systems biology: can mathematics lead experiments?

jeremy gunawardena

department of systems biology
harvard medical school
200 longwood avenue
boston, ma 02115

jeremy@hms.harvard.edu

http://vcp.med.harvard.edu/

broad institute, MIA

10 february 2016
what is systems biology?

how do the collective interactions of dead molecules give rise to the physiology of the living organism?
“The lack of real contact between mathematics and biology is either a tragedy, a scandal, or a challenge, it is hard to decide which.”


“Eugene Wigner wrote a famous essay on the unreasonable effectiveness of mathematics in natural sciences. He meant physics, of course. There is only one thing which is more unreasonable than the unreasonable effectiveness of mathematics in physics, and this is the unreasonable ineffectiveness of mathematics in biology.”

Some lessons about models from Michaelis and Menten

Biology is more theoretical than physics
Gunawardena, Mol Biol Cell, 24:1827-9 2013

Beware the tail that wags the dog: informal and formal models in biology

Models in biology: ‘accurate descriptions of our pathetic thinking’
Gunawardena, BMC Biology, 12:29 2014
eukaryotic gene regulation

joint work with:

Angela DePace  Javier Estrada  Felix Wong  Tobias Ahsendorf  Roland Eils

Estrada, Wong, DePace, Gunawardena, “Higher order cooperativity and energy dissipation can sharpen switching of eukaryotic genes”, submitted, 2015

eubacterial gene regulation

- transcription factor (TF) binding motifs ~16bp on average
- pairwise cooperative interactions TF-DNA, TF-TF, TF-RNAP
- regulation takes place without energy expenditure
eukaryotic gene regulation

- TF binding motifs ~8bp on average
- information integration over long distances
- many forms of energy expenditure
  - chromatin reorganisation
  - nucleosome remodelling
  - PTM of histone tails, TFs, co-regulators, RNAP
  - DNA methylation
eukaryotic gene regulation II

sharpness in gene regulation

consistent with the idea that Hb transcription is activated by cooperative binding of effectively five Bcd molecules

Gregor, Tank, Wieschaus, Bialek, “Probing the limits to positional information”, Cell 130:153-64 2007
the Hill function

$$\mathcal{H}_a(x) = \frac{x^a}{1 + x^a}$$


Archibald Vivian Hill 1886 - 1977

“The Hill equation remains what Hill intended it to be: an empirical descriptor”


“Despite its appealing simplicity, the Hill equation is not a physically realistic reaction scheme, raising the question of whether it should be abandoned in favor of realistic schemes; at the very least, its limitations should be more widely recognized”

macroscopic interpretation

one-dimensional chemistry

labelled directed graph (no self-loops)

\[
\begin{align*}
\begin{pmatrix}
 x_1 \\
 x_2 \\
 x_3 \\
 x_4 \\
 x_5 \\
 x_6
\end{pmatrix}
&= 
\begin{pmatrix}
 -(b + c + f) & a & 0 & 0 & 0 & 0 \\
 b & -(a + d) & 0 & 0 & 0 & 0 \\
 c & 0 & -(g + e) & 0 & 0 & 0 \\
 0 & d & e & 0 & 0 & 0 \\
 f & 0 & 0 & 0 & -i & h \\
 0 & 0 & g & 0 & i & -h
\end{pmatrix}
\begin{pmatrix}
 x_1 \\
 x_2 \\
 x_3 \\
 x_4 \\
 x_5 \\
 x_6
\end{pmatrix}
\end{align*}
\]

Laplacian dynamics

\[
\frac{dx}{dt} = \mathcal{L}(G).x
\]

conservation law

\[1.\mathcal{L}(G) = 0\]

microscopic interpretation

let \( X(t) \) be a time-homogeneous Markov process on the states \( 1, \cdots, n \) for which infinitesimal transition rates exist -

\[
\lim_{\Delta t \to 0} \frac{\Pr(X(t + \Delta t) = i \mid X(t) = j)}{\Delta t} = a_{ij}
\]

define the graph, \( G_X \), with vertices \( 1, \cdots, n \) and an edge \( j \to i \) iff \( a_{ij} \neq 0 \) give this edge the label \( a_{ij} \)

the master equation (Kolmogorov forward equation), for the probability of \( X(t) \) being in state \( i \) at time \( t \), is identical to Laplacian dynamics on \( G_X \)

\[
x_i(t) = \Pr(X(t) = i) \quad \frac{dx}{dt} = \mathcal{L}(G_X).x
\]

basic results

- for any graph, $G$, the dynamics always tends to a stable steady state

$$x(t) \to x^* \quad \frac{dx}{dt} \bigg|_{x=x^*} = 0 \quad x^* \in \ker \mathcal{L}(G)$$

- if $G$ is strongly connected, there is an unique steady state up to scalars

$$\ker \mathcal{L}(G) = \langle \rho \rangle$$

- and a canonical basis element is given by the Matrix-Tree Theorem

$$\rho_i = \sum_{T \in \Theta_i(G)} \left( \prod_{j \xrightarrow{a_{jk}} k \in T} a \right)$$

