

Modeling gene expression in biological networks

Leo Lahti¹, Juha Knuuttila², and Samuel Kaski¹

1. Department of Information and Computer Science, TKK Helsinki University of Technology, Finland;

2. Neuroscience Center, University of Helsinki, Finland

Contact: leo.lahti@tkk.fi



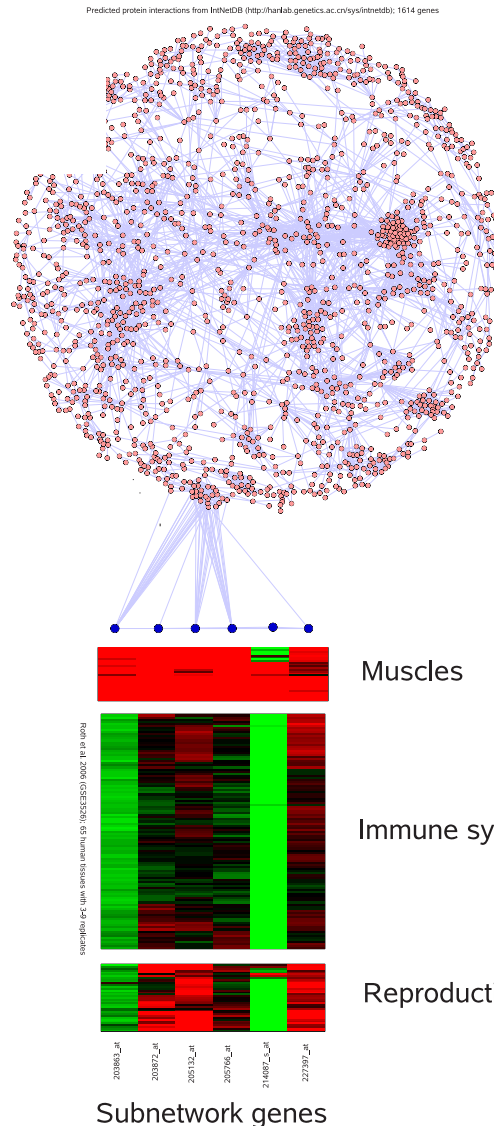
Motivation

Biological processes are regulated by coordinated activation and inactivation of genes in a context-specific manner. Different contexts can induce different co-activation patterns for the same genes if the involved genes participate in multiple biological processes. Identification of gene groups with alternative, context-sensitive co-activation patterns can lead to novel functional hypotheses of gene function in previously unexplored contexts.

We have developed a novel computational approach, NEMO, for modelling co-expression in interaction networks. NEMO detects functionally coordinated subnetworks that have context-sensitive co-expression patterns.

Computational approach

NEMO performs agglomerative subspace clustering on gene expression data, which is constrained by an interaction network between the genes. If neighboring genes have coordinated changes in expression, they are merged into larger subnetworks. Gene expression in each subnetwork is modeled with a joint Gaussian mixture model. Each mixture component characterizes co-expression of the subnetwork genes in a group of physiological conditions.



Agglomerative subspace clustering proceeds along the network and **detects connected regions** in the network **that have context-sensitive co-activation patterns**.

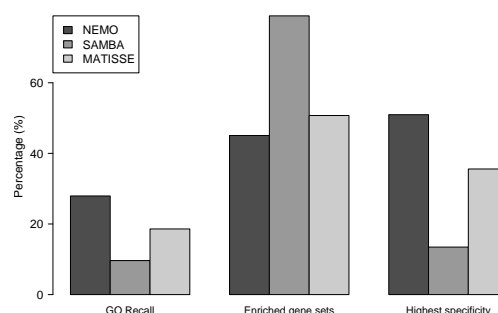
Performance

Gene expression was studied in a network of 1614 genes (IntNetDB) across 65 distinct physiological conditions (Roth et al., 2006).

NEMO could retrieve the largest number of known biological processes in Gene Ontology. The findings were also more specific than in the alternative approaches, while smaller proportion of NEMO subnetworks corresponds to known processes.

Many of the distinct co-expression patterns were associated with coherent physiological contexts such as muscles, immune system, or the brain.

Gaussian mixture model **reveals associations between co-expression and the activating contexts** in each subnetwork. Although tissue labels were not used in modeling, NEMO could detect physiologically coherent groups of biological conditions with distinct co-expression patterns.



Acknowledgements

This poster at ECCB'08 has been supported by The International Society for Computational Biology (ISCB), and Graduate School in Computational Biology, Bioinformatics, and Biometry (ComBi).