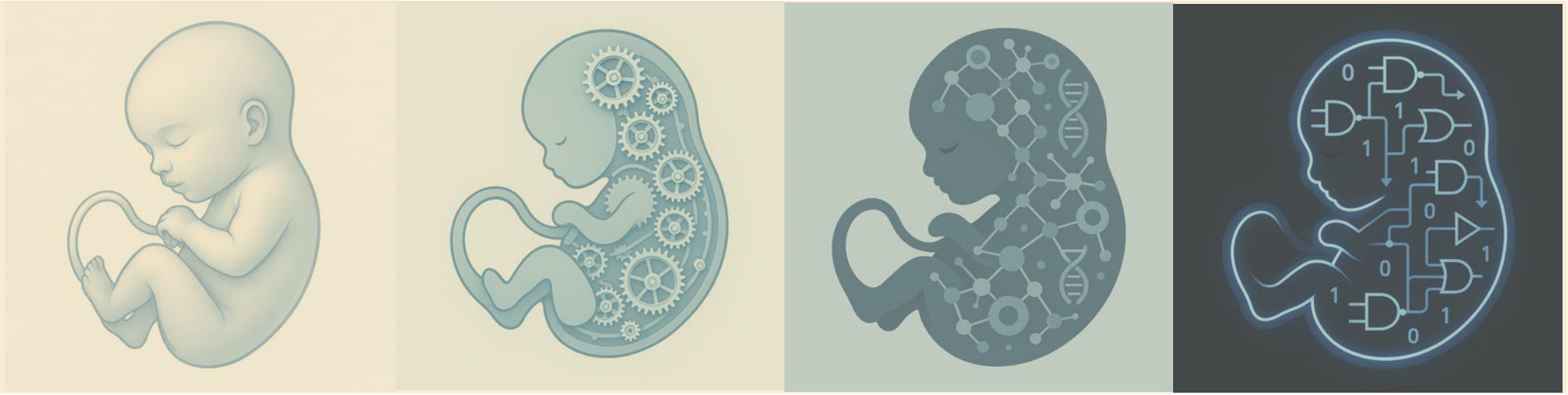
What is biological information? (II)



Course 3: A logic view of information processing: Network motifs

Thomas Lecuit

chaire: Dynamiques du vivant

COLLÈGE DE FRANCE
1530

Logic and function of biological networks

Information provides a language to decipher the *meaning* and *logic of living systems*

Use Computational metaphor to account for:

- Purpose & Function
- Rules & Logic

Biological processes can be usefully described as **information processing**: they involve the manipulation of symbolic representations according to formal rules to achieve a functional outcome.

This perspective shifts the focus from the specific chemical substrates (e.g., DNA, proteins) to the logical and algorithmic principles governing the system.

From Molecules to Function

One gene one function hypothesis

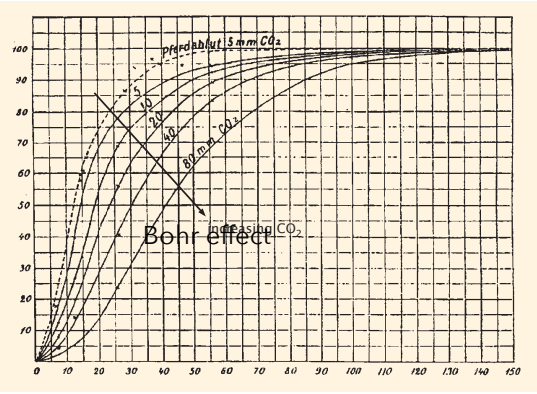
– Ephrussi and Beadle: *Drosophila* eye colour phenotypes and mutants suggested 1 gene - 1 enzyme relationship

– Beadle and Tatum (amino acids synthesis mutant in *Neurospora crassa*): 1 gene - 1 enzyme. Nobel Prize, 1958.

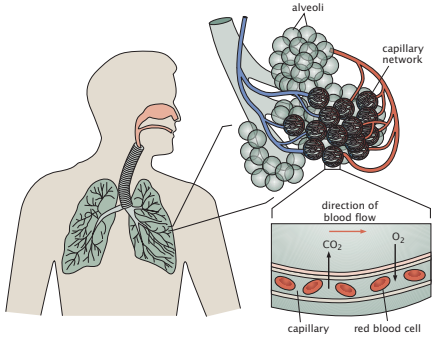
1 molecule – 1 function

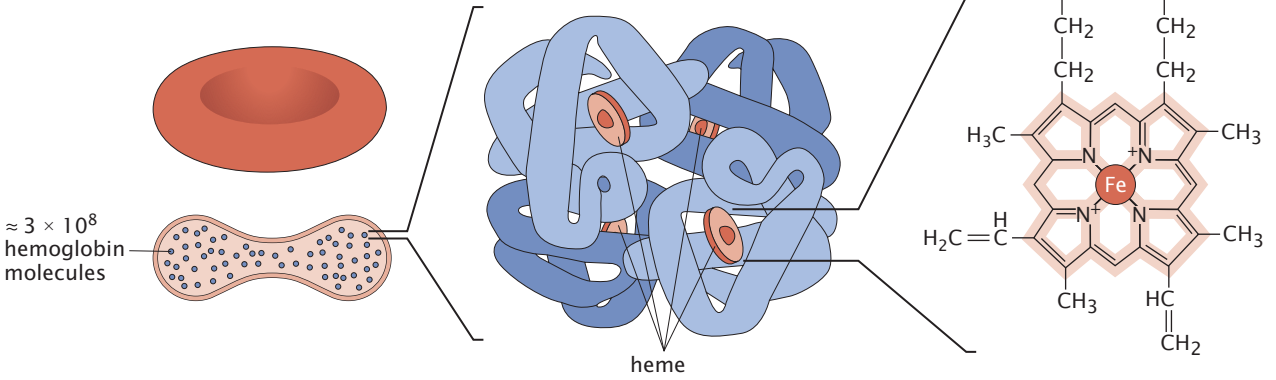
Oxygen transport by haemoglobin

Fractional binding of hemoglobin



O_2 concentration





R. Phillips, The Molecular Switch: signaling and allostery. *Princeton Univ. Press*. 2020

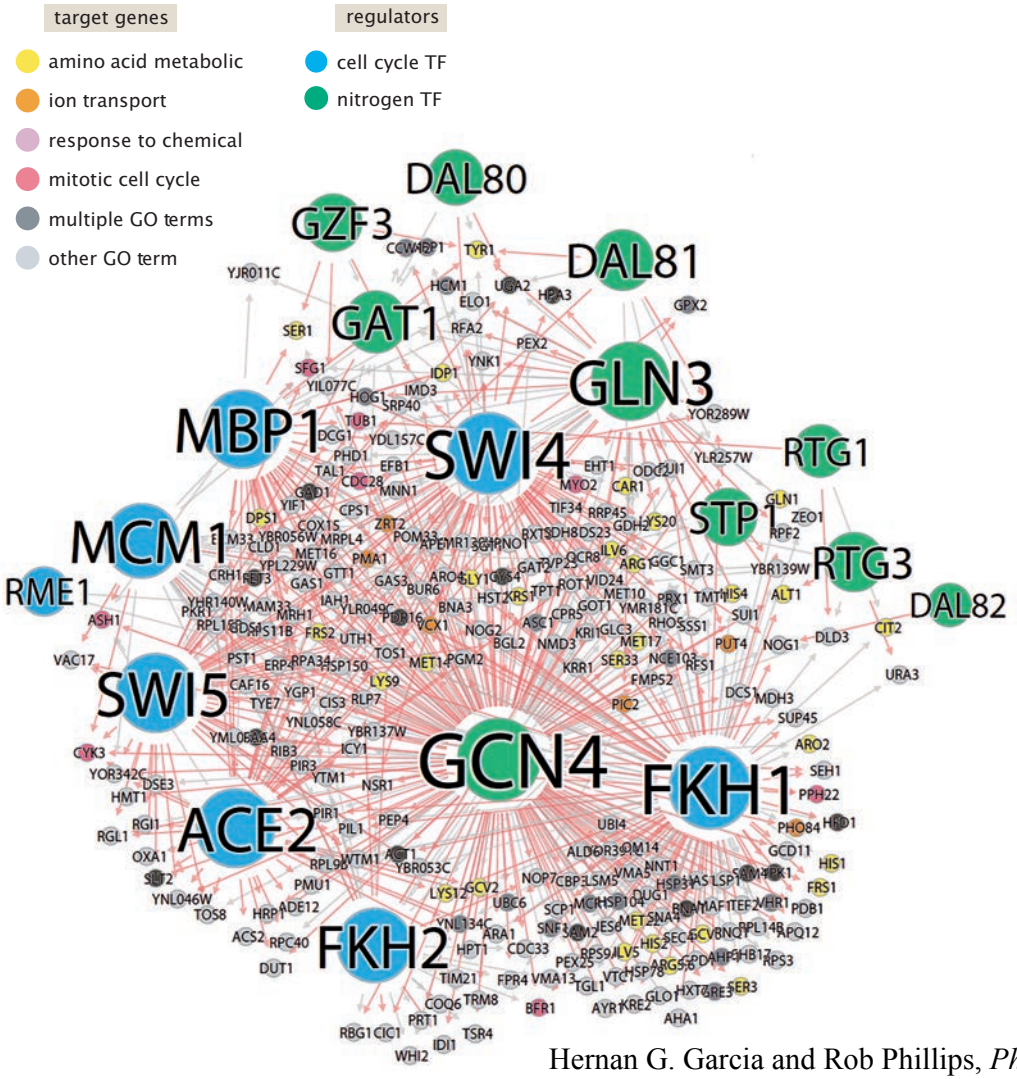
From Molecules to Function

1 network – 1 function

Nitrogen response in yeast and coordination with cell cycle

Target genes that are regulated by >one nitrogen-related transcription factor (green) or >one cell cycle-related transcription factor (blue).

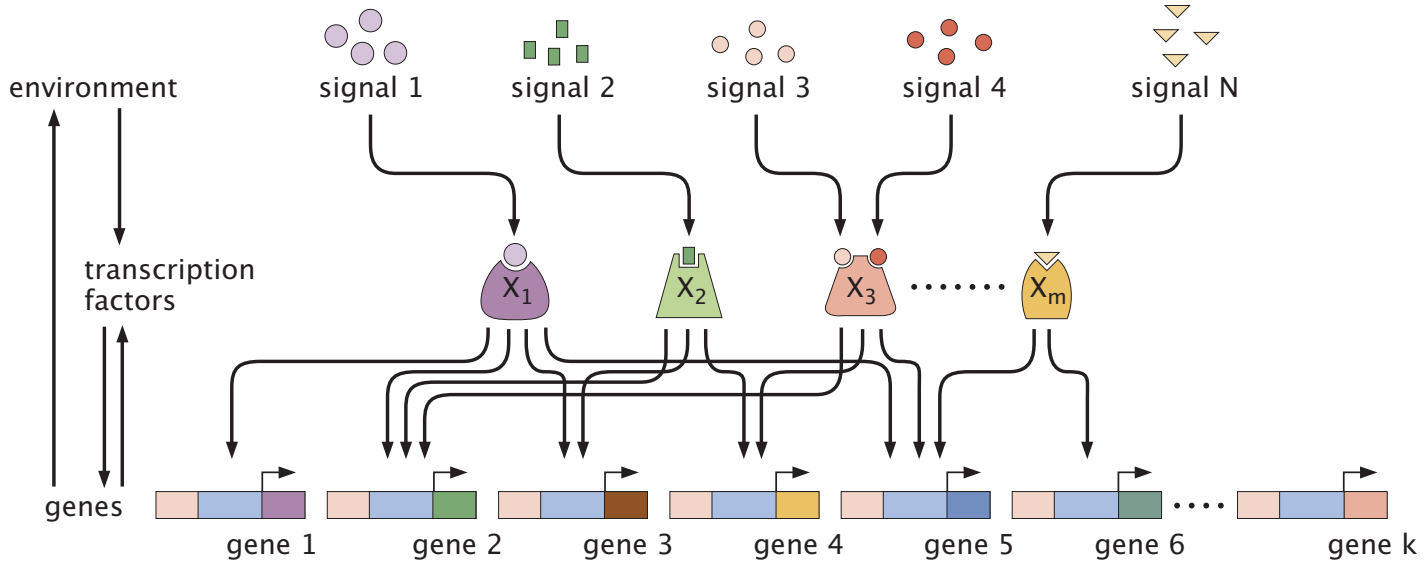
C. A. Jackson at al., *eLife* 9:e51254, 2020.



