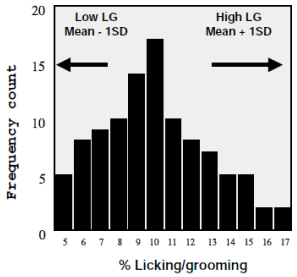
dynamic processes in cells (a systems approach to biology)

jeremy gunawardena
department of systems biology
harvard medical school

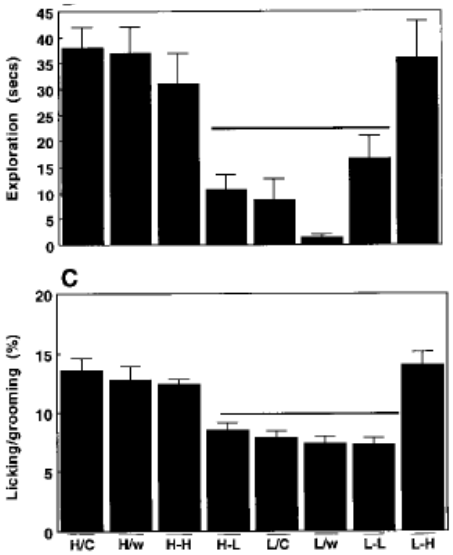
lecture 9

29 september 2016

cellular behaviour can also be re-programmed



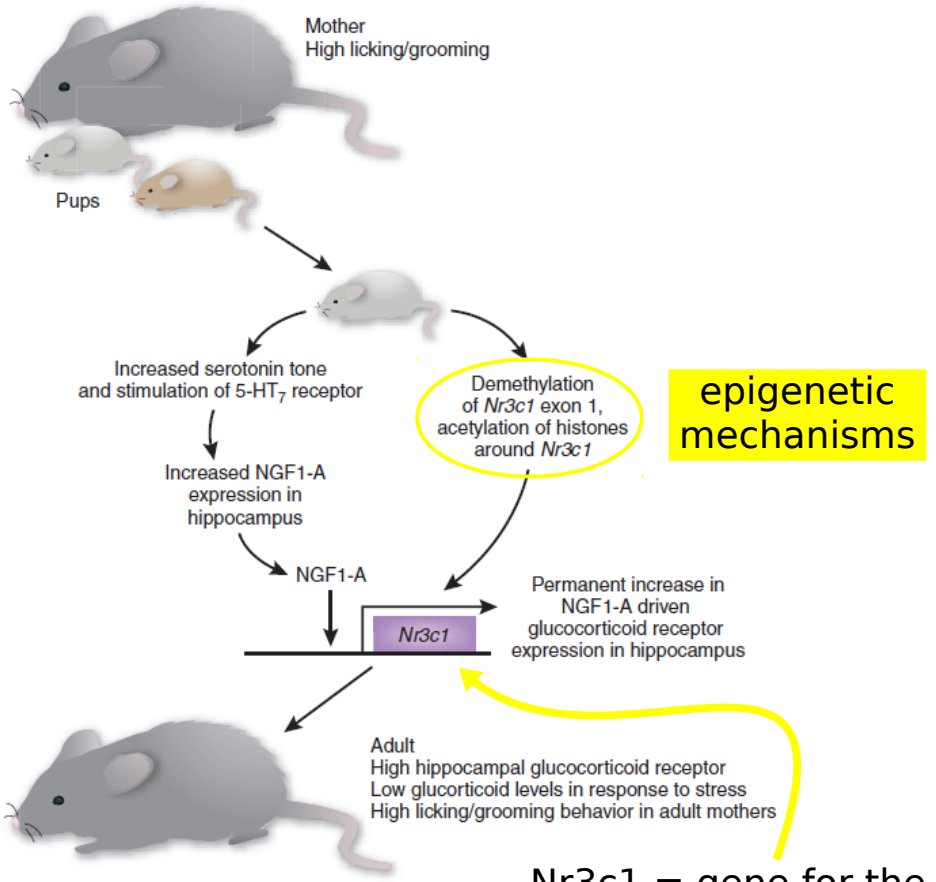
adult female offspring



open field exploration

licking/grooming

cross-fostering group



Nr3c1 = gene for the glucocorticoid receptor

Francis, Diorio, Liu, Meaney, *"Nongenomic transmission across generations of maternal behaviour and stress responses in the rat"*, Science **286**:1155-8 1999; Szyf, Weaver, Champagne, Diorio, Meaney, *"Maternal programming of steroid receptor expression and phenotype through DNA methylation in the rat"*, Frontiers in Neuroendocrinology **26**:139-62 2005

