

NetGestalt: integrating multidimensional omics data over biological networks

Zhiao Shi, Jing Wang, Bing Zhang

ZS: Advanced Computing Center for Research and Education; Department of Electrical Engineering and Computer Science, Vanderbilt University; Nashville, TN 37235 USA.

zhiao.shi@accre.vanderbilt.edu.

JW: Department of Biomedical Informatics, Vanderbilt University School of Medicine; Nashville, TN 37232 USA, jing.wang.2@vanderbilt.edu.

BZ: Department of Biomedical Informatics; Department of Cancer Biology, Vanderbilt University School of Medicine, Nashville, TN 37232 USA, bing.zhang@vanderbilt.edu.

Biological networks provide excellent functional contexts for exploring and interpreting omics data; however, node-link diagrams quickly become inadequate as network size and data complexity increase. In this paper¹, we introduced NetGestalt (<http://www.netgestalt.org>), a web application that exploits the inherent hierarchical modular architecture of biological networks to achieve high scalability. NetGestalt orders the nodes of a network along the horizontal dimension of a webpage based on the underlying hierarchical organization of the network. Visualization in the horizontal dimension conveys the functional relationship between different nodes (*i.e.*, genes) as encoded in the network. The linear layout makes it easy to scale up to thousands of genes. Moreover, node-related information from different data sources and existing knowledge can be rendered as “tracks” along the vertical dimension of the webpage for visual comparison and integration. Since the publication of the paper, we have implemented many new features in NetGestalt and have redesigned the user interface to allow more intuitive access to different features. The new version is scheduled for release in June 2014. Recent completion of the global proteomic characterization of the TCGA colorectal cancer (CRC) cohort by the Clinical Proteomic Tumor Analysis Consortium (CPTAC) provides the first tumor data set with complete molecular measurements at DNA, RNA and protein levels². Using CRC as a paradigm, we demonstrate that NetGestalt not only provides easy data query and visualization but also enables efficient data integration within a single omics data type, across multiple omics data types, and over biological networks.

References:

1. Shi, Z., Wang, J. & Zhang, B. NetGestalt: integrating multidimensional omics data over biological networks. *Nat Methods* **10**, 597-598 (2013).
2. Zhang, B. et al. Proteogenomic characterization of human colon and rectal cancer. *Nature* (Provisionally accepted).