- in general, the kernel is determined by the terminal strongly connected components $C_1, \cdots, C_k$

$$\ker \mathcal{L}(G) = \langle \rho_{C_1}, \cdots, \rho_{C_k} \rangle$$

Mirzaev & Gunawardena, Bull Math Biol 75:2118-49 2013
elimination of internal complexity

strongly-connected graph

if there is a steady state

\[ \frac{dx}{dt} = 0 \]

\[ x = \lambda \rho \]

\[ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \lambda \begin{pmatrix} \rho_1 \\ \vdots \\ \rho_n \end{pmatrix} \]

if the \( x_i \) are probabilities, then

\[ x_1 + \cdots + x_n = 1 \]

the probability of being in state \( i \), is then

\[ x_i = \left( \frac{\rho_i}{\rho_1 + \cdots + \rho_n} \right) \]
linear framework in gene regulation

thermodynamic equilibrium

**principle of detailed balance**: at thermodynamic equilibrium, every reaction is reversible and each pair of reversible reactions is separately at equilibrium, irrespective of any other reactions in which the components participate.

\[
k_1[A] = k_2[B] \\
k_5[C] = k_6[A] \\
k_3[B] = k_4[C]
\]

\[k_1k_3k_5 = k_2k_4k_6\]

*cycle condition*

Gilbert Lewis, “*A new principle of equilibrium*”, PNAS 11:179-83 1925
away from thermodynamic equilibrium

Kinetic Proofreading: A New Mechanism for Reducing Errors in Biosynthetic Processes Requiring High Specificity
(protein synthesis/DNA replication/amino-acid recognition)
J. J. HOPFIELD Proc. Nat. Acad. Sci. USA
Vol. 71, No. 10, pp. 4135-4139, October 1974

“THE HOPFIELD BARRIER”

thermodynamic equilibrium sets an upper bound to how well information processing tasks can be undertaken by a biochemical system.

the only way to exceed this barrier is to dissipate energy and maintain the system away from equilibrium
gene regulation model

graph $G_n$ for $n = 3$ sites
higher-order cooperativity at equilibrium

\[ K_{i,S} = \frac{a_{i,S}}{b_{i,S \cup \{i\}}} \]

\[ \omega_{i,S} = \frac{K_{i,S}}{K_{i,\emptyset}} \]

exchange formula

\[ \omega_{i,S \cup \{j\}} \omega_{j,S} = \omega_{j,S \cup \{i\}} \omega_{i,S} \]

independent parameters

\[ \omega_{i,S} (i < S) \quad \kappa_i = \frac{K_{i,\emptyset}}{K_{1,\emptyset}} \]
gene regulation function

\[ f_n(x) = \frac{\sum_i r_i \rho_i}{\sum_i \rho_i} \]

\[ x = [T] \]

\[ f_n(x) = \frac{c_n x^n}{1 + c_1 x + \cdots + c_n x^n} \]

\[ c_k = \left( \sum_{1 \leq i_1 < \cdots < i_k \leq n} \left( \prod_{j=1}^{k} \kappa_{i_j, \omega_{i_j, \{i_{j+1}, \ldots, i_k\}}} \right) \right) (K_{1,0})^k \]
sharpness

\[ f_n(x) \]

\[ g_n(y) = \frac{x}{x_{0.5}} \]

\[ \rho(g_n) = \max_{y \geq 0} \frac{dg_n}{dy} \]

\[ \gamma(g_n) = z \text{ such that } \left. \frac{dg_n}{dy} \right|_{y = z} = \rho(g_n) \]

“steepness”

“position”
pairwise cooperativity at equilibrium

\[ \omega_{i,S} = 1 \text{ for } \# S > 1 \]
all higher-order cooperativities at equilibrium
the Hill function, revisited yet again

\[ f_n(x) = \frac{c_n x^n}{1 + c_1 x + \cdots + c_n x^n} \]

\[ \approx \frac{x^k}{1 + x^k} = H_k(x) \quad k < n \]

the Hill function appears to be more "realistic" than previously thought but it lies on the boundary of what is possible
sharpness away from equilibrium

\[ f_{ne}(x) = \frac{d_n x^n + \cdots + d_{2n-1} x^{2n-1}}{e_0 + e_1 x + \cdots + e_{2n-1} x^{2n-1}} \quad d_{2n-1} = e_{2n-1} \]

n = 3 sites

"Hopfield barrier"

Parameter ranges:
- \([10^{-3}, 10^3]\) at equilibrium
- \([10^{-2}, 10^2]\) away from equilibrium
history dependence

equilibrium GRF:

\[
f_n(x) = \frac{c_n x^n}{1 + c_1 x + \cdots + c_n x^n}
\]

non-equilibrium GRF:

\[
f_{ne}^n(x) = \frac{d_n x^n + \cdots + d_{2n-1} x^{2n-1}}{e_0 + e_1 x + \cdots + e_{2n-1} x^{2n-1}}
\]

n = 2 sites \quad 4 spanning trees

n = 3 sites \quad 384 spanning trees

n = 4 sites \quad ?
history dependence!