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

"controlling elements"
in a dynamic genome

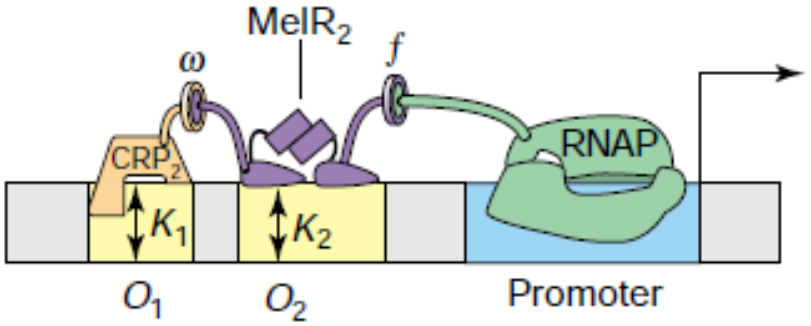
1910 - 1976

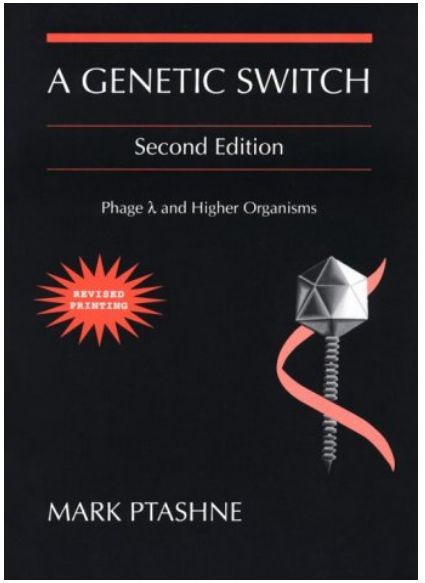
1920 - 2013

operon theory of gene
regulation

Barbara McClintock, "*Chromosome organization and genic expression*", Cold Spring Harbor Symp Quat Biol **16**:13-47 1951; Jacques Monod, Francois Jacob, *"Genetic regulatory mechanisms in the synthesis of proteins"*, J Mol Biol **3**:318-56 1961

bacterial gene regulation

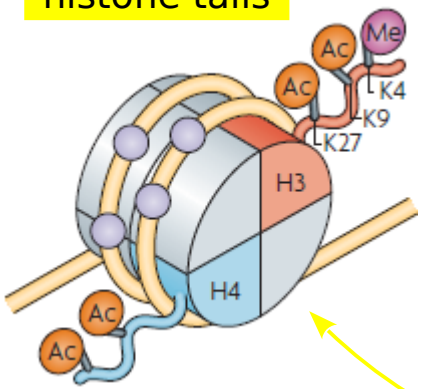




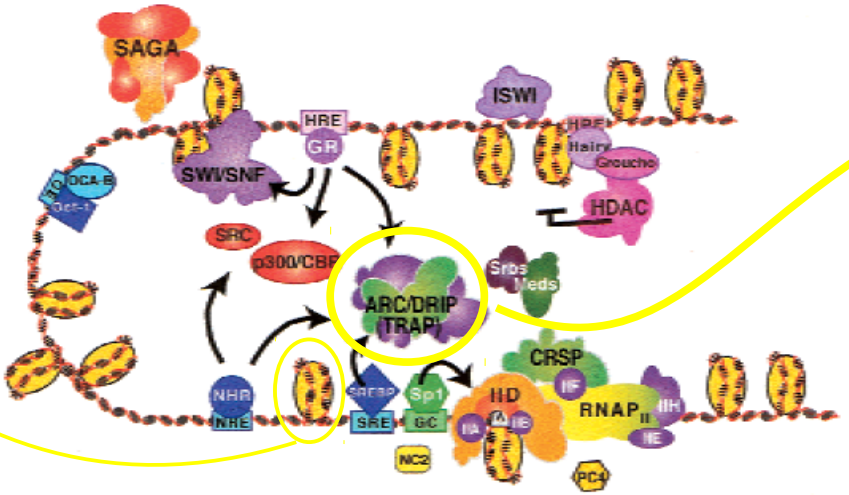
- transcription factor (TF) binding motifs ~16bp on average
- pairwise cooperative interactions TF-DNA, TF-TF, TF-RNAP
- regulation takes place without energy expenditure

