

Poster J-35

FunCoup: data integration and the network of functional coupling in Eukaryotes



Authors:

Andrey V Alexeyenko (*Stockholm Bioinformatics Center, Stockholm University*)

Erik LL Sonnhammer (*Stockholm Bioinformatics Center, Stockholm University*)

Short Abstract: Functional coupling of genes is predicted from high-throughput data of multiple sources. Optimized pattern recognition techniques employ information on orthologous genes of model organisms to improve accuracy and sensitivity of the method. FunCoup is a flexible tool transferring information between genomes and automatically optimizing for different data sources.

Long Abstract:

FunCoup is a method aimed to discover novel functional links between proteins - which, in turn, can be used for focussing work on individual proteins of interest or creating sophisticated gene networks. The method uses Bayesian and multivariate techniques to integrate multiple genomics and proteomics features of various types (from relatively loose associations like coexpression to physical interactions between proteins) and from different sources. The crucial novelty is involving data on orthologous genes in well-studied model organisms. Orthologs are provided by InParanoid, which has the advantage of including inparalogs. FunCoup, as InParanoid, is mainly focused on eukaryotic genomes.

A key feature of FunCoup is that each data source from each organism is weighted by its reliability and relevance. Earlier functional coupling predictors often used naive Bayesian networks that simply summed up individual features' likelihood scores. Methods have been devised for removing redundancy between dependent datasets, but few attempts have been made to exploit interactions between features. This can be important as the effect of the interaction between two features can be very non-additive. For example, if two genes are co-expressed but not co-localized, then the likelihood of functional coupling should be less than the sum of the individual likelihoods. The latest version of FunCoup implements a framework for feature interactions.

The output of FunCoup is the likelihood that two proteins are functionally coupled compared to an expected background probability. The framework efficiently uses available data and, together with rich information sources, has a prediction potential for different classes of functional coupling.

On average, FunCoup discovers up to 10-20 % of existing links that are e.g. protein-protein interactions or metabolic relations, while 50 % of the predicted functional links are thought true. This figures depend on the quality and amount of available (mainly high-throughput) data and would only improve in future.

Querying of particular proteins for functional coupling is available at <http://www.sbc.su.se/~andale/>