

Protein Pathways RIPS 2003 Project

Background

The transcriptional rates of genes are regulated by enhancer and repressor proteins. Regulated genes may in turn regulate the transcription of other genes. This situation leads to a complex network of transcriptional regulation that is present within a cell.

Current technology, in the form of expression microarrays, allows one to simultaneously measure the mRNA levels of all genes in a cell (see Schena 1995 and Eisen 1999). By perturbing a cell, and measuring the changes in mRNA levels, it is possible to study transcriptional networks.

However, despite the availability of expression microarrays, there are still no models that allow one to predict changes in gene expression that result from specific perturbations. The aim of this project is to construct such a model and develop an optimal algorithm to identify free energy minima of the model.

A Model of Transcriptional Regulation in a Cell

Under normal conditions a gene is transcribed at a steady rate. When the cell is perturbed by external stimuli, or by genetic modifications, the transcription rate of a gene may be altered. We can quantify the change in terms of a ratio

$R = \text{observed mRNA concentration} / \text{steady state mRNA concentration}.$

To make the absolute value of the Ratio symmetric with respect to increases and decreases we can take the log of the ratio.

In the model we consider here, each gene's expression level is coupled to that of all other genes through symmetric weights w_{ij} :

$$E = \sum_{i=1}^N \sum_{j=1}^N w_{ij} \log R_i \log R_j + \sum_{i=1}^N (\log R_i)^2$$

The second term produces a penalty for each gene to deviate from its steady state expression level.

The weights of this function may be set so that the free energy minima correspond to observed states under perturbations. That is, we can measure using expression microarrays the log ratios of mRNA concentrations when the cell is subject to a specific perturbation. We can then set the weights so that this state is a free energy minimum of the energy function when the perturbation is applied.

For instance, a genetic perturbation may involve the deletion of a single gene. We can simulate this by setting the logR value to a very negative number for the deleted gene. If

an expression microarray has been measured under these conditions, we know the values of the log ratios for all genes in this condition. We then must solve for the weights so that this new state represents a free energy minimum.

One possible solution (see Hopfield 1982) is to set the weights to the product of the observed values of the log ratios:

$$w_{ij} = \sum_{k=1}^N \log R_i^k \log R_j^k$$

where the sum is over the values of the ratios in various experiments.

Once the weights are set one must develop an efficient algorithm to find the free energy minima of the system under other perturbations. One possible solution to this problem is to use the Metropolis algorithm for Monte Carlo sampling (see Metropolis 1953).

It is hoped that if we can find values of the weights that produce minima for many measured states of perturbed cells, then the model will be predictive of new unmeasured perturbations. In other words, the aim of this project is to produce a model that allows us to predict the effect of new perturbations of the mRNA concentrations of genes.

References

Hopfield JJ. Neural networks and physical systems with emergent collective computational abilities. Proc Natl Acad Sci U S A. 1982 Apr;79(8):2554-8.

Eisen MB, Brown PO. DNA arrays for analysis of gene expression. Methods Enzymol. 1999;303:179-205.

Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science. 1995 Oct 20;270(5235):467-70.

Metropolis N, Rosenbluth MN, Rosenbluth A, Teller H, Teller E. (1953). Equations of state calculations by fast computing machines. J. Chem. Phys, **21**, 1087.