Design principles of complex cell-fate decision networks: Examples in development & cancer – *Part 1*

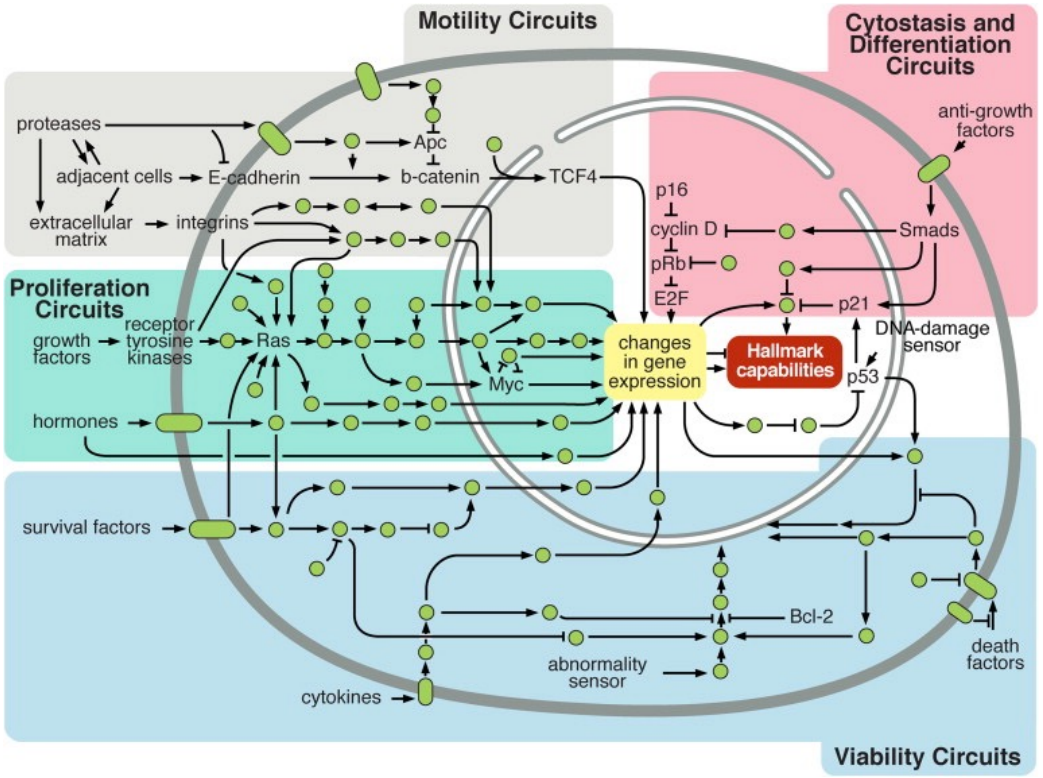
Mohit Kumar Jolly, PhD
Cancer Systems Biology Laboratory
Associate Professor,
Department of Bioengineering,
Indian Institute of Science (IISc), Bangalore, India
mkjolly@iisc.ac.in

Editor-in-Chief, NPJ Systems Biology & Applications

Workshop on Flags, Landscapes and Signals | IMSc Chennai

mkjolly15

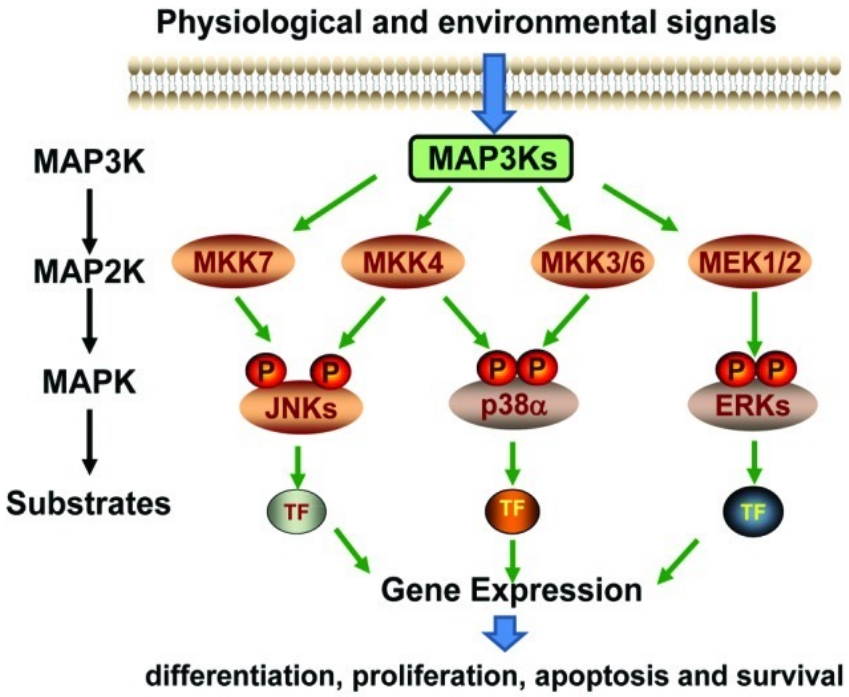
Cellular decision-making



- Cells receive diverse biophysical/chemical signals varying in (x, t).
- Cells in a population can respond differently to the same signals.
- Cellular decision-making is driven by interconnected complex networks.

Hanahan & Weinberg, Cell 2011

How do we understand cellular decision-making?



What information does it lack?

- Timescale(s)
- Strength of regulation
- Direct/indirect
- Spatial scale(s)
- Nonlinearity of interaction
- Combinatorial effects

Assumptions are implicit or hidden in a “black box” and can have unknown logical consequences

Wang *et al.* Comm Integr Biol 2012

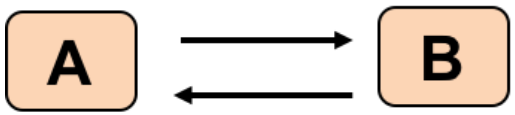
Outline for today

- Dynamics of simplest 2-node decision-making network motifs
- Dynamics of 3-node, 4-node networks, and its implications in T-cell differentiation
- Impact of embedding 2-node, 3-node network motifs in larger networks

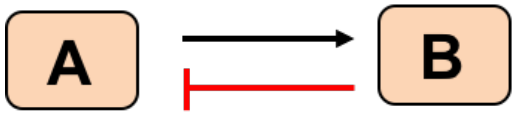
Simplest two-gene circuits

Consider two transcription factors A and B that regulate each other:

Double negative feedback loop



Double positive feedback loop



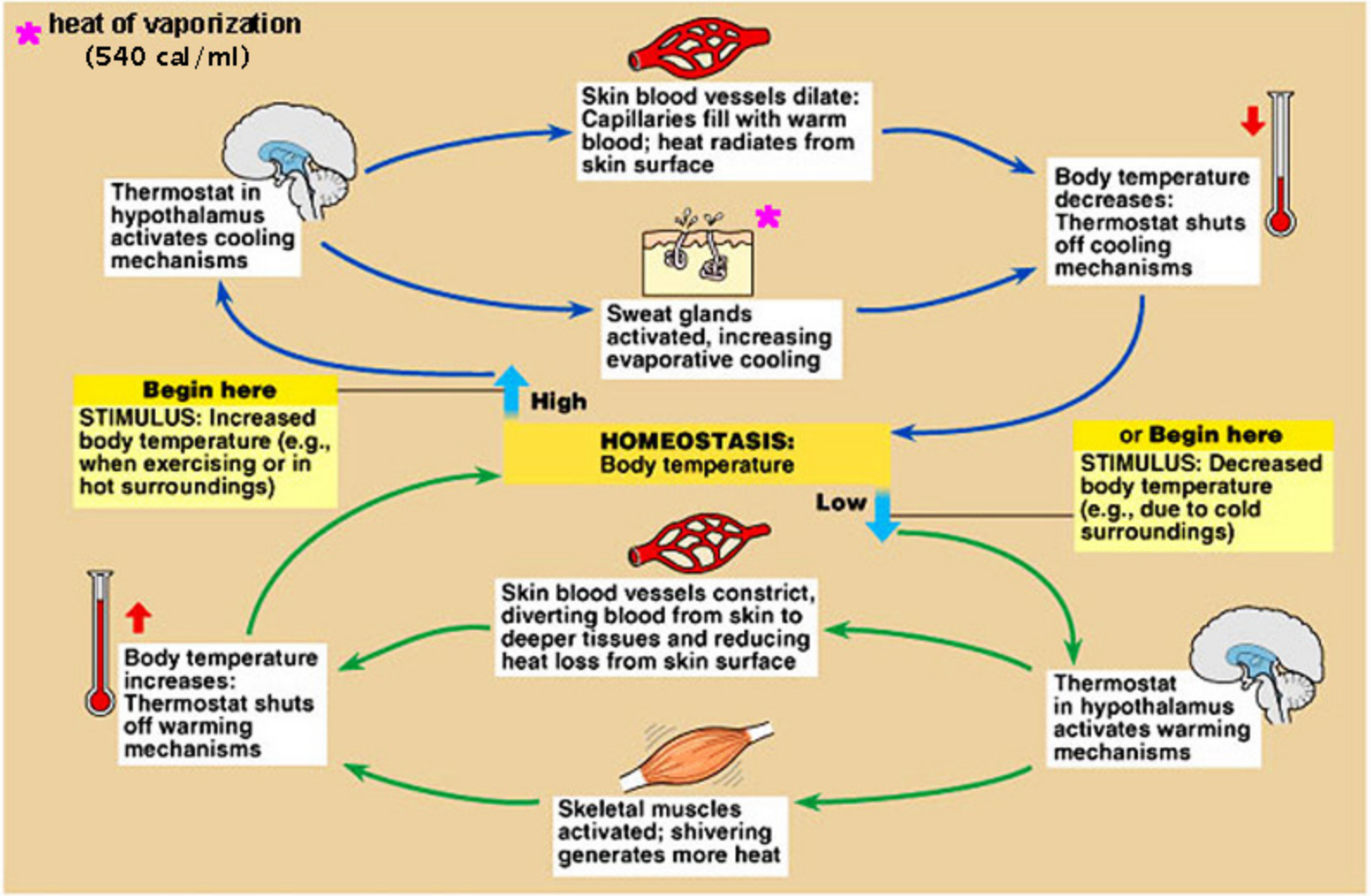
Negative-positive feedback loop

Let us first understand the basic 'design principles' of a negative and a positive feedback loop:

- How does the system work?
- Why is the system designed the way it is?

"Nothing in biology makes sense except in the light of evolution"

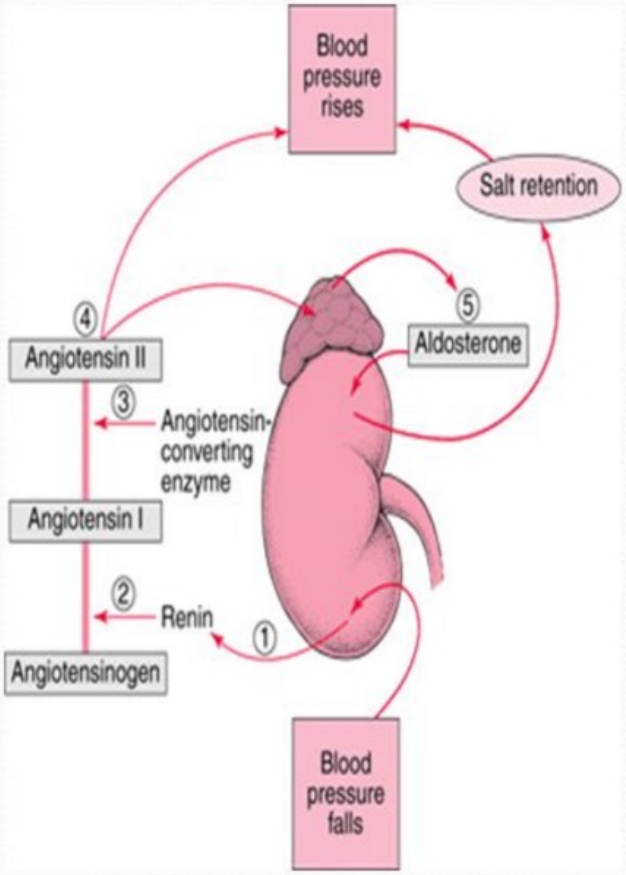
Positive-negative feedback loop: body temperature



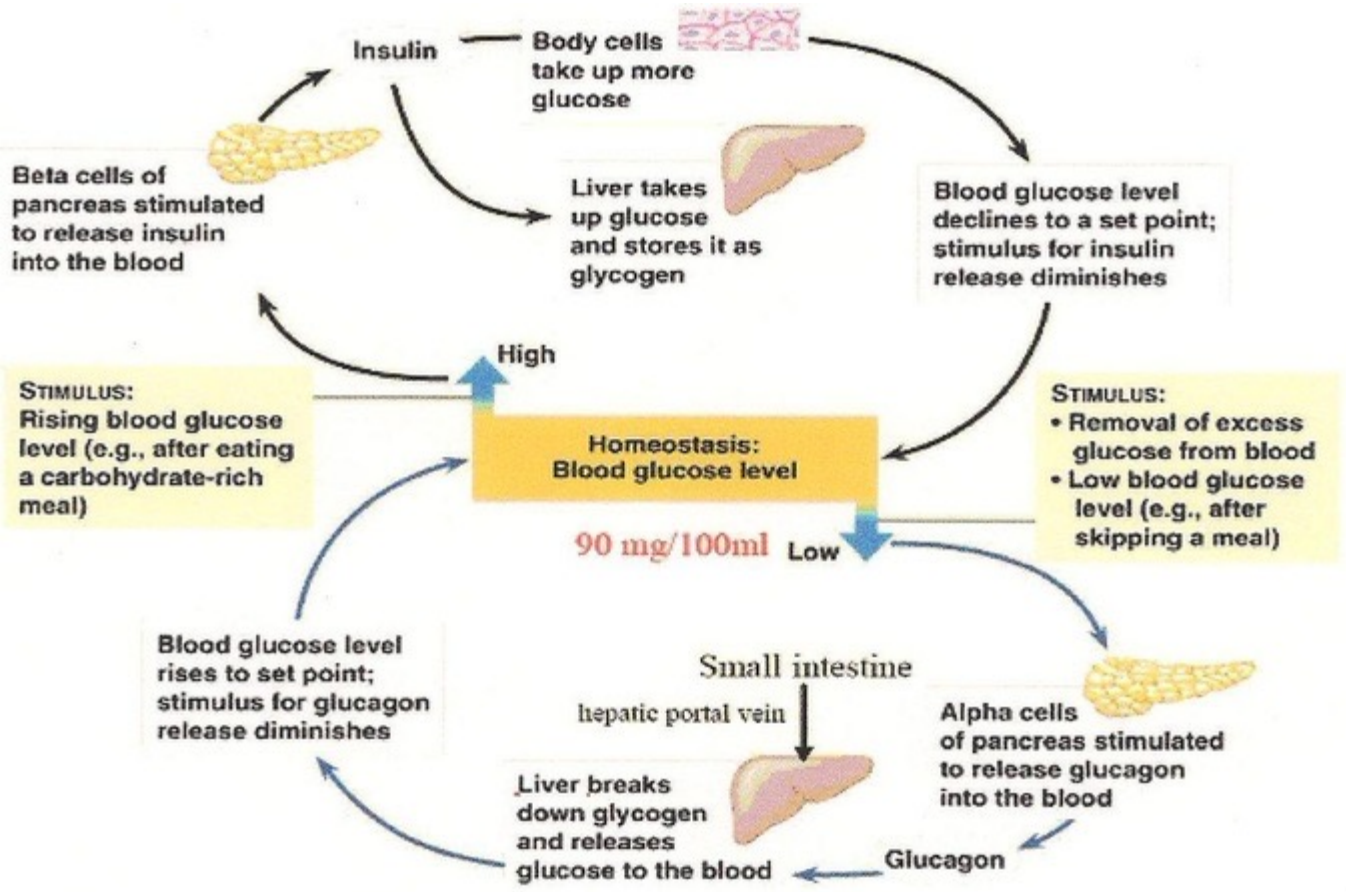
Positive-negative feedback loop: blood pressure

2. Blood Pressure

- Blood volume is regulated by the hormone *aldosterone*
- Aldosterone affects the rate of sodium ion reabsorption, which in turn affects the rate of water reabsorption
- Increased aldosterone ➔ increased water reabsorption ➔ higher blood pressure
- Decreased aldosterone ➔ decreased water reabsorption ➔ lower blood pressure



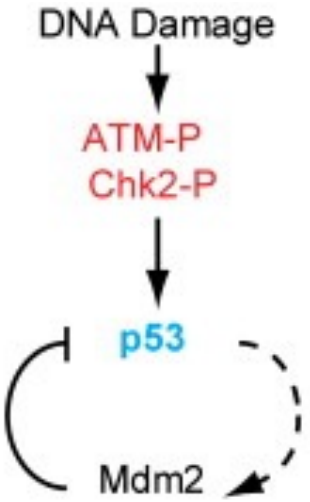
Positive-negative feedback loop: glucose levels



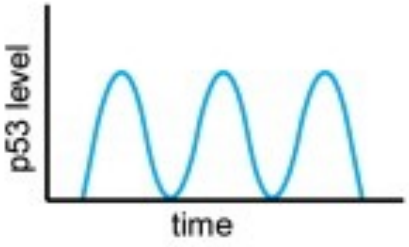
http://www.psychology4a.com/uploads/3/0/2/1/30214259/1203971.jpg?674

Why are vital signs mostly regulated by negative feedback loops?

