

Chapter XXIX

Methods for Reverse Engineering of Gene Regulatory Networks

Hendrik Hache

Max Planck Institute for Molecular Genetics, Germany

ABSTRACT

In this chapter, different methods and applications for reverse engineering of gene regulatory networks that have been developed in recent years are discussed and compared. Inferring gene networks from different kinds of experimental data are a challenging task that emerged, especially with the development of high throughput technologies. Various computational methods based on diverse principles were introduced to identify new regulations among genes. Mathematical aspects of the models are highlighted, and applications for reverse engineering are mentioned.

INTRODUCTION

Deciphering the structure of gene regulatory networks by means of computational methods is a challenging task that emerged during last decades. Large scale experiments, not only gene expression measurements from microarrays but also promoter sequence search for transcription factor binding sites and investigations of protein-DNA interactions have spawned various computational algorithms to infer the structure of underlying gene regulatory networks. Identifying gene interactions yields to an understanding of the topology of gene regulatory networks and the functional role of any gene in a particular pathway. Of interest are genes with a certain impact on pathways. These can be therefore putative drug targets. Once a network is obtained, *in silico* experiments can be performed to test hypotheses and generate predictions on disease states or on the behavior of the system under different conditions (Wierling et al., 2007).

Quantitative gene expression measurements using microarrays were first performed by Schena et al. (1995) for 45 *Arabidopsis thaliana* genes and shortly after, for thousands of genes or even a whole

genome (DeRisi et al., 1996; DeRisi et al. 1997). Since that time various methods for the analysis of such large-scale data have been developed. First of all, clustering algorithms were used to partition genes into subsets of co-regulated genes regarding to their expression profiles (Eisen et al., 1998). It is found that genes belonging to the same cluster have similar biological functions. But this does not imply information about any directed regulatory interactions among these genes. For this purpose, more sophisticated methods were employed to reverse engineer gene networks and regulation causality from such data. Therefore, reverse engineering constitutes an intermediate step from correlative to causative data analysis.

Gardner and Faith (2005) classified reconstruction algorithms into two general strategies: “physical” approaches and “influence” approaches. Algorithms of the first group seek to identify interactions between transcription factors and DNA and reveal protein factors that physically control RNA synthesis. These methods, such as promoter binding analysis, as performed by Lee et al. (2002), uses the genomic sequence information directly. The second strategy, the “influence” approach, aims to identify causal relationships between RNA transcripts by the examination of expression profiles. The regulation of the transcription machinery can be effected on multiple levels. For instance, it can be regulated on the DNA, transcriptional, or translational level. The regulation on the DNA or transcriptional level is mainly due to binding of transcription factors on specific parts of DNA or by chemical or structural modifications of DNA. A regulation on the translational level might be due to microRNAs resulting in the decay of the respective mRNA target (Ruvkun, 2001). Since the quality of current measured protein concentration is not sufficiently high enough for reconstruction purposes, it is assumed in mathematical models that changes in expression as measured by mRNAs concentrations can explain changes of other gene transcripts. However, a recent study by Newman et al. (2006) showed that many changes in the measured protein levels at single-cell resolution are not observable by DNA microarray experiments. In computational models many regulation effects are neglected or included as hidden factors. In recent years, a combination of both approaches were employed by integrating multiple data sources for the construction of priors of networks or parameters (Imoto et al., 2003; Bernard and Hartemink, 2005; Werhli and Husmeier, 2007).

In this chapter different reverse engineering algorithms following an “influence approach” that have been proposed in the last decades are discussed. Moreover, relevant mathematical aspects are briefly described and applications to reveal gene regulatory networks or part of them are highlighted. Another crucial point that is discussed in this chapter is the validation of algorithms. The reverse engineering methods have to cope with noisy, high dimensional, and incomplete data. But the quality and amount of useful measurements for reverse engineering is increasing.

COMPUTATIONAL MODELS AND METHODS

The following models can be considered as graphical models that define a mathematical structure $\mathbf{M} = (G, \mathbf{F}, \mathbf{Q})$ with a graph G , a set of functions \mathbf{F} , and a set of parameters \mathbf{Q} of these functions. The graph $G = (\mathbf{X}, \mathbf{E})$ consists of nodes $\mathbf{X} = \{X_1, \dots, X_n\}$ with real values $\mathbf{x} = \{x_1, \dots, x_n\}$ which are interconnected via edges \mathbf{E} , see Figure 1. An edge can be directed or undirected. For a directed graph the k_i direct regulators of node X_i are represented by $\mathbf{X}_{pa[i]} = \{X_{pa[i]}^1, \dots, X_{pa[i]}^{k_i}\}$ otherwise it represents all connected nodes. The associated values are depicted by $\mathbf{x}_{pa[i]} = \{x_{pa[i]}^1, \dots, x_{pa[i]}^{k_i}\}$. A function $f_i \in \mathbf{F}$ with $f_i : \mathbf{x}_{pa[i]} \rightarrow x_i$ is

17 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/methods-reverse-engineering-gene-regulatory/21551

Related Content

Classification of Sleep Apnea Types Using Clustering with SVM Classifier

Faiza Charfiand Ali Kraiem (2012). *International Journal of Biomedical and Clinical Engineering* (pp. 39-48).

www.irma-international.org/article/classification-sleep-apnea-types-using/73692

The Core Governmental Perspectives of E-Health

Antti Syväjärvi (2009). *Medical Informatics: Concepts, Methodologies, Tools, and Applications* (pp. 153-162).

www.irma-international.org/chapter/core-governmental-perspectives-health/26213

ICT in Medical Education in Trinidad and Tobago

Marilyn Lewis (2009). *Medical Informatics: Concepts, Methodologies, Tools, and Applications* (pp. 1180-1185).

www.irma-international.org/chapter/ict-medical-education-trinidad-tobago/26289

Intelligent Stethoscope

B Buvaneshwari, NA Rohinee, Sahana Roopkumaran and Prabhu Ravikala Vittal (2014). *International Journal of Biomedical and Clinical Engineering* (pp. 73-80).

www.irma-international.org/article/intelligent-stethoscope/115887

Telehealth Organizational Implementation Guideline Issues: A Canadian Perspective

Maryann Yeo (2009). *Medical Informatics: Concepts, Methodologies, Tools, and Applications* (pp. 1186-1208).

www.irma-international.org/chapter/telehealth-organizational-implementation-guideline-issues/26290