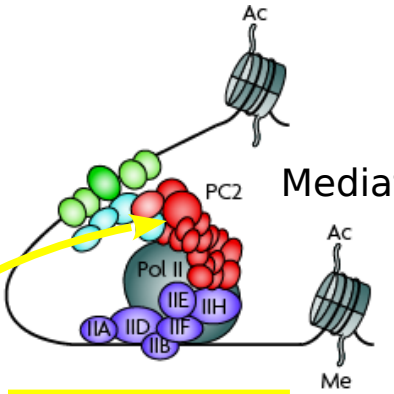
eukaryotic gene regulation I

PTM of histone tails

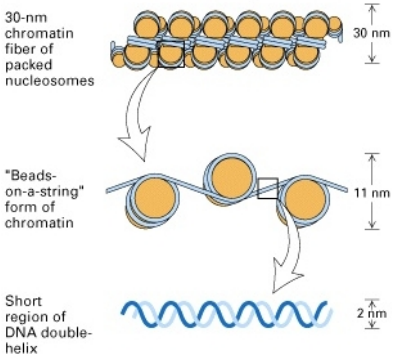


nucleosome

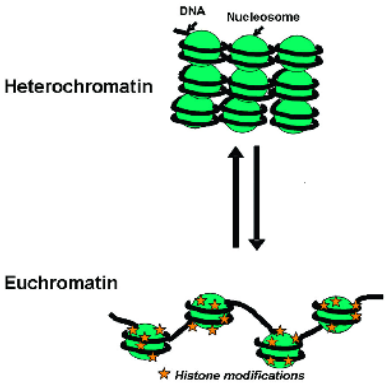




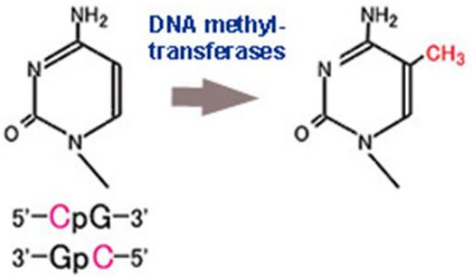
multi-protein coregulators



chromatin organisation



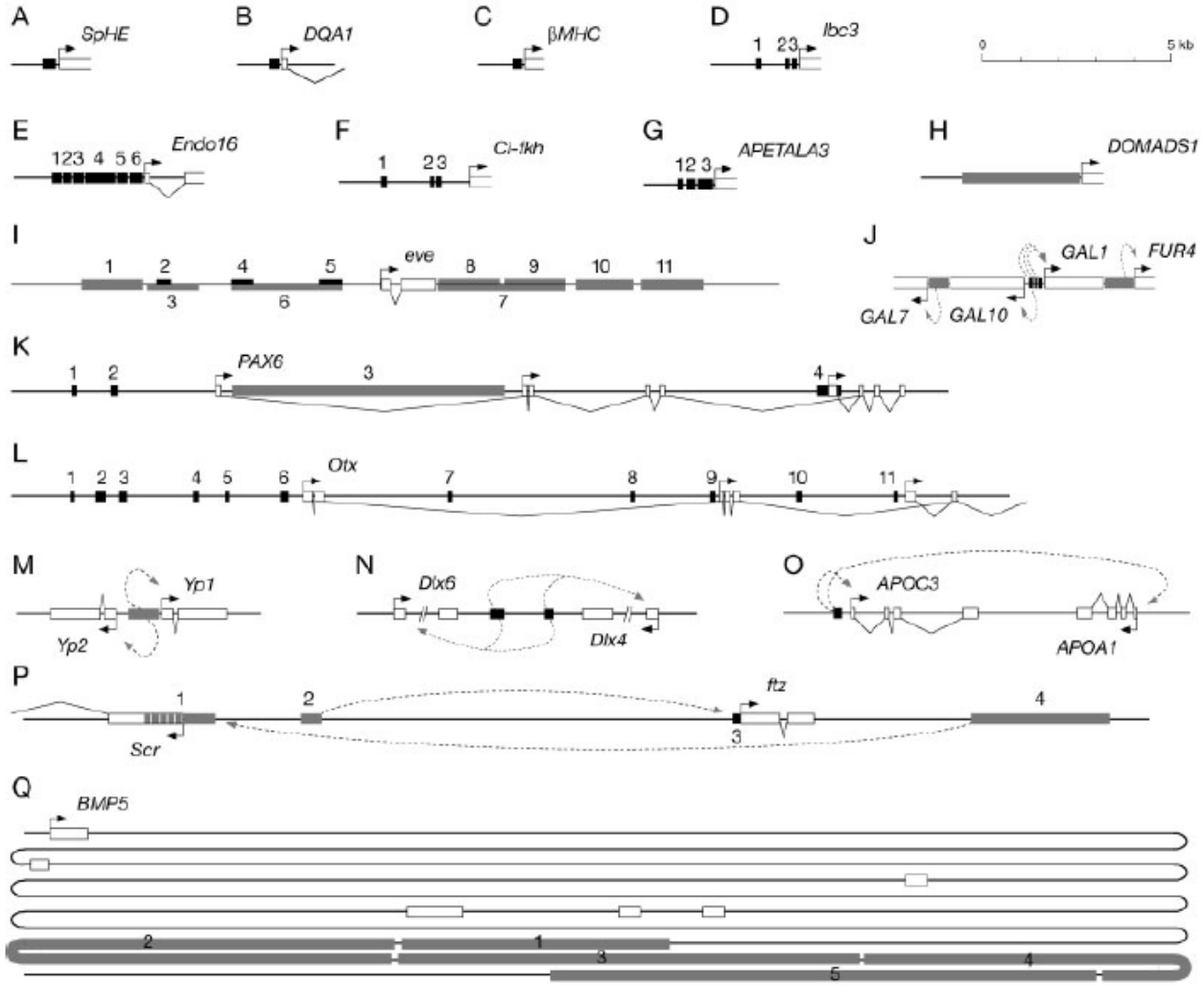
DNA methylation



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



exon

defined regulatory region

region containing regulatory sequences

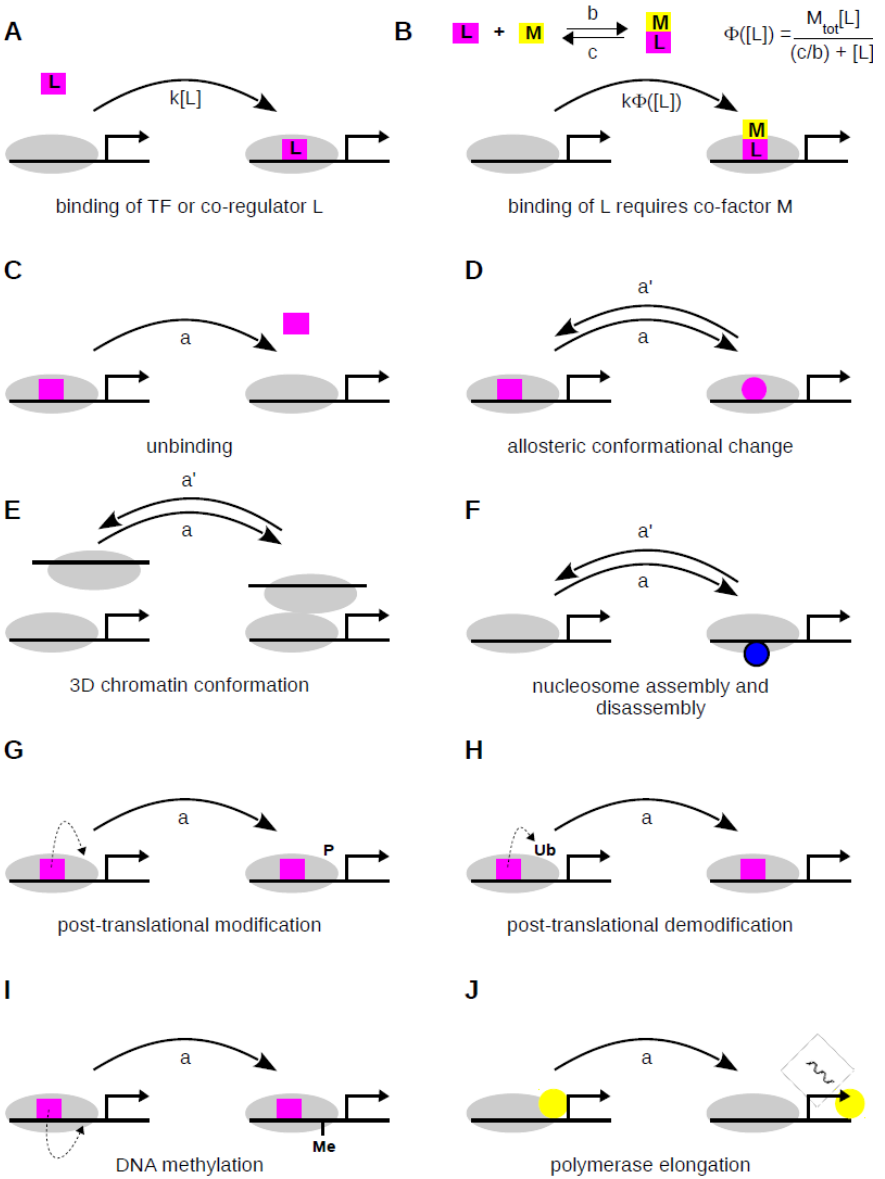
Wray, Hahn, Abouheif, Balhoff, Pizer, Rockman, Romano, *"The evolution of transcriptional regulation in eukaryotes"*, Mol Biol Evol **20**:1377-419 2003

gene regulation in the linear framework

Ahsendorf, Wong, Eils, Gunawardena, "A *framework for modelling gene regulation which accommodates non-equilibrium mechanisms*", BMC Biology **12**:102 2014

vertex = "snapshot" of DNA state
or "microstate"

edge = transition between states



calculating gene expression

$$\rho \in \ker \mathcal{L}(G) \qquad \frac{\rho_\mu}{\sum_\mu \rho_\mu}$$ probability of microstate μ

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

