

MEETING ABSTRACT

Open Access

Mining the human proteome for conserved mechanisms

Stefan Naulaerts^{1,2*}, Pieter Meysman^{1,2}, Wim Vanden Berghe³, Kris Laukens^{1,2}

From Third International Society for Computational Biology (ISCB) European Student Council Symposium 2014

Strasbourg, France. 6 September 2014

Background

All cells are subject to ever-changing environments to which they have to adapt, using their sensory system to provide input for the regulatory systems that integrate the information and trigger the eventual effectors. These cascades constitute a very complex cellular wiring that is highly relevant due to its medical importance. The omni-present application of high-throughput analysis techniques has resulted in an unprecedented level of available detail about gene expression and various aspects of cellular proteins, such as abundance, function and localization, often captured in well-curated compendia that are publicly available.

Although these information-rich inventories exist, the adaptive nature of protein complexes and signalling cascades remain poorly understood, as the current predominant approaches are not always suited to describe the associations between proteins. For example, binary protein interactions do not necessarily occur *in vivo* as the proteins could be expressed in different compartments of the cell or at different time points. This severely complicates the analysis of any protein interaction data. It thus remains a challenge to find out how biological entities cooperate to regulate cellular response to stimuli.

Methods

We used an integrative method, reliant on advanced pattern mining approaches to gain a deeper understanding of protein network dynamics. To this end, we created a compendium consisting of a large amount of proteomics papers for *Homo sapiens* that report differentially expressed proteins in cell lines. Next, we analysed this collection with frequent itemset mining to

identify proteins that are often co-occurring in publications and used these patterns as the backbone structure of our further analysis. These patterns of co-occurring proteins were enriched with additional attributes, such as gene expression correlation, protein localization and functional coherence metrics derived from the Gene Ontology tree [1] and used as a filter on top of an integrated binary protein interaction network, obtained by fusing several of the most popular resources.

Results

We found that several proteins and GO-functions, such as transcriptional regulation, are consistently reported and deemed significant regardless of the research topic. Furthermore, we were able to find associations across the various “omics” levels that are conserved in a wide range of human cancers and managed to identify lists of frequently occurring patterns that can be used to classify between pre- and post-metastatic tumour development.

Conclusions

Pattern-based analysis on multiple “omics” levels can be used to identify the cellular logic circuits and holds many promising applications in the biotechnological and biomedical areas.

Authors' details

¹ADReM research group, Department of Mathematics and Computer Science, University of Antwerp, Belgium. ²Biomedical Informatics Research Center Antwerp (biomina), University of Antwerp/Antwerp University Hospital, Belgium. ³Laboratory of Protein Science, Proteomics and Epigenetic Signaling (PPES), Department of Biomedical Sciences, University of Antwerp, Belgium.

Published: 13 February 2015

¹ADReM research group, Department of Mathematics and Computer Science, University of Antwerp, Belgium

Full list of author information is available at the end of the article

Reference

1. Ashburner M, Ball C, Blake J, Botstein D, Butler H, Cherry J, Davis A, Dolinski K, Dwight S, Eppig J, *et al.*: **Gene ontology: tool for the unification of biology.** *Nat Genet* 2000, **25**(1):25-29.

doi:10.1186/1471-2105-16-S3-A6

Cite this article as: Naulaerts *et al.*: Mining the human proteome for conserved mechanisms. *BMC Bioinformatics* 2015 **16**(Suppl 3):A6.

**Submit your next manuscript to BioMed Central
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

