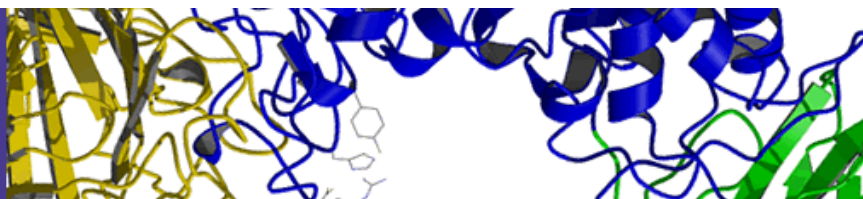




Human Proteome Organisation

HUPO 5th Annual World Congress, Long Beach 2006
October 28th to November 1st, 2006



From disparate multi-disciplinary data to coherent knowledge: Informatics for systems biology-based disease profiling

Erich Gombocz, Robert Stanley, Michael Travers
IO Informatics, Inc., Emeryville, CA 94608, U.S.A.

Contact:

Dr. Erich A. Gombocz
IO Informatics, Inc.
2000 Powell Street, Suite 520
Emeryville, CA 94608, U.S.A.
(+1) 510-420-8400
egombocz@io-informatics.com

Abstract

To undertake meaningful disease profiling, translational proteomics needs to deal with large sets of data that cross disciplinary and organizational boundaries to capture, interpret and use functional relationships on a systems biology level for modeling, screening and analysis. The objectives of such efforts are to gain the ability, once coherence is established, to define profiles based on common but mostly unidentified characteristics and their relationships for a defined disease from a controlled case study. Such profiles then can be applied as models to larger unclassified datasets for diagnostic or therapeutic decision support.

This talk will review examples of such approaches used in a corporate partnership under a NIST advanced technology project applying integration informatics software to capture, define and model gene function and metabolism biomarkers within a systems environment for development of diagnostic methods. It also will examine a use case for combining public data sets from different sources and complimentary data types into a meaningful knowledgebase for pre-symptomatic disease detection in neurodegenerative conditions such as Alzheimer's. The presented informatics method and software provides a user-centric approach to define, curate and analyze complex, multi-dimensional relationships, allowing differentiation of subsets within large datasets based on function.

We show that a single system can be used in a systems biology scenario for queries across normalized data subsets from MRI images, quantitative tissue analysis, gel electrophoretic and gene array data as well as LC/MS identified metabolites. Examples of profiles as results from interactive, form-based and graphical queries in a semantic framework and their value in modeling complex diseases are discussed.

(257 words)