

POSTER PRESENTATION

Open Access

WebGestalt2: an updated and expanded version of the Web-based Gene Set Analysis Toolkit

Dexter Duncan, Naresh Prodduturi, Bing Zhang*

From UT-ORNL-KBRIN Bioinformatics Summit 2010
Cadiz, KY, USA. 19-21 March 2010

Background

Functional enrichment analysis has become a popular approach for the interpretation of gene lists derived from large scale genomic, transcriptomic, and proteomic studies. Many software tools have been developed for this analysis, including our previously published Web-based Gene Set Enrichment Analysis Toolkit (WebGestalt) [1]. Major differences among existing tools include: 1) the statistical test used for the enrichment analysis; 2) supported organisms; 3) supported input ID types; 4) coverage of functional categories; and 5) presentation of the output results. We have made improvements in the above areas in the updated and expanded version, WebGestalt2.

Materials and methods

First, we have implemented methods for the multiple-test adjustment. Because we are testing many functional categories at the same time, multiple-test adjustment methods have been implemented to correct p-values generated by the hypergeometric test. Second, although the original version only supports human and mouse, the new version has been expanded to cover other organisms including rat, worm, fly, yeast, dog, and zebrafish. Third, WebGestalt2 supports more input ID types from various databases and different technology platforms. For example, 32 ID types are supported for human and 29 ID types are supported for mouse. The addition of protein ID types such as the Ensembl peptide ID, the IPI ID, the refseq_peptide ID, and the UniProt ID makes this tool directly available to the fast growing proteomics community. One of the most important improvements in WebGestalt2 is the increased coverage of functional categories in various biological contexts, including chromosomal location,

Gene Ontology (GO), KEGG pathways, Pathway Commons, Wikipathways, Transcription factor targets, microRNA targets, and functional modules in protein interaction networks. WebGestalt2 retains the enriched directed acyclic graph (DAG) representation for the GO enrichment analysis results and uses tables for results from other analyses. WebGestalt2 can also highlight genes in KEGG pathway maps and Wikipathway maps.

Results and conclusion

WebGestalt2 does not require registration and does not require a login to upload gene lists. As a result, analysis results will not be stored on the server. Instead, users are given options to save their results locally. The code has been optimized so that the new version is much faster than the old one and can scale to large input lists with thousands of genes.

In conclusion, significant improvements have been made in WebGestalt2 compared to the original version. The new version provides a simple but powerful platform for the efficient interpretation of gene lists.

Acknowledgments

This work was supported by the National Institutes of Health (NIH)/ National Institute on Alcohol Abuse and Alcoholism (NIAAA) through grant U01AA016662, the NIH/National Institute of General Medical Sciences (NIGMS) through grant R01GM088822, and the NIH/ National Institute of Mental Health (NIMH) through grant P50MH078028.

Published: 23 July 2010

Reference

1. Zhang B, Kirov S, Snoddy J: WebGestalt: an integrated system for exploring gene sets in various biological contexts. *Nucleic Acids Res* 2005, **33**(Web Server issue):W741-8.

doi:10.1186/1471-2105-11-S4-P10

Cite this article as: Duncan et al.: WebGestalt2: an updated and expanded version of the Web-based Gene Set Analysis Toolkit. *BMC Bioinformatics* 2010 **11**(Suppl 4):P10.

* Correspondence: bing.zhang@vanderbilt.edu
Department of Biomedical Informatics, Vanderbilt University School of Medicine, Nashville, TN 37232, USA