Hernan G. Garcia and Rob Phillips, *Physical genomics*

From Molecules to Function

- Transcription factors (TFs) are « symbols » that provide an internal representation of the environment. Combinatorial activity enriches this representation
- In *E. coli*, ~300 TFs provide a 300-dimensional representation of the world to regulate ~4500 genes.
- GRNs allow « predictions » of environmental changes: statistics of changes, ie. fluctuations and persistant change etc



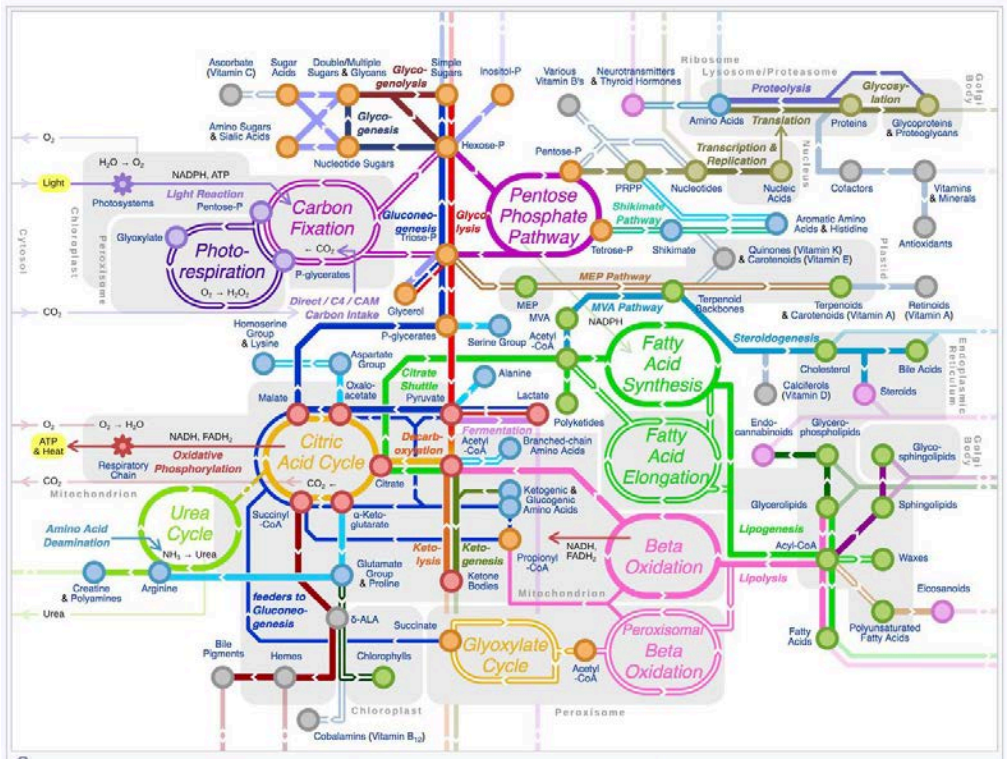
Hernan G. Garcia and Rob Phillips, *Physical genomics*

Chemical signalling pathways are complex

Metabolic pathways

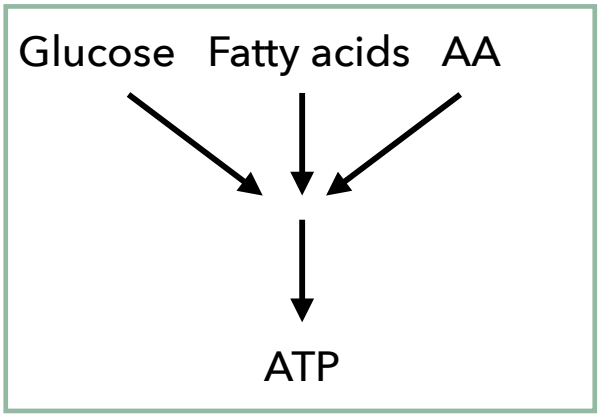
~10^4 different proteins per cell

Enzyme/substrate, Regulator/Enzyme

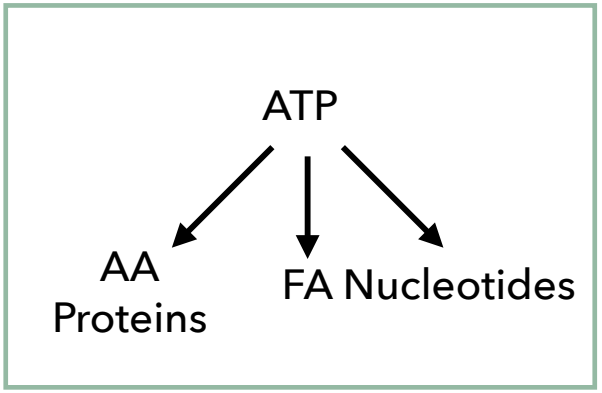


https://en.wikipedia.org/wiki/Template:Metabolic_metro

Catabolism

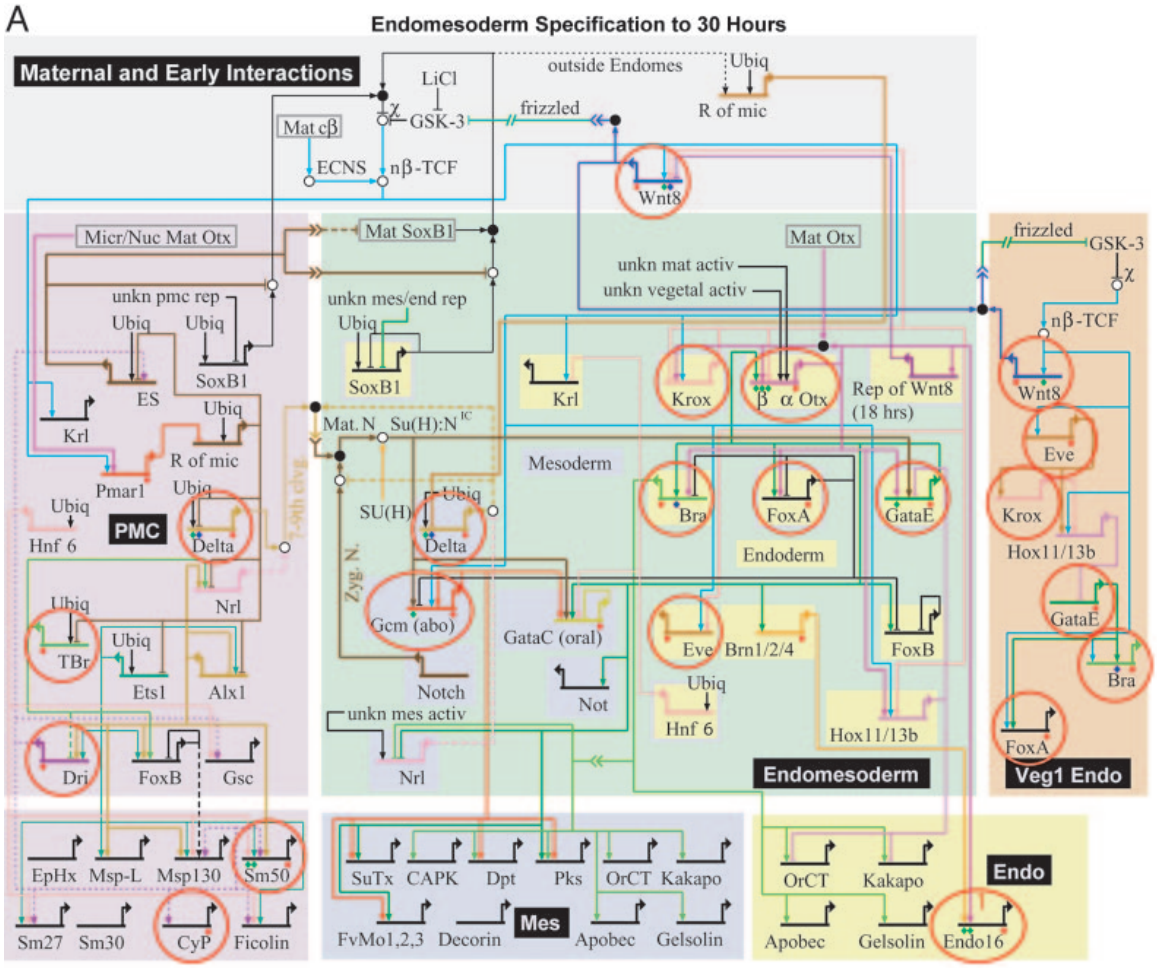


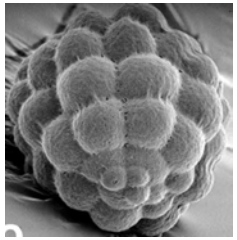
Anabolism



Chemical signalling pathways are complex

Developmental Gene regulatory networks

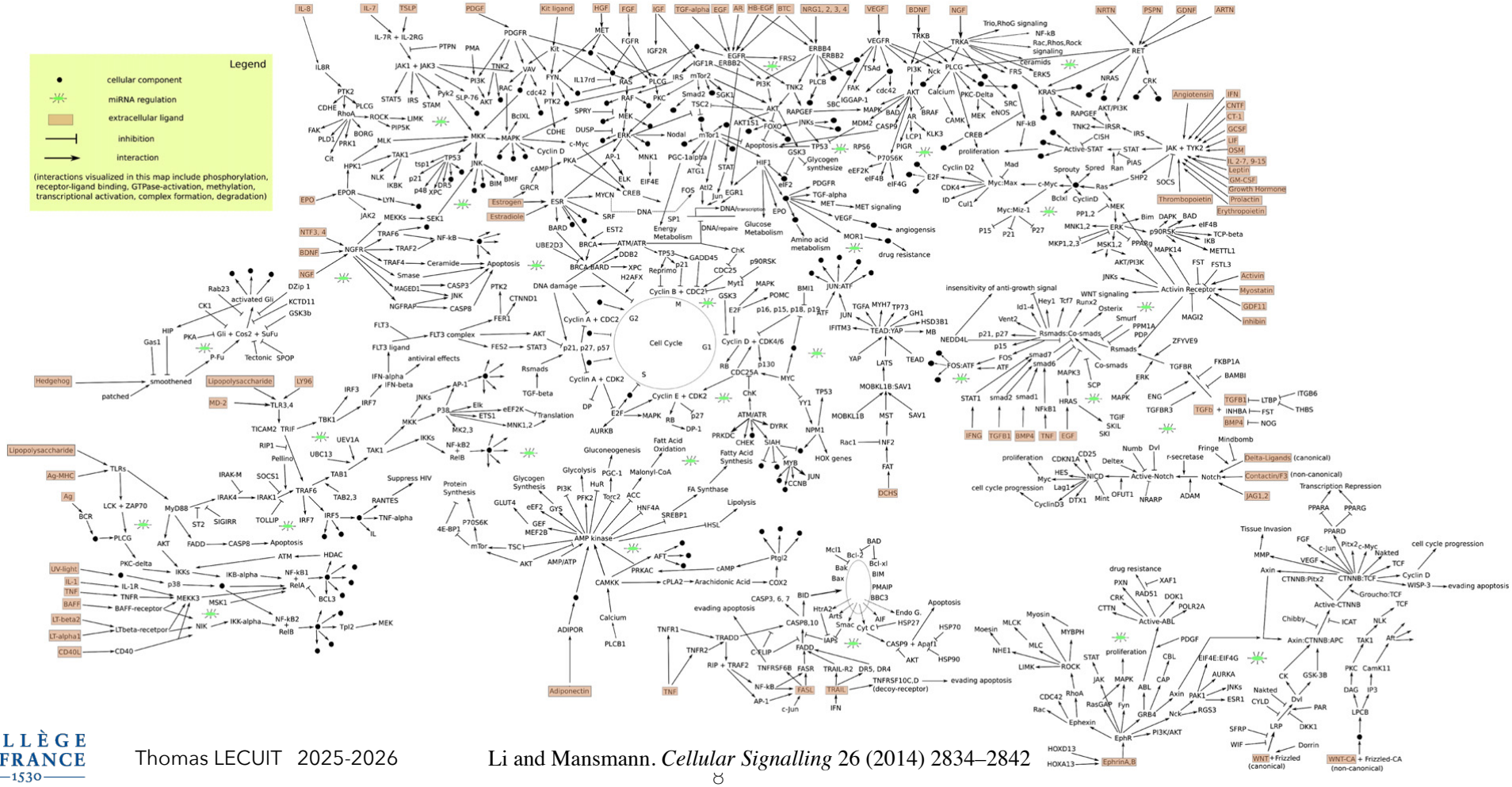




L. Bodenstein. *Mechanisms of Development*, 162 (2020)
https://doi.org/10.1016/j.mod.2020.103606

Chemical signalling pathways are awfully complex

Signalling pathways: from cell surface to gene regulation



Li and Mansmann. *Cellular Signalling* 26 (2014) 2834–2842

How to make sense of such complexity?

