

Towards system level modeling of functional modules and context-specific pathways using genome-scale data

Tom Michoel, Anagha Joshi, Eric Bonnet, Vanessa Vermeirssen, Yves Van de Peer

Division of Bioinformatics and Systems Biology, Department of Plant Systems Biology, VIB, Department of Plant Biotechnology and Genetics, Ghent University, Technologiepark 927, B-9052 Gent, Belgium

Introduction

Understanding complex biological processes such as development and pathology of multicellular organisms at a system level requires the study of dynamic networks of interacting DNA, RNAs, proteins and metabolites. High-throughput methods have generated large-scale, static networks of physical interactions which may occur or not depending on spatial, temporal or context-specific variation. On the other hand, genome-wide microarray datasets measure expression levels precisely under such varying conditions. In this talk I will present our recent efforts towards integrative, system level modeling of functional modules and context-specific pathways using diverse genome-scale data.

Results

We have developed an algorithm, called LeMoNe, to infer regulatory modules and their condition-specific regulators from gene expression data. The algorithm computes a centroid-like solution extracted from an ensemble of possible statistical models to explain the data and automatically selects a subset of most informative genes and builds a quantitatively better model for them [1]. We used LeMoNe to generate hypotheses for differential gene expression in *C. elegans* [2] and to predict experimentally confirmed microRNA regulatory modules in human cancer cells [3]. A detailed analysis of these modules shows that their predicted regulators often act indirectly through context-specific pathways of protein-DNA and protein-protein interactions that cannot be inferred from expression data alone. We use topological motifs to represent functional relationships within and between regulatory and protein-protein interaction networks and introduced 'regulatory path motifs', short paths in integrated physical networks which occur significantly more often than expected by chance between regulators and their targets in perturbational expression data [4]. Furthermore, we developed a novel computational method for the automated identification of clusters of topological motifs in integrated networks which provides additional insight into the organization of functional modules and overlapping pathways [5].

Conclusions

The availability of ever increasing amounts of cellular interaction data and context-specific expression data enables the study at a system level of dynamic regulatory networks. We have developed several computational methods to address the challenging problem of how complex processes like development and disease are regulated at a system level.

References

- [1] A. Joshi, R. De Smet, K. Marchal, Y. Van de Peer, T. Michoel. Module networks revisited: computational assessment and prioritization of model predictions, *Bioinformatics* 25, 490 - 496 (2009).
- [2] V. Vermeirssen, A. Joshi, T. Michoel, E. Bonnet, T. Casneuf, Y. Van de Peer. Transcription regulatory networks in *Caenorhabditis elegans* inferred through reverse-engineering of gene expression profiles constitute biological hypotheses for metazoan development, *Molecular BioSystems* 5, 1817 - 1830 (2009).
- [3] E. Bonnet, M. Tatari, A. Joshi, T. Michoel, K. Marchal, G. Berx, Y. Van de Peer. Network inference from a cancer gene expression data set identifies microRNA regulatory modules, *PLoS One* (under review) (2010).
- [4] A. Joshi, T. Van Parys, Y. Van de Peer, T. Michoel. Characterizing regulatory path motifs in integrated networks using perturbational data, *Genome Biology* (under review) (2010).
- [5] T. Michoel, B. Nachtergaele, Y. Van de Peer. Enrichment and aggregation of topological motifs are independent organizational principles of integrated interaction networks, *Genome Biology* (under review) (2010).