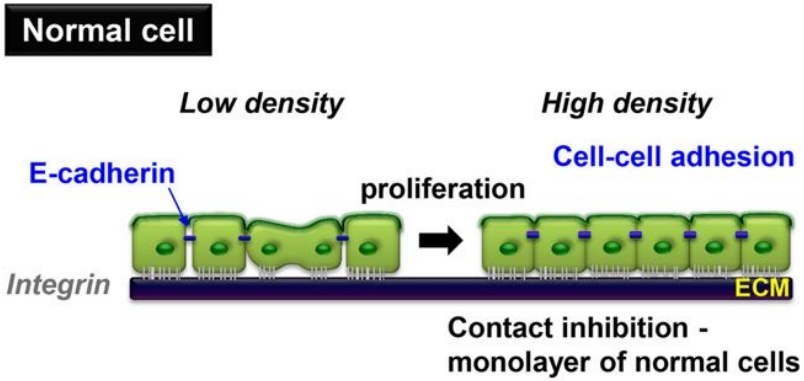
(Broken) Negative feedback loops in cancer

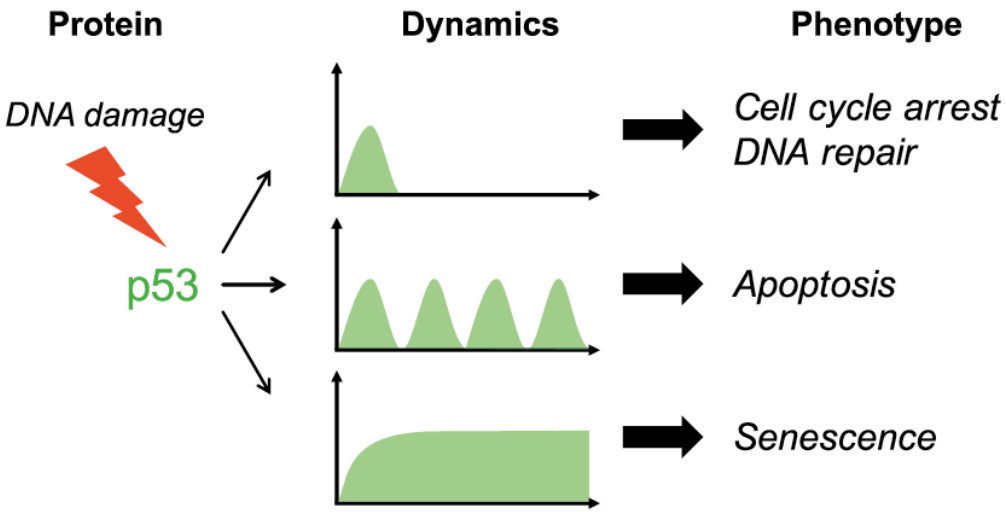


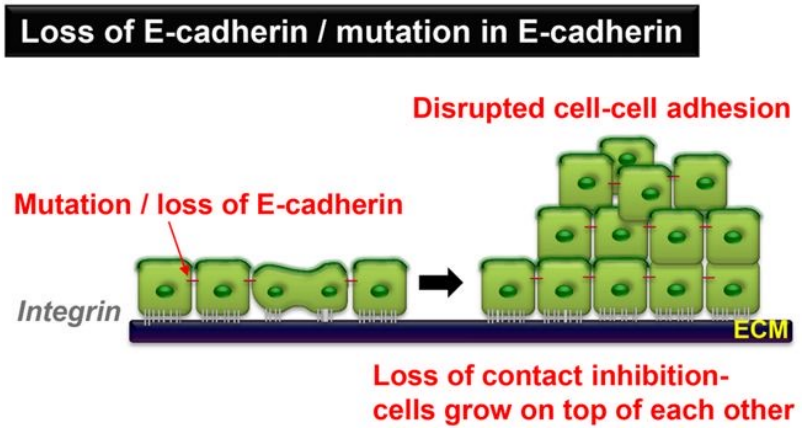
Intracellular NFL involves p53 - most common mutated gene in cancer



Intercellular NFL - Contact inhibition of proliferation

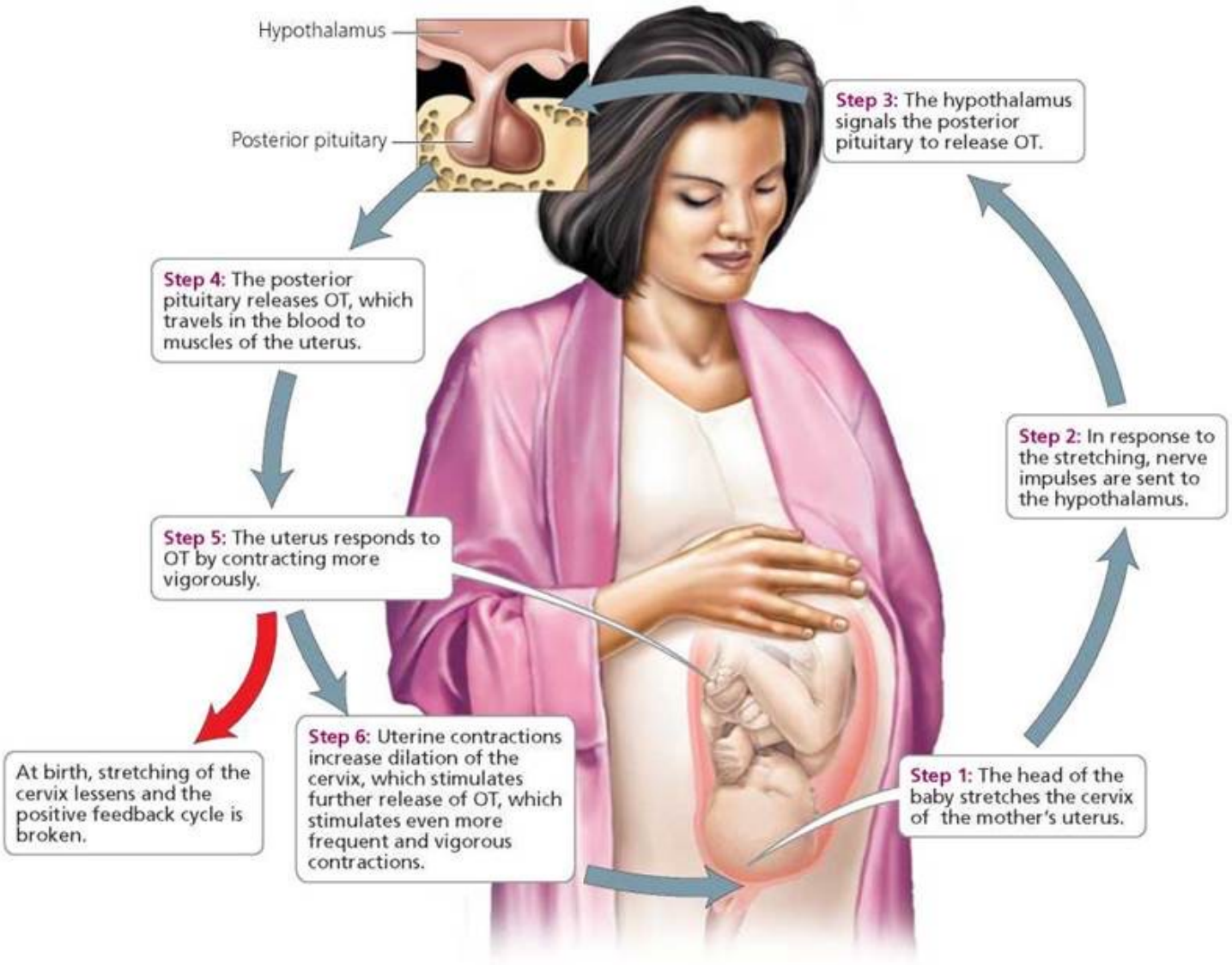






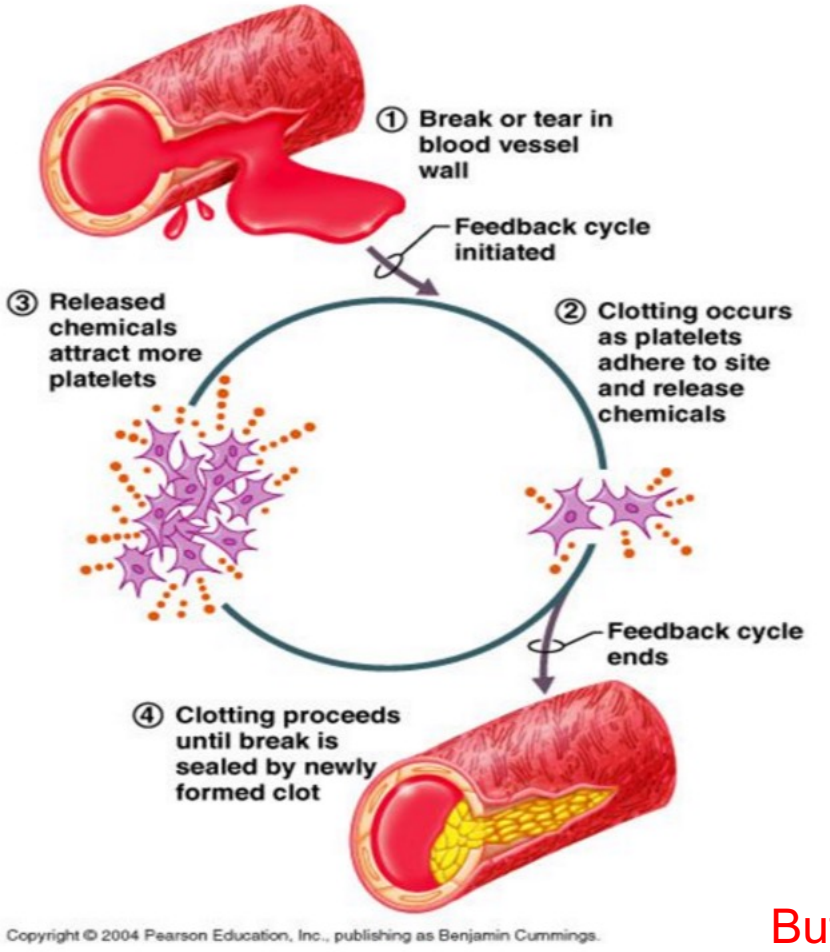
Purvis *et al.* Science 2012
Magi *et al.* Curr Opin Sys Biol 2018

Amplification of response: Biological examples



http://schoolbag.info/biology/humans/12.html

Amplification of response: Biological examples



Examples:

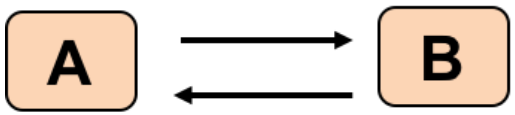
- **Blood clotting:**
  - Blood clotting is a normal response to a break in the lining of a blood vessel
  - 1. Once vessel damaged has occurred
  - 2. Blood elements called **platelets** immediately begin to cling to the injured site
  - 3. Platelets release chemical that attract more platelets
  - 4. This rapidly growing pileup of platelets initiates the sequence of events that finally forms a **clot**

But can over-amplification be dangerous?
Unwanted clots – cause of heart attacks

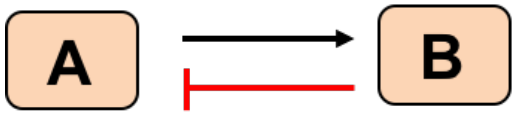
Simplest two-gene circuits

Consider two transcription factors A and B that regulate each other:

Double negative feedback loop



Double positive feedback loop



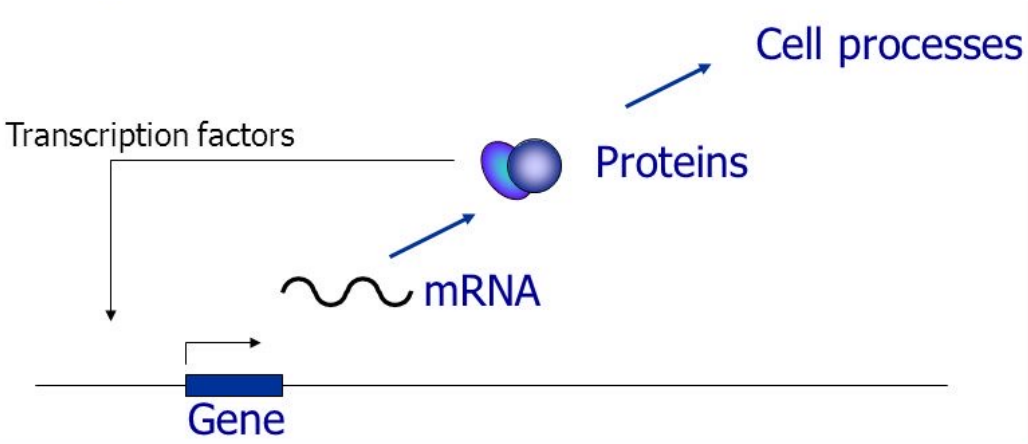
Negative-positive feedback loop

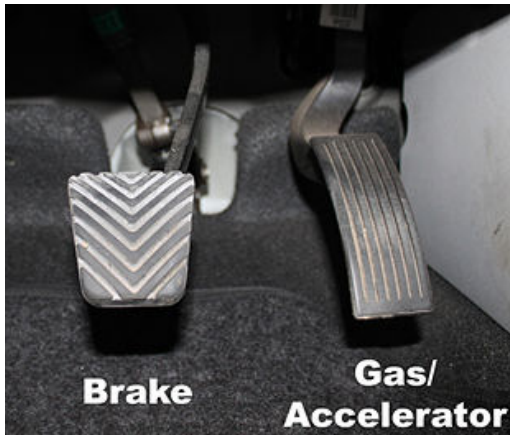
Let us first understand the basic 'design principles' of a negative and a positive feedback loop:

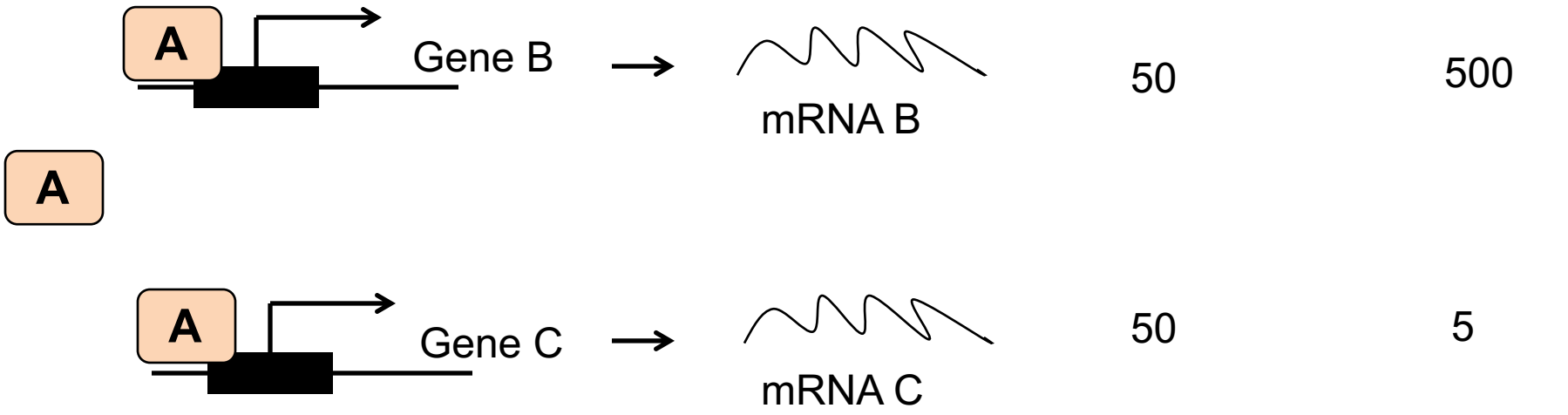
- How does the system work?
- Why is the system designed the way it is?

"Nothing in biology makes sense except in the light of evolution"

Transcription factors: Activators and inhibitors

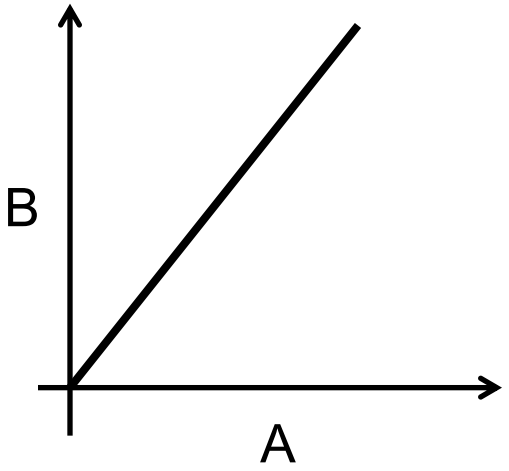


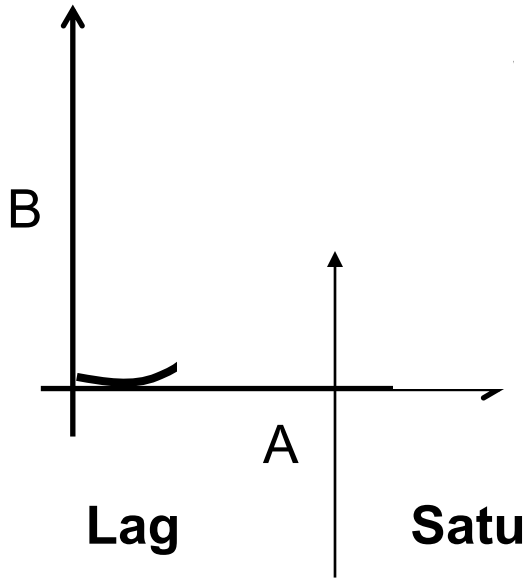




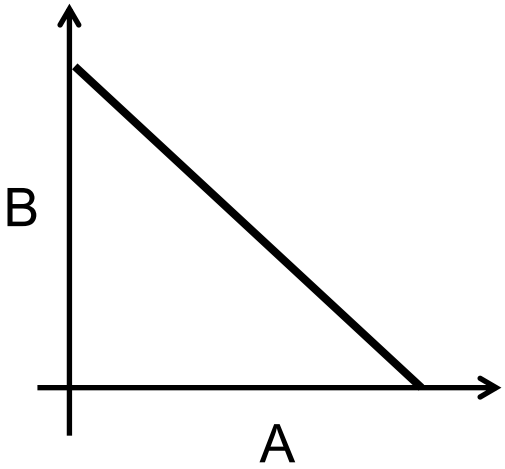
All units in molecules/unit time

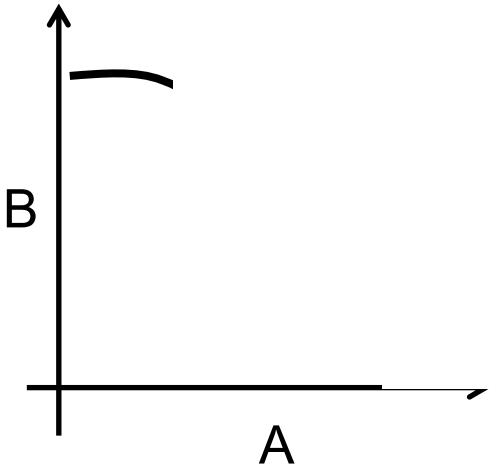
Transcriptional regulation





Slope of this curve determines how strongly A affects B





Transcriptional regulation

$$\text{gene expression} = f([\text{TF}])$$

$$\text{TF} + \text{DNA} \underset{K_r}{\overset{K_f}{\rightleftarrows}} \text{DNA-TF} \implies A + B \underset{K_r}{\overset{K_f}{\rightleftarrows}} AB$$

$$\begin{aligned}
\frac{d[AB]}{dt} &= K_f[A][B] - K_r[AB] \\
\frac{d[A]}{dt} &= -K_f[A][B] + K_r[AB] \\
\frac{d[B]}{dt} &= -K_f[A][B] + K_r[AB]
\end{aligned}$$

DNA-protein binding happens much faster (seconds) as compared to actual transcription (minutes/hours)

$$\begin{aligned}
&\frac{d[AB]}{dt} = 0 \\
\implies\ & k_f[A][B] = K_r[AB] \\
\implies\ & K_{eq} \equiv \frac{K_f}{K_r} = \frac{[AB]}{[A][B]}
\end{aligned}$$

Transcriptional regulation

$$\text{Fraction of Bound DNA} = \theta \equiv \frac{[AB]}{[B_{total}]} = \frac{[AB]}{[B]+[AB]}$$

$$\overset{Eq}{=} \frac{[AB]}{\frac{[AB]}{[A]K_{eq}}+[AB]} = \frac{1}{\frac{1}{[A]K_{eq}}+1}$$

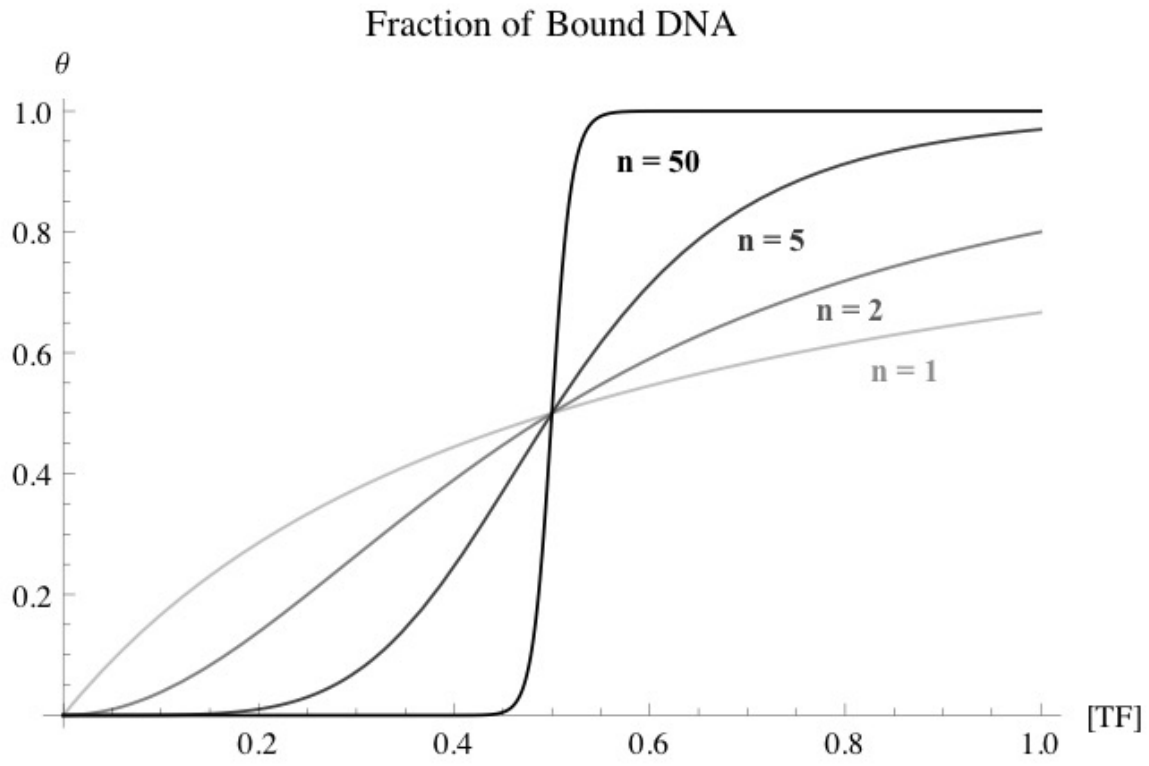
$$= \frac{1}{\frac{1+[A]K_{eq}}{[A]K_{eq}}} = \frac{[A]K_{eq}}{1+[A]K_{eq}}$$

$$= \frac{[A]}{\frac{1}{K_{eq}}+[A]} \overset{K_D \equiv \frac{1}{K_{eq}}}{=} \frac{[A]}{K_D+[A]} = \frac{[\text{TF}]}{K_D+[\text{TF}]}$$

What's the underlying assumption in the following case?

$$n\text{TF}+\text{DNA} \underset{K_r}{\overset{K_f}{\rightleftarrows}} \text{DNA-nTF} \implies nA+B \underset{K_r}{\overset{K_f}{\rightleftarrows}} nAB$$

Transcriptional regulation



Construction of a genetic toggle switch in *Escherichia coli*

Timothy S. Gardner*†, Charles R. Cantor* & James J. Collins*†

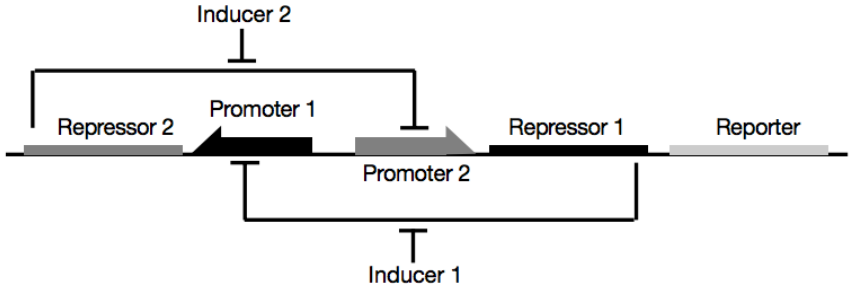


Figure 1 Toggle switch design. Repressor 1 inhibits transcription from Promoter 1 and is induced by Inducer 1. Repressor 2 inhibits transcription from Promoter 2 and is induced by Inducer 2.