- Modular design of biochemical networks
- « Design principles » reflect evolutionary contraints

(ie. Marr's computational level: function/task)

From molecular
to modular cell biology

Leland H. Hartwell, John J. Hopfield, Stanislas Leibler and Andrew W. Murray

Cellular functions, such as signal transmission, are carried out by 'modules' made up of many species of interacting molecules. Understanding how modules work has depended on combining phenomenological analysis with molecular studies. General principles that govern the structure and behaviour of modules may be discovered with help from synthetic sciences such as engineering and computer science, from stronger interactions between experiment and theory in cell biology, and from an appreciation of evolutionary constraints.

Hartwell et al. *Nature* 402, 1999

- Modules are composed of many different **interacting molecules**.
- **Modules have discrete functions** that arise from interactions among their components (proteins, DNA, RNA etc).
- **Modules may be related by shared « design principles », even if they are not related by genomic sequences and genetics** (evolutionary descent).
- Conservation of algorithmic logic

How to make sense of such complexity?

- Modular design of biochemical networks
- 'Design principles' reflect evolutionary contraints

- Examples of modules from engineering:
  - *Positive feedback*: rapid transitions
    - Cell cycle transitions: G2/M or exit from mitosis
  - *Negative feedback*: homeostasis and adaptation
    - Adaptation in bacterial chemotaxis and sensitivity: reset over large range of input values.
  - *Coincidence detector*: requires 2 input for response
    - Transcription in eukaryotes requires multiples co-factors.
  - *Amplifier*: minimise effect of noise on signal
  - *Parallel circuits* (fail-safe): survive failures of one circuit

Hartwell et al. *Nature* 402, 1999

Marr's Tri-level of analysis

A framework to disentangle in complex processes:

- The *purpose/function* (why): *computational* level
- The *strategy* (how): *algorithmic* level
- The *biology/physics* (what): *implementational* level

Need to characterise the Algorithmic information

VISION

A Computational Investigation
into the Human Representation
and Processing of Visual Information

David Marr (1945-1980)

1982, Vision, David Marr
W. H. Freeman and Company
2010: *MIT press* (re-published)

Logic using Boolean networks

- Genes are in binary state (ON or OFF)
- Genome forms nets of genes
- Study the logic of network dynamics (attractors, cycles etc)

J. Theoret. Biol. (1969) **22,** 437–467

Metabolic Stability and Epigenesis in Randomly Constructed Genetic Nets

S. A. KAUFFMAN

Department of Anatomy, University of California Medical School, San Francisco, California, U.S.A.

and

Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts, U.S.A.†

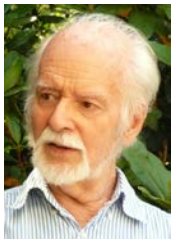
(*Received* 19 *March* 1968, *and in revised form* 8 *July* 1968)

"The world is either the effect of cause or chance. If the latter, it is a world for all that, that is to say, it is a regular and beautiful structure."

Marcus Aurelius

Proto-organisms probably were randomly aggregated nets of chemical reactions. The hypothesis that contemporary organisms are also randomly constructed molecular automata is examined by modeling the gene as a binary (on–off) device and studying the behavior of large, randomly constructed nets of these binary "genes". The results suggest that, if each "gene" is directly affected by two or three other "genes", then such random nets: behave with great order and stability; undergo behavior cycles whose length predicts cell replication time as a function of the number of genes per cell; possess different modes of behavior whose number per net predicts roughly the number of cell types in an organism as a function of its number of genes; and under the stimulus of noise are capable of differentiating directly from any mode of behavior to at most a few other modes of behavior. Cellular differentation is modeled as a Markov chain among the modes of behavior of a genetic net. The possibility of a general theory of metabolic behavior is suggested.

J. theor. Biol. (1973) **42,** 563–585



Boolean Formalization of Genetic Control Circuits

RENÉ THOMAS

Laboratoire de Génétique, Faculté des Sciences, et Laboratoire des systèmes logiques et numériques, Faculté des Sciences appliquées, Université libre de Bruxelles

(*Received* 25 *January* 1973, *and in revised form* 21 *June* 1973)

This paper is an attempt to formalize in Boolean terms genetic situations, from simple concepts like recessitivity and *cis*-dominance, to models describing complex control circuits.

A primary objective was to provide a language describing in a compact and unambiguous way, systems which become more and more difficult to describe as their complexity is being unravelled. Expression of a gene is

J. theor. Biol. (2001) **211,** 115–141
doi:10.1006/jtbi.2001.2335, available online at http://www.idealibrary.com on IDEAL®

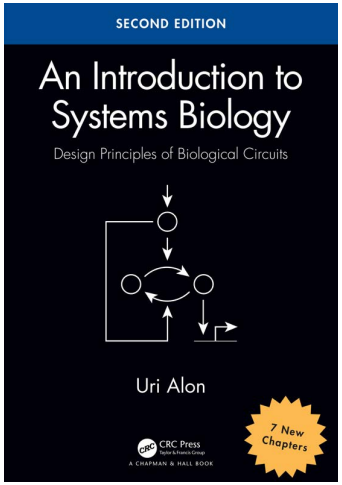
A Logical Analysis of the *Drosophila* Gap-gene System

LUCAS SÁNCHEZ*† AND DENIS THIEFFRY‡§

**Centro de Investigaciones Biológicas, Velázquez* 144, 28006 *Madrid, Spain, and* ‡*IBMM-ULB, B-*6041 *Gosselies, Belgium*

R. Milo et al and U. Alon. *Science*, 298: 824-827 (2002)

Network motifs



Uri Alon

Ron Milo

Weizmann Institute, Rehovot

Network Motifs: Simple Building Blocks of Complex Networks

R. Milo,[1] S. Shen-Orr,[1] S. Itzkovitz,[1] N. Kashtan,[1] D. Chklovskii,[2] U. Alon[1]*

R. Milo et al and U. Alon. *Science*, 298: 824-827 (2002)

Network motifs

Recurrent themes in biological networks: Positive Feedback

- Biological networks have internal structures and motifs that distinguish them from random networks
- Number of autoregulations in a random network :
  Arrows are distributed randomly among pairs of nodes
  Probability of self-activation among N nodes with A arrows

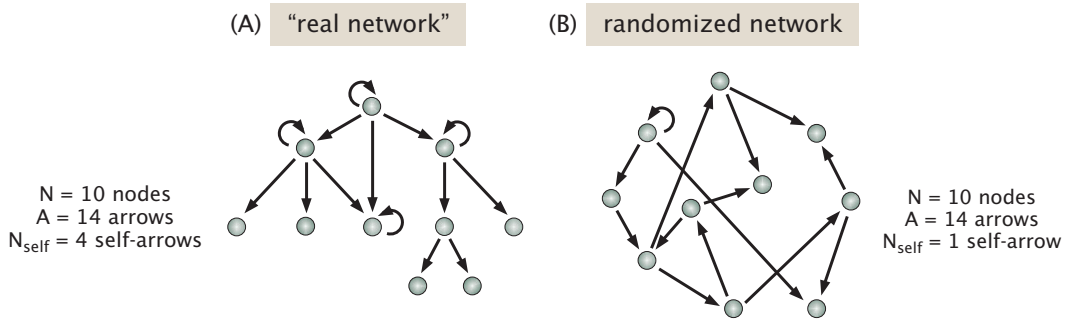
$$p_{self} = \frac{1}{N} \qquad \langle N_{self} \rangle = A p_{self} = \frac{A}{N}$$

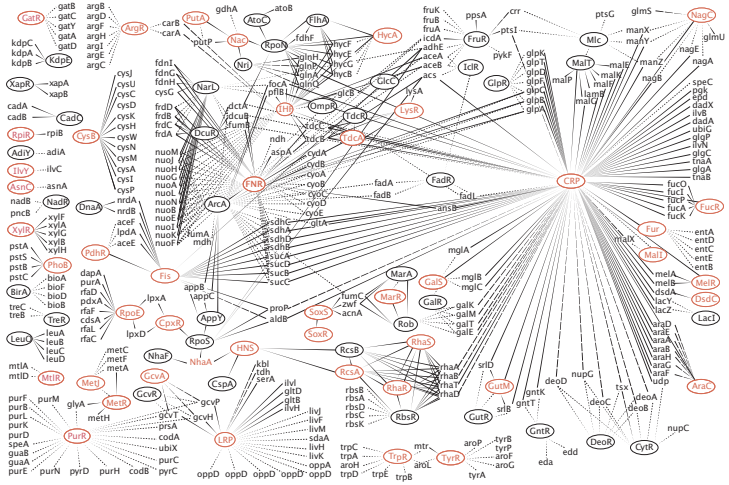
$$\sigma_{self} = \sqrt{\text{variance}} = \sqrt{\langle N_{self} \rangle} = \sqrt{\frac{A}{N}}$$

- Calculation from *E.coli* network:
  N=424, A=514

$$\langle N_{self} \rangle \approx 1.2 \quad \text{and} \quad \sigma_{self} \approx 1.1$$

Real network: $N = 44$
(~40 std compared to random network)





M. W. Covert, Fundamentals of Systems Biology: From Synthetic Circuits to Whole-cell Models, CRC Press, 2014.

Hernan G. Garcia and Rob Phillips, *Physical genomics*

Network motifs

Sensory system networks

- 1-2 nodes networks:
  - Simple regulation and autoregulation: response accelerator
- 3 nodes networks: Feedforward loops: Persistence detector, pulse generator, response accelerator
- Programmes (temporal, combinatorial)
  - Single input modules (SIMs)
  - Feedforward circuits (multi FFLs)
  - Dense overlapping regulons (DORs)

Developmental networks:

- Bistable networks: Memory
- Fold change detector

Most simple motifs/architectures

simple activation

NAR

negative autoregulation

PAR

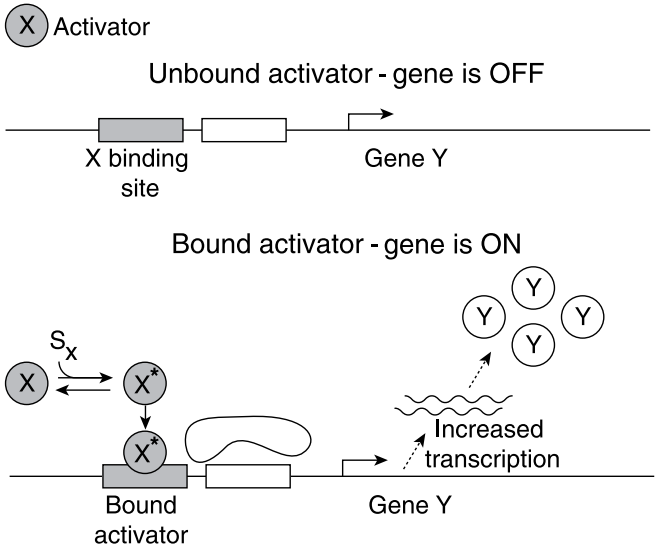
positive autoregulation

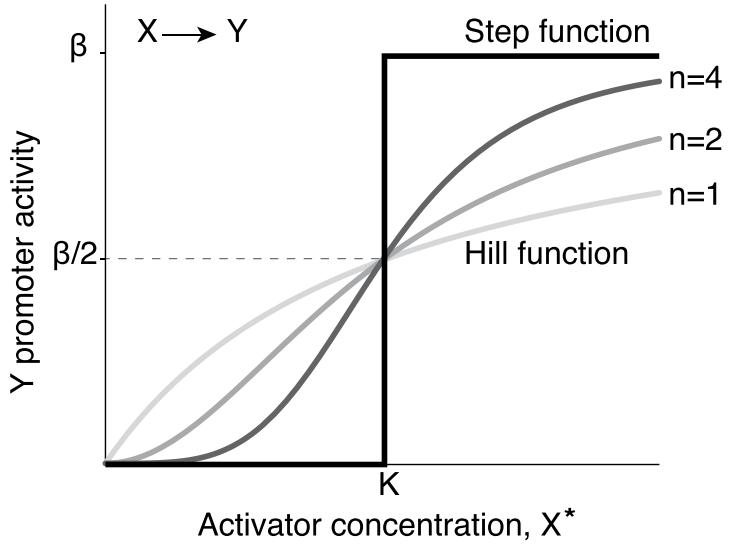
Simple regulation

X → Y

$$Y = f(X^*) = \beta \frac{X^{*n}}{K^n + X^{*n}}$$

- Three parameters characterise the production (transcription) of Y as a function of X
- K: Activation coefficient (M unit), defines the concentration at half production of Y
- β: maximal promoter activity of Y gene
- n: Hill coefficient, reflects cooperative effects (non-linearity) in the production of Y





U. Alon. *An introduction to systems biology*. 2020
Design principles in biological networks (2nd edition, CRC Press)

Simple regulation

X → Y

- Dynamics of response time.

