Genetic Regulatory Network Simulations for Data Analysis

Daniel E. Zak*
*Presenter
Department of Chemical Engineering
University of Delaware
Newark, DE 19716
Daniel Baugh Institute
Thomas Jefferson University
Philadelphia, PA 19107
zak@che.udel.edu

Francis J. Doyle III
Department of Chemical Engineering
University of Delaware
Newark, DE 19716
doyle@che.udel.edu

James S. Schwaber
Daniel Baugh Institute
Department of Pathology, Anatomy and Cell Biology
Thomas Jefferson University
1020 Locust Street
Philadelphia, PA 19107
james.schwaber@mail.tju.edu

ABSTRACT

The explosion of methods for analyzing biological data since the advent of the microarray need not be emphasized. There have been many publications concerning the biological insights that may and may not be gained from this increasingly common data source.

One important aspect that is missing from the proposed microarray data analysis methods is validation through simulation. This concept is well established in the field of process control, where controllers are routinely developed and tested on realistic mathematical models of actual processes. There are a number of mechanistic models of small biological networks ([3], for example) but these are not nearly complex enough to represent the systems that the microarray analysis methods are designed to investigate. The cause is the obvious lack of sufficient mechanistic details for any but the simplest systems. Despite the lack of mechanistic details, the development of computational models of genetic regulatory networks for the testing and validation of microarray data analysis methods is useful. Models for validation purposes need only to be constrained by our concept of possible architectures, and a lack of some mechanistic details is not fatal. The most obvious benefits of such models are the ease with which *in silico* experiments can be performed and that the underlying system architecture is known *a priori*, making validation possible. Another benefit is the clarification of unknowns since the models must include all *interesting* and *obscuring sources of variation* [2]. This process generates a quantitative and computational framework for the evaluation of analytical approaches.

In previous work [7] we developed a ten gene genetic regulatory network based on a simple steroid receptor mechanism and simple genetic networks from the literature. We used this model to demonstrate the need for biological information in addition to microarray data to make network identification possible. In the present work, we expand this model to include cell to cell stochastic variability through stochastic simulations [1] and hybrid deterministic/stochastic simulations that we have developed. We also present an additional model that includes a signaling mechanism based on the EGF receptor [3] and more complex forms of transcriptional regulation. As before, we initially use these simulation models to evaluate linear [6], log-linear [7] and linear + squashing [5] approximations to the gene network dynamics. In addition, we also use these simulation models as a basis for various gray-box modeling approaches [4] that attempt to combine biological insights with abundant but incomplete measurement data. For example, in attempting to relate experimentally controllable ligand inputs to changes in gene expression levels, it is possible to develop detailed mechanistic models relating ligand inputs to signalling activities, but downstream relations between these activities and gene expression must be developed empirically. Given the potential for combinatorially complex interconnection searches in these networks, we are particularly interested in exploring and developing analytical methods that combine biologically-motivated constraints with coarse prior information (e.g., certain interactions known to be inhibitory, excitatory, or absent).

ADDITIONAL AUTHORS

Ronald K. Pearson, Tampere International Center for Signal Processing, Tampere University of Technology, Tampere, Finland (`pearson@cs.tut.fi`)

REFERENCES

[1] D. T. Gillespie. *Journal of Physical Chemistry*, 81:2340–23611, 1977.

[2] A. J. Hartemink, D. K. Gifford, T. S. Jaakkola, and R. A. Young. *SPIE BiOS*, 2001.

[3] B. N. Kholodenko, O. V. Demin, G. Moehren, and J. B. Hoek. *Journal of Biological Chemistry*, 274:30169–30181, 1999.

[4] R. K. Pearson. *Discrete-Time Dynamic Models*. Oxford, 1999.

[5] D. C. Weaver, C. T. Workman, and G. D. Stormo. *Pacific Symposium on Biocomputing*, 4:102–111, 1999.

[6] L. F. Wessels, E. P. V. Someren, and M. J. Reinders. *Pacific Symposium on Biocomputing*, 6:508–519, 2001.

[7] D. E. Zak, F. J. D. III, G. E. Gonye, and J. S. Schwaber. *Proceedings from the Second International Conference on Systems Biology*, 2:231–238, 2001.