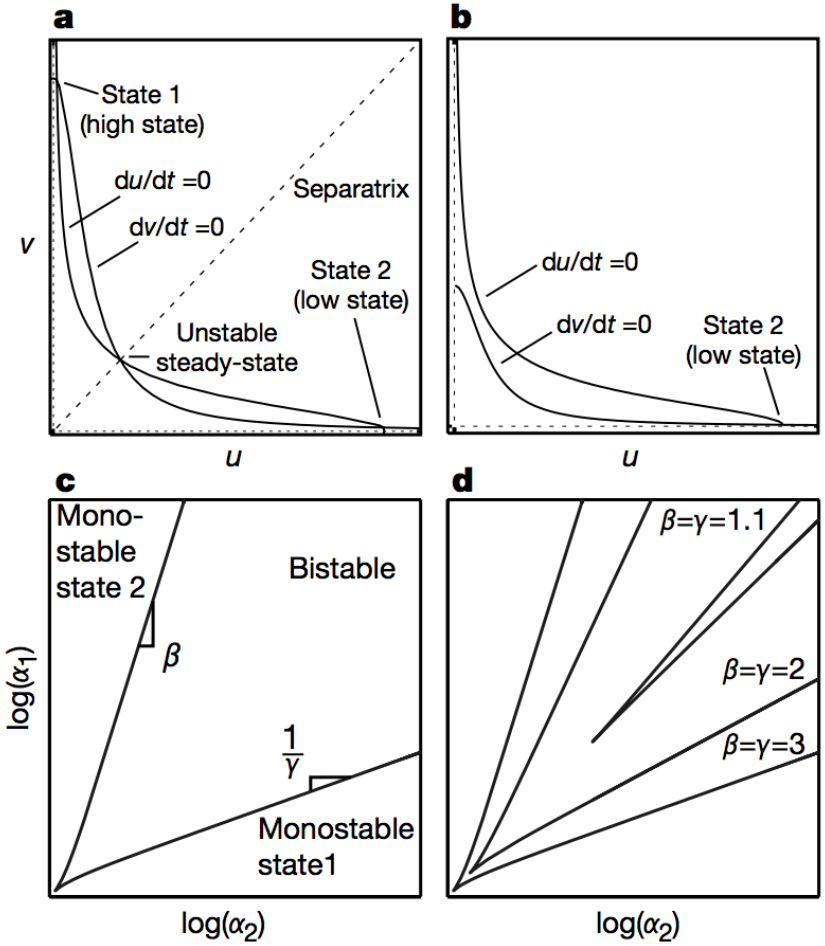
The behaviour of the toggle switch and the conditions for bistability can be understood using the following dimensionless model for the network:

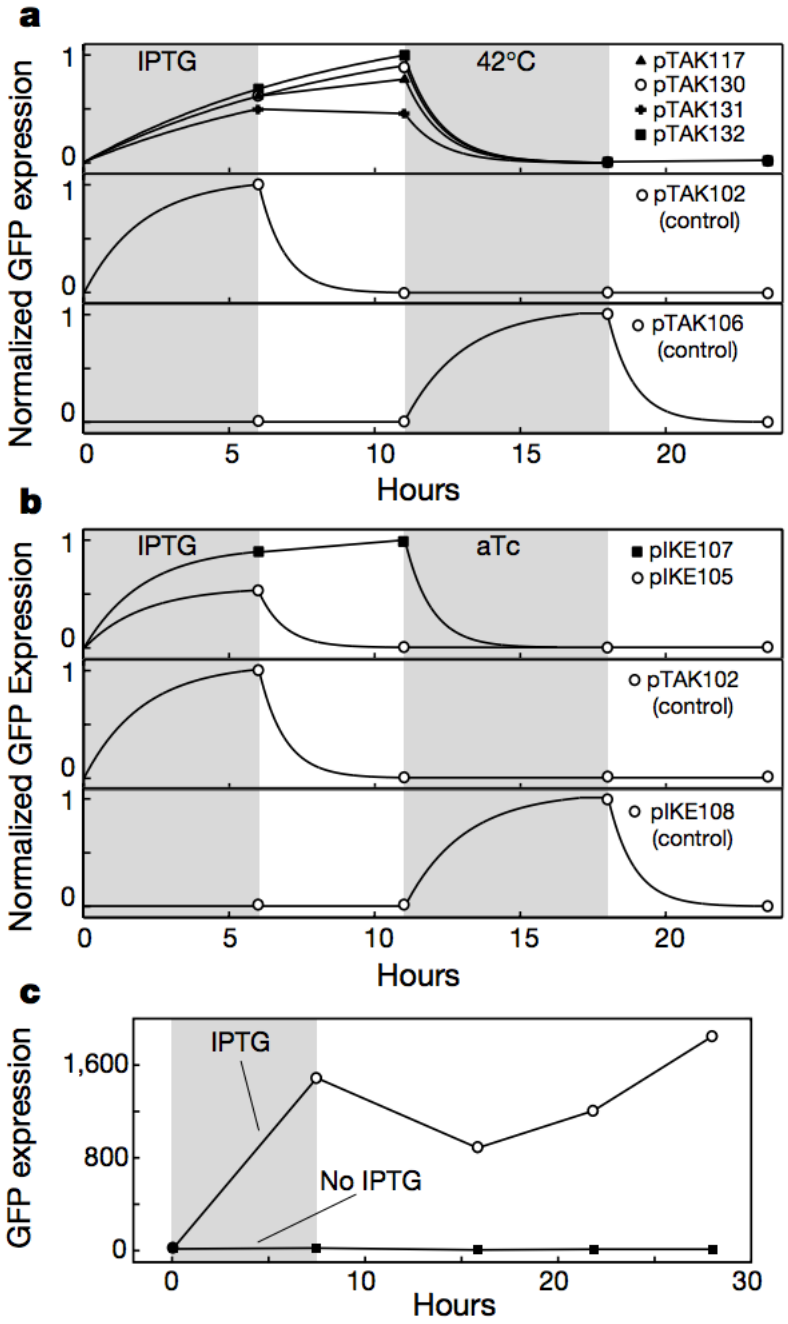
$$\frac{du}{dt} = \frac{\alpha_1}{1 + v^{\beta}} - u \tag{1a}$$

$$\frac{dv}{dt} = \frac{\alpha_2}{1 + u^{\gamma}} - v \tag{1b}$$

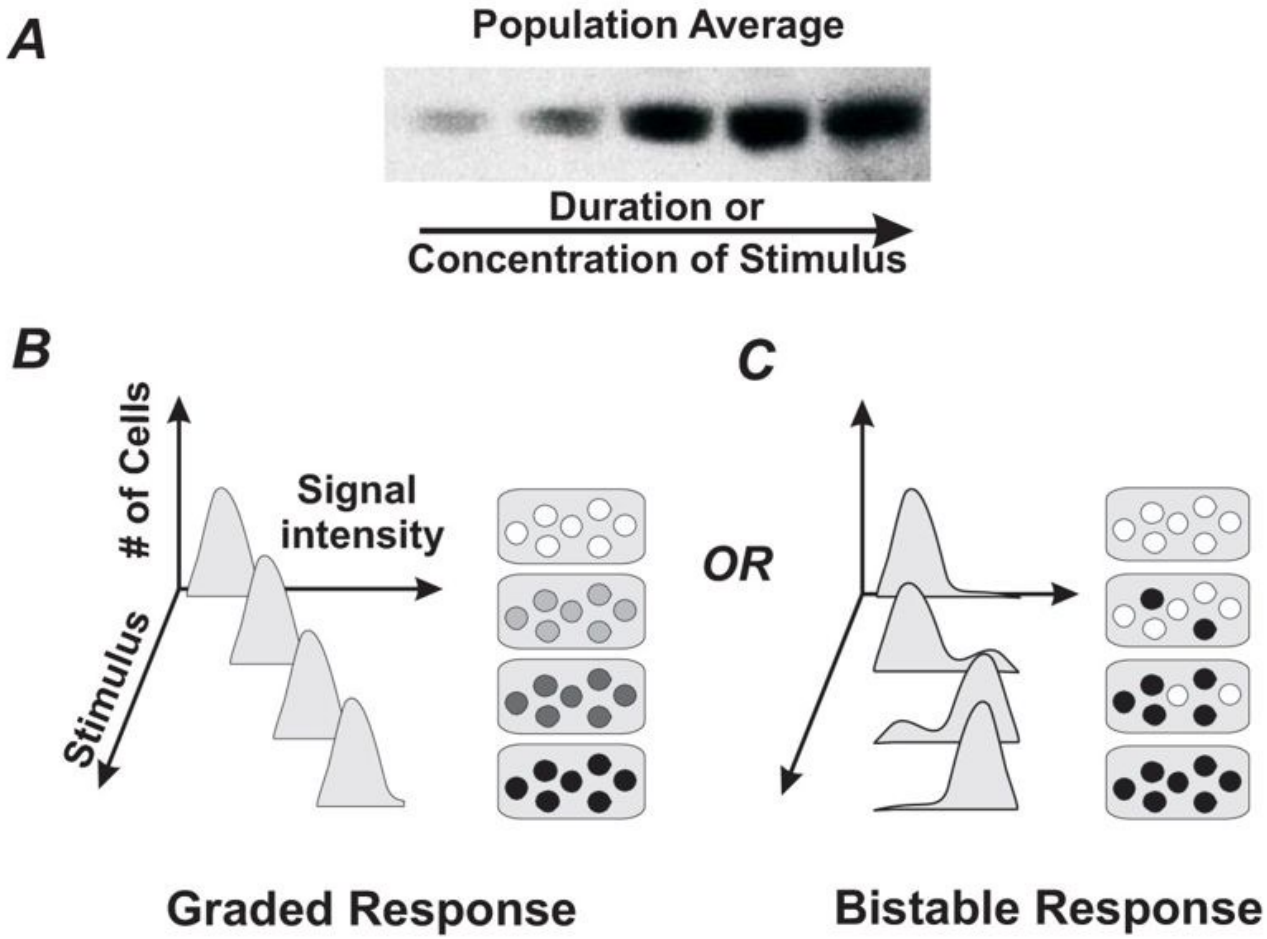
where u is the concentration of repressor 1, v is the concentration of repressor 2, α_1 is the effective rate of synthesis of repressor 1, α_2 is the effective rate of synthesis of repressor 2, β is the cooperativity of repression of promoter 2 and γ is the cooperativity of repression of promoter 1. The above model is derived from a biochemical rate equation formulation of gene expression[24–27]. The final form of the toggle equations preserves the two most fundamental aspects of the network: cooperative repression of constitutively transcribed promoters (the first term in each equation), and degradation/dilution of the repressors (the second term in each equation).

The parameters α_1 and α_2 are lumped parameters that describe the net effect of RNA polymerase binding, open-complex formation, transcript elongation, transcript termination, repressor binding, ribosome binding and polypeptide elongation. The cooperativity described by β and γ can arise from the multimerization of the repressor proteins and the cooperative binding of repressor multimers to multiple operator sites in the promoter. An additional modification to equation (1) is needed to describe induction of the repressors (Fig. 5).




a
IPTG
42°C
pTAK117
pTAK130
pTAK131
pTAK132
pTAK102
(control)
pTAK106
(control)
Normalized GFP expression
Hours
b
IPTG
aTc
pIKE107
pIKE105
pTAK102
(control)
pIKE108
(control)
Normalized GFP Expression
Hours
c
IPTG
No IPTG
GFP expression
Hours

Bistability: Digital vs. analog response



Chang *et al.* BMC Cell Biol 2006

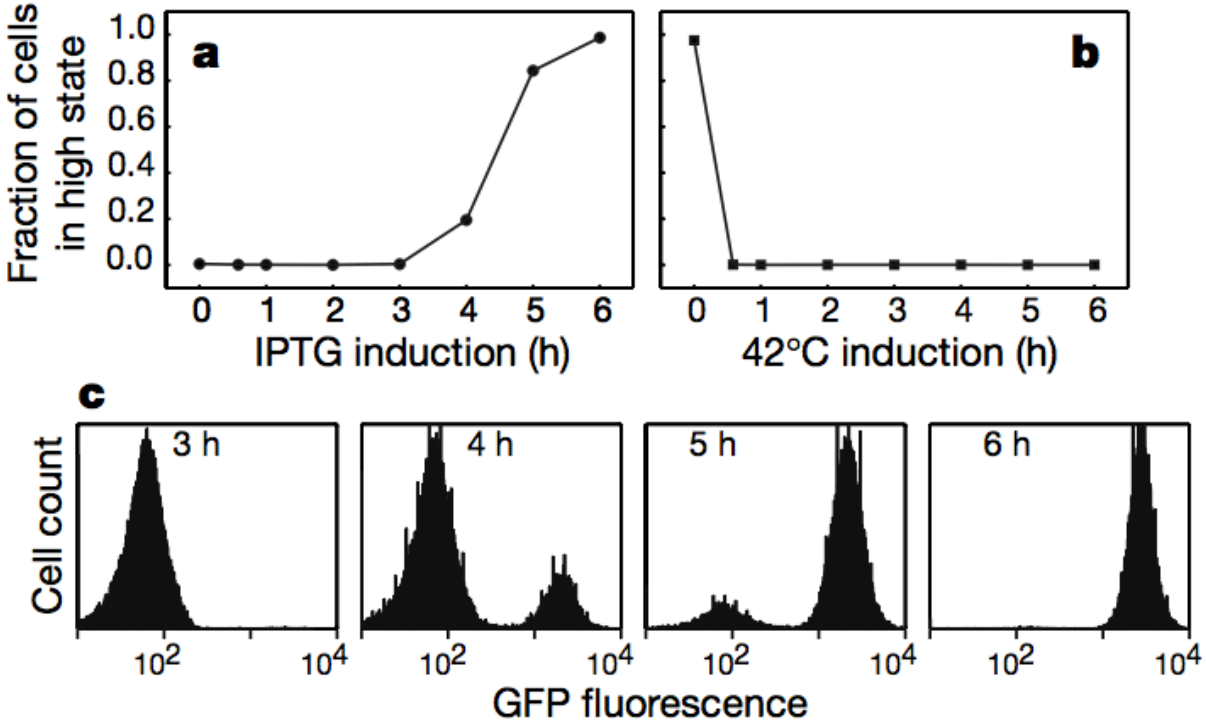
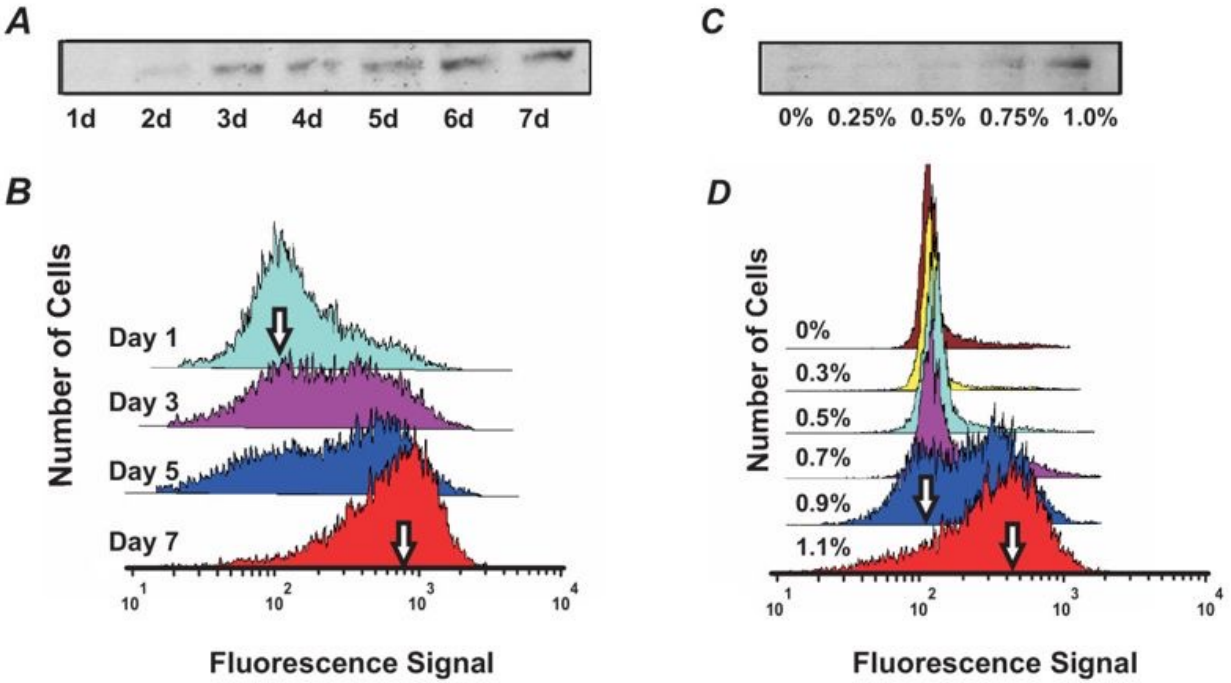
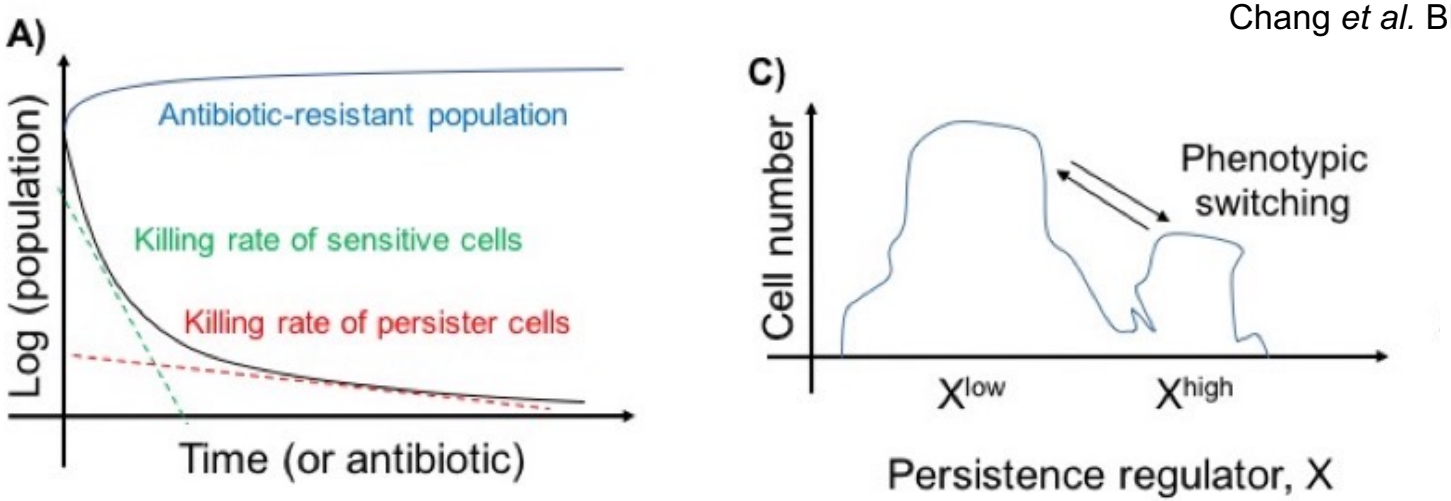


Figure 6 pTAK117 switching time. **a**, **b**, The fraction of cells in the high state is plotted as a function of the induction time. Cells were divided between high and low states as in Fig. 5c. **c**, Switching of pTAK117 cells from the low to the high state by IPTG induction. The cell population is illustrated at four time points. Cells begin switching between 3 and 4 h as shown by the appearance of a bimodal distribution. The switching is complete by 6 h.