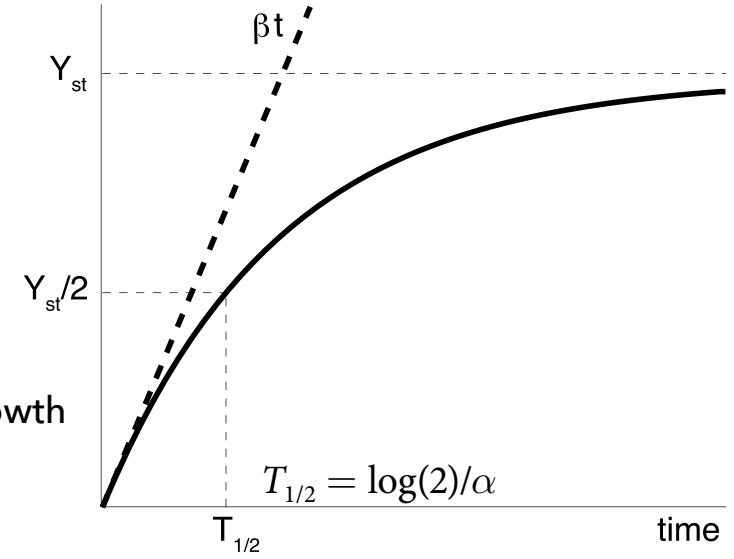
$$dY/dt = \beta - \alpha Y \qquad \text{where} \quad \alpha = \alpha_{dil} + \alpha_{deg} \qquad \text{(1/time)}$$

dilution due to cell growth

degradation

- At steady state: $Y_{st} = \beta/\alpha$

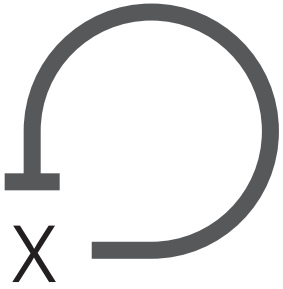
$$Y(t) = Y_{st}\,(1 - e^{-\alpha t})$$

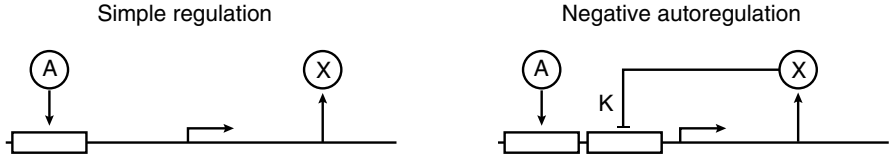


- The response time depends on α: $T_{1/2} = \log(2)/\alpha$
- In *E. coli* proteins are stable so α reflects dilution by growth
- Therefore the response time is the cell generation time which can be limiting in some conditions.
  (slower growth -> lower α_{dil} -> higher $T_{1/2}$).

U. Alon. *An introduction to systems biology*. CRC Press (2020)

Negative auto-regulation (NAR)

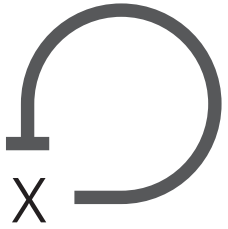




Negatively autoregulating transcription factors	Additional transcription regulation	Function
AraC	Crp	Arabinose utilization
ArgR		Arginine biosynthesis
Crp		Catabolite repression, global regulator
CysB		Cysteine biosynthesis
DsdC		Regulator of D-serine dehydratase
EmrRAB		Multidrug resistance pump
ExuR		Carbon utilization
Fis		rRNA and tRNA operons and DNA replication
Fnr		Aerobic, anaerobic respiration, osmotic balance
Fur	Crp	Iron transport
GalS	GalR	Galactose utilization
GcvA		Cleavage of glycine
GlnA	Crp, RpoN	Glutamine synthesis
Hns		Global regulator
Ihf		Integration host factor, global regulator
IlvY		Isoleucine and valine synthesis
LexA		SOS DNA repair
Lrp		Leucine response, amino acid limitation, global regulator
LysR		Lysine biosynthesis
MarA	Rob	Multiple antibiotic resistance
ModE		Molybdate transport
MtlADR		Mannitol utilization
Nac	RpoN	Histidine utilization/nitrogen assimilation
NadR		NAD biosynthesis, other roles
NagC	Crp	Repressor of genes for catabolic enzymes
OxyR		Oxidative stress
PhdR	Fis	Activator of hca cluster, other roles
PurR		Purine biosynthesis
PutAP	Nac	Proline synthesis, other roles
RpiR		Ribose catabolism
SoxS	SoxR	Superoxide stress
SrlA-D		Glucitol/sorbitol utilization
TrpR		Tryptophan biosynthesis
UxuABR	ExuR	Mannonate utilization

N. Rosenfeld, M. Elowitz and U. Alon 1 *J. Mol. Biol.* 323, 785–793 (2002)

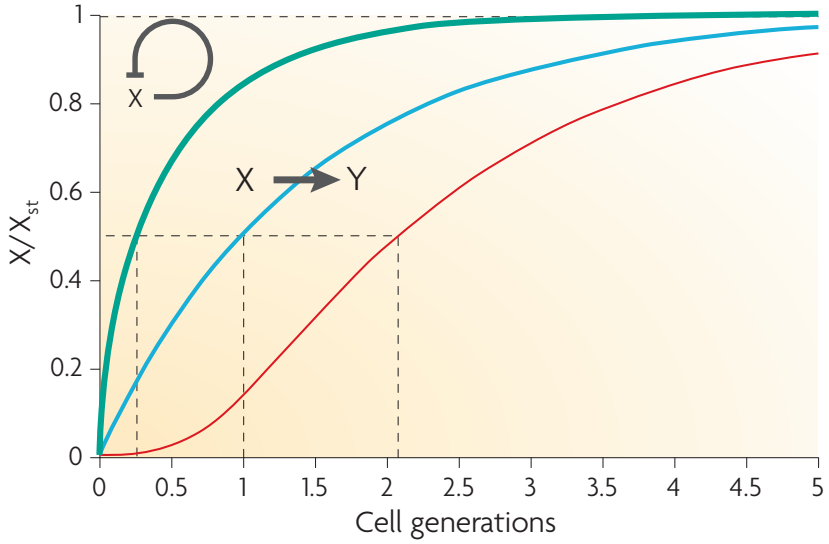
Negative auto-regulation (NAR)



Features (computation):

- *Acceleration of Response time*
- *Robustness* to promoter activity and degradation rate and growth rate

- Uses a strong promoter for constitutive activation leading to rapid activation of the promoter.
- When X reaches the negative feedback threshold, then X production slows down and finally reaches a steady state.
- To reach the same steady state, in simple regulation, the promoter must be weaker.

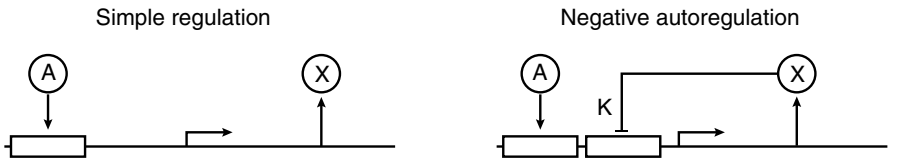


U. Alon *Nature Genetics* 8, 450-461 (2007)

Negative auto-regulation (NAR)

- *Acceleration* of Response time

$$\frac{dX}{dt} = f(X) - \alpha X \qquad f(X) = \frac{\beta K^n}{K^n + X^n}$$



K : concentration at which X represses at 50%

- **Logic approximation** (step function):

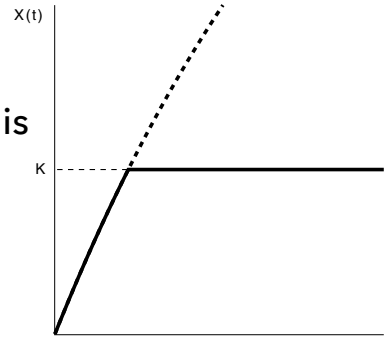
When $X < K$, $f(X) = \beta\theta(X < K)$ so $\frac{dX}{dt} = \beta - \alpha X$

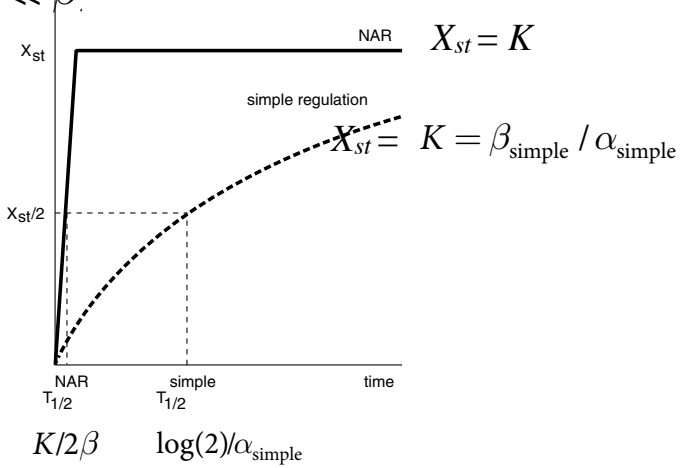
At early times, we can neglect degradation (or dilution due to growth) $\alpha X \ll \beta$

$X(t) \sim \beta t$

When $X = K$, X reaches self-repression threshold, production is zero and the steady state value of X is set by K (repression coefficient)

Response time in NAR is short and set by promoter strength: $T_{1/2}^{NAR} = K/2\beta$





For same steady state value X_{st} $\frac{T_{1/2}^{\mathrm{NAR}}}{T_{1/2}^{\mathrm{simple}}} = \frac{1}{2\log(2)} \frac{\beta_{\mathrm{simple}}}{\beta}$

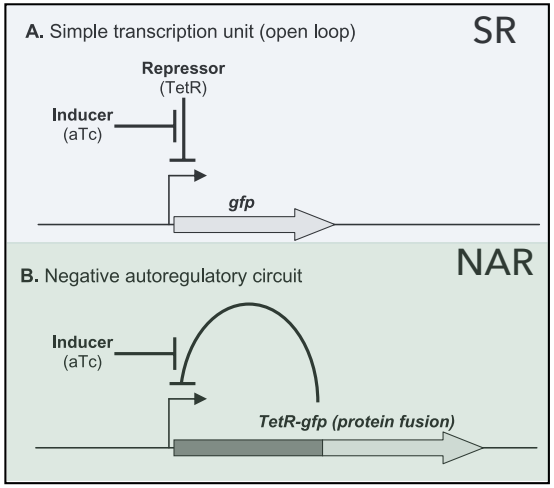
U. Alon. *An introduction to systems biology*. CRC Press (2020)

