Poster J-41

Identification of transcription logic-gate dynamics in Escherichia coli

Authors:
Ilaria Mogno (*University of Rome La Sapienza*)
Lorenzo Farina (*University of Rome La Sapienza*)
Timothy Gardner (*Boston University*)

Short Abstract: We develop a method to characterize combinatorial regulation in bacteria. Time series data is collected using GFP as a reporter for gene expression. We use nonlinear regression and an iterative Kalman filter to identify the logic gate and dynamics of the regulatory function. Our method provides a simple way to characterize promoters in an effort to control the cellular behavior.

Long Abstract:

Bacterial gene expression is typically controlled by transcription factors that bind a gene's promoter. In a many cases a single gene is regulated by multiple transcription factors. In order to characterize the promoter and to develop methods for controlling the dynamic behavior of the cell, it becomes important to understand the regulatory logic gate of the transcription factors at the promoter. An AND gate occurs when the presence of all the transcription factors is necessary to regulate their target gene. An OR gate occurs when any of the transcription factors is sufficient for the regulation. More complicated behaviors are possible, for example a single transcription factor might be sufficient to induce its target, while the presence of another one increases the rate of the induction. Our aim in this work is to identify the gene regulatory function (the function that describes the relationships between transcription factors and their target genes) from time series gene expression data. The use of time series data is a natural choice for dynamic systems like gene regulation. Time series data also let us capture transient behaviors in gene expression that are lost when using steady state data only.

We generate time series data using green fluorescent protein (GFP) as a reporter of gene expression. We clone the promoter of interest into a low copy plasmid driving the gene coding for GFPmut2. The measured fluorescence indicates the amount of GFP protein in the cells at a certain time. By measuring optical density of the cells and fluorescence at the same time we determine the relative amount of GFP per cell. With a dynamical model, we estimate the GFP mRNA level from the amount of GFP protein in each cell. This mRNA level is proportional to the expression level of the gene whose promoter was cloned in front of the GFP transcription start site. We use this technique to obtain time series gene expression data for a gene and each of its transcription factors under different conditions known to perturb them.

We use a nonlinear dynamic model to describe how the expression level of the transcription factors influences the expression level of their target. We model the change of mRNA abundance as depending on its degradation and on its production. Production of mRNA is modeled with a Hill function that includes multiple transcription factors. The function

describes the occupancy of the transcription factors at the promoter.

The regulatory logic gate is identified with nonlinear regression techniques. With this knowledge of the regulatory gate, we learn the dynamics of the gene regulatory function using an iterative Extended Kalman filter approach to determine promoter activity and key parameters such as threshold of expression initiation and the degree of nonlinearity of the reaction. A Kalman filter is commonly used in system theory to identify hidden states in a stochastic dynamical system. In our model, the states represent quantities such as protein concentration and mRNA expression level, which are inferred from our fluorescence data. We define as hidden states the parameters of the regulatory function. By setting them to invariant we can use the Extended Kalman filter to identify them.

Our method was successfully applied and validated on the lac and arabinose operons in Escherichia coli. The lac operon is regulated by transcription factors LacI and CRP; we find, as expected, that the regulation depends primarily on the level of LacI. while CRP increases the rate of transcription. The expression of the arabinose operon, on the other hand, results from a AND gate between the transcription factors AraC and CRP. Our method was also used to reconstruct the gene regulatory function for genes involved in iron uptake and genes recently discovered to be part of the class III flagella network.

Our method is a valuable tool for synthetic and systems biology, as it provides a simple way to characterize newly created promoters or newly discovered regulatory interactions allowing the potential for more informed control over cellular dynamics.