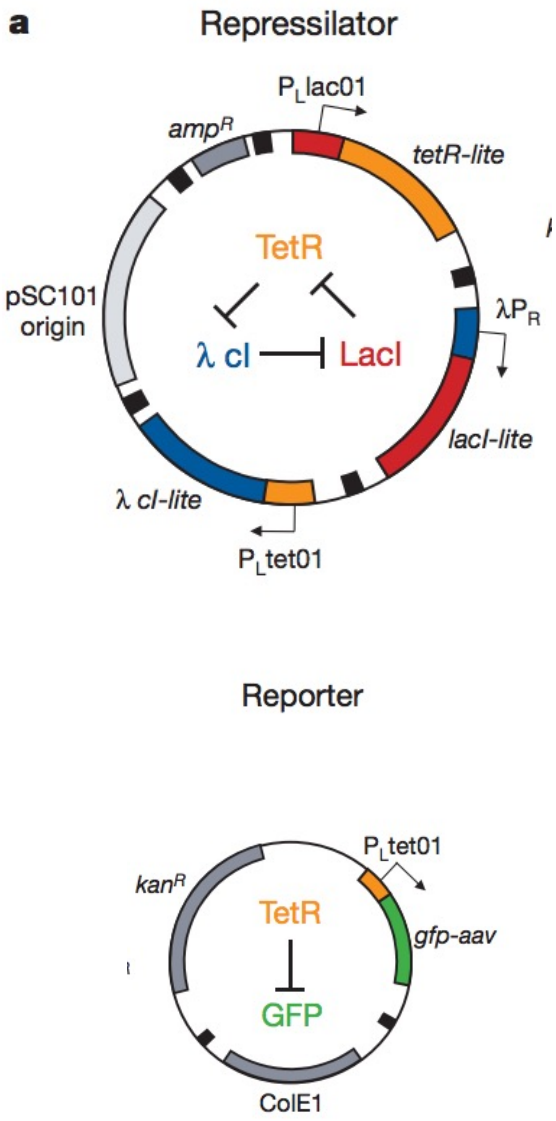
Consequences of bistability in drug resistance



Chang *et al.* BMC Cell Biol 2006



Jolly *et al.* Front Oncol 2018



Deterministic, continuous approximation

Three repressor-protein concentrations, p_i, and their corresponding mRNA concentrations, m_i (where i is *lacI*, *tetR* or *cI*) were treated as continuous dynamical variables. Each of these six molecular species participates in transcription, translation and degradation reactions. Here we consider only the symmetrical case in which all three repressors are identical except for their DNA-binding specificities. The kinetics of the system are determined by six coupled first-order differential equations:

$$\frac{dm_i}{dt} = -m_i + \frac{\alpha}{(1+p_j^n)} + \alpha_0 \qquad \begin{pmatrix} i = lacI, tetR, cI \\ j = cI, lacI, tetR \end{pmatrix}$$

$$\frac{dp_i}{dt} = -\beta(p_i - m_i)$$

where the number of protein copies per cell produced from a given promoter type during continuous growth is α_0 in the presence of saturating amounts of repressor (owing to the 'leakiness' of the promoter), and $\alpha + \alpha_0$ in its absence; β denotes the ratio of the protein decay rate to the mRNA decay rate; and n is a Hill coefficient. Time is rescaled in units of the mRNA lifetime; protein concentrations are written in units of K_M, the number of repressors necessary to half-maximally repress a promoter; and mRNA concentrations are rescaled by their translation efficiency, the average number of proteins produced per mRNA molecule. The numerical solution of the model shown in Fig. 1c

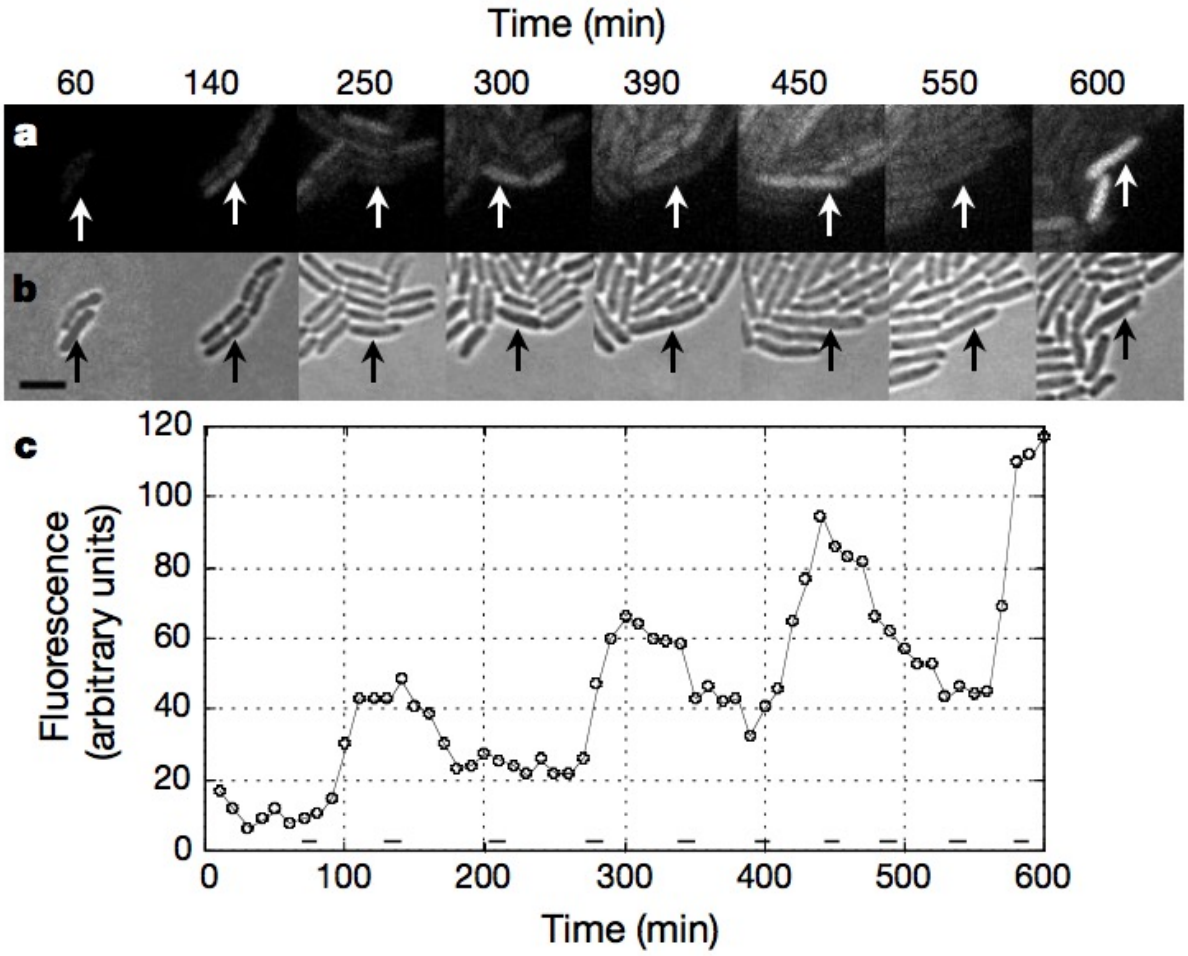
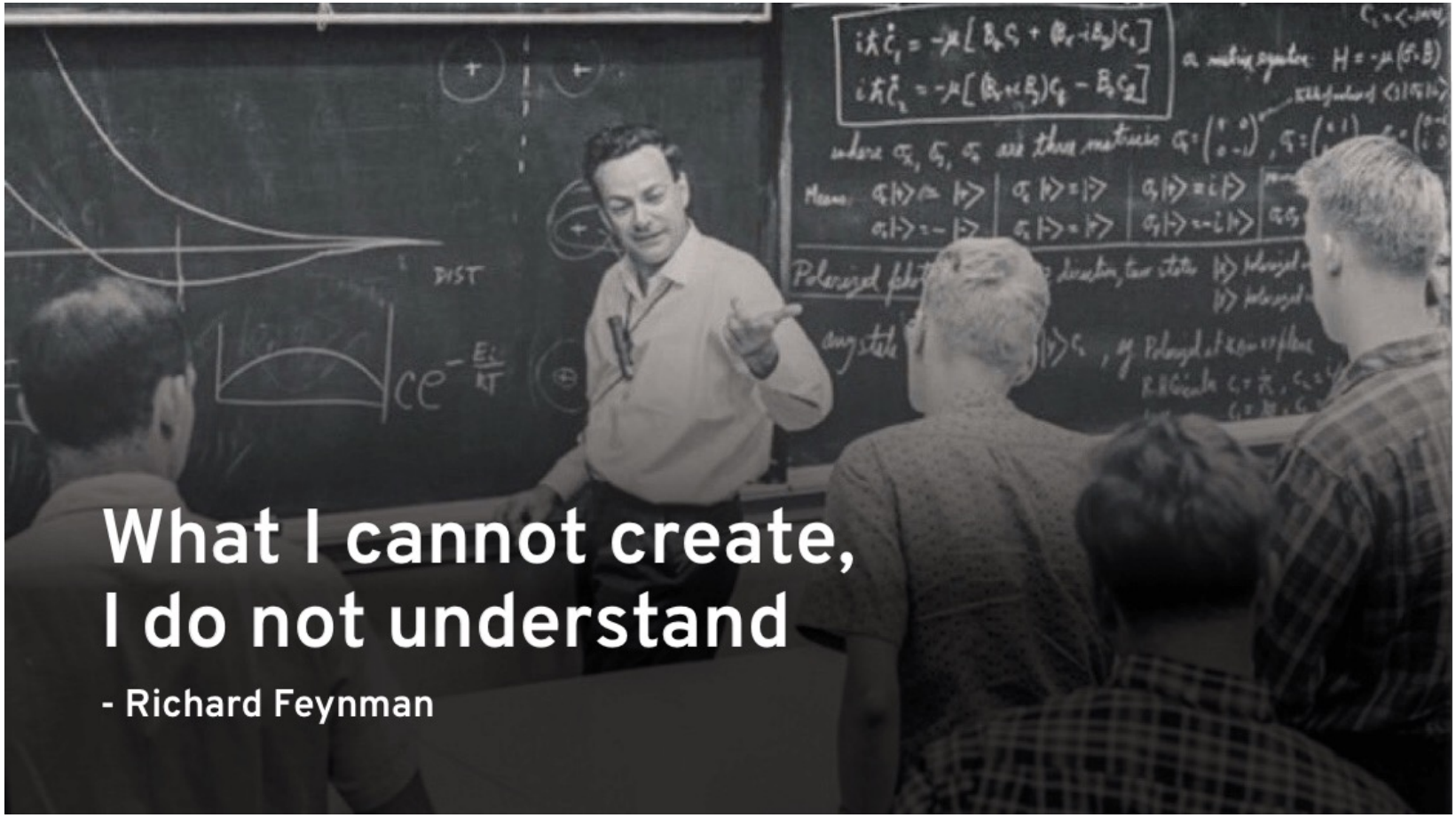


Figure 2 Repressilation in living bacteria. **a**, **b**, The growth and timecourse of GFP expression for a single cell of *E. coli* host strain MC4100 containing the repressilator plasmids (Fig. 1a). Snapshots of a growing microcolony were taken periodically both in fluorescence (**a**) and bright-field (**b**). **c**, The pictures in **a** and **b** correspond to peaks and troughs in the timecourse of GFP fluorescence density of the selected cell. Scale bar, 4 μm. Bars at the bottom of **c** indicate the timing of septation events, as estimated from bright-field images.



Beginning of synthetic biology: two Nature papers in 2000

Construction of a genetic toggle switch in *Escherichia coli*

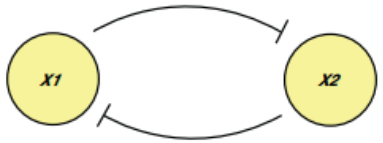
Timothy S. Gardner*†, Charles R. Cantor* & James J. Collins*†

A synthetic oscillatory network of transcriptional regulators

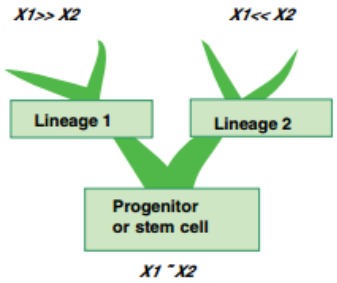
Michael B. Elowitz & Stanislas Leibler

Toggle switch: a ubiquitous network motif

(a)



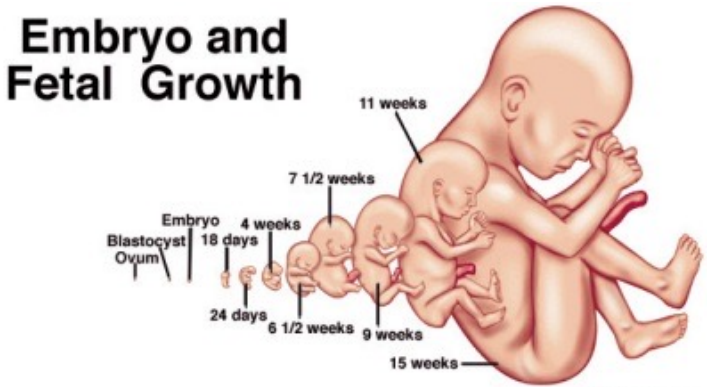
(b)

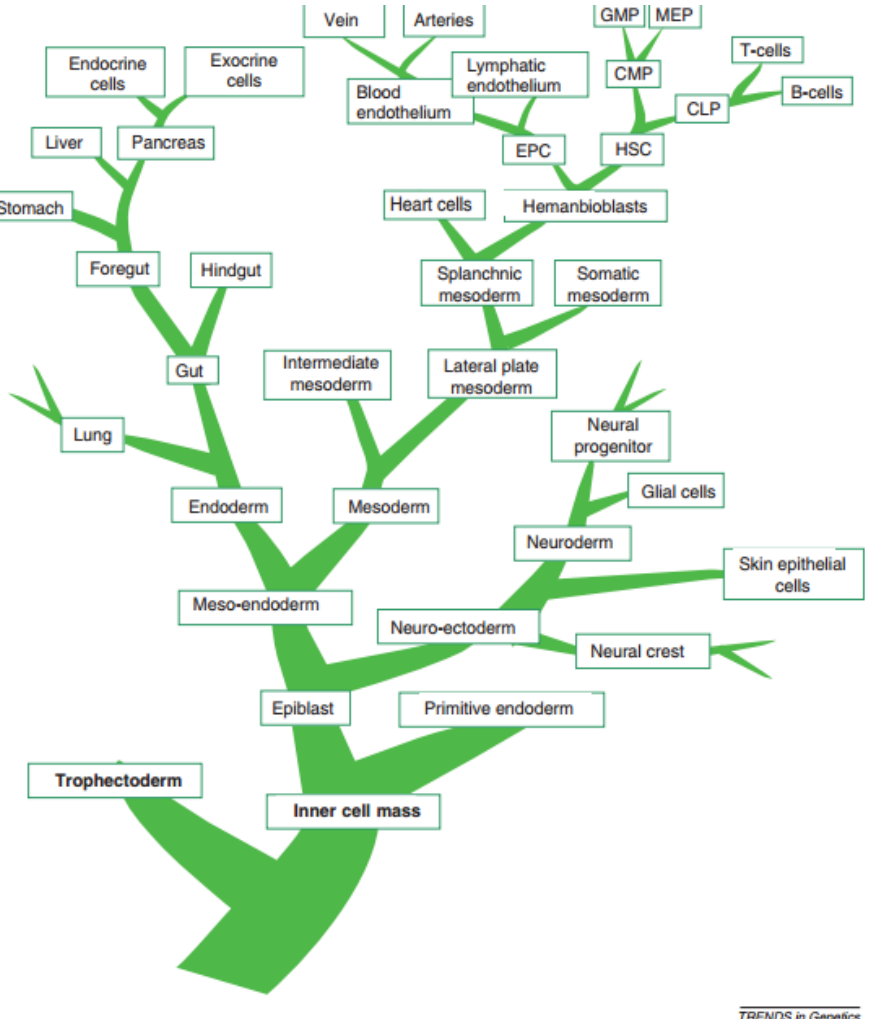


(c)

Cells	Transcription factors		Cell fates		
	X1	X2	X1> X2	X1~ X2	X1< X2
Early embryo	Cdx2	Oct4	Trophectoderm	Totipotent embryo	Inner cell mass
Embryo ICM	GATA6	Nanog	Primitive endoderm	Inner cell mass	Epiblast
Blood	GATA1	PU.1	Erythroid cells	Common myeloid progenitor	Myeloid cells
Pancreas	Ptf1a	Nkx6	Exocrine cells	Pancreatic progenitor	Endocrine cells
Somite	Pax3	Foxc2	Myogenic cells	Dermomyotome progenitor	Vascular cells

TRENDS in Genetics





TRENDS in Genetics

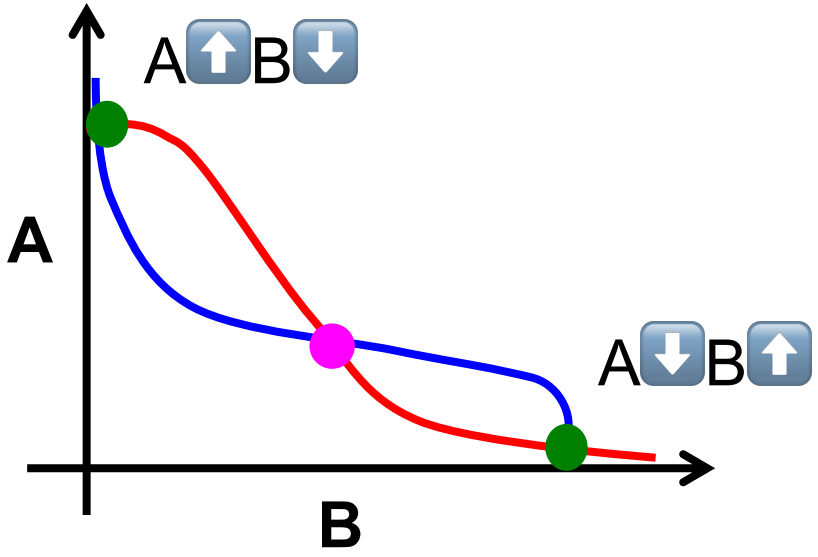
Zhou & Huang, Trends Genet 2012

Toggle switch allows for multiple mutually exclusive cell-states

Toggle switch

Bistability
(A, B) = (high, low)
(A, B) = (low, high)

Huang, PloS Biology 2013
Gardner *et al.* Nature 2000



$$\frac{dA}{dt} = g_A \frac{(B_0)^{n_B}}{(B_0)^{n_B} + B^{n_B}} - k_A A$$

$$\frac{dB}{dt} = g_B \frac{(A_0)^{n_A}}{(A_0)^{n_A} + A^{n_A}} - k_B B$$

Production

Degradation

Regulation

A_0 , B_0 = Threshold concentrations

- Hallmark of cell-fate decision making during embryonic development
- One of the first synthetic bio circuits designed

Is a toggle switch always bistable?

$$\frac{dA}{dt} = g_A \frac{(B_0)^{n_B}}{(B_0)^{n_B} + B^{n_B}} - k_A A$$

$$\frac{dB}{dt} = g_B \frac{(A_0)^{n_A}}{(A_0)^{n_A} + A^{n_A}} - k_B B$$

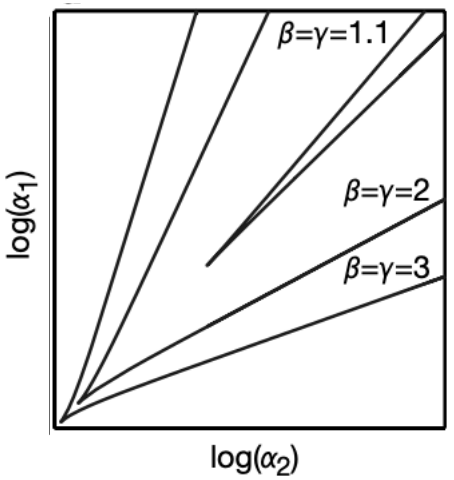
$$\frac{du}{dt} = \frac{\alpha_1}{1 + v^{\beta}} - u$$

$$\frac{dv}{dt} = \frac{\alpha_2}{1 + u^{\gamma}} - v$$

At steady-state, du/dt = dv/dt = 0

What happens at β = γ = 1?