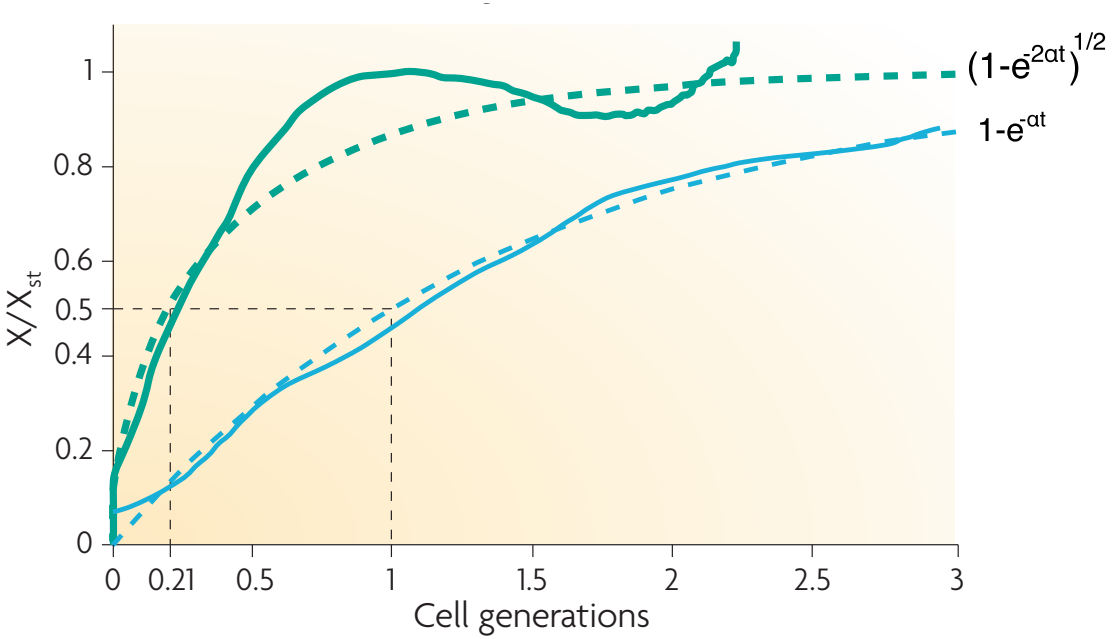
Positive and Negative auto-regulations

- *Acceleration* of Response time
  - Overshoot near X= K if delays, repression reduces X to below K, X is induced again etc (~oscillations)

  - Experimental test:

    TetR::GFP autoregulation by negative feedback





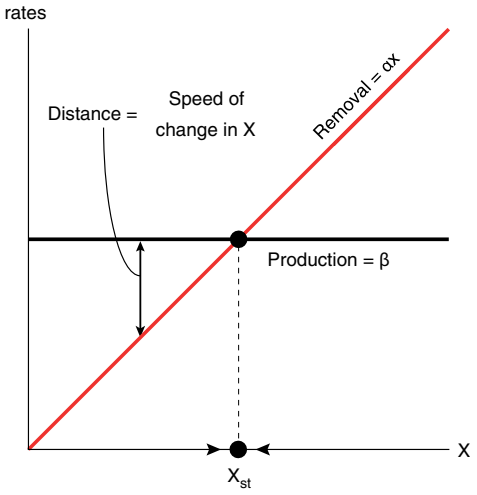
U. Alon *Nature Genetics* 8, 450-461 (2007)
N. Rosenfeld, M. Elowitz and U. Alon 1 *J. Mol. Biol.* 323, 785–793 (2002)

Negative auto-regulation (NAR)

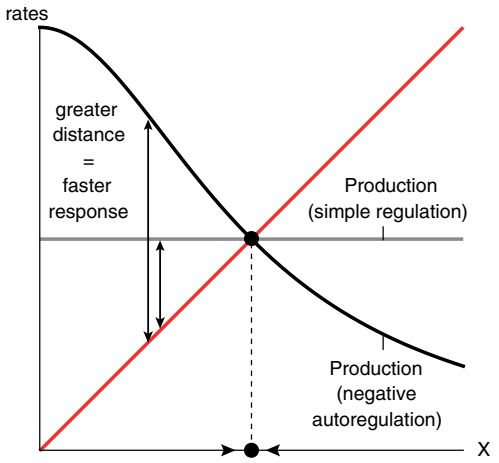
- *Acceleration* of Response time $\frac{dX}{dt} = f(X) - \alpha X$

- Simple regulation (SR):
  - Rate plot identifies a unique fixed point
  - the rate of approach to steady state is set by $\beta - \alpha X$
- More general case where f is *any* decreasing function of X
- NAR: whatever the shape of f, there is a unique fixed point and the rate of approach is greater than in SR

$$\frac{dX}{dt} = \beta - \alpha X$$



$$\frac{dX}{dt} = f(X) - \alpha X$$



U. Alon. *An introduction to systems biology*. CRC Press (2020)

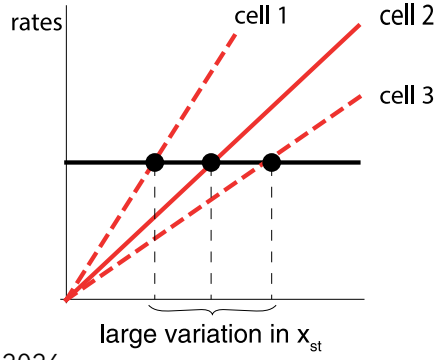
Negative auto-regulation (NAR)

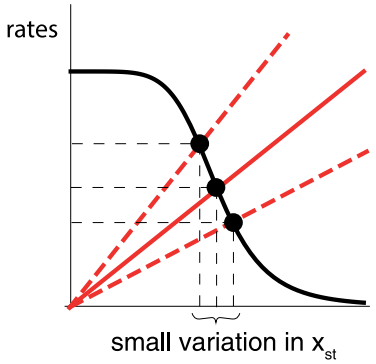
- *Robustness* to production rate and degradation & dilution/growth rate

$$\frac{dX}{dt} = f(X) - \alpha X \qquad \alpha = \alpha_{dil} + \alpha_{deg}$$

- For simple regulation SR, great sensitivity to promoter strength (β) since steady state value is β/α (Moreover the metabolic activity in *E. coli* has strong impact on promoter strength β)
- In Negative autoregulation (NAR), the steady state value only depends on *K* (eg. Binding strength of X to its own promoter). *K* which is robust to environmental perturbations/growth rate.

- In SR, the degradation rate and/or dilution rate (ie. growth rate) have a strong impact on the steady state value of *X*.
- But in NAR, if the negative feedback is non linear, degradation has a more limited impact on the steady state value.
- The greater the non linearity the greater the robustness to degradation/dilution/growth.

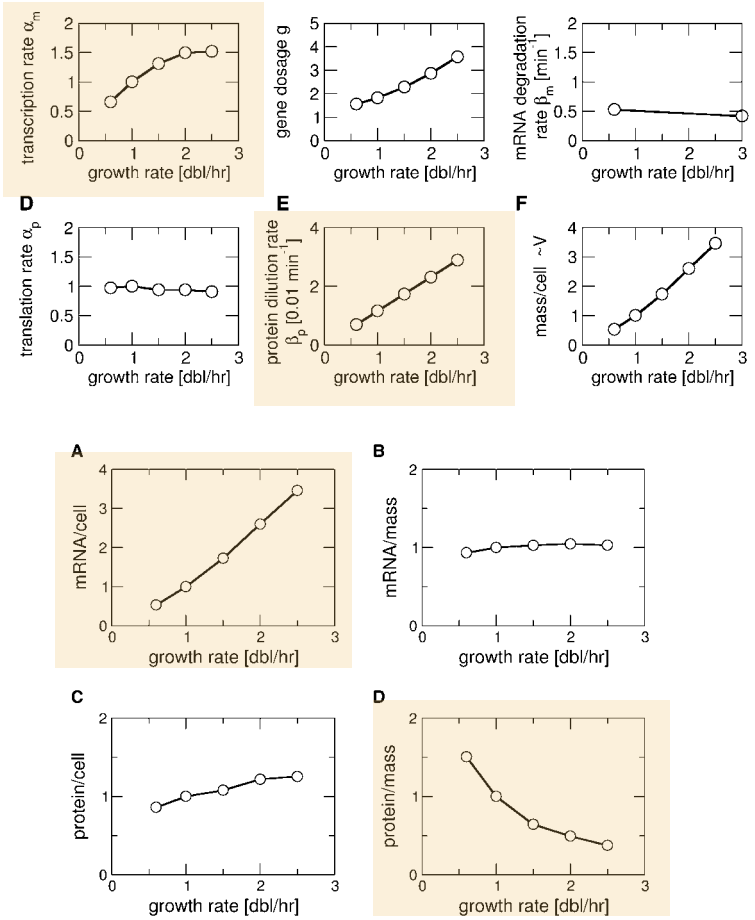




U. Alon. *An introduction to systems biology*. CRC Press (2020)

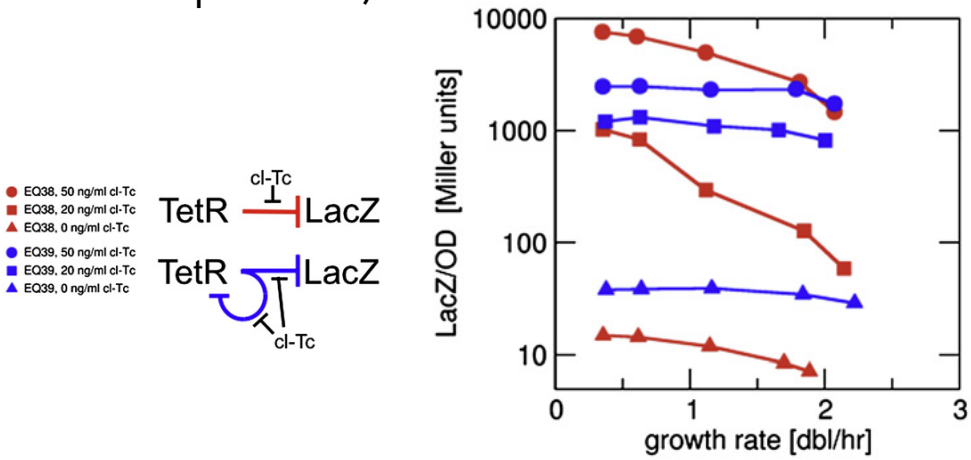
Negative auto-regulation (NAR)

- *Robustness* to production rate and degradation & dilution/growth rate



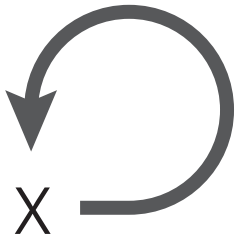
- Growth rate is dependent on the environment (eg. nutrients) and temperature: doubling time can vary from 20 min to several hours.
- <u>Cellular parameters are sensitive to growth rate.</u>

- Simple repression and negative auto regulation (NAR) are differently sensitive to changes in growth rate.
- NAR is robust to growth rate changes (dilution and transcription rate)

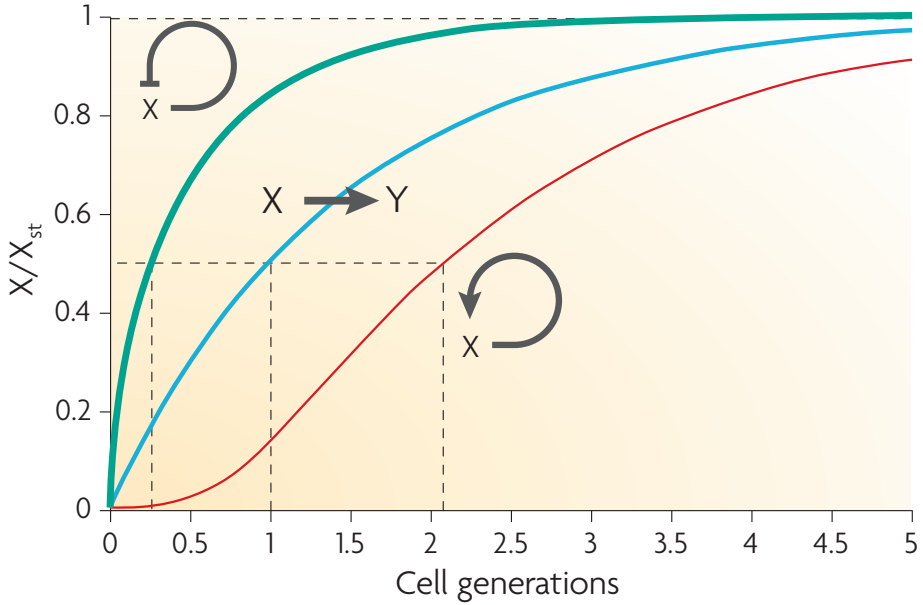


S. Klumpp et al and T. Hwa *Cell* 139, 1366–1375 (2009)

Positive auto-regulation (PAR)



PAR slows the response time because at early stages, when levels of X are low, production is slow.
Production increases when X concentration approaches the activation threshold for its own promoter.



