulated annealing heuristic to identify high-scoring subgraphs in integrated networks. Recently, we have devised an algorithm (heinz, heaviest induced subgraph) that computes provably optimal and suboptimal solutions to the maximal-scoring subgraph (MSS) problem in reasonable running time using integer linear programming (ILP) (Dittrich et al., 2008). In extension to this, we present an R package for (i) integrating multiple P-values obtained from different experiments, (ii) scoring the nodes of the network by a modular scoring function, (iii) calculating P-values for differential expression and risk association,

© The Author 2010. Published by Oxford University Press. All rights reserved. For Permissions, please email: journals.permissions@oxfordjournals.org
respective. Then we aggregate both $P$-values by the second-order statistics using BioNet.

\begin{verbatim}
> data(dataLym)
> pvals <- cbind(t=dataLym$t.pval, s=dataLym$s.pval)
> pval <- aggrPvals(pvals, order=2, plot=FALSE)

We now fit a BUM model to the distribution of aggregated $P$-values and score the nodes using an FDR threshold of 0.001.

\begin{verbatim}
> fb <- fitBumModel(pval, plot=FALSE)
> scores <- scoreNodes(network, fb=fb, fdr=0.001)
> writeHeinz(network, file="lym_001
+ node.scores=scores)
\end{verbatim}

The exact search algorithm can be started from R by runHeinz if the CPLEX library is installed (Dittrich et al., 2008). Alternatively, the fast heuristic search algorithm (runFastHeinz) often delivers a close approximation. Finally, the resulting modules can be visualized in 2D or 3D.

\begin{verbatim}
> module <- readHeinzGraph(node.file=
+ "lym_001_n.txt.0.hnz", network)
> plot3DModule(module)
\end{verbatim}

BioNet captures an interaction module that has been described to play major biological roles in the GCB and ABC DLBCL subtypes (Fig. 1C). The combination of biological and clinical data with PPI networks generates a meaningful biological context in terms of functional association for differentially expressed, survival-relevant genes.

\section*{Conflict of Interest:} none declared.

\section*{REFERENCES}