A quadratic equation in u or v =>
At most two real distinct solutions



Gardner *et al.* Nature 2000

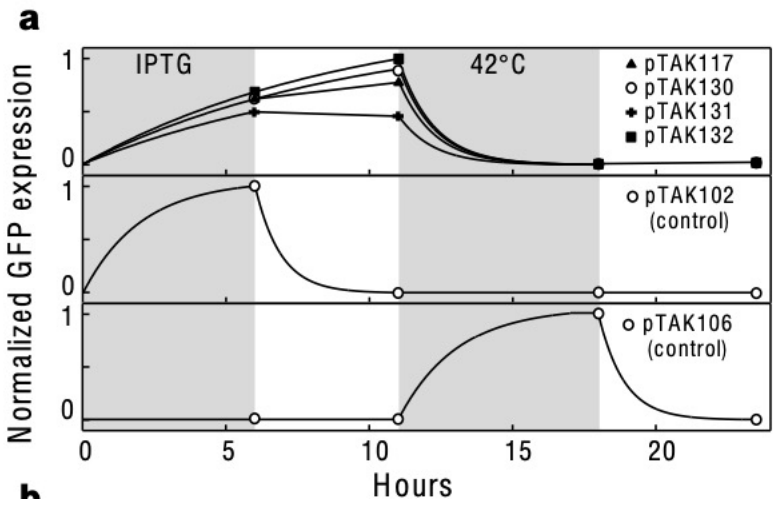
Summary (Part 1)

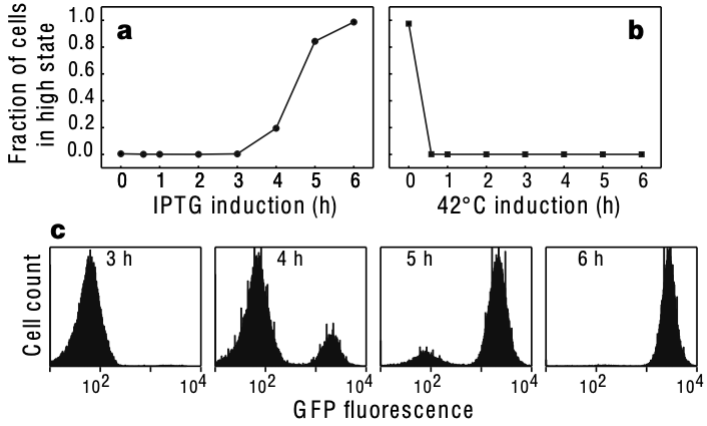
Toggle switch (positive feedback loop) can allow for:

a) Cells in a population having multiple steady states (phenotypes) => Bi-stability – (A low, B high) and (A high, B low)

b) Cells being pushed from one state to another => Plasticity

c) Cells requiring different duration/ extents of signal to switch => Heterogeneity



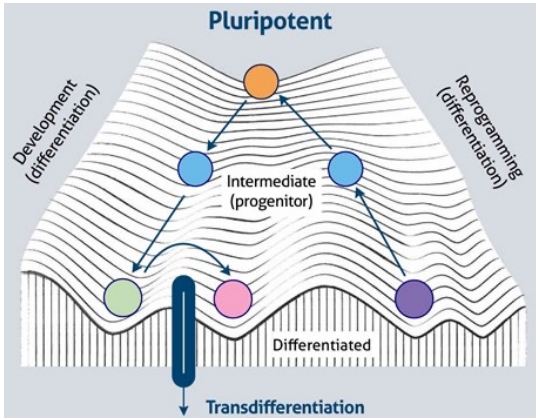


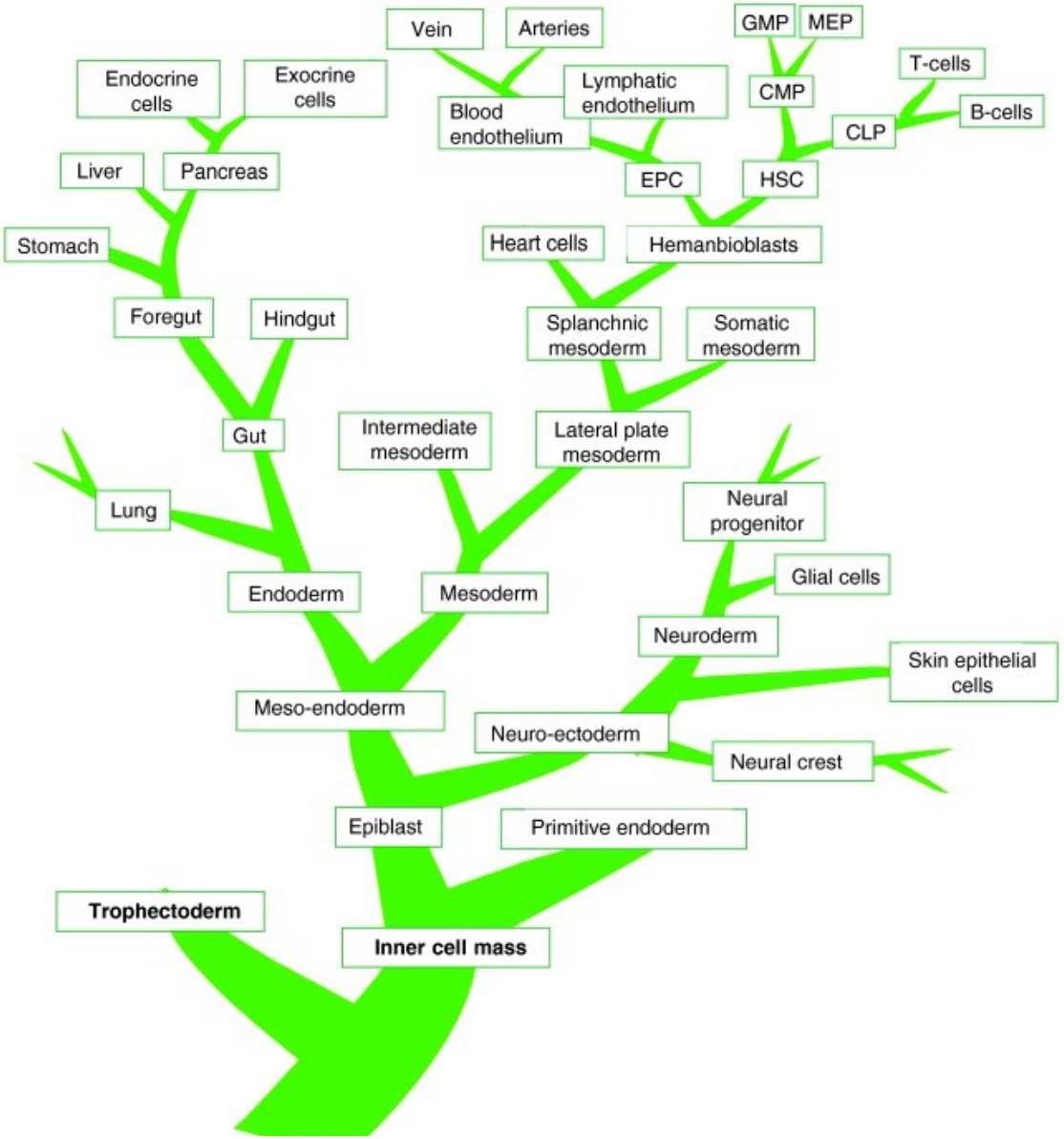
Negative feedback loop can allow for sustained or decaying oscillations.

Double positive feedback loop between A and B can also allow for bistability – (A low, B low) and (A high, B high)

Gardner et al. Nature 2000

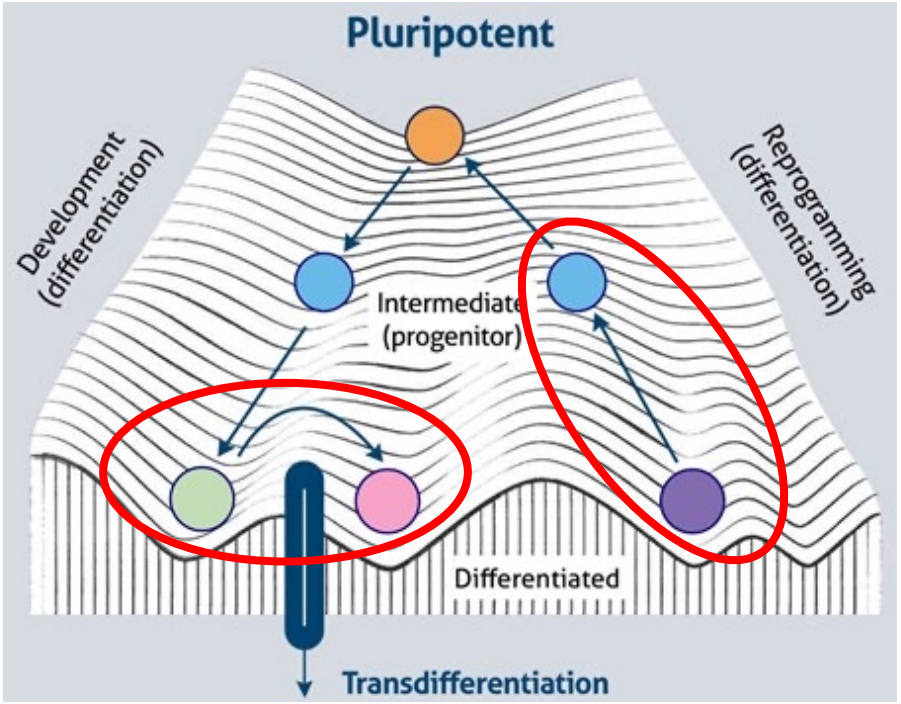
The ‘bifurcating’ cell-state tree





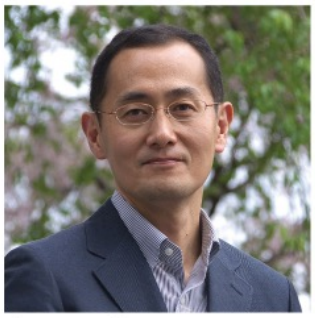
Zhou & Huang, Trends Genet 2011

Cell-state changes: bidirectional, reversible



Granados *et al.* Int J Mol Sci 2020

2012 Nobel Prize in Physiology or Medicine



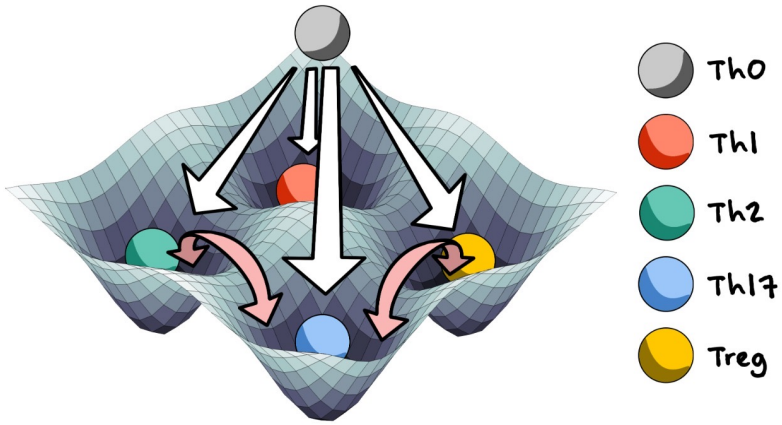
Shinya Yamanaka
University of Kyoto, Japan

John B. Gurdon
Gurdon Institute in Cambridge, UK

Cells can also reversibly change their identity => "Controlled enthusiasm"

What rules/principles cells follow in decision-making?

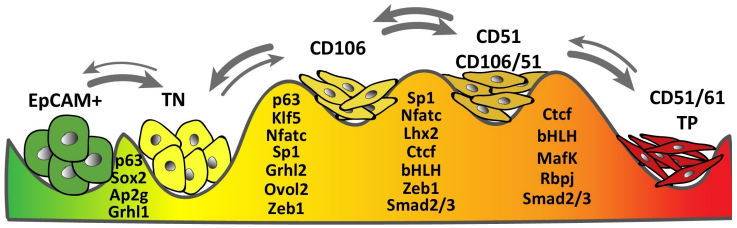
CD4+ T-cell differentiation



Atchuta S Duddu

Duddu *et al.* J R Soc Interface 2020
Duddu *et al.* Mol Biol Cell 2022
Duddu *et al.* bioRxiv 2024

Epithelial-Mesenchymal Transition

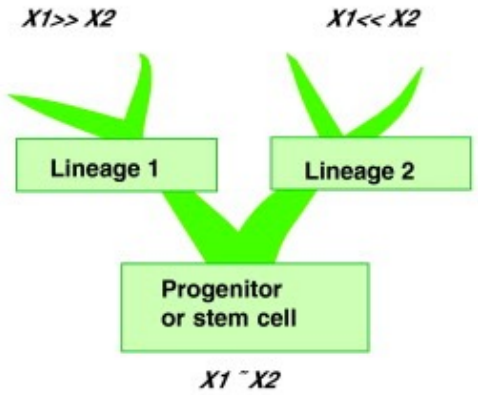


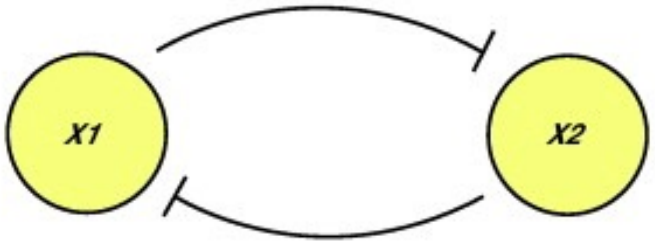
	Epithelial tumor cells	Early hybrid EMT state	Hybrid EMT state	Late hybrid EMT state	Mesenchymal tumor cells
Proliferation	+++++	++++	+++	++	+
Invasion	+	++	+++	++++	+++++
Plasticity	+	++	+++	++++	++
Stemness	+	+++	+++	+++	+++
Metastasis	+	++++	++++	++	+

Kishore Hari

Hari *et al.* eLife 2022
Hari*, Rashid* *et al.* PLoS Comp Biol 2022
Hari *et al.* bioRxiv 2023

Toggle switch: a motif for bifurcating cell-states

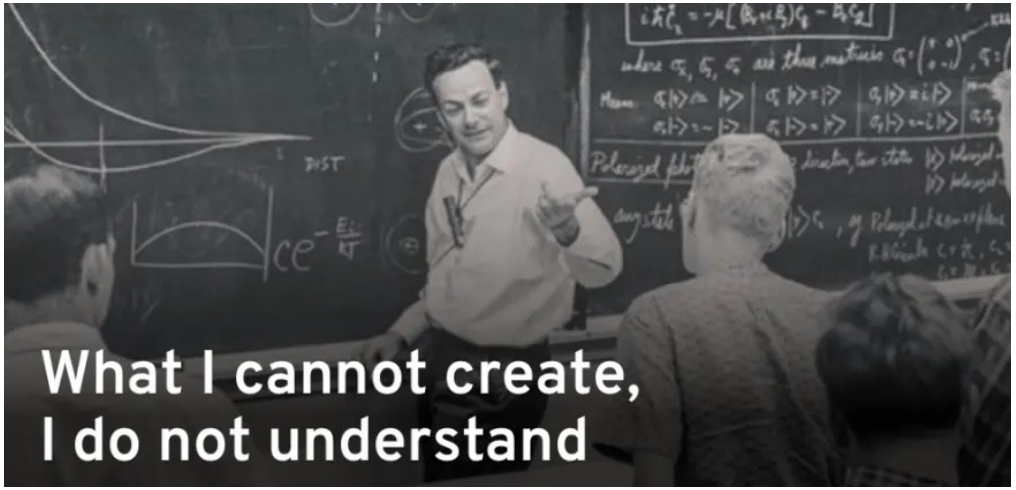




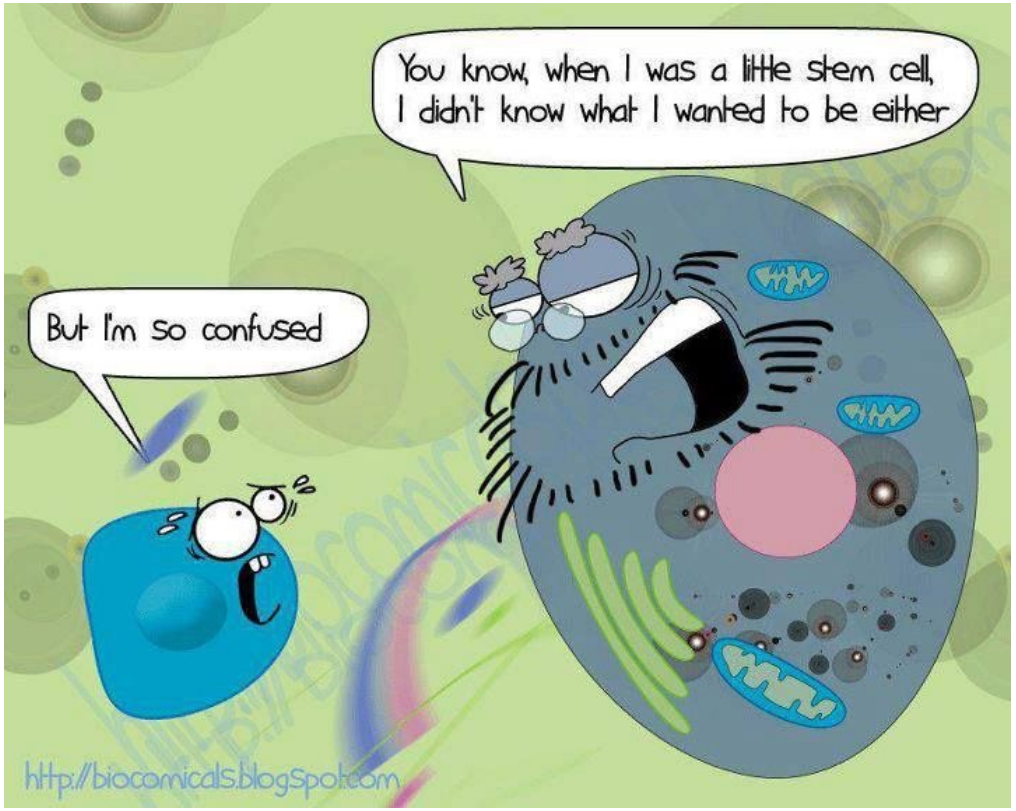
Cells	Transcription factors		Cell fates		
	X1	X2	X1> X2	X1~ X2	X1< X2
Early embryo	Cdx2	Oct4	Trophectoderm	Totipotent embryo	Inner cell mass
Embryo ICM	GATA6	Nanog	Primitive endoderm	Inner cell mass	Epiblast
Blood	GATA1	PU.1	Erythroid cells	Common myeloid progenitor	Myeloid cells
Pancreas	Ptf1a	Nkx6	Exocrine cells	Pancreatic progenitor	Endocrine cells
Somite	Pax3	Foxc2	Myogenic cells	Dermomyotome progenitor	Vascular cells