Network motifs

Sensory system networks

- 1-2 nodes networks:
  - Simple regulation and autoregulation: response accelerator
- 3 nodes networks: Feedforward loops: Persistence detector, pulse generator, response accelerator
- Programmes (temporal, combinatorial)
  - Single input modules (SIMs)
  - Feedforward circuits (multi FFLs)
  - Dense overlapping regulons (DORs)

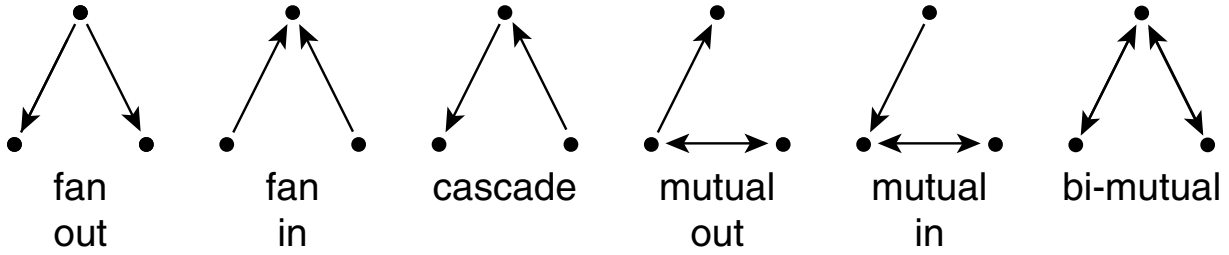
Developmental networks:

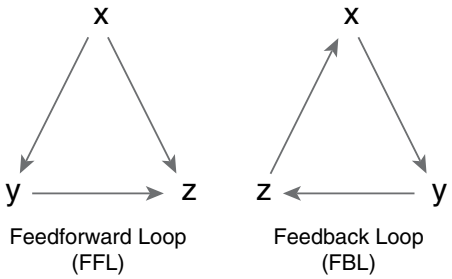
- Bistable networks: Memory
- Fold change detector

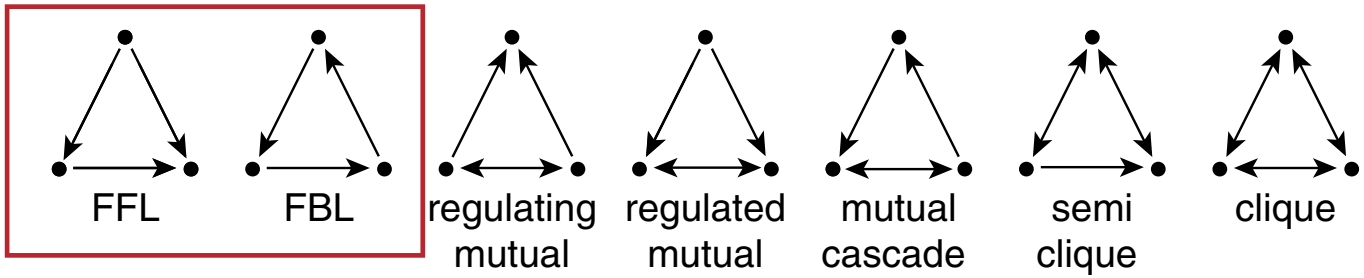
Network motifs

3-node connected subgraphs

13 possible networks







U. Alon. *An introduction to systems biology*. CRC Press (2020)

Network motifs

Overrepresentation of FFL subgraphs: feedforward loops.

- **Randomised networks:**

The mean number of feedforward loops $<N_{FFL}>$ is equal to the mean connectivity λ (mean number of arrows per node, $\lambda = A/N$) raised to the third power:

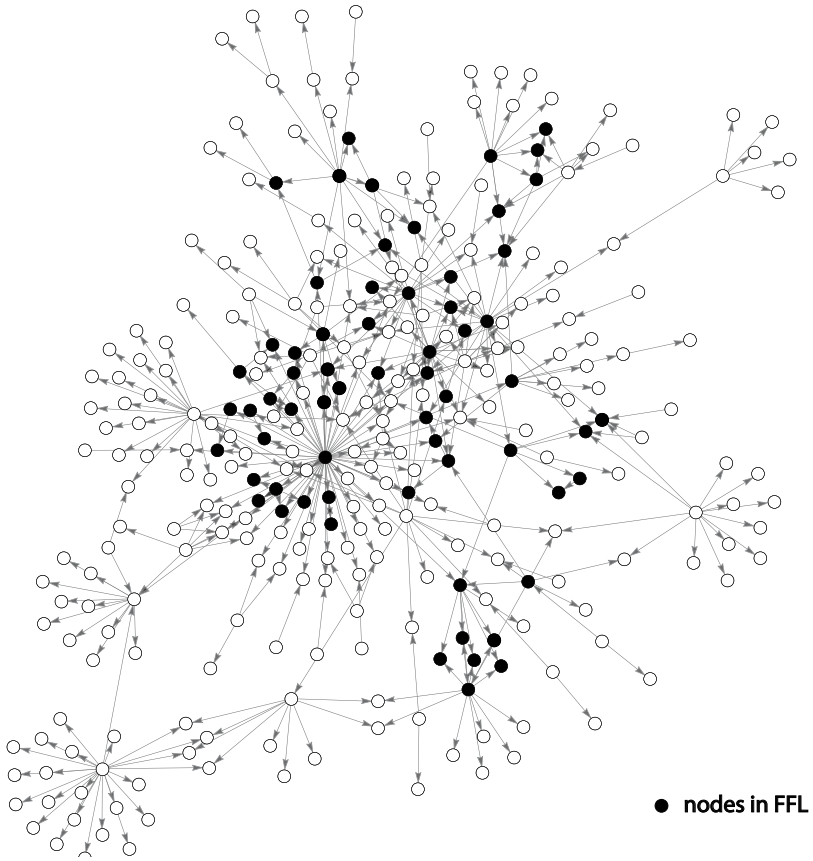
$$<N_{FFL}> = \lambda^3$$

A arrows and N nodes $\quad \langle N_{FFL} \rangle_{\text{rand}} = \lambda^3 \sim 1.7$

$\lambda = 500/400 \sim 1.2$

- **Biological networks** (*E. coli*): 42 FFL motifs, 0 FBL

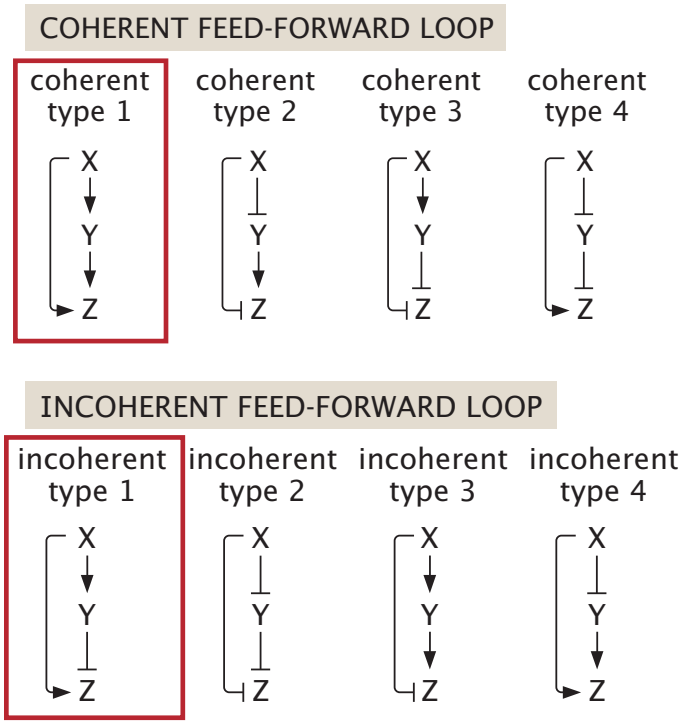
with same mean connectivity $\lambda = 1.2$

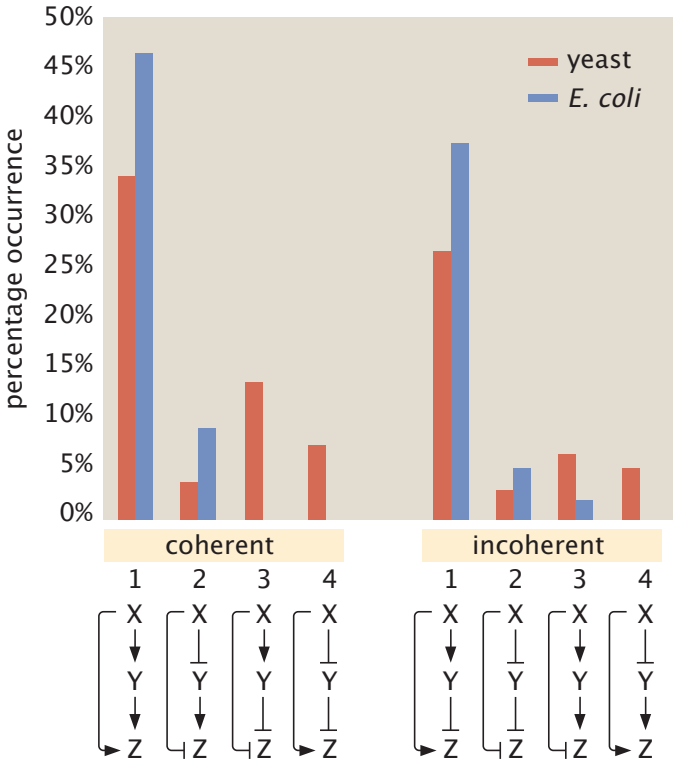


U. Alon. *An introduction to systems biology*. CRC Press (2020)

Feedfoward (FFL) motifs

- Arrows may represent activation or repression
- Direct and indirect branches in subgraph
- The 2 branches may be synergistic (coherent) or antagonistic (incoherent)

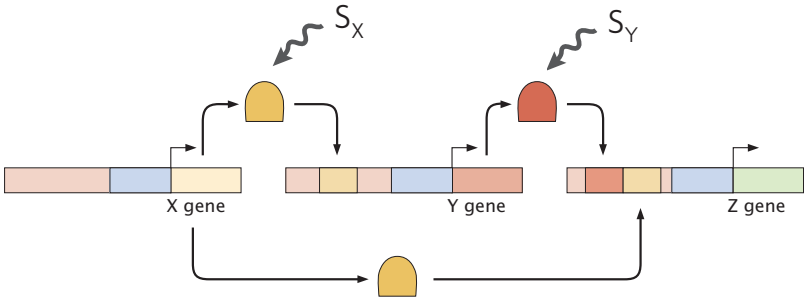


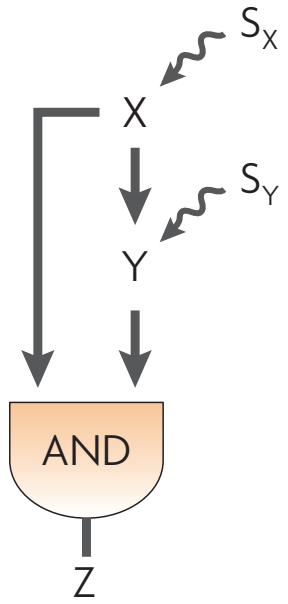


U. Alon. *An introduction to systems biology*. CRC Press (2020)

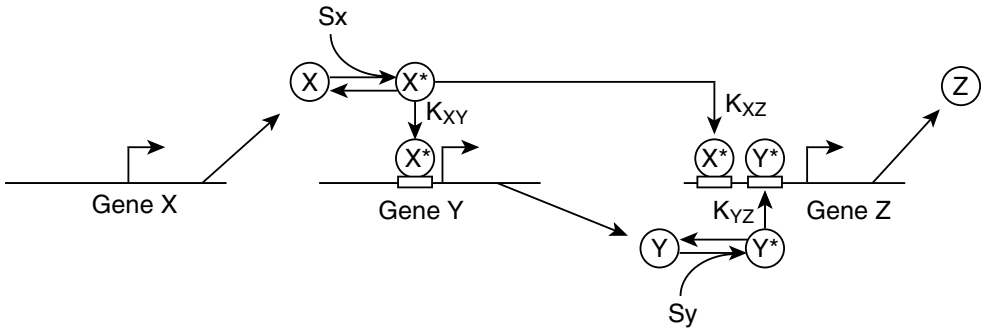
Hernan G. Garcia and Rob Phillips, *Physical genomics*