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

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

$$\frac{d[\text{mRNA}]}{dt} = \alpha \left(\frac{\sum_{\mu, P=1} \rho_\mu}{\sum_\mu \rho_\mu} \right)$$

detailed balance at thermodynamic equilibrium

if the system reaches **thermodynamic equilibrium**, then **detailed balance** holds and ρ 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."* (*)

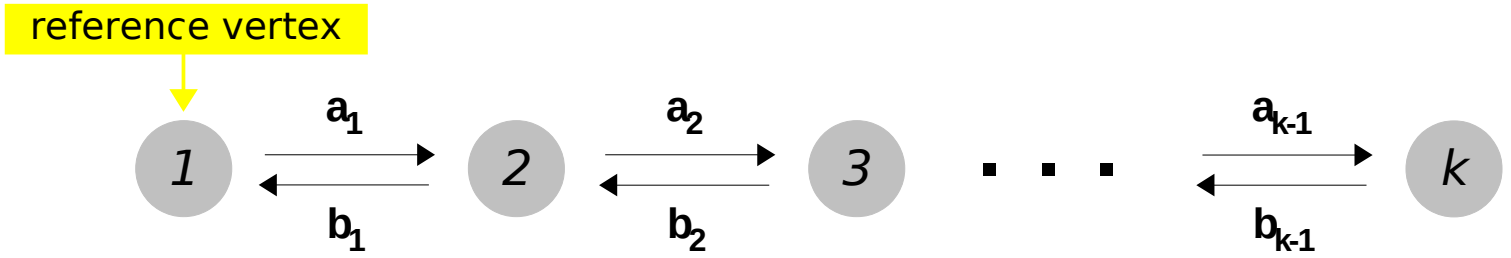
detailed balance is a consequence of **microscopic reversibility**: the fundamental laws of physics, whether classical newtonian mechanics or quantum mechanics, exhibit time-reversal symmetry (**+**)

(*) Gilbert Lewis, *"A new principle of equilibrium"*, PNAS **11**:179-83 1925

(**+**) Bruce Mahan, *"Microscopic reversibility and detailed balance"*, J Chem Edu **52**:299-302 1975

calculating ρ at equilibrium

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



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,

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

$$x_k^* = \left(\frac{a_1}{b_1}\right)\left(\frac{a_2}{b_2}\right)\cdots\left(\frac{a_{k-1}}{b_{k-1}}\right) x_1^*$$