Zhou & Huang, Trends Genet 2011

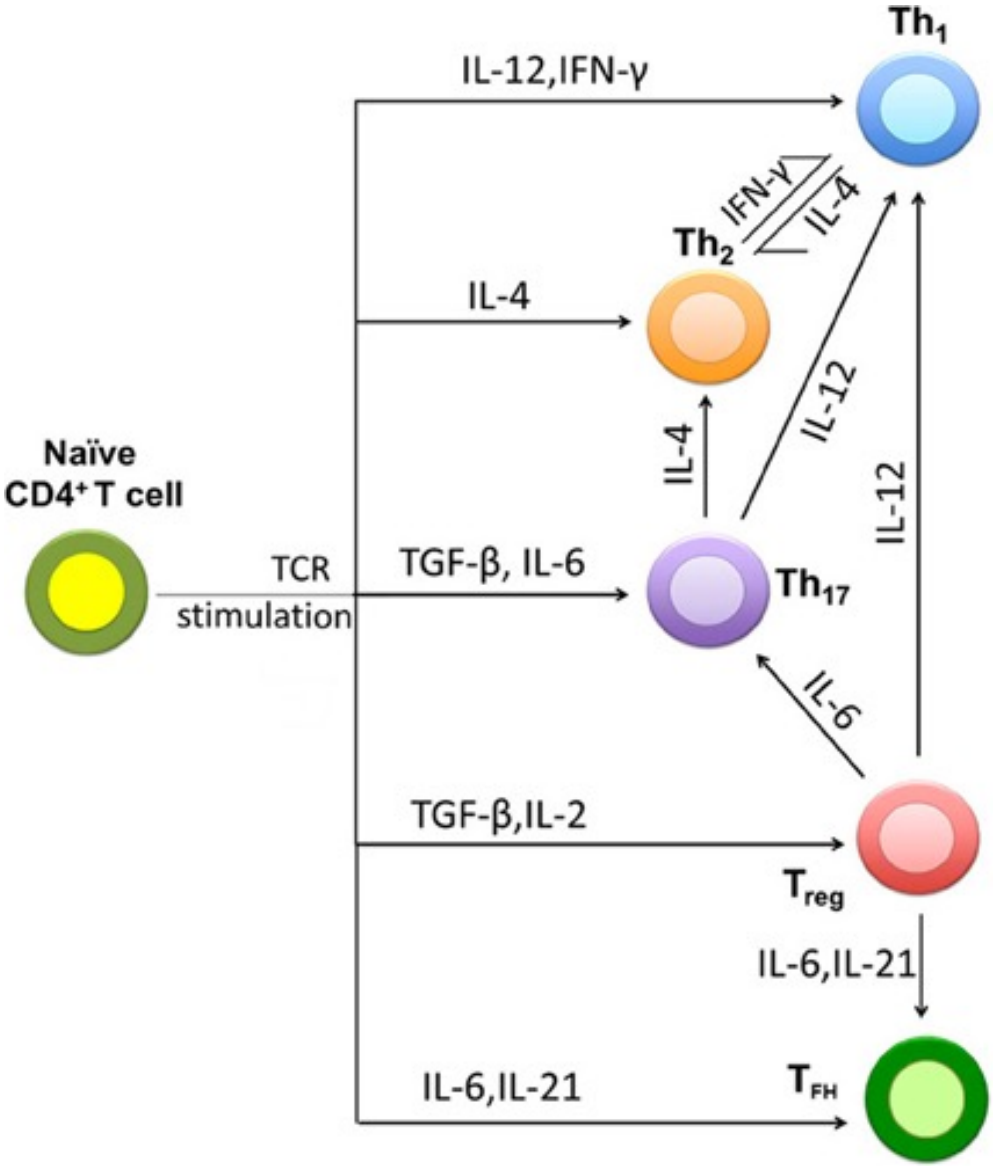
OK, so we understand a toggle switch (2 cell-states)!



What if a progenitor cell can give rise to more than 2 cell-states together?



Differentiation of naïve CD4+ T cells into > 2 subsets



Magombedze *et al.* Front Physiol 2013

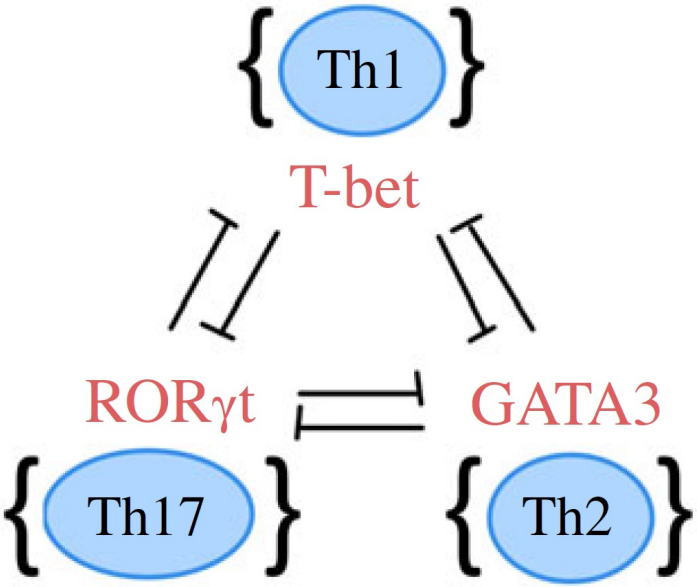
Network governing Th1, Th2, Th17 cell-states

Th1 | { **T-bet high**, GATA3 low, RORγT low }

Th2 | { T-bet low, **GATA3 high**, RORγT low }

Th17 | { T-bet low, GATA3 low, **RORγT high** }

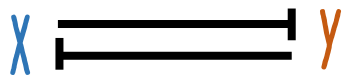
Can this toggle triad explain
T-cell differentiation?



Atchuta

Duddu *et al.* J R Soc Interface 2020

ODEs governing the dynamics



$$\frac{d[X]}{dt} = g_X\left(\frac{1+\lambda_{YX}[Y/T_{YX}]^{n_{YX}}}{1+[Y/T_{YX}]^{n_{YX}}}\right) - k_X[X]$$

$$\frac{d[Y]}{dt} = g_Y\left(\frac{1+\lambda_{XY}[X/T_{XY}]^{n_{XY}}}{1+[X/T_{XY}]^{n_{XY}}}\right) - k_Y[Y]$$

Here,

g_X and g_Y are the production rates of A and B

k_X and k_Y are the degradation rates of A and B

λ_{XY} and λ_{YX} is the fold change in interaction

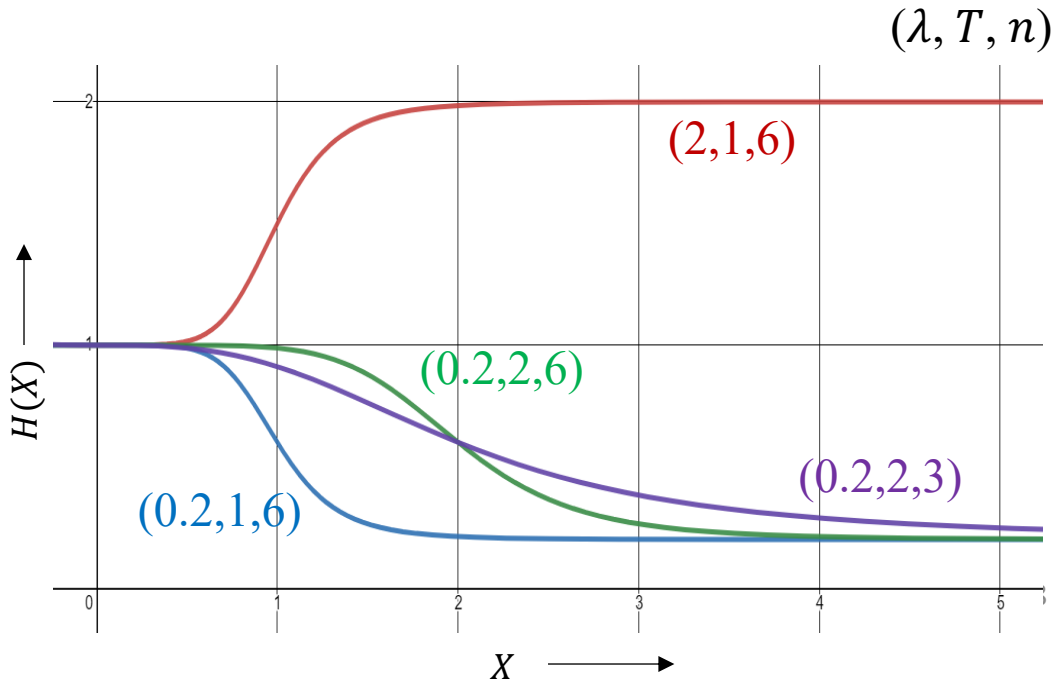
n_{XY} and n_{YX} is the hill's coefficient in interaction

T_{XY} and T_{YX} is threshold value of the interaction

$$\frac{dY}{dt} = gH(X) - kY \qquad H(X) = \frac{1+\lambda\left(\frac{X}{T}\right)^n}{1+\left(\frac{X}{T}\right)^n}$$

$$\frac{1+2\left(\frac{X}{1}\right)^6}{1+\left(\frac{X}{1}\right)^6} \qquad \frac{1+0.2\left(\frac{X}{1}\right)^6}{1+\left(\frac{X}{1}\right)^6}$$

$$\frac{1+0.2\left(\frac{X}{2}\right)^6}{1+\left(\frac{X}{2}\right)^6} \qquad \frac{1+0.2\left(\frac{X}{2}\right)^3}{1+\left(\frac{X}{2}\right)^3}$$



Dynamics over a kinetic parameter set ensemble

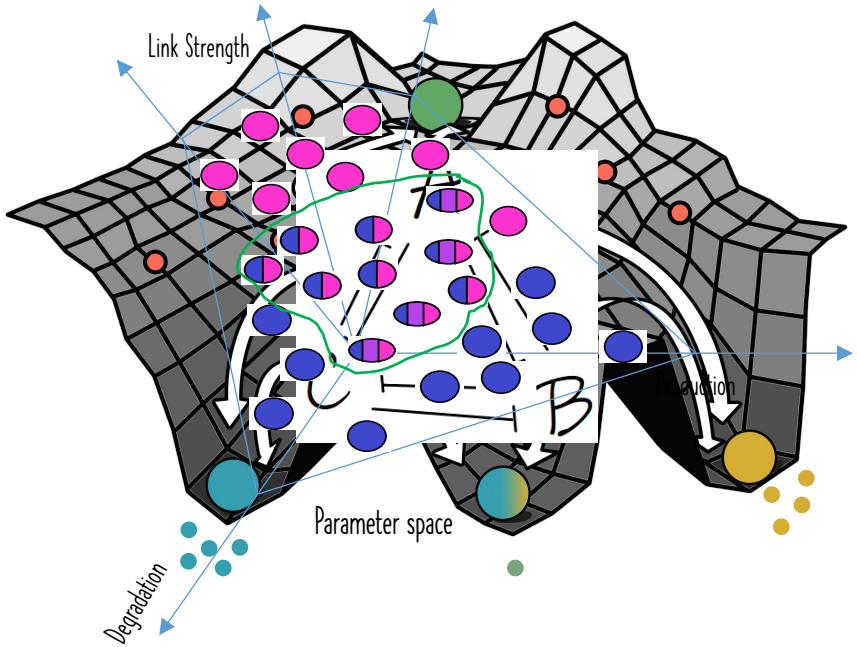
RACIPE - RAndomized CIrcuit PErturbation

Topology of network

User defined parametric space

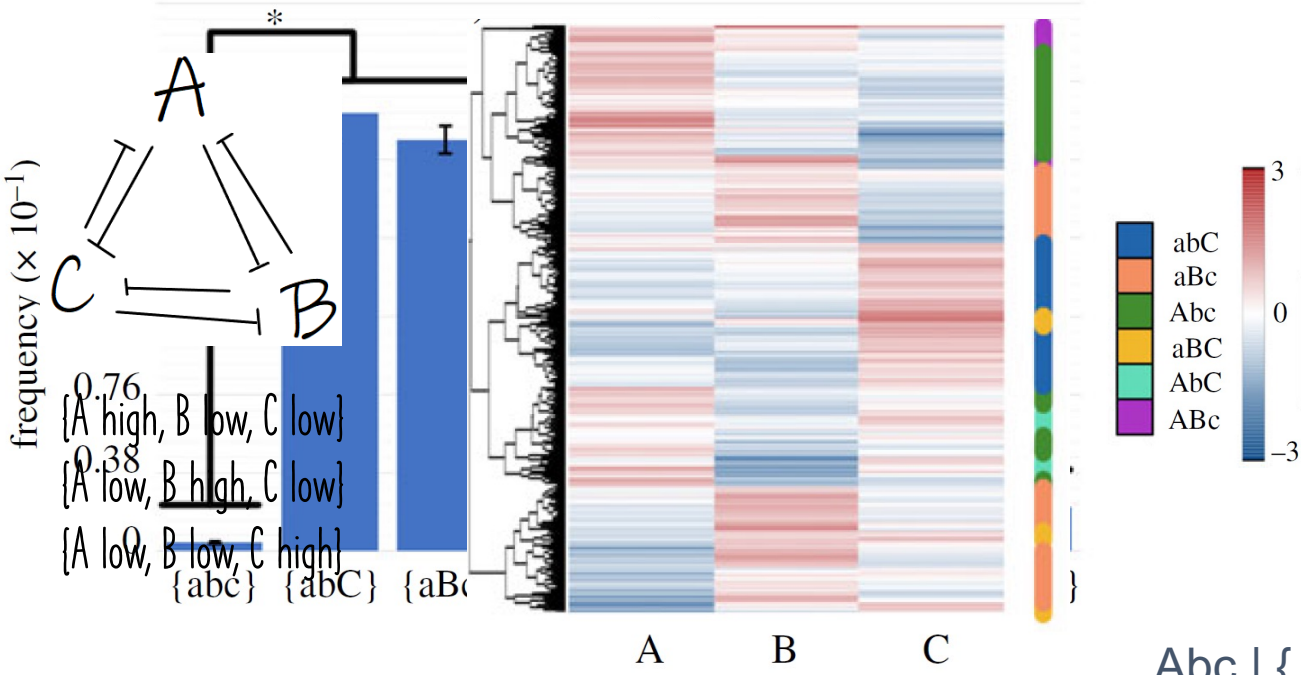
Ensemble of kinetic models

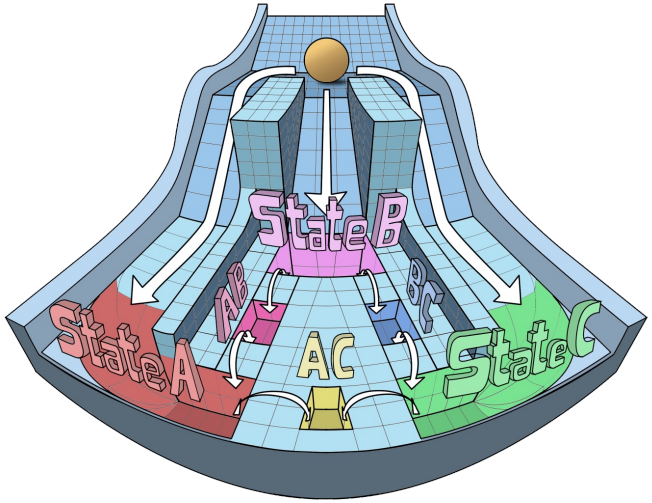
z-scores calculated over solutions
across parameter sets * initial conditions



Huang *et al.* PloS Comp Biol 2017

Toggle triad enables tristability





Abc | { **A high**, B low, C low }

aBc | { A low, **B high**, C low }

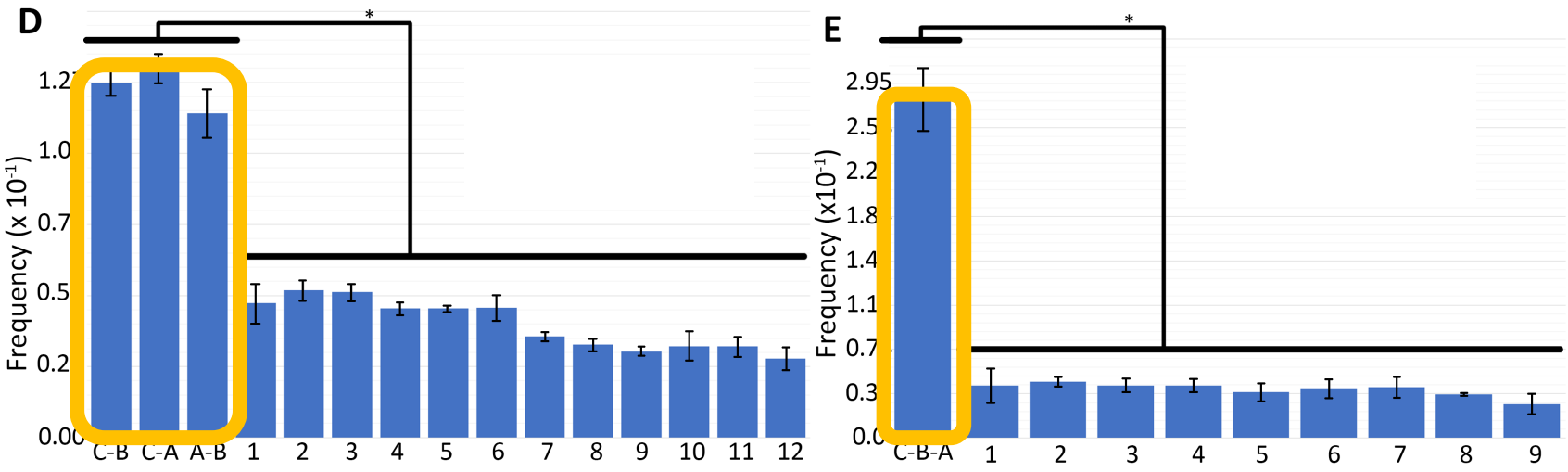
ABc | { A low, B low, **C high** }

ABc | { **A high**, **B high**, C low }

AbC | { **A high**, B low, **C high** }

aBC | { A low, **B high, C high** }

These three states can co-exist



What are the defining traits of the topology? Are the traits unique?