Coherent type 1 FFL



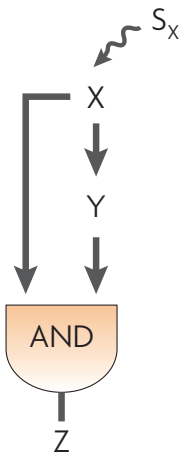


- X and Y are activated by stimuli S_X and S_Y to yield X* and Y*.
- **Let's consider an AND gate to control activation of Z by X and Y:**
  Z is transcribed if X AND Y are present above the activation thresholds K_{XZ} and K_{YZ}



U. Alon. *An introduction to systems biology*. CRC Press (2020)
Hernan G. Garcia and Rob Phillips, *Physical genomics*

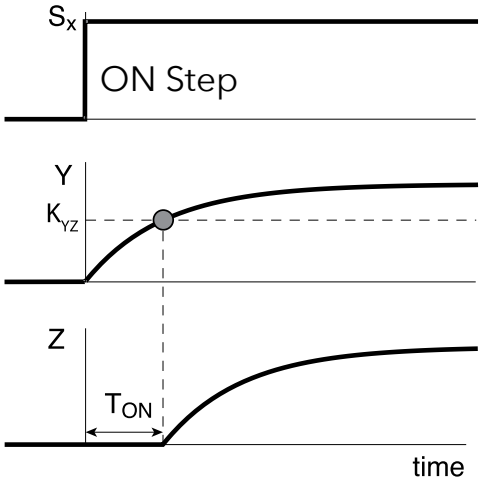
Coherent type 1 FFL



- Mathematical formulation of C1-FFL dynamics:
- ON Step function of S_X which activates X instantaneously.

$$\frac{dY}{dt} = \beta_Y \theta(X^* > K_{XY}) - \alpha_Y Y = \beta_Y - \alpha_Y Y$$

$$\frac{dZ}{dt} = \beta_Z \theta(X^* > K_{XZ}) \theta(Y^* > K_{YZ}) - \alpha_Z Z \quad \text{(AND gate)}$$



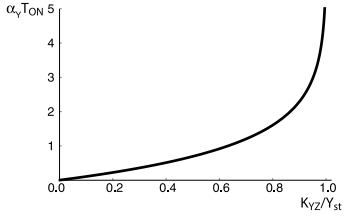
- Simple induction of Y by X:

$$Y^* = Y_{st}(1 - e^{-\alpha_Y t})$$

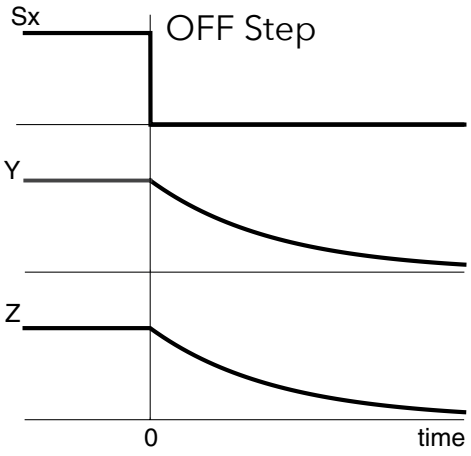
- Simple induction of Z by Y shows a delay:
- Z induction requires that Y accumulates to pass the activation threshold K_{YZ}
- We can calculate the delay T_{ON} until $Y = K_{YZ}$:

$$Y^*(T_{\mathrm{ON}}) = Y_{st}(1 - e^{-\alpha_Y T_{\mathrm{ON}}}) = K_{YZ}$$

$$T_{\mathrm{ON}} = \frac{1}{\alpha_Y} \log\left(\frac{1}{1 - K_{YZ}/Y_{st}}\right)$$



- Delay T_{ON} can evolve independently as a function of α_Y and K_{YZ}
- OFF Step of S_X : there is **no delay in downregulation of Y and Z.**

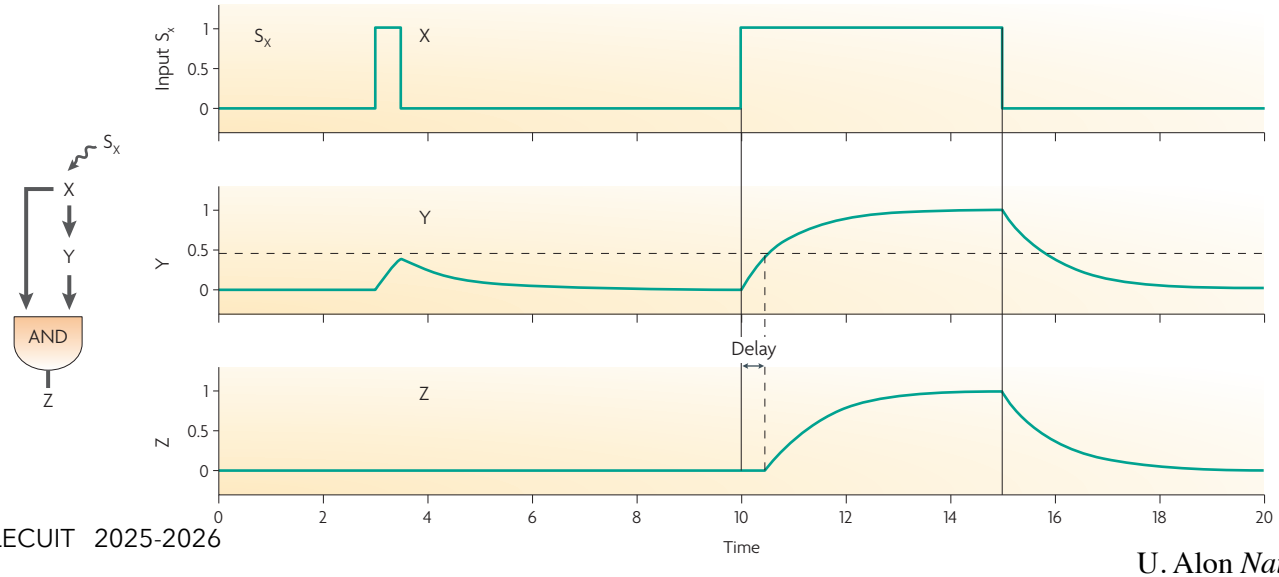


Coherent type 1 FFL

- Computation/Function:
- The C1-FFL with AND gate is a persistence detector and an asymmetric filter.

- A short pulse of S_X or X, ie. shorter than the delay $T_{ON,}$ does not induce Z (because Y is below the activation threshold K_{XY})
- Only a sufficiently persistent signal will activate Z.
- Therefore the C1-FFL works as an asymmetric filter, so that only meaningful inputs cause a response.

- Bacteria are exposed to a fluctuating environment. Cells respond to long enough changes in nutrient concentrations.
- The delay T_{ON} is an internal representation of environmental fluctuations that help the system make good predictions.

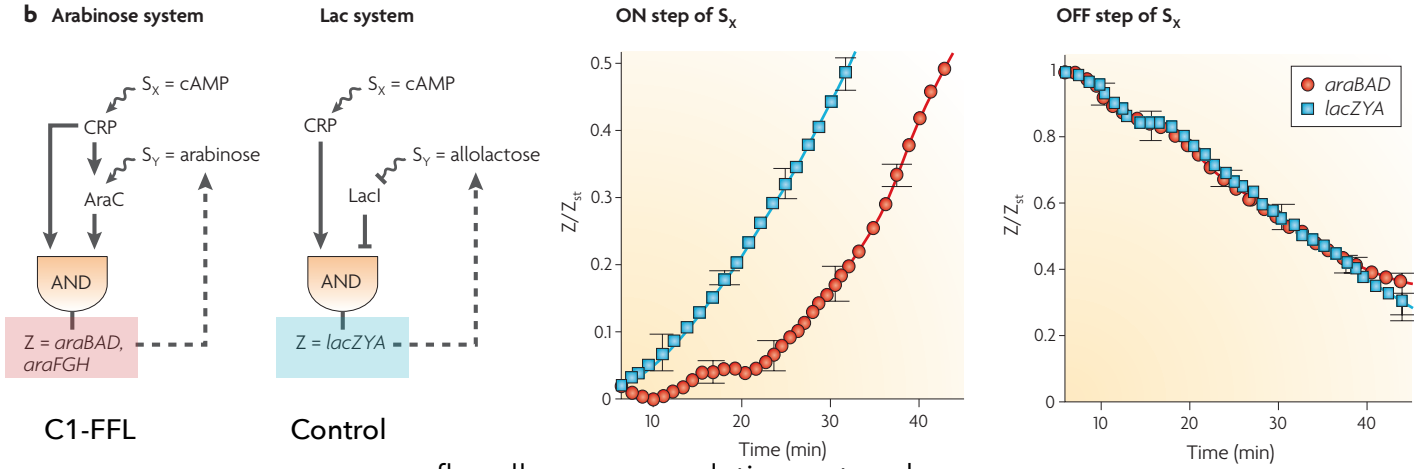


U. Alon *Nature Genetics* 8, 450-461 (2007)

Coherent type 1 FFL

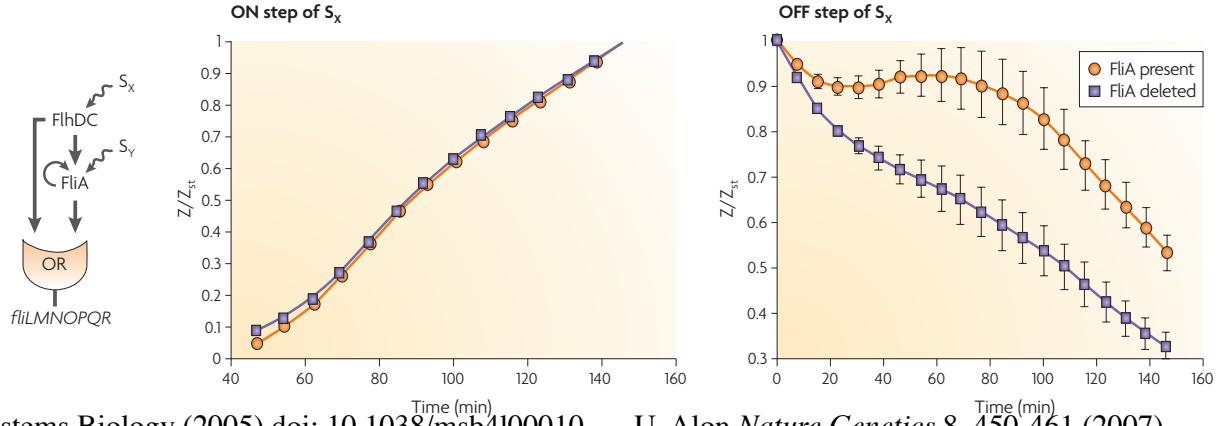
• The C1-FFL with AND gate is a persistence detector

- **Experimental test:**
- The Arabinose system: arabinose is an energetically unfavourable nutrient compared to glucose.
- Activation of enzymes to digest arabinose requires the double detection of **absence of glucose (presence of cAMP :X) AND the presence of arabinose** (Y).