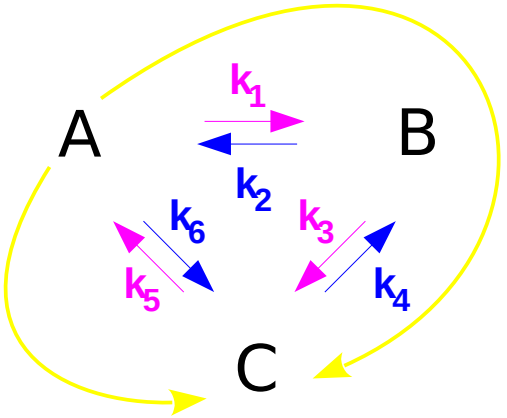
so we can take

$$\boxed{\rho_k = \left(\frac{a_1}{b_1}\right)\left(\frac{a_2}{b_2}\right)\cdots\left(\frac{a_{k-1}}{b_{k-1}}\right)} \qquad \rho \in \ker \mathcal{L}(G)$$

the cycle condition

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

$$[C] = \left(\frac{k_6}{k_5}\right)[A]$$



$$[C] = \left(\frac{k_1}{k_2}\right)\left(\frac{k_3}{k_4}\right)[A]$$

$$k_1 k_3 k_5 = k_2 k_4 k_6$$

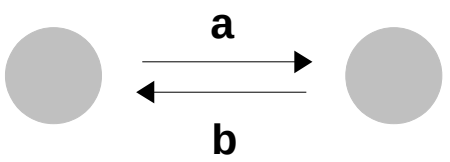
the rate constants of a system that can reach equilibrium cannot be chosen arbitrarily but are constrained by the cycle condition: in any cycle, the product of the rate constants going clockwise equals the product going counterclockwise.

Gunawardena, *"A linear framework for time-scale separation in nonlinear biochemical systems"*, PLoS ONE **7**:e36321 2012

equivalent to equilibrium statistical mechanics

the **van't Hoff equation** links the two methods

1852-1911



change in Gibbs free energy

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

so, for any microstate, k

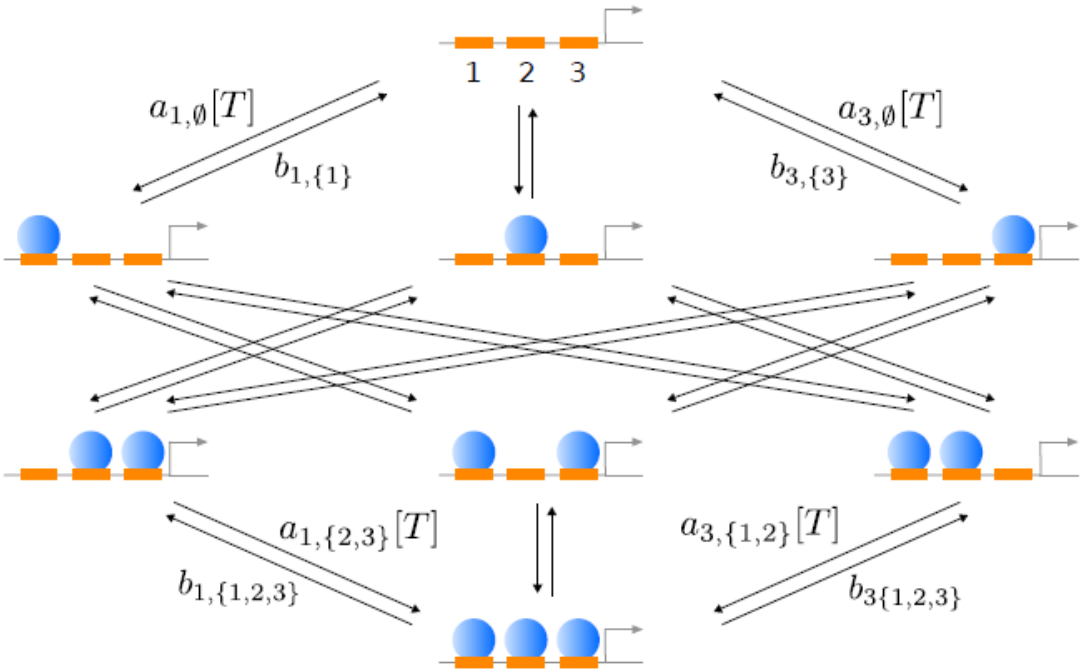
$$\ln \rho_k = -\sum (\text{interaction energies})/\text{RT}$$

TF-DNA, TF-TF, TF-RNAP, ...

the partition function is given by $\sum_\mu \rho_\mu = \sum_\mu e^{-(\Delta G_\mu / RT)}$

Bintu et al, *"Transcriptional regulation by the numbers"*, Curr Opin Gen Dev, **15**:116-24 & 125-35, 2005; Segal, Widom, *"From DNA sequence to transcription behaviour"*, Nat Rev Genetics **100**:443-56 2009; Sherman, Cohen, *"Thermodynamic state-ensemble models of cis-regulation"*, PLoS Comp Biol **8**:e1002407 2012.

