ScholarWorks@UMass Amherst

OPEN ACCESS DISSERTATIONS

Title

Discovery of Complex Regulatory Modules from Expression Genetics Data

Author

Manjunatha Jagalur, *University of Massachusetts Amherst*

Date of Award

5-2010

Document Type

Open Access Dissertation

Degree Name

Doctor of Philosophy (PhD)

Degree Program

Computer Science

First Advisor

David C. Kulp

Second Advisor

Gary A. Churchill

Third Advisor

Erik G. Learned-Miller

Subject Categories

Computer Sciences

Abstract

Mapping of strongly inherited classical traits have been immensely helpful in understanding many important traits including diseases, yield and immunity. But some of these traits are too complex and are difficult to map. Taking into consideration gene expression, which mediates the genetic effects, can be helpful in understanding such traits. Together with genetic variation data such data-set is collectively known as expression genetics data. Presence of discrete and continuous variables, observed and latent variables, availability of partial causal information, and under-specfied nature of the data make expression genetics data computationally challenging, but potentially of great biological importance. In this dissertation the underlying regulatory processes are modeled as Bayesian networks consisting of gene expression and genetic variation nodes. Due to the underspecified nature of the data, inferring the complete regulatory network is impractical. Instead, the following techniques are proposed to extract interesting subnetworks with high confidence. The network motif searching technique is used to recover instances of a known regulatory mechanism. The local network inference technique is used to identify immediate neighbors of a given transcript. Application of these two techniques often results in identification of hundreds of individual networks. The network aggregation technique extracts the most common subnetwork from those networks, and identifies its immediate neighbors by collapsing them into a common network. In all the above tasks, simulation studies were carried out to estimate the robustness of the proposed methods and the results suggest that these techniques are capable of recovering the correct substructure with high precision and moderate recall. Moreover, manual biological review shows that the recovered regulatory network substructures are typically biologically sensible.

Recommended Citation

Download

DOWNLOADS

Since September 23, 2010

Included in

Computer Sciences Commons

Share

COinS