- Experiments show a delay of ~20min for the induction (ON step). No delay for the OFF step.

flagella gene-regulation network



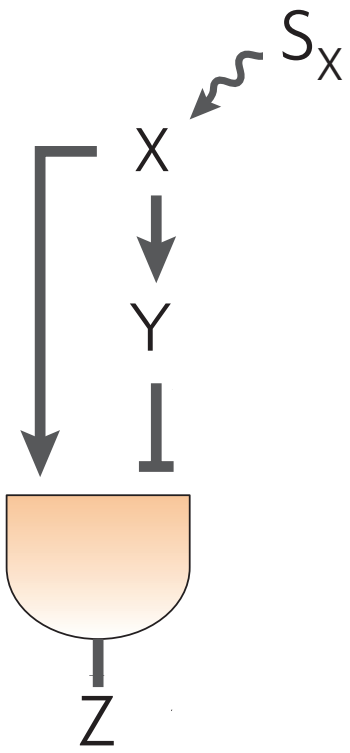
- C1-FFL with OR gate is also an asymmetric filter. The removal of S_X causes a delayed response.
- In the simple regulation, there are no delay.
- It protects against transient loss in input signal

Molecular Systems Biology (2005) doi: 10.1038/msb4100010 U. Alon *Nature Genetics* 8, 450-461 (2007)

Incoherent type 1 FFL

- Computation/Function: a pulse generator and response accelerator

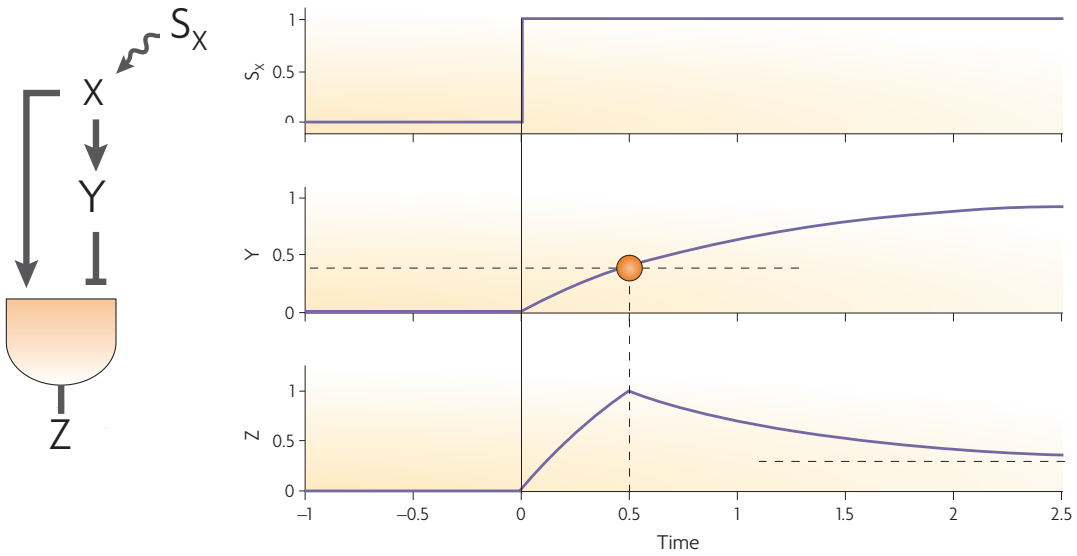
- The two arms are in opposition



Incoherent type 1 FFL

- Computation/Function: a pulse generator and response accelerator

- This motif can use a strong activation dynamics of Z by X with a strong promoter/synthesis rate.
- Y is also induced (at a slower rate). **When Y passes the activation threshold, it represses Z to a steady state.**
- This results in a **pulse of expression** of Z.

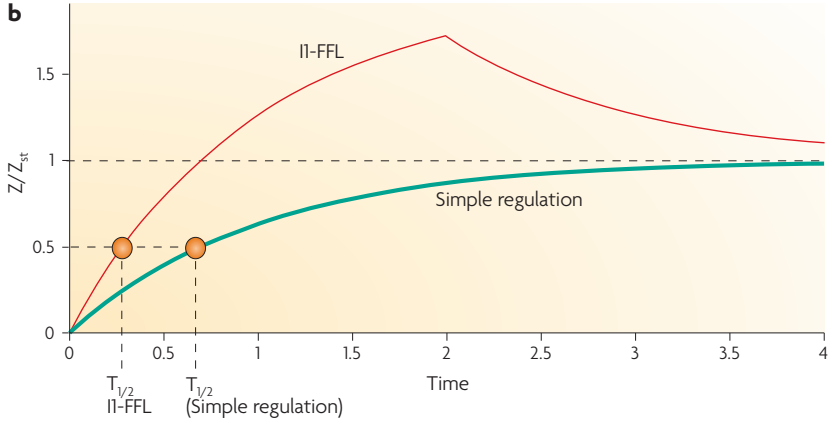


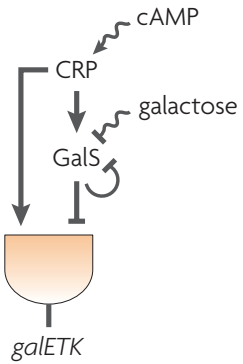
- The shape of the pulse depends on tuning of the activation threshold of Y (K_{XY}) and activation strength of Z by X/S_X (β_Z)

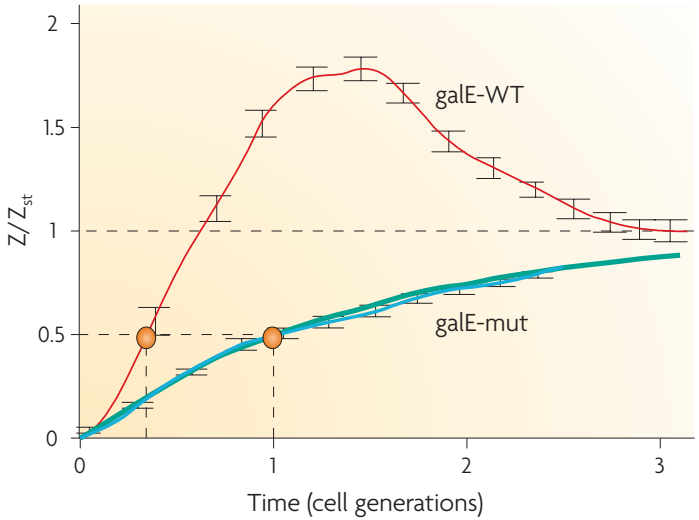
Incoherent type 1 FFL

- Computation/Function: a pulse generator and response accelerator

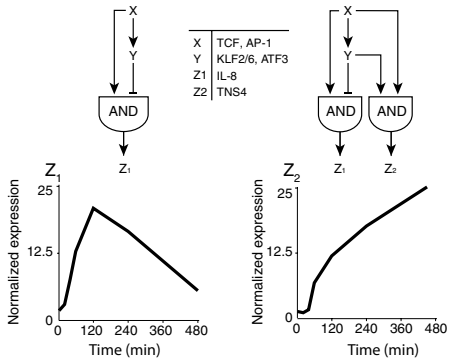
- **Response acceleration:**
- Rapid initial activation of Z by X that is later turned off by a delayed repression by Y.
- Simple regulation is slower to reach same steady state.







- **Experimental test:**
- The galactose system: galactose is an energetically unfavourable carbon source
- The absence of glucose (presence of cAMP) AND of galactose are detected to induce enzymes to metabolise galactose. When galactose is present, enzymes are available.



I. Amit et al, and Y. Yarden. *Nature Genetics*, 39:503-512 (2007)

Network motifs

Sensory system networks

- 1-2 nodes networks:
  - Simple regulation and autoregulation: response accelerator
- 3 nodes networks: Feedforward loops: Persistence detector, pulse generator, response accelerator
- Programmes (temporal, combinatorial)
  - Single input modules (SIMs)
  - Feedforward circuits (multi FFLs)
  - Dense overlapping regulons (DORs)

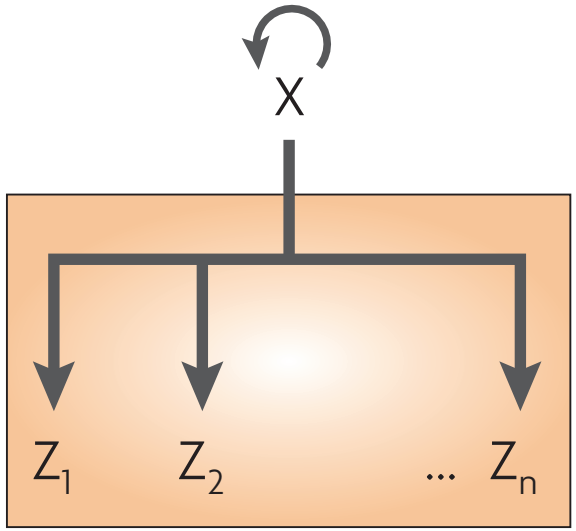
Developmental networks:

- Bistable networks: Memory
- Fold change detector

Single input module (SIM): temporal programme

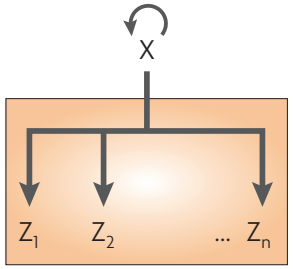
- Computation/Function: coordinated expression of different genes with shared function

- All genes are regulated by common «master » regulator X
- X is often autoregulatory.

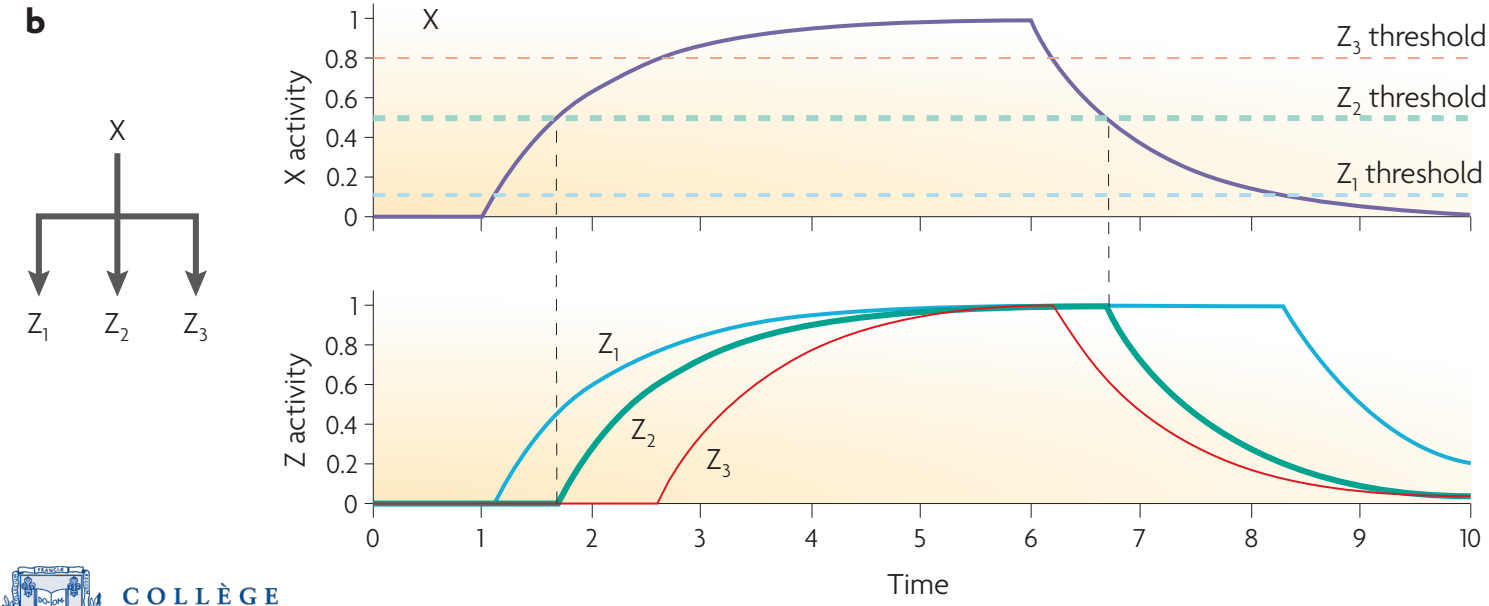


Single input module (SIM): temporal programme

- Computation/Function: coordinated expression of different genes with shared function

