

Poster presentation

Open Access

Gene regulatory network of human adipocyte differentiation

Judit Kumuthini^{*1}, Conrad Bessant¹, Paul Wilson² and Daniel Crowther²

Address: ¹Cranfield University, Barton Road, Silsoe, Bedfordshire, MK454DT, UK and ²GSK, Bioinformatics Research and Development, Stevenage, Hertfordshire, SG1 2NY, UK

Email: Judit Kumuthini^{*} - s.j.kumuthini.s02@cranfield.ac.uk

^{*} Corresponding author

from BioSysBio 2007: Systems Biology, Bioinformatics and Synthetic Biology
Manchester, UK. 11–13 January 2007

Published: 8 May 2007

BMC Systems Biology 2007, 1(Suppl 1):P57 doi:10.1186/1752-0509-1-S1-P57

This abstract is available from: <http://www.biomedcentral.com/1752-0509/1?issue=S1>

© 2007 Kumuthini et al; licensee BioMed Central Ltd.

Background

In this article we demonstrate novel pre-processing methods to reduce data dimensionality of human adipocyte differentiation microarray data. Genetic networks of the insulin receptor family, ppar family, fox family, cebp family, mef2, fabp, add1 and klf, and probes with highly significant change in gene expression level were learned separately using a Bayesian framework. The extracted networks were validated against many publicly available and as well as in-house interaction and literature databases available at GSK.

Methods

Multidimensional, hMAD microarray data provided by GSK was used to generate additional artificial experiments using a novel technique and the differentially expressed probes were filtered. Through Gaussian clustering 45 clusters plus the outliers extracted were used to learn the genetic network using a tabu search algorithm using BayesiaLab[®].

Results

See Figure 1 and Figure 2.

Conclusion

The choices of pre-processing methods and dimensionality reduction techniques applied in this work have a major impact on Bayesian network extraction. The Bayesian networks extracted were validated against a proprietary Network warehouse database at GSK. Many novel genetic

interactions were identified. We suggest that these pre-processing methods can be widely used for genetic network extraction, even in the absence of many experiments but replicates. This improves the prediction of significant changes in gene expression for microarray experiments and reduces both false positives and negatives.

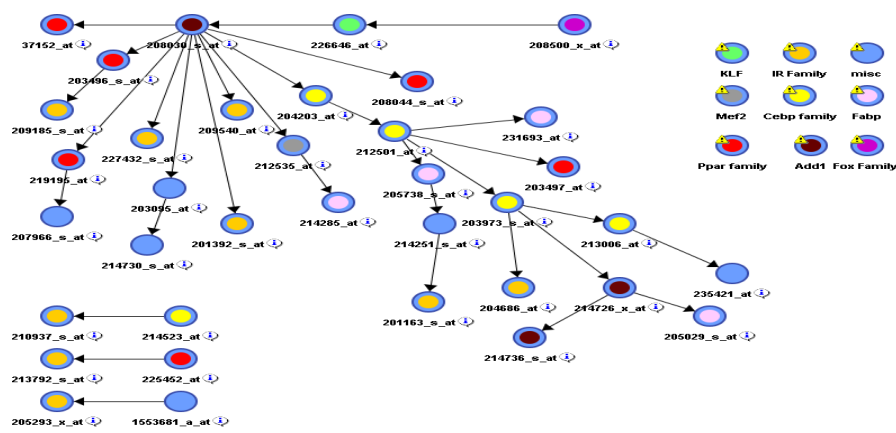


Figure 1
GRN of genes of interest, colour coded according to the biological family they belong to. The respective Affymetrix probe ID annotation from Tiger database is associated with each node as indicated by the sign.

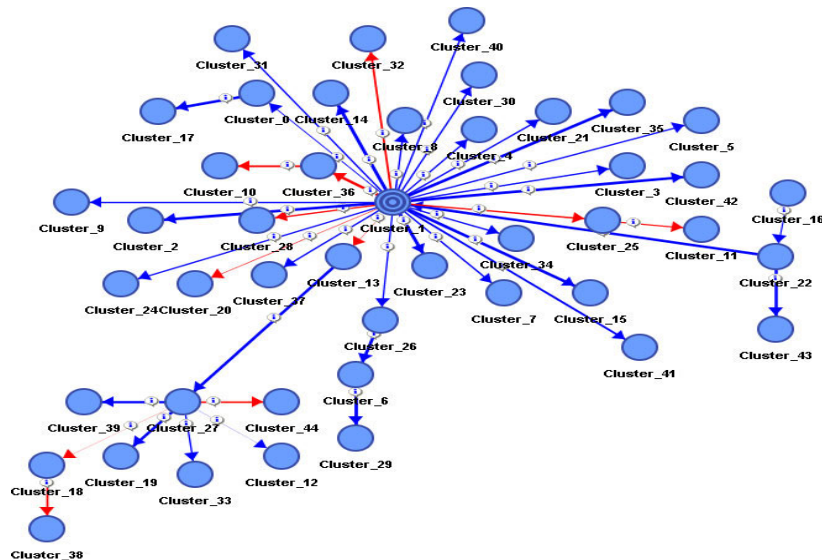


Figure 2
Pearson correlation coefficient of optimised 45 Clusters TN represented by the wavy circular node. The thickness of the arc represents the information contributed to the global structure. The strength of the arc is proportional to the strength of the probabilistic relation. The association of marginal probability to the modalities (outcome) of nodes is computed R, the Pearson's linear correlation coefficient between two nodes linked by an arc. The thickness of an arc is directly proportional to the absolute value of R. The blue indicated positive correlation and red a negative correlation. The exact value of the correlation for each arc is temporary displayed in the comment of the arc as indicated by the sign. The association of values to nodes' modalities and R, the Pearson's linear correlation coefficient between two nodes linked by an arc. If the modalities don't have associated values, default values are defined in order to compute R (from 0 to n-1 for a node with n modalities). The thickness of an arc is directly proportional to the absolute value of R, its colour represents the sign of R (blue if positive and red if not). The exact value of the correlation for each arc is temporary displayed in the comment of the arc.