single transcription factor, multiple sites – I



labelled, directed graph for n = 3 sites

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 – $[T] \approx T_{tot}$

Estrada, Wong, DePace, Gunawardena, *"Information integration and energy expenditure in gene regulation"*, Cell **166**:234-44 2016

single transcription factor, multiple sites – II

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** $\omega_{i,S} = \dfrac{K_{i,S}}{K_{i,\emptyset}}$ higher-order cooperativity

detailed balance and the cycle condition leads to

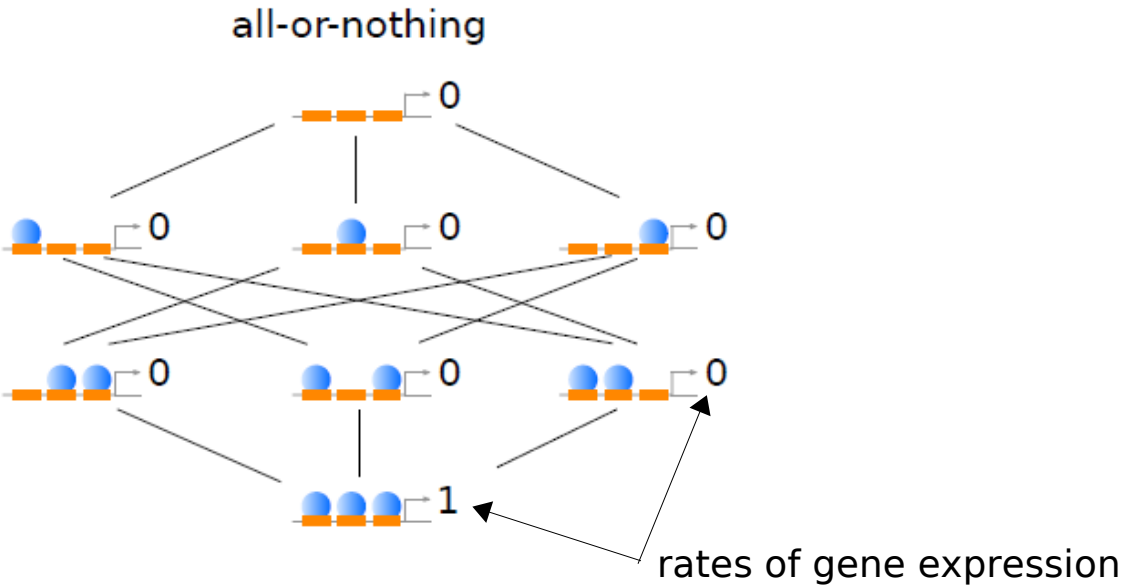
$$\omega_{i,S\cup\{j\}}\omega_{j,S} = \omega_{j,S\cup\{i\}}\omega_{i,S}$$ exchange formula