equilibrium GRF:

\[ f_n(x) = \frac{c_n x^n}{1 + c_1 x + \cdots + c_n x^n} \]

non-equilibrium GRF:

\[ f_{ne}^n(x) = \frac{d_n x^n + \cdots + d_{2n-1} x^{2n-1}}{e_0 + e_1 x + \cdots + e_{2n-1} x^{2n-1}} \]

n = 2 sites \quad 4 \text{ spanning trees}

n = 3 sites \quad 384 \text{ spanning trees}

n = 4 sites \quad 42,467,328 \text{ spanning trees}
# Systems Biology Theory Lunch

"Chalk" talks (whiteboard only, no slideware) on conceptual issues in biological systems. Open to all.
To join the mailing list or for more information, contact jeremy (at) hms.harvard.edu

## Schedule of talks for 2015-16

<table>
<thead>
<tr>
<th>Date</th>
<th>Speaker</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>25 Sep 2015</td>
<td>Dagmar Iber</td>
<td>Scaled read-out of morphogen gradients on growing domains</td>
</tr>
<tr>
<td>2 Oct 2015</td>
<td>Eric Kramer</td>
<td>Regulating cell polarity in trees and other plants</td>
</tr>
<tr>
<td>9 Oct 2015</td>
<td>Gordon Hager</td>
<td>Transcription factor/chromatin interactions: integrating genome-wide and real-time dynamic datasets</td>
</tr>
<tr>
<td>16 Oct 2015</td>
<td>Harold Zakon</td>
<td>Evolution of molecular complexity in the earliest nervous system(s)</td>
</tr>
<tr>
<td>23 Oct 2015</td>
<td>Thomas Höfer</td>
<td>Eukaryotic gene regulation as frequency modulation</td>
</tr>
<tr>
<td>30 Oct 2015</td>
<td>Ariel Amir</td>
<td>Simultaneous regulation of cell size and chromosome replication in bacteria</td>
</tr>
<tr>
<td>6 Nov 2015</td>
<td>Eve Marter</td>
<td>Robustness, degeneracy and neuromodulation of neurons and networks</td>
</tr>
<tr>
<td>13 Nov 2015</td>
<td></td>
<td>SB PhD PROGRAMME RETREAT</td>
</tr>
<tr>
<td>20 Nov 2015</td>
<td>Johan Paulesson</td>
<td>Exploiting single-cell fluctuations</td>
</tr>
<tr>
<td>27 Nov 2015</td>
<td></td>
<td>THANKSGIVING</td>
</tr>
<tr>
<td>4 Dec 2015</td>
<td>Harinder Singh</td>
<td>Analyzing cell fate choice and dynamics using single-cell RNA-seq</td>
</tr>
<tr>
<td>11 Dec 2015</td>
<td>Jané Kondev</td>
<td>Building a proportional cell: the physics of size control</td>
</tr>
</tbody>
</table>

**THEORY LUNCH TAKES A BREAK FROM 18 DEC 2015 - 5 FEB 2016 INCLUSIVE**

<table>
<thead>
<tr>
<th>Date</th>
<th>Speaker</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>12 Feb 2016</td>
<td>Stanislav Shvartsman</td>
<td>Towards quantitative biology of developmental abnormalities</td>
</tr>
<tr>
<td>19 Feb 2016</td>
<td>David Zwicker</td>
<td>Recognition of natural odors by receptor arrays</td>
</tr>
<tr>
<td>26 Feb 2016</td>
<td>Ahmad Khalil</td>
<td>Naturally-inspired engineering of biological systems</td>
</tr>
<tr>
<td>11 Mar 2016</td>
<td>Stirling Churchman</td>
<td>Dissecting gene expression at high resolution</td>
</tr>
<tr>
<td>18 Mar 2016</td>
<td>Celeste Nelson</td>
<td>Complexity and the mechanics of folding in developing tissues</td>
</tr>
<tr>
<td>25 Mar 2016</td>
<td>Philip Maini</td>
<td>Title to be announced</td>
</tr>
<tr>
<td>1 Apr 2016</td>
<td>Suri Vaikuntanathan</td>
<td>Title to be announced</td>
</tr>
<tr>
<td>8 Apr 2016</td>
<td>Luca Peiti</td>
<td>Evolution strategies in fluctuating environments</td>
</tr>
<tr>
<td>15 Apr 2016</td>
<td>Margaret Gardel</td>
<td>Title to be announced</td>
</tr>
<tr>
<td>29 Apr 2016</td>
<td>Elena Rivas</td>
<td>Title to be announced</td>
</tr>
<tr>
<td>27 May 2016</td>
<td>Sean Eddy</td>
<td>Phylostratigraphy</td>
</tr>
<tr>
<td>3 June 2016</td>
<td></td>
<td>SB DEPARTMENT RETREAT</td>
</tr>
</tbody>
</table>
With thanks to HMS, Armenise, NSF, NIH, CDP@MIT, HFSP, Novartis