Inference of a Gene Regulatory Network Using Steady-State Expression Data

Koji M. Kyoda, Mineo Morohashi, Shuichi Onami
Kitano Symbiotic Systems Project, ERATO, JST.
E-mail: {**kyoda,moro,sonami**}**@symbio.jst.go.jp**

Key Word: Inference method, Steady-state gene expression data, Gene regulatory network

Recent progress in the field of molecular biology enables us to obtain whole-genome expression data, facilitated by the development of DNA microarray. For identifying the biological functions of genes from those data, several systematic approaches have been proposed, such as gene clustering. Moreover, gene knockouting is currently performed in a large scale for several species, from which a huge number of gene expression data for many different genetic background will be available with a unified manner. Given those data, analytical methods are required to infer underlying regulatory relationships in the gene network.

Here, we present a method for inferring a gene network using steady-state gene expression data. Steady-state gene expression data are typical data of large scale gene expression analysis and each of those data represents a set of gene expression level in cells or animals in which specific gene activities are altered by e.g., gene knockouting or overexpressing. Our method determines a regulatory structure consistent with an observed set of steady-state expression data, generated from wild-type and single gene knockouting mutant of the target network. The advantage of our method is to be able to deal with the original continuous values of steady-state data, while most of the methods proposed in the past require conversion of the original data into binary data since they are based on the Boolean network model. The binary conversion results in negligible loss of the original information. Performance of our method is evaluated on simulated artificial networks with varying the size of networks, indegree of each gene, and the data characteristics (continuous/binary), indicating the superiority of using continuous values to binary values and the superiority of our method to already reported methods. We also report our attempt to apply our method

to real gene expression data -- data obtained by DNA microarray for a number of single gene knockout mutants --, and to analyze the applicability of our method.