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

$$\omega_{i,S}\,(i < S) \qquad \kappa_i = K_{i,\emptyset}/K_{1,\emptyset}$$

single transcription factor, multiple sites – III

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

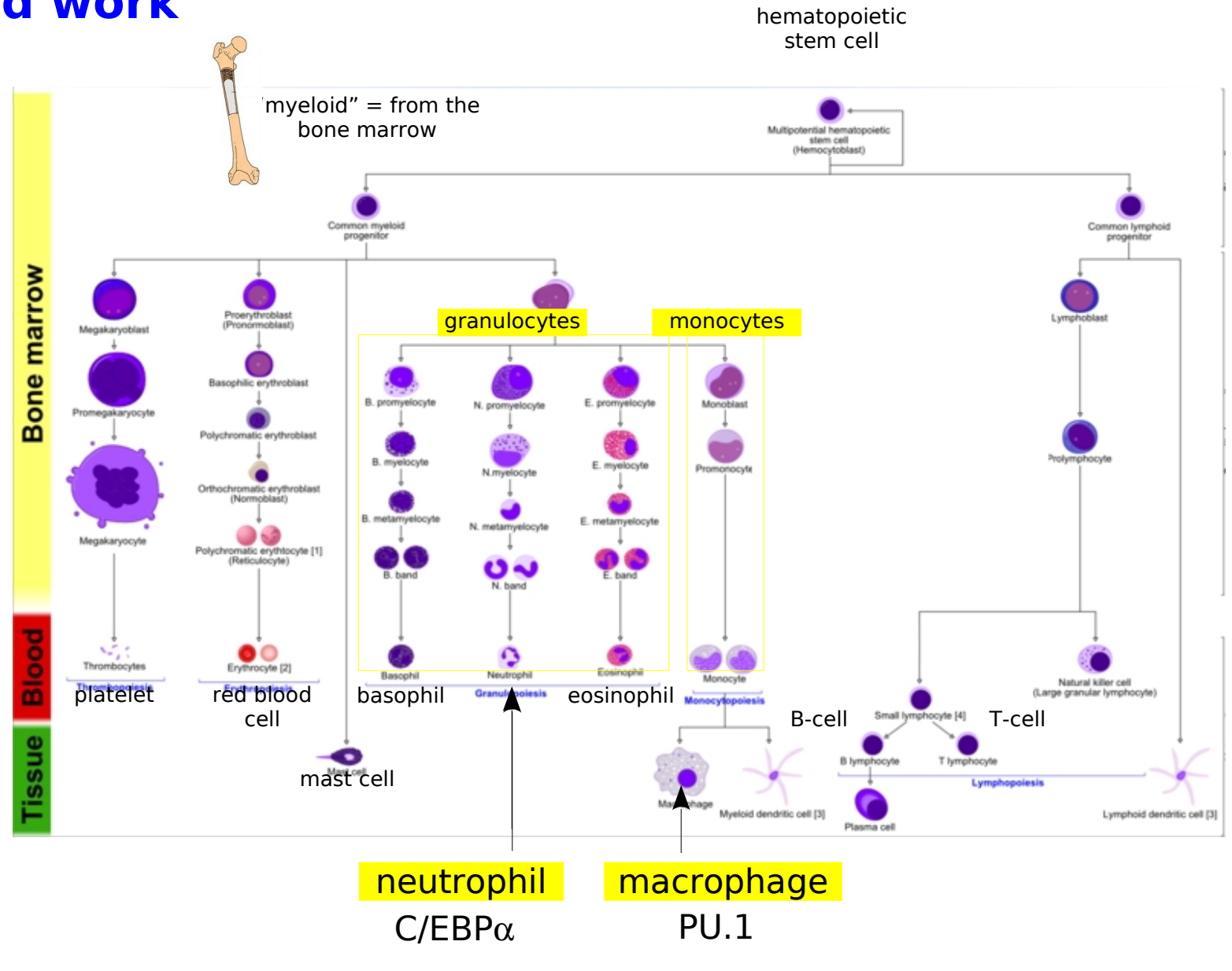


gene regulation function

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

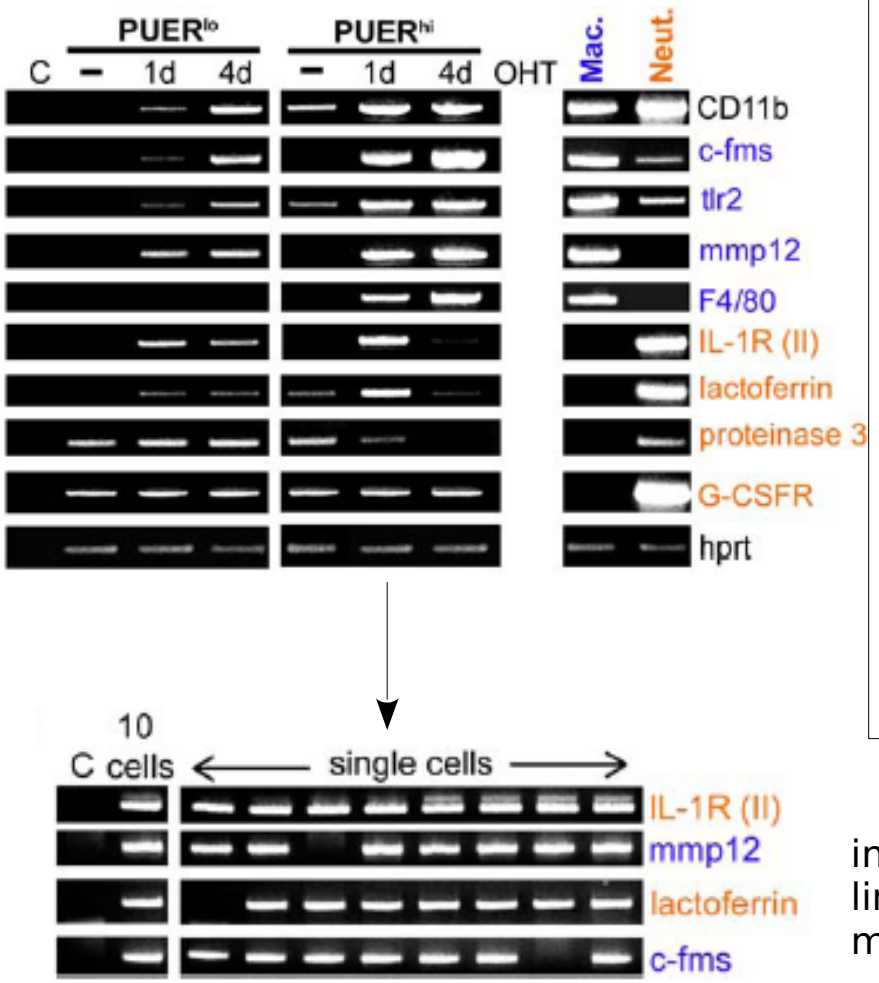
$$c_k = \left(\sum_{1 \le i_1 < \cdots < i_k \le n} \left(\prod_{j=1}^{k} \kappa_{i_j} \omega_{i_j, \{i_{j+1}, \cdots, i_k\}} \right) \right) (K_{1,\emptyset})^k$$

blood work

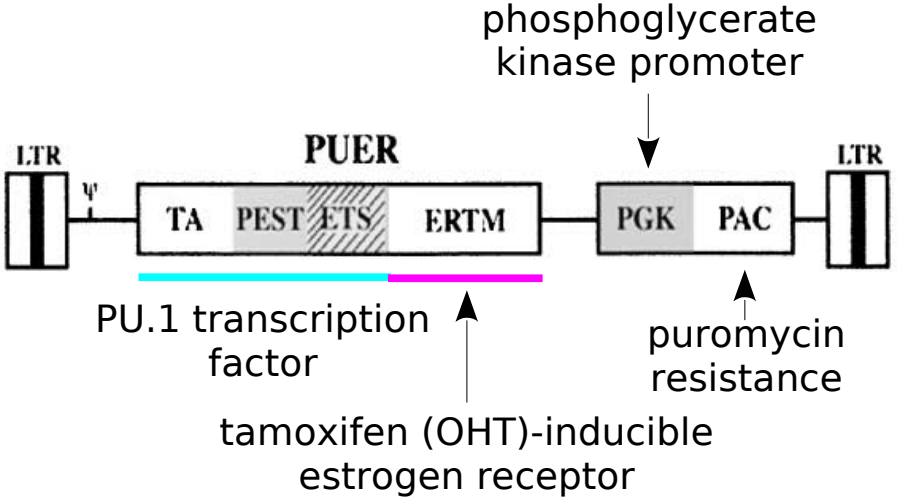


Laszlo, Spooner, Warmflash, Lancki, Lee, Sciammas, Gantner, Dinner, Singh, "*Multilineage transcriptional priming and determination of alternate hematopoietic cell fates*", Cell **126**:755-66 2006

