

# Inferring Functional Relationships Using Genome-wide Dependency Networks

Johan Rung  
EMBL - European  
Bioinformatics Institute  
Wellcome Trust Genome  
Campus, Hinxton  
Cambridge, CB10 1SD  
United Kingdom  
johan@ebi.ac.uk

Patrick Kemmeren  
EMBL - European  
Bioinformatics Institute  
Wellcome Trust Genome  
Campus, Hinxton  
Cambridge, CB10 1SD  
United Kingdom  
patrick@ebi.ac.uk

Thomas Schlitt  
EMBL - European  
Bioinformatics Institute  
Wellcome Trust Genome  
Campus, Hinxton  
Cambridge, CB10 1SD  
United Kingdom  
schlitt@ebi.ac.uk

## ABSTRACT

We have developed a method to build gene dependency networks from mRNA expression data, incorporating information about effects from gene disruptions. Each gene is represented as a node in the network, and each dependency found as a weighted directed edge between two genes. Rather than using machine learning techniques, we have chosen an approach where we simply gather the different types of dependencies and present them in a network format without any structure restrictions. In particular, this allows for cyclic structures within the networks. This means that the networks we build have a different meaning than “traditional” networks, and should not be interpreted as classical gene regulatory pathways. Instead, we exploit the information in our networks for other tasks. For instance, we can analyse domains of the network for presence of biological signals, such as regulatory patterns in promotor regions or common functional properties. We can use this to generate hypotheses that can be tested in a wet lab, for example about functions of unknown genes or putative transcription factor binding sites.

Using previously published data [1], we have built such networks for *Saccharomyces cerevisiae* on a genome-wide scale. The networks are complex and contain many cyclic structures. Most genes have only a few incoming or outgoing edges, but some genes are highly connected either with many outgoing edges or many incoming edges. We show that the genes with high numbers of outgoing edges often have regulatory functions, whereas the genes with high numbers of incoming edges often have functions subject to complex regulation, for example in metabolism. Functionally related genes tend to show up close to each other in the networks.

We present the method used for building gene dependency networks, along with a genome-wide dependency network for *S. cerevisiae*. The structure properties of the network are described, along with a functional analysis of network domains and a validation using literature data. Domains of the network surrounding transcription factors are analysed for common DNA patterns in upstream promotor regions. We show that we indeed find such patterns, and that we can validate already known transcription factor binding sites this way.

## ADDITIONAL AUTHORS

Alvis Brazma (EMBL - EBI, email address: [brazma@ebi.ac.uk](mailto:brazma@ebi.ac.uk)) and Jaak Vilo (EMBL - EBI, email: [vilo@ebi.ac.uk](mailto:vilo@ebi.ac.uk)).

## REFERENCES

- [1] A. R. Hughes *et al.* Functional discovery via a compendium of expression profiles. *Cell*, 102:109–126, 2000.