Duddu *et al.* J R Soc Interface 2020

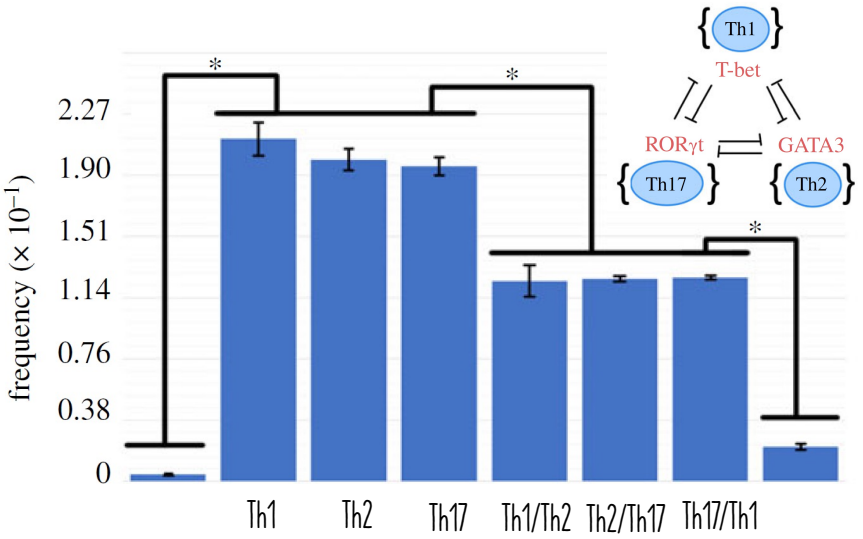
Dynamical traits of Toggle Triad

Negligible frequency of ‘all-high’ or ‘all-low’ monostable states

Similar frequency of three ‘single-positive’ states

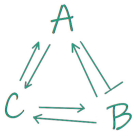
Similar frequency of three ‘double-positive’ states

‘Single-positive’ states are more frequent than ‘double-positive’ states

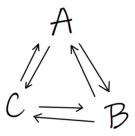


Duddu *et al.* J R Soc Interface 2020

Dynamical traits of Toggle Triad are unique to this topology

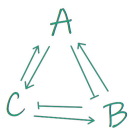


Circuit C1

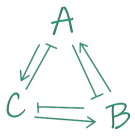


Circuit C2

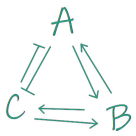
the case of all activations



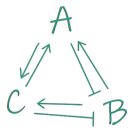
Circuit C3



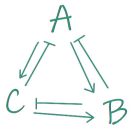
Circuit C4



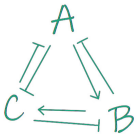
Circuit C5



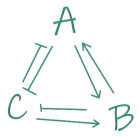
Circuit C6



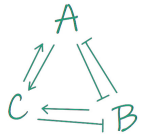
Circuit C7



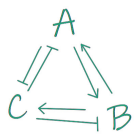
Circuit C8



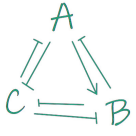
Circuit C9



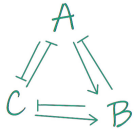
Circuit C10



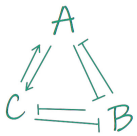
Circuit C11



Circuit C12



Circuit C13



Circuit C14

Negligible frequency of 'all-high' or 'all-low' monostable states

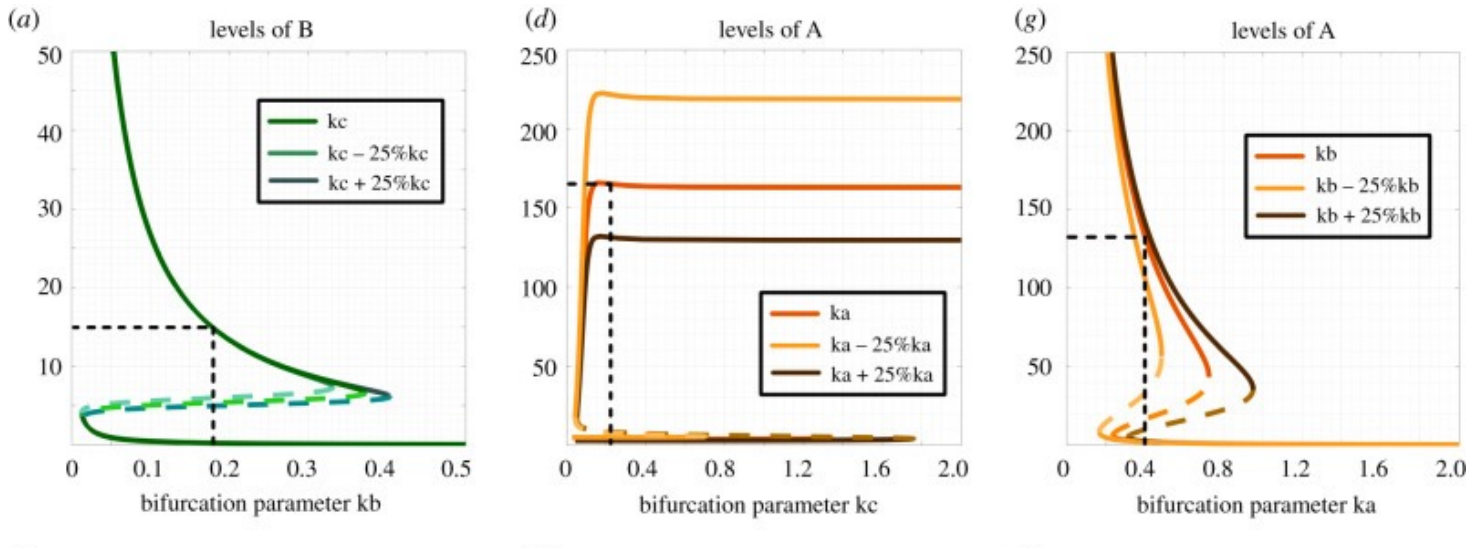
Similar frequency of three 'single-positive' states

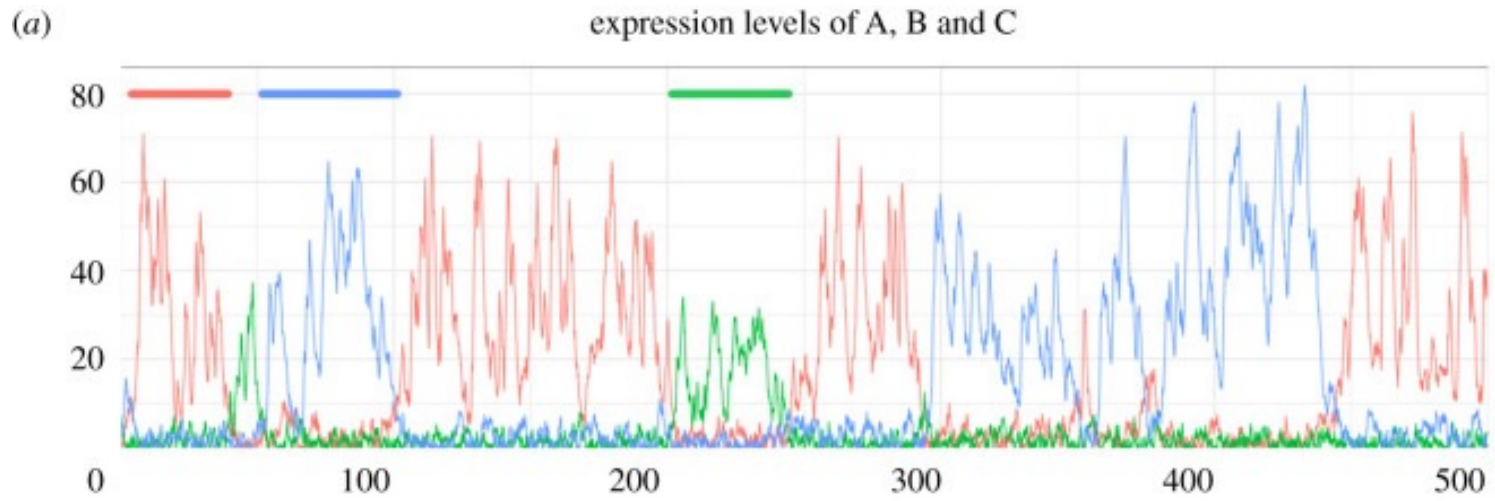
Similar frequency of three 'double-positive' states

'Single-positive' states are more frequent than 'double-positive' states

Duddu *et al.* J R Soc Interface 2020

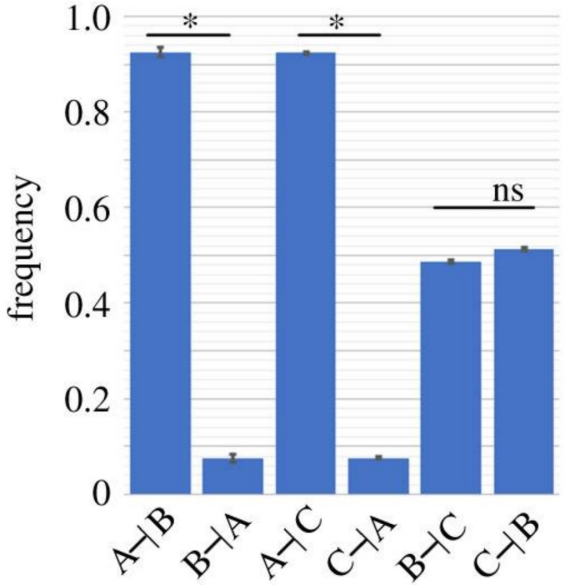
Stochastic switching among multiple T cell-states





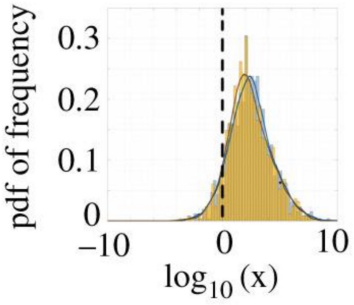
Duddu *et al.* J R Soc Interface 2020

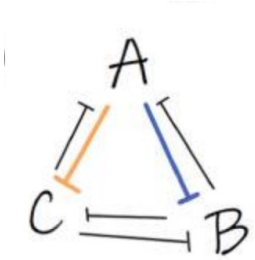
How is frustration resolved in toggle triad?

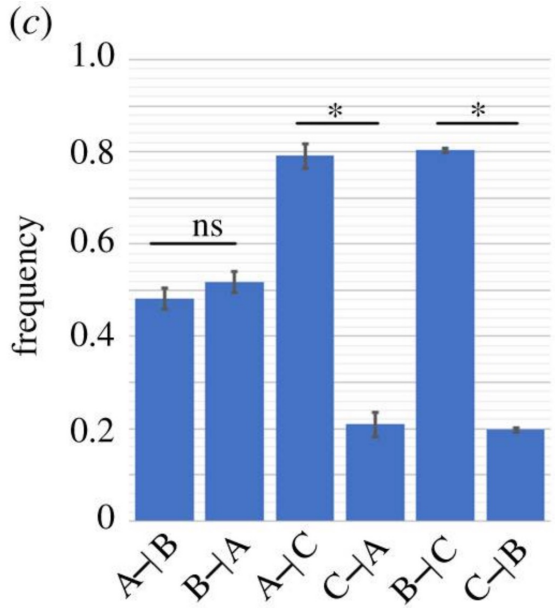


Monostable: Abc | { **A high**, B low, C low }

Link strength ~ $n/(\lambda^{*}H_0/(g/k))$







Bistable: Abc | { **A high**, B low, C low } and
aBc | { A low, **B high**, C low }

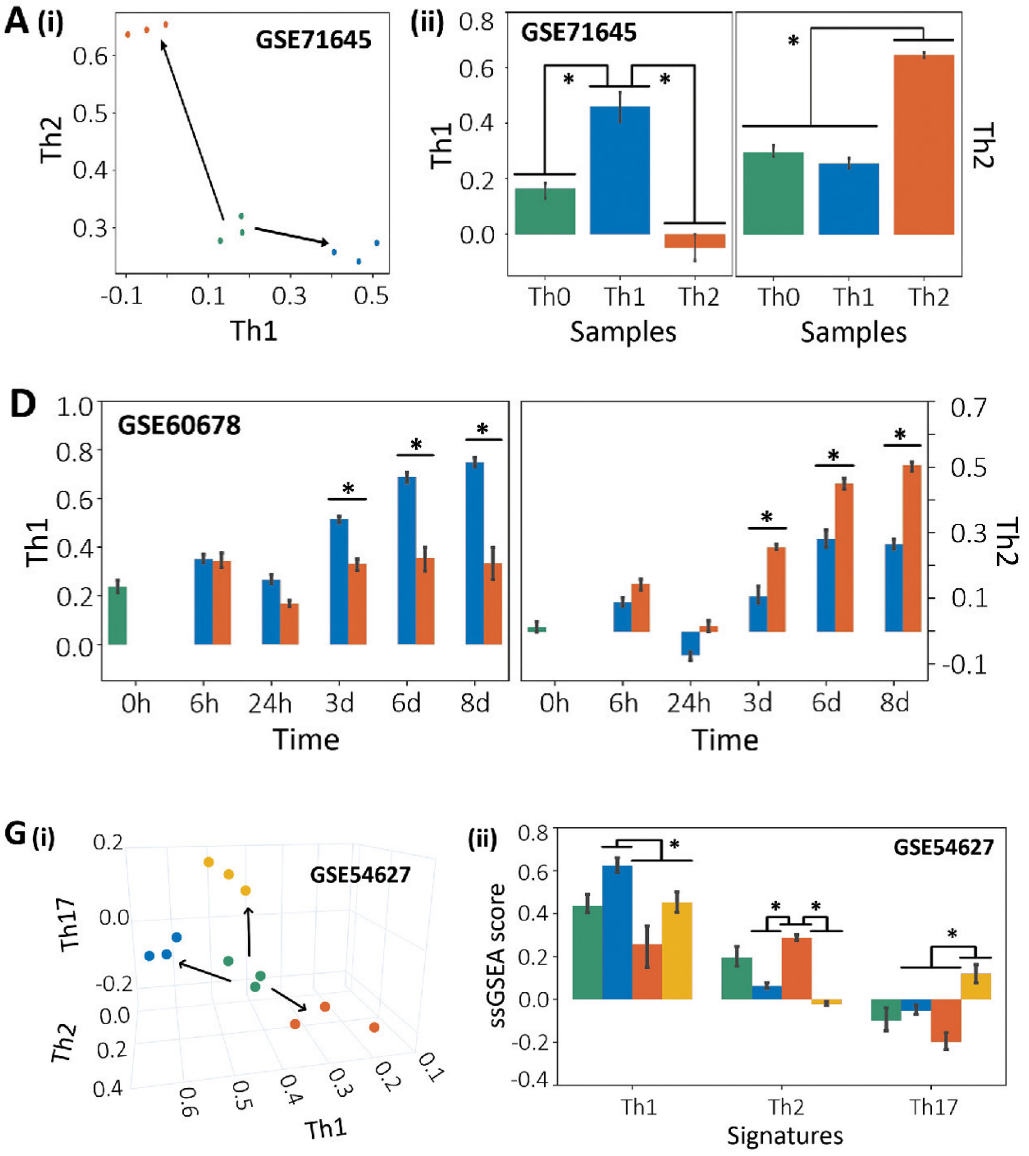
Duddu *et al.* J R Soc Interface 2020

Toggle triad in CD4+ T-cell differentiation

Model predictions	Experimental validations	
Existence of three 'single-positive' states	Th1 (T-bet high, GATA3 low, RORγT low)	The three states have been well established
	Th2 (T-bet low, GATA3 high, RORγT low)	
	Th17 (T-bet low, GATA3 low, RORγT high)	
Existence of three 'double-positive states	Th1/2 (T-bet high, GATA3 high, RORγT low)	Antebi *et al.*, PLoS Biol. 2013
	Th1/17 (T-bet high, GATA3 low, RORγT high)	Chatterjee *et al.*, Cell Metab. 2018
	Th2/17 (T-bet low, GATA3 high, RORγT high)	Tortola *et al.*, Immunity 2020

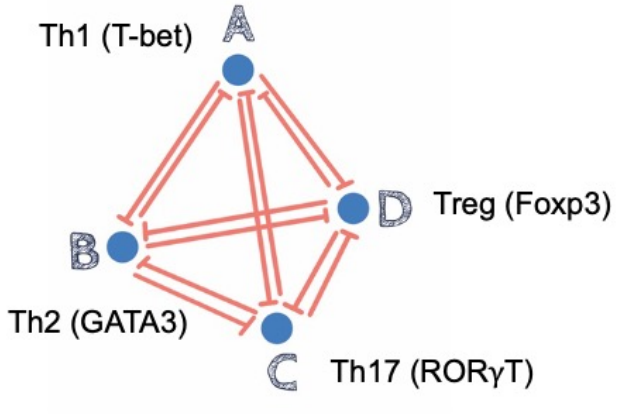
Duddu *et al.* J R Soc Interface 2020

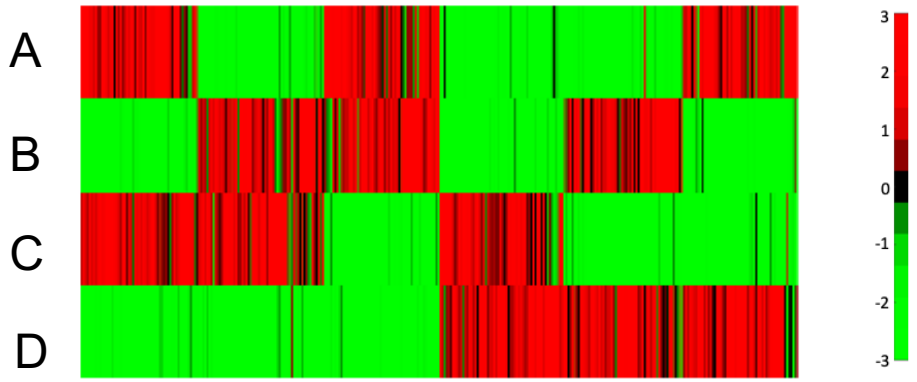
RNA-seq data validates our model predictions

