Population-Expression Dynamics

Sean P. Stromberg[1], Rustom Antia[1], and Ilya Nemenman[1,2]

***Short Abstract* — The expression profile of a population of cells is governed by both the within cell genetic regulatory networks and the population scale dynamics of division and death. The integration of these two scales is essential in analyzing differentiating cell populations undergoing selection. Here we present population-expression models that combine these scales, incorporating the chemical kinetics including stochastic fluctuations, dilution from cell division, and expression dependent rates of division and death. When translated into equations these models generate non-conservative, non-local advection-diffusion equations that model the population dynamics of heterogeneous differentiating populations.**

***Keywords* — Heterogeneous Populations, Population-Expression, Fokker-Planck, Non-local PDE, Advection-Diffusion**

I. Motivation

Flow-cytometry yields up to fifteen simultaneous protein expression measurements, per-cell, for tens of thousands of cells from each patient. Measuring an ensemble of cells in this way yields the distribution of expression levels. These distributions are then measured at multiple time points. In comparison, while gene-arrays yield data for tens of thousands of genes from each patient, their typical use returns only the average cellular expression, giving no other information about the distribution.

Even with these limitations the data produced in typical gene-array experiments is capable of inferring the connectivity of genetic networks [1]. With knowledge of the full distribution of expression levels, rather than just the average, we can infer more than just connectivity: we can infer directionality in the gene network, account for spurious correlations resulting from cell division, fit models of the chemical kinetics to estimate parameters [2], and discover the drivers of selection in a multi-cellular population [3].

II. Method

Typical models of within cell systems biology give a set of ODE rate laws $\vec{\gamma}(\vec{A})$ for a set of chemicals $\vec{A}$. Treating the chemical quantities A_i as spatial variables in a high-dimensional phase space, the state of a cell is given by a point in the space and the differentiation of cells gives a trajectory governed by the vector field $\vec{\gamma}(\vec{A})$. Formulating the problem this way allows us to model the trajectory of an ensemble of cells $\rho(\vec{A},t)$, describing the abundance of cells with expression level $\vec{A}$ at time t. The dynamics of $\rho(\vec{A},t)$ are then given by an advection equation:

$$\frac{\partial\rho(\vec{A},t)}{\partial t} = -\nabla\cdot\left[\vec{\gamma}(\vec{A})\rho(\vec{A},t)\right].$$

This approach in ecology is often referred to as a structured population model with continuous structuring variable [4].

This equation is appended with additional terms [5] to capture the effects of stochasticity in expression (as a Fokker-Planck like equation [6]), sources of new cells $\Gamma(\vec{A})$, cell death dependent on expression $\nu(\vec{A})\rho(\vec{A},t)$, and cell division $2^{d+1}\mu(2\vec{A})\rho(2\vec{A},t) - \mu(\vec{A})\rho(\vec{A},t)$. Cell division generates non-local terms as cells cut their chemical quantities roughly in half as they divide. There is a source of new cells added to $\vec{A}$ coming from the density at $2\vec{A}$. (The factor of 2^{d+1} comes from a subtlety of the calculus of non-local equations where we are adding to a region $\vec{A}+\delta\vec{A}$ from a region $2\vec{A}+2\delta\vec{A}$, which is twice as wide in every dimension d modeled, with an additional factor of two for doubling on division.)

III. Examples

These methods have their greatest utility in immunology, developmental biology, and cancer biology, where cell division and death have time scales similar to differentiation. We present the following examples:

- Where ODE models of population dynamics succeed and fail,
- The skewed expression profile of dividing populations,
- Dilution by cell division compared with changes in transcriptional regulation,
- Spurious correlations resulting from cell division,
- Biases of selection on expression profiles,
- Inference of directionality in gene-networks.

References

[1] Margolin AA , et al. (2006) "ARACNE: An Algorithm for the Reconstruction of Gene Regulatory Networks in a Mammalian Cellular Context." *BMC Bioinformatics* **7**(Suppl 1).

[2] Munsky B, Trinh B, Khammash M. (2009) "Listening to the noise: random fluctuations reveal gene network parameters." *Molecular systems biology* **5**(1).

[3] Lehner B, Kunihiko K. (1985) "Fluctuation and response in biology." *Cellular and Molecular Life Sciences* **68**(1005-1010).

[4] Metz JAJ and Diekmann O. (1986) The Dynamics of Physiologically Structured Populations. In *Lecture Notes in Biomathematics* **68**, Springer

[5] Stromberg S, Antia R, Nemenman I. "Population-expression models of immune response" in preparation.

[6] Van Kampen NG. (2007) Stochastic Processes in Physics and Chemistry, Third Edition, North Holland.

[1]Department of Biology, Emory University. E-mail: stromberg@emory.edu

[2]Department of Physics and Computational Life Sciences Initiative, Emory University.