Pathway Reconstruction via Analysis of Constitutive Fluctuations in Molecular Activities

Hunter Elliott*, Khuloud Jaqaman, Gaudenz Danuser
Cell Biology, The Scripps Research Institute, La Jolla, USA
*email: helliott@scripps.edu

Abstract
In homeostasis cellular processes fluctuate about their respective equilibriums. These variations propagate through cellular pathways, inducing similar fluctuations in downstream component processes. Analysis of fluctuations in several component processes may thus elucidate the pathway structure and hierarchy of interaction.
Critical to this concept of pathway reconstruction are methodologies for extracting relationships and hierarchies from concurrent time series of cellular activities. To test the feasibility of pathway reconstruction by fluctuation analysis software was developed to simulate and analyze the propagation of fluctuations through cellular pathways. It has been shown previously that constitutive fluctuations of individual cellular processes in steady state can often be characterized by autoregressive moving average (ARMA) models[1]. To simulate interactions among multiple processes, ARMA models were generalized to allow exogenous inputs to each process, enabling the description of multi-nodal networks of arbitrary topology. Simulated time series can be generated from networks with predetermined topologies of any kind, i.e. they may contain multiple and nested feed-back and feed-forward loops (Figure 1).

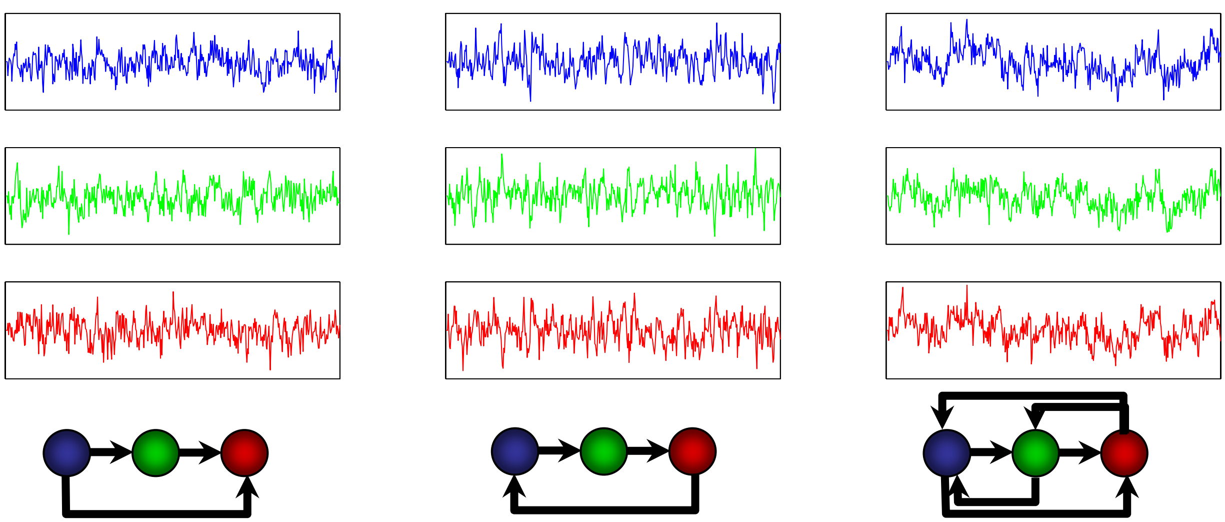

Figure 1 : Feed-forward, feed-back and complex ARMA networks with corresponding simulated time series.

The inter-process connections in these networks are flexible, allowing them to accurately represent cellular pathways. Their strength may vary, representing either a direct interaction or interaction through intermediaries. Information passes from process to process in a time-dependent fashion. Connections may be delayed or immediate. Together with autoregressive models in each process, which reflect local feed-back loops, the variability of the connections implicitly accounts for unobserved or unknown processes not present in the model.

The multivariate time-series generated by these simulations were regressed[2] to evaluate the conditions under which network topologies can be recovered. For small networks, an exhaustive topology search can be performed by regressing with models that cover the full space of possible connections. The quality of each fit is judged via the Bayesian information criterion and the topology with the lowest BIC assumed to be most likely. This method correctly recovered network topology up to approximately a 5% signaling rate, i.e. when at least 5% of the activity at one process is communicated to a downstream or back to an upstream process. However, this method is limited in applicability due to the combinatorial complexity that arises with an increasing number of pathway components. At signaling rates of ~10% a second method allows a non-exhaustive topology search to be performed, utilizing the stratification of BIC scores into layers of "topology motifs" (Figure 2).

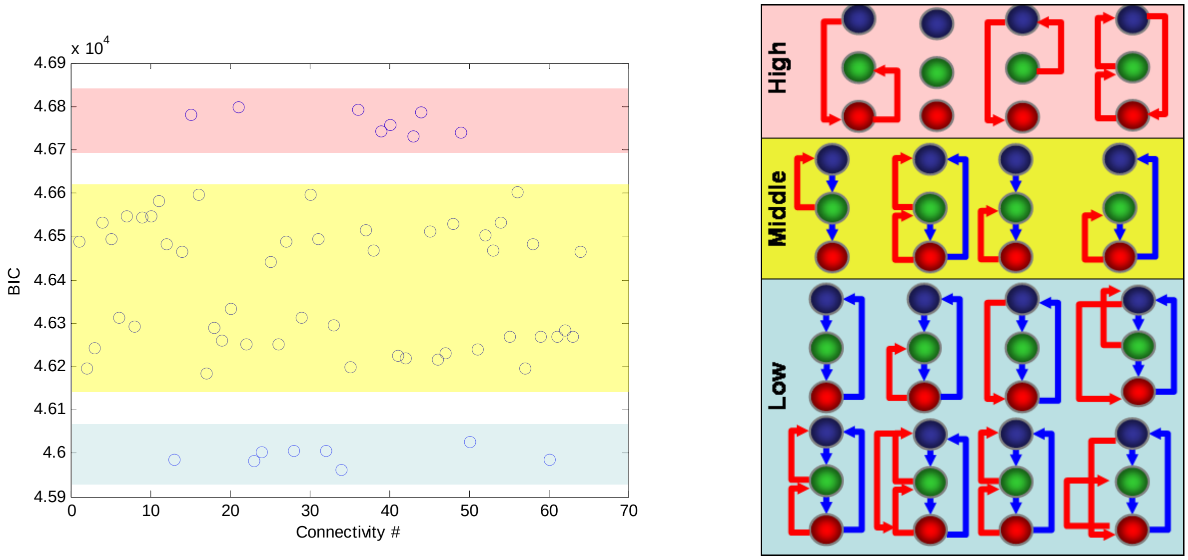

Figure 2 - BIC strata and examples from corresponding network topology motifs. Blue connections are present in the actual model while red are not.

The BICs from different models take on distinct strata, corresponding to the number and type of connections they share with the correct model. In combination with covariance data this allows a more directed search, decreasing the number of required regressions and allowing reconstruction of networks containing a large number of processes.

Thus, this work establishes the initial steps towards reconstructing complex molecular pathways solely from constitutive cellular fluctuations. This provides a potent alternative to the probing of pathways by perturbation, in which the connection between observed phenotype and primary function of the disrupted process is obfuscated by cellular compensation mechanisms.

1. Jaqaman, K. et al. Comparative Autoregressive Moving Average Analysis of Kinetochore Microtubule Dynamics in Yeast. Biophys. J. 91, 2312-2325 (2006).
2. Jones, R. H. Maximum Likelihood Fitting of ARMA Models to Time Series with Missing Observations. Technometrics 22, 389-395 (1980).