Chapter 10
Inference of Gene Regulatory Networks by Topological Prior Information and Data Integration

David Correa Martins Jr.
Federal University of ABC (UFABC), Brazil

Fabricio Martins Lopes
Federal University of Technology – Paraná (UTFPR), Brazil

Shubhra Sankar Ray
Indian Statistical Institute, India

ABSTRACT

The inference of Gene Regulatory Networks (GRNs) is a very challenging problem which has attracted increasing attention since the development of high-throughput sequencing and gene expression measurement technologies. Many models and algorithms have been developed to identify GRNs using mainly gene expression profile as data source. As the gene expression data usually has limited number of samples and inherent noise, the integration of gene expression with several other sources of information can be vital for accurately inferring GRNs. For instance, some prior information about the overall topological structure of the GRN can guide inference techniques toward better results. In addition to gene expression data, recently biological information from heterogeneous data sources have been integrated by GRN inference methods as well. The objective of this chapter is to present an overview of GRN inference models and techniques with focus on incorporation of prior information such as, global and local topological features and integration of several heterogeneous data sources.

1. INTRODUCTION

Systems Biology is an interdisciplinary research field that aims at the study of complex interactions occurring in living organisms (Snoep & Westerhoff, 2005). Research in this field focuses on the study of biological processes such as cell cycles and the conditions for the origin of certain diseases. The ultimate goal of these studies is to help the development of new treatments and drugs against diseases, biofuel production techniques, among many other applications.

The genome of an organism has a central role in the control of cell processes such as cell response to environmental stimuli, cell differentiation in its respective functional groups, DNA replication for cell division, and many others. An organism can be seen as a network of molecules connected by biochemical reactions (Voet, Voet & Pratt, 2005). Proteins synthesized from genes may work as transcription factors which bind to regulatory sites of other genes, such as enzymes which catalyze metabolic reactions or components of signal transduction pathways. Such regulatory mechanism forms a complex system of sending and receiving signals (RNAs) which can be investigated to identify the control mechanisms of the cell and the relationships among various biological entities like genes, RNAs and proteins. However, there is still much to be discovered about the functional relationships of control mechanisms, e.g., transcription levels and proteins, in the regulatory system (Barabasi,2002; Fall, Marland, Wagner & Tyson, 2002; Shmulevich & Dougherty, 2007).

With few exceptions, all cells of an organism contain the same genetic material, although cells of different tissues are functionally different. The cell function is partially determined and controlled by gene expression profiles. With the aim of understanding how genes are involved in control of intra and inter cell processes, the scope of the molecular biology studies needs to be enlarged to include not only the discovery of nucleotide sequences that codes for proteins, but also the unraveling of the regulatory systems which determine what genes are expressed, when, where, and to how much extent (Snoep & Westerhoff, 2005). The explanation of these regulatory networks functioning, by means of sending and receiving signals, is currently one of the main objectives of the systems biology studies.

One of the most challenging research problems of Systems Biology is the inference (or reverse-engineering) of gene regulatory networks (GRNs) from expression profiles (Werhli, Grzegorczyk & Husmeier, 2006; Marbach et al, 2012). This research issue became important after the development of high-throughput technologies for extraction of gene expressions such as DNA microarrays (Schena, Shalon, Davis & Brown, 1995) or SAGE (Velculescu, Zhang, Vogelstein, & Kinzler, 1995), and more recently RNA-Seq (Wang, Gerstein, & Snyder, 2009). The importance of GRN reconstruction can be seen through initiatives taken for this purpose such as DREAM (Dialogue for Reverse Engineering Assessments and Methods) (Marbach et al, 2012). The inference problem involves discovery of complex regulatory relationships among biological molecules which can describe not only diverse biological functions, but also the dynamics of molecular activities. Once the network is recovered, intervention studies can be conducted to control the dynamics of the biological systems aiming to prevent or treat diseases (Shmulevich & Dougherty, 2007).

In general, it is not possible to recover GRNs very accurately based only on gene expression profiles for several reasons, including the presence of significant noise in the data, limited number of samples and large dimensionality. Also, GRN inference is considered an ill-posed problem, meaning that many networks may be able to explain the data in hand. Besides, the lack of information about the biological organism and the high complexity of the networks are additional challenges involved in GRN inference.

Related Content

Macroalgae-Based Bioethanol

Sakshi Kolpe, Shankar Mukundrao Khadeand Shivraj Nile (2023). *Biomass and Bioenergy Solutions for Climate Change Mitigation and Sustainability (pp. 206-216).*

www.irma-international.org/chapter/macroalgae-based-bioethanol/314365

Design of a Prosthetic Ankle Complex: A Study in Biomimetic System Design

Dheeman Bhuyanand Kaushik Kumar (2019). *Design, Development, and Optimization of Bio-Mechatronic Engineering Products (pp. 101-125).*

www.irma-international.org/chapter/design-of-a-prosthetic-ankle-complex/223410

Classification of Surface Constructs

 (2021). *Inspiration and Design for Bio-Inspired Surfaces in Tribology: Emerging Research and Opportunities (pp. 108-130).*

www.irma-international.org/chapter/classification-of-surface-constructs/257598

Biofuel Policies in India: An Assessment of Policy Barriers

Sunil Kumar Vermaand Prashant Kumar (2023). *Biomass and Bioenergy Solutions for Climate Change Mitigation and Sustainability (pp. 44-64).*

www.irma-international.org/chapter/biofuel-policies-in-india/314357

Transcriptomics to Metabolomics: A Network Perspective for Big Data

Ankush Bansaland Pulkit Anupam Srivastava (2019). *Biotechnology: Concepts, Methodologies, Tools, and Applications (pp. 361-379).*

www.irma-international.org/chapter/transcriptomics-to-metabolomics/228630