Bioinformatics Framework for the Analysis and Interpretation of Metabolomic Data

Alla Karnovsky1, Jing Gao2, Glenn Tarcea3, Christopher Beecher3, Charles Burant4, Barbara Mirel5, H.V. Jagadish2, Gilbert S. Omen4,1,2

1Center for Computational Medicine and Biology, 2National Center for Integrative Biomedical Informatics, 3Michigan Center for Translational Pathology, 4Departments of Internal Medicine and Human Genetics, University of Michigan, Ann Arbor, MI 48109-2218

BACKGROUND AND MOTIVATION
Metabolomics is a rapidly emerging field that is joining other high-throughput "omics", such as proteomics and transcriptional profiling. It promises to be a powerful systems approach for studying metabolic profiles pertinent to a variety of normal and disease states. Transcriptional profiling and proteomics have established data analysis tools; a metabolomics analytical toolkit is yet to be developed. We are creating a set of tools that will allow the user to examine experimental metabolomic data in the context of human metabolic networks and to combine them with other high throughput data.

A number of public sources contain information about human metabolic networks consisting of compounds, chemical reactions, pathways, enzymes and genes: KEGG (Kanehisa et al., 2008), BIGG (Duarte et al., 2007), INtRet (Mia et al., 2007). These networks are indispensable for the interpretation of experimental metabolomic data (Beecher, 2003).

In this project, we used KEGG and EHMN data to trace the connections between metabolites and genes. Compounds, reactions, enzymes, genes and the relationships between them provide an initial framework for the analysis of metabolomic data. The Michigan Molecular Interactions database [MMI (plug-in)] developed by the National Center for Integrative Biomedical Informatics (NCIBI) integrates protein interactions data from a number of public sources and thus can supply broader context for the analysis of the experimental data (Tarcea et al., 2009).

The data are stored locally in a Microsoft SQL Server database. They can be accessed either via a web-based query interface, or via Metscape, our new plug-in for Cytoscape (http://www.cytoscape.org/). Users can upload a list of metabolites or genes to the web interface, identify reactions, genes and pathways that are associated with these, and explore their relationships. In a parallel workflow users can import normalized experimental metabolite data directly into Metscape and display them within the network of metabolic reactions (Fig. 1).

METHODS
We developed a set of tools that has core functionality for visualizing experimental metabolomic data in context of human metabolism. Since metabolites are linked to genes, we can take advantage of multiple data sources (e.g. GO, enzyme database) to enhance our current toolset.

- We plan to fully implement the interconnectivity between MMI Metabolome web interface and Metscape.
- We are currently working on enhancing the existing pathway layouts.
- We plan to use dynamic expression plug-in for Cytoscape (dynaxmipxr - http://www.integratedgenomics.com/egvisplugins/metscape) to visualize normalized expression profiling data side by side with metabolomic data.
- We are working on enhancing the compound search.

ACKNOWLEDGEMENTS
This work is supported by NIH grant 1R15 DK54222-01 for the National Center for Integrative Biomedical Informatics.

REFERENCES