neutrophil/macrophage differentiation – experiment



single-cell multiplex RT-PCR



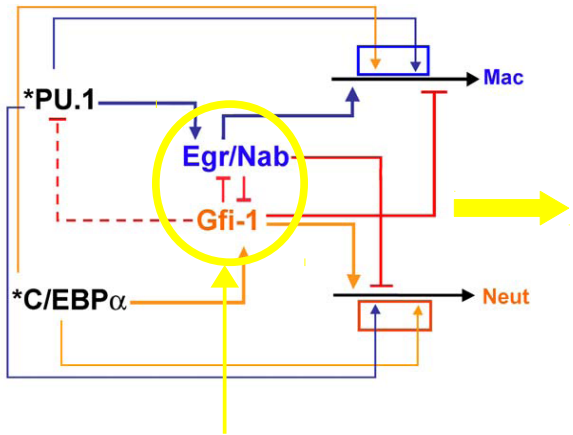
PU.1-/- myeloid progenitor cells were transformed and clones selected which expressed PUER at low ($PUER^{lo}$) and high ($PUER^{hi}$) levels

individual $PUER^{hi}$ cells exhibit a "mixed lineage" of macrophage and neutrophil markers 1 day after OHT treatment

"transcriptional priming"

how does mixed lineage transcriptional priming arise?

neutrophil/macrophage differentiation – theory

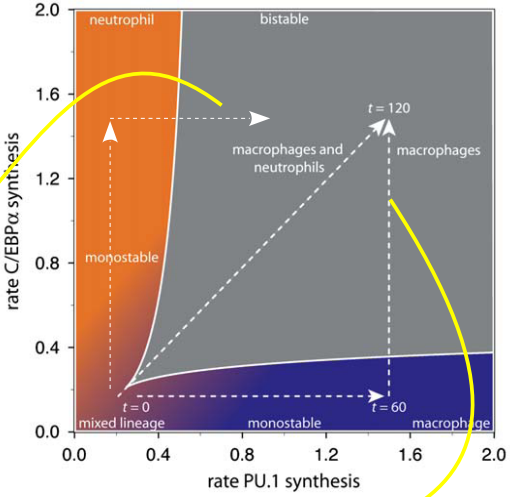


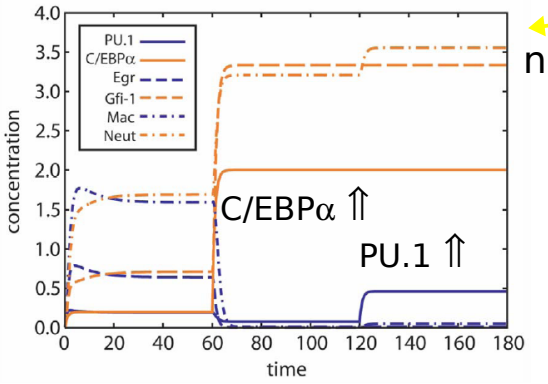
mutual repression

nonlinear dynamical system

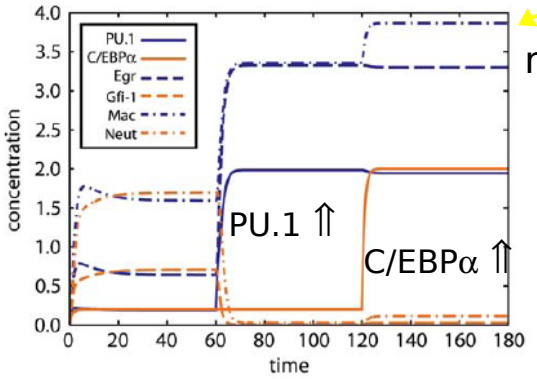
$$\frac{dx_1}{dt} = f_1(x_1, x_2, \cdots, x_n; k_1, \cdots, k_m)$$
$$\frac{dx_2}{dt} = f_2(x_1, x_2, \cdots, x_n; k_1, \cdots, k_m)$$
$$\vdots \quad \vdots \quad \vdots$$
$$\frac{dx_n}{dt} = f_n(x_1, x_2, \cdots, x_n; k_1, \cdots, k_m)$$

BISTABILITY





neutrophil

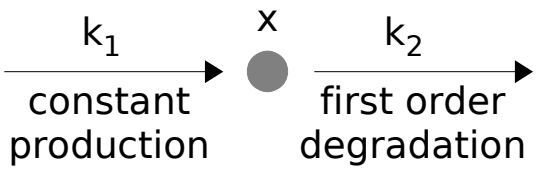


macrophage

primary determinants can remain high despite lineage resolution

Strogatz, **Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry and Engineering**, Westview Press, 2001

quantitative to qualitative dynamics



$$\frac{dx}{dt} = k_1 - k_2 x$$

from a quantitative analytical solution

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

to a qualitative geometric perspective