

Experimental Design in Genetic Regulatory Network Identification: Results From *In Silico* Studies

Daniel E. Zak, Francis J. Doyle, III, Ronald K. Pearson, James S. Schwaber

As stated by Leroy Hood (2001), "the difference between man and monkey is gene regulation." One of the great prospects in the post-genomic era is that the genetic regulatory networks that govern such varied processes as development, differentiation, and neuronal adaptation will be uncovered through the analysis of gene expression and related perturbation time courses. In previous work we developed a 10-gene genetic regulatory network hybrid stochastic simulator for the purpose of developing and validating genetic regulatory network identification methods. Using this simulator we have demonstrated the value of including datasets that constrain possible gene-gene interactions in addition to gene expression data for the purpose of identifying genetic regulatory networks (Zak et al., 2001).

In the present study we use our genetic regulatory network simulator to explore the impact of several design variables on the ability of identification methods to identify genetic regulatory networks from gene expression data. Specifically, we consider how excitation in the perturbation sequence (e.g., pulse trains versus steps) can influence the results of the identification methods. We also consider the impact of sampling rate and number of samples, which are especially important due to the varied time scales present in genetic regulatory networks and the difficulty in obtaining large numbers of gene expression measurements. Finally, treating individual trajectories of the stochastic simulator as individual "cells", we explore the impact of cell sample size on the success of genetic regulatory network identification methods. The network identification methods we use are based on the linear discrete dynamic formulation originally used by D'haeseleer et al. (1999) as well as qualitative methods that we have been developing.

References

D'Haeseleer, P., Wen, X., Fuhrman, S., and Somogyi, R. (1999). Linear modeling of mRNA expression levels during CNS development and injury. *Pacific Symposium on Biocomputing* 4: 41-52.

Hood L., *Bio2001*, San Diego, CA. 24-28 June, 2001.

Zak D., Doyle F., Gonye G. and Schwaber J. "Simulation studies for the identification of genetic networks from cDNA array and regulatory activity data." *Proceedings of the Second International Conference on Systems Biology*. 231-238, 2001.