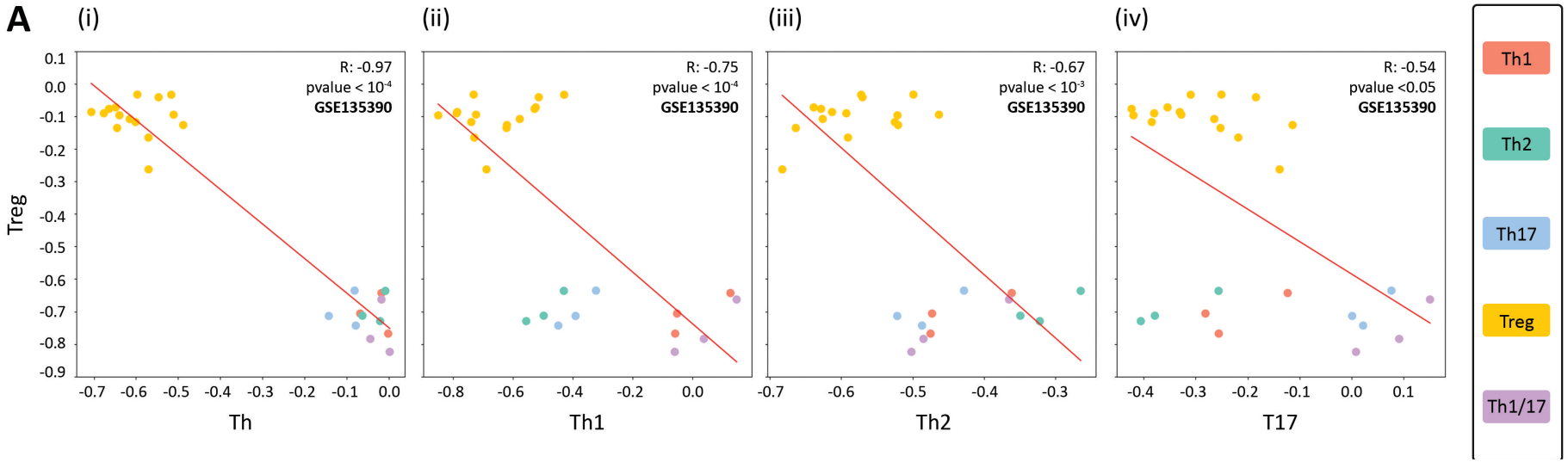


Duddu *et al.* Mol Biol Cell 2022

Toggle tetrahedron => Predominant ‘hybrid’ states



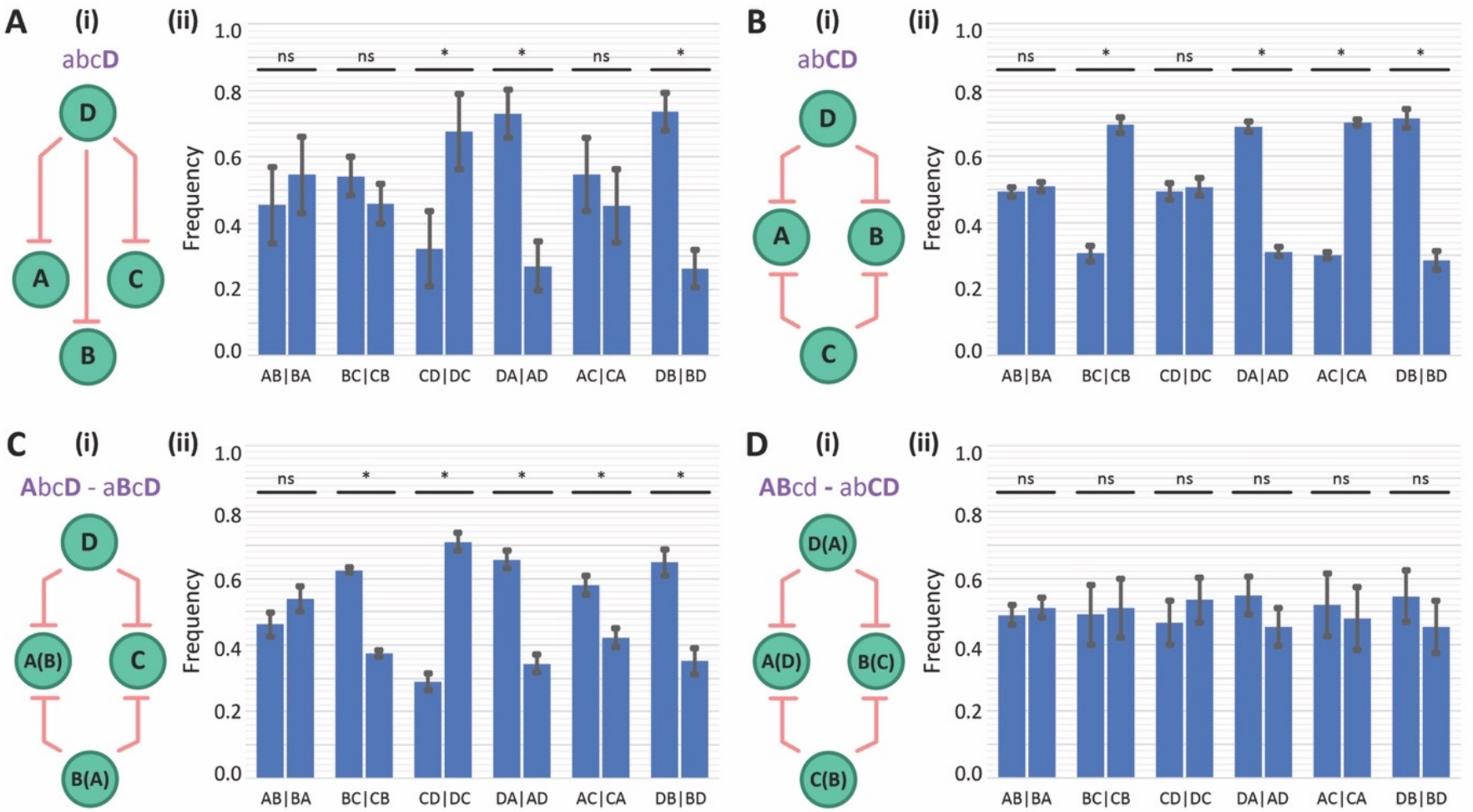




None of the 4 node network structures allows for (1,0,0,0) states as prevalent.

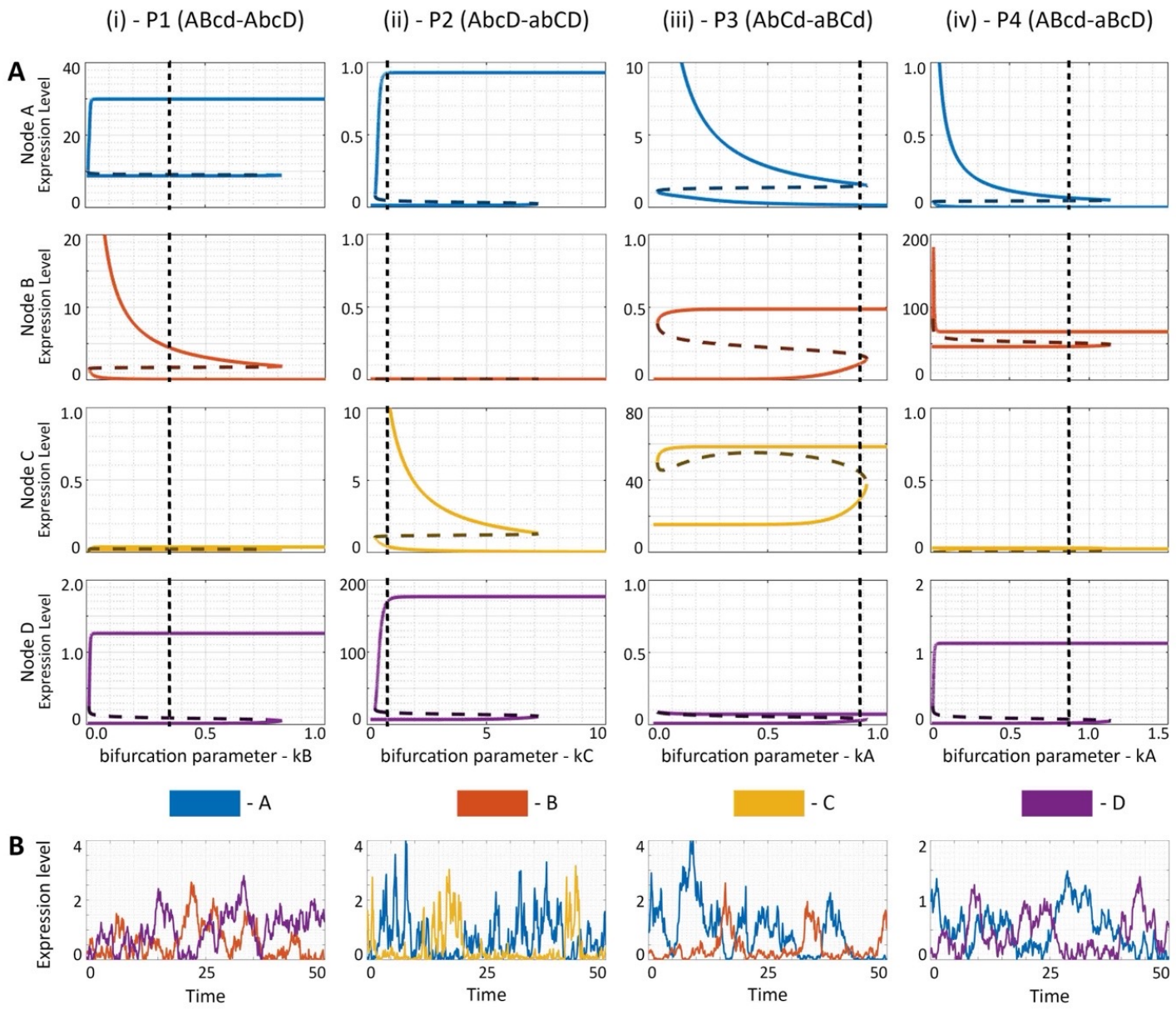
Duddu *et al.* bioRxiv 2024

Link strength analysis for toggle tetrahedron



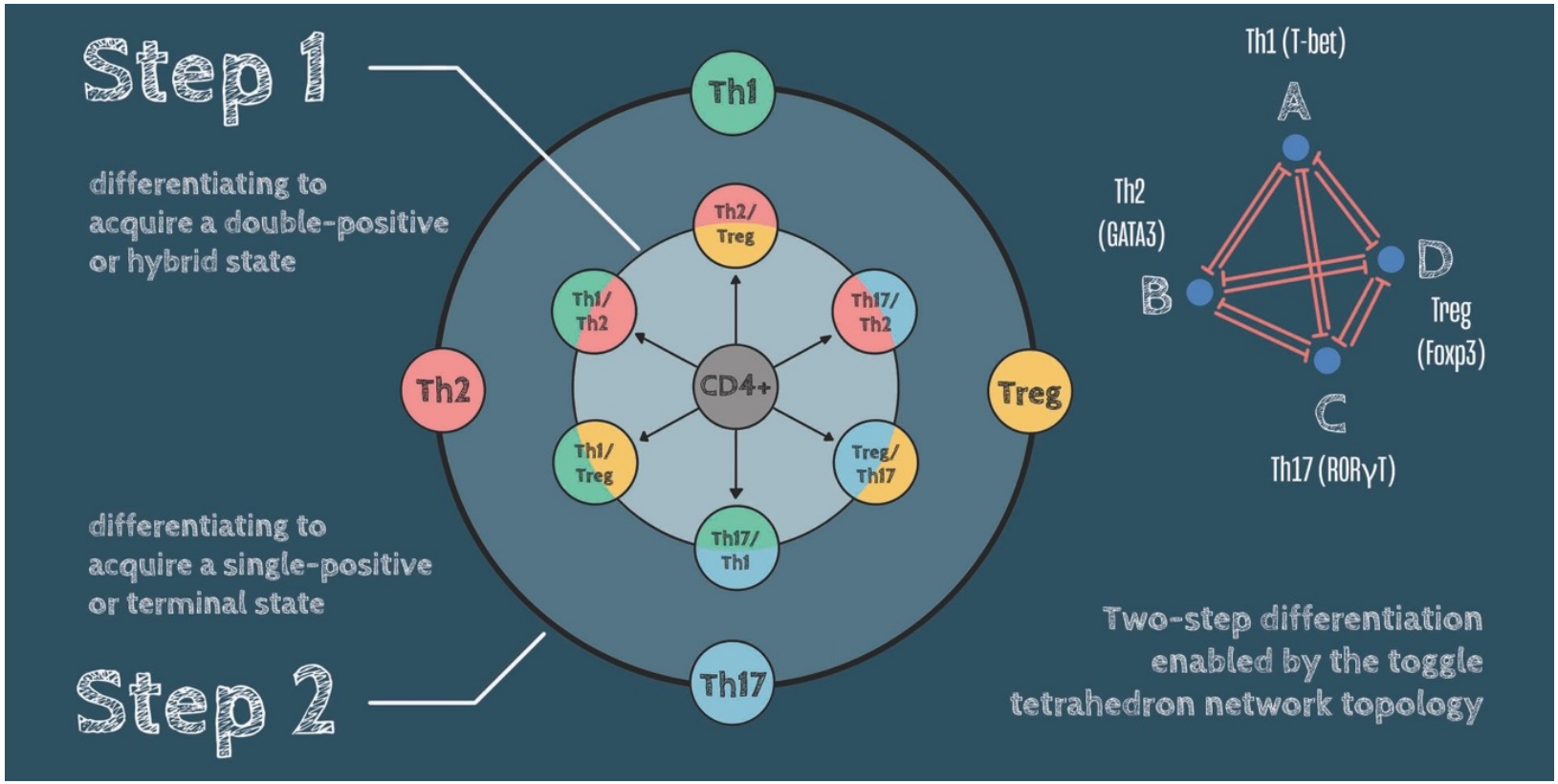
Duddu *et al.* bioRxiv 2024

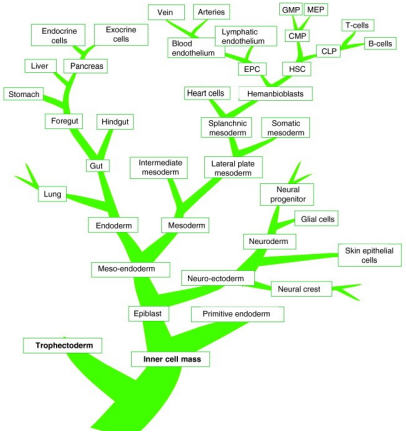
Stochastic state-switching in toggle tetrahedron



Duddu *et al.* bioRxiv 2024

Two-step decision-making in toggle tetrahedron

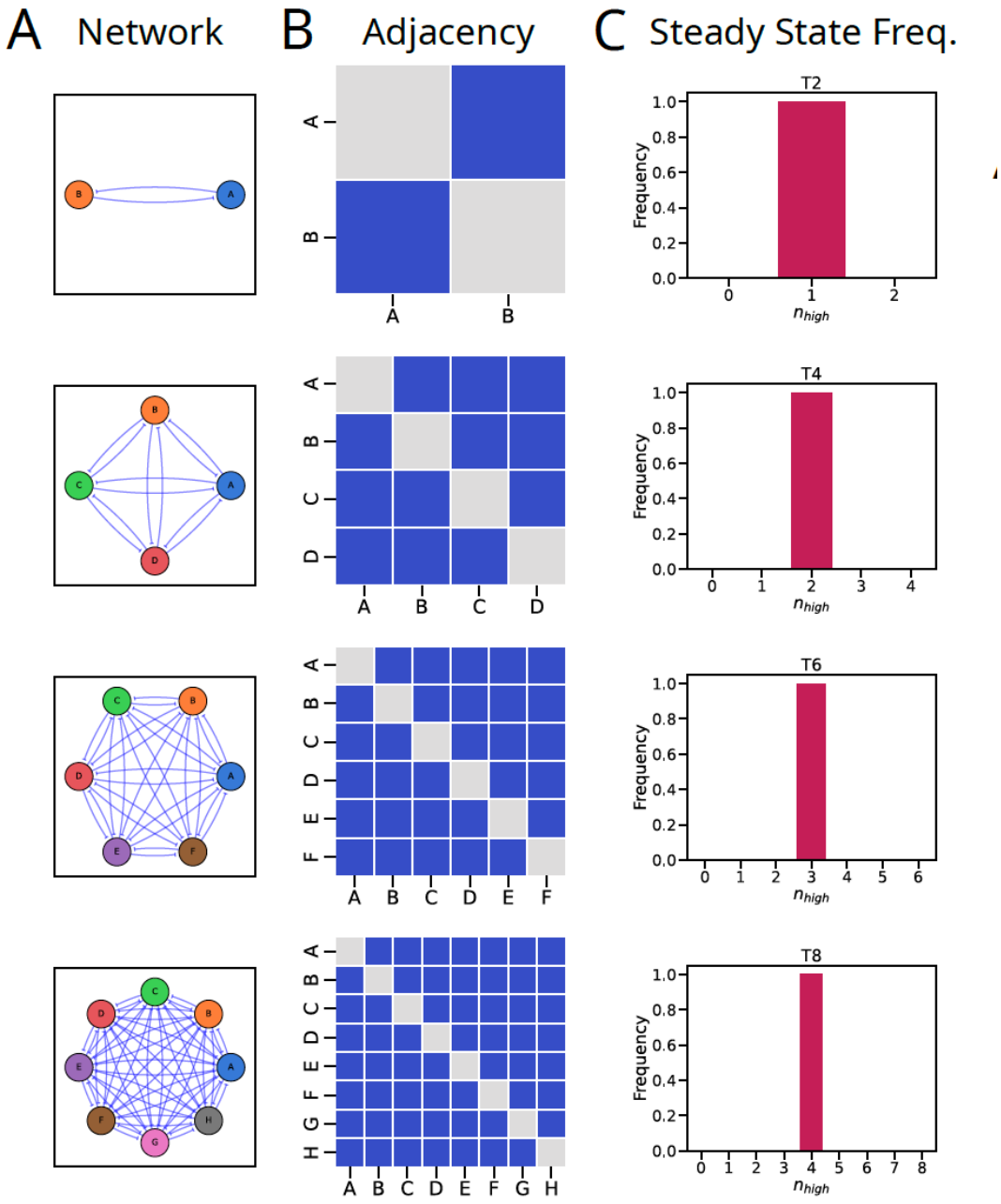




Could this possibly explain why often binary/ternary branches are seen in developmental decision-making?

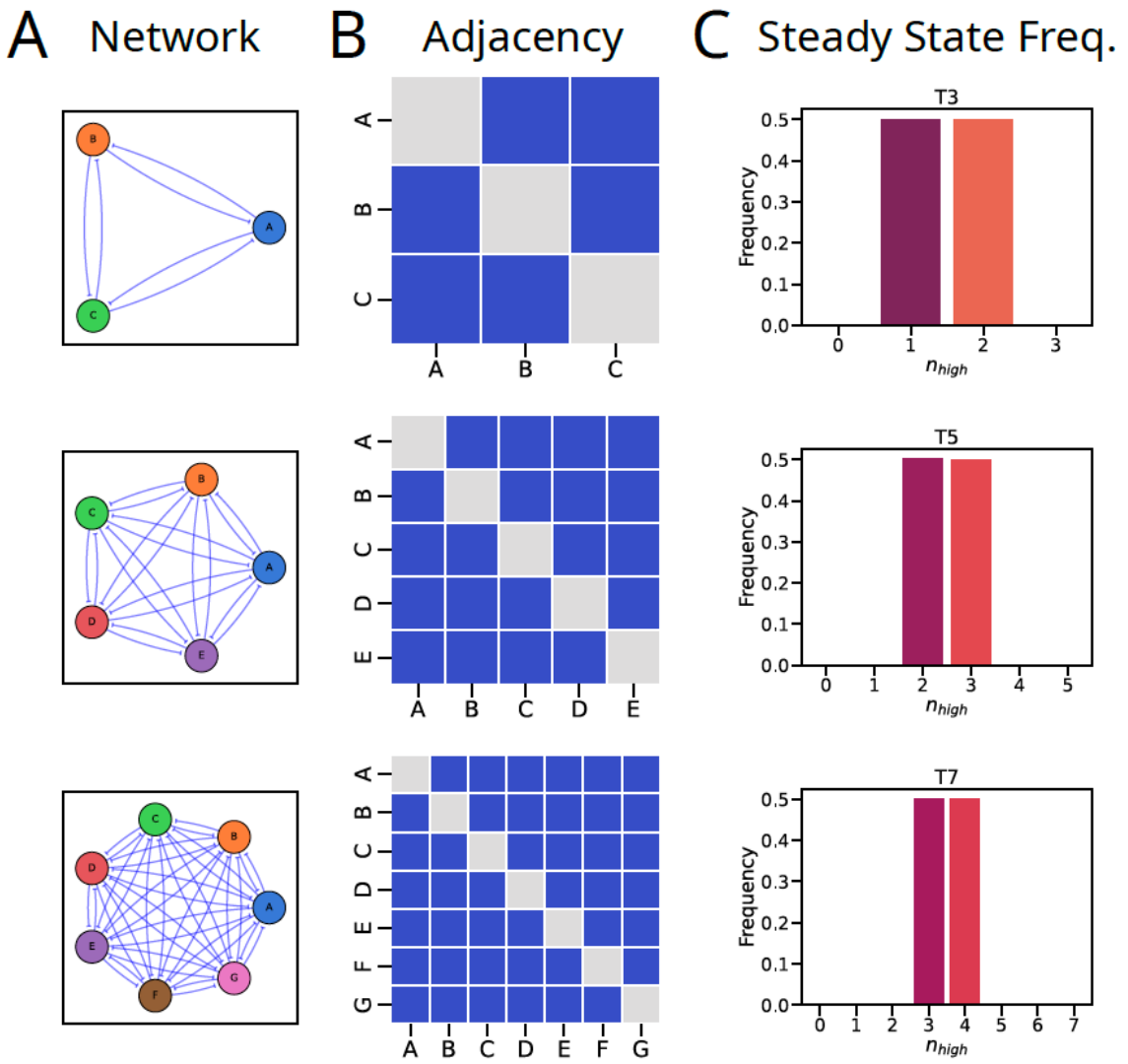
Duddu *et al.* bioRxiv 2024

Extending toggle switch and toggle tetrahedron results to larger n-node mutually repressive networks



Harshavardhan*, Billakurthi* & Jolly; manuscript in preparation

Extending toggle triad results to larger n-node mutually repressive networks



Harshavardhan*, Billakurthi* & Jolly; manuscript in preparation

Summary (Part 2)

- Toggle triad explains the co-existence and switching among differentiated (Th1, Th2, Th17) and hybrid (Th1/Th2, Th2/Th17, Th1/Th17) T-cell states.

- Toggle tetrahedron reveals 6 hybrid states (Th1/Th2, Th2/ Th17, Th1/Th17, Th1/Treg, Th2/Treg, Th17/Treg) as the most frequent states, suggesting a two-step decision.

- Beyond toggle triad, no mutually repressive network allows for single-step decision-making.