dynamic processes in cells
(a systems approach to biology)

jeremy gunawardena
department of systems biology
harvard medical school

lecture 9
29 september 2016
cellular behaviour can also be re-programmed

molecular basis of cellular identity

how does a single genome give rise to many different cellular identities & behaviours?

different identities are represented by different patterns of gene expression

expression of a protein from a gene is regulated (weak linkage!)

1902 - 1992
1910 - 1976
1920 - 2013

"controlling elements"
in a dynamic genome
operon theory of gene regulation

bacterial 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 I

PTM of histone tails

nucleosome

chromatin organisation

DNA methylation

multi-protein coregulators
eukaryotic gene regulation II

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

gene regulation in the linear framework


vertex = “snapshot” of DNA state or “microstate”

dge = transition between states
calculating gene expression

\[ \rho \in \ker \mathcal{L}(G) \]

\[ \frac{\rho_\mu}{\sum_\mu \rho_\mu} \quad \text{probability of microstate } \mu \]

each microstate has a characteristic rate of gene expression and the overall rate is the average rate over all microstates

\[ \frac{d[mRNA]}{dt} = \frac{\sum_\mu r_\mu \rho_\mu}{\sum_\mu \rho_\mu} \quad \text{rate of gene expression in microstate } \mu \]

this formula is convenient but artificial – it separates “regulation” from “expression” – and an alternative approach is to include RNAP in the microstate and to calculate the overall rate as proportional to the probability that RNAP is present

\[ \frac{d[mRNA]}{dt} = \alpha \left( \frac{\sum_{\mu, P=1} \rho_\mu}{\sum_\mu \rho_\mu} \right) \]
detailed balance at thermodynamic equilibrium

if the system reaches **thermodynamic equilibrium**, then **detailed balance** holds and $\rho$ can be calculated in a particularly simple way.

**principle of detailed balance (in the linear framework):** every edge in the graph has a complementary reverse edge and, in any steady state, each pair of such reversible edges is independently at steady state, irrespective of any other edges reaching those vertices.

“For if this were not the case we could add a minute amount of some catalyst which would increase the rate of the reaction and its inverse along one of the paths, without affecting the two rates in the other path. This would disturb the existing equilibrium, contrary to the results of observation and of thermodynamics.” (*)

detailed balance is a consequence of **microscopic reversibility**: the fundamental laws of physics, whether classical newtonian mechanics or quantum mechanics, exhibit time-reversal symmetry (+)


**calculating ρ at equilibrium**

choose any path of reversible edges from a reference vertex to a given vertex, k

![Reference vertex diagram](image)

suppose given a steady state of the Laplacian dynamics \( x^* \in \ker \mathcal{L}(G) \)

then, by detailed balance applied to each reversible pair of edges,

\[
x_2^* = \left( \frac{a_1}{b_1} \right) x_1^* \\
x_3^* = \left( \frac{a_2}{b_2} \right) x_2^* \\
\ldots \\
x_k^* = \left( \frac{a_{k-1}}{b_{k-1}} \right) x_{k-1}^*
\]

so we can take

\[
\rho_k = \left( \frac{a_1}{b_1} \right) \left( \frac{a_2}{b_2} \right) \ldots \left( \frac{a_{k-1}}{b_{k-1}} \right) \\
\rho \in \ker \mathcal{L}(G)
\]
the cycle condition

for this construction to work, the result must be independent of the chosen path

\[
[C] = \left( \frac{k_6}{k_5} \right) \begin{bmatrix} A \end{bmatrix}
\]

\[
[C] = \begin{pmatrix} k_1 & k_3 & k_5 \end{pmatrix} \begin{pmatrix} k_2 & k_4 & k_6 \end{pmatrix} \begin{bmatrix} A \end{bmatrix}
\]

the rate constants of a system that can reach equilibrium cannot be chosen arbitrarily but are constrained by the cycle condition: in any cycle, the product of the rate constants going clockwise equals the product going counterclockwise.

equivalent to equilibrium statistical mechanics

the van't Hoff equation links the two methods

\[
\ln \left( \frac{a}{b} \right) = -\frac{\Delta G}{RT}
\]

so, for any microstate, \( k \)

\[
\ln \rho_k = -\sum \frac{\text{(interaction energies)}}{RT}
\]

the partition function is given by

\[
\sum \rho_\mu = \sum e^{-\left(\Delta G_\mu / RT\right)}
\]

single transcription factor, multiple sites - 1

other mechanisms (nucleosomes, co-regulators, chromosome structure, etc) are not explicitly modelled and their influence is assumed to be exerted through the on-rates and off-rates

to deal with the nonlinearity in the labels, assume that -  \([T] \approx T_{\text{tot}}\)

Estrada, Wong, DePace, Gunawardena, “Information integration and energy expenditure in gene regulation”, Cell 166:234-44 2016
assuming thermodynamic equilibrium, the relevant parameters are (slide 11)

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

define nondimensional **higher-order cooperativities**

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

detailed balance and the cycle condition leads to

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

a set of non-dimensional **independent parameters** is given by

\[ \omega_{i,S} \ (i < S) \quad \kappa_i = \frac{K_{i,\emptyset}}{K_{1,\emptyset}} \]
single transcription factor, multiple sites - III

assume the “all-or-nothing” (AN) strategy of gene expression – no expression unless all sites are bound

\[
\begin{align*}
\text{rates of gene expression} & \\
& \quad 0 \\
& \quad 0 \\
& \quad 1 \\
\end{align*}
\]

gene regulation function

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

\[
c_k = \left( \sum_{1 \leq i_1 < \ldots < i_k \leq n} \left( \prod_{j=1}^{k} \kappa_{i_j} \omega_{i_j,\{i_{j+1},\ldots,i_k\}} \right) \right) (K_{1,0})^k
\]
neutrophil/macrophage differentiation - experiment

individual Puer\textsuperscript{hi} cells exhibit a “mixed lineage” of macrophage and neutrophil markers 1 day after OHT treatment

PU.1-/- myeloid progenitor cells were transformed and clones selected which expressed Puer at low (Puer\textsuperscript{lo}) and high (Puer\textsuperscript{hi}) levels

how does mixed lineage transcriptional priming arise?
neutrophil/macrophage differentiation - theory

mutual repression

nonlinear dynamical system

\[
\begin{align*}
\frac{dx_1}{dt} &= f_1(x_1, x_2, \ldots, x_n; k_1, \ldots, k_m) \\
\frac{dx_2}{dt} &= f_2(x_1, x_2, \ldots, x_n; k_1, \ldots, k_m) \\
& \vdots \\
\frac{dx_n}{dt} &= f_n(x_1, x_2, \ldots, x_n; k_1, \ldots, k_m)
\end{align*}
\]

BISTABILITY

primary determinants can remain high despite lineage resolution

quantitative to qualitative dynamics

$$\frac{dx}{dt} = k_1 - k_2x$$

from a quantitative analytical solution

$$x(t) = \frac{k_1}{k_2} + \left(x(0) - \frac{k_1}{k_2}\right) \exp(-k_2t)$$

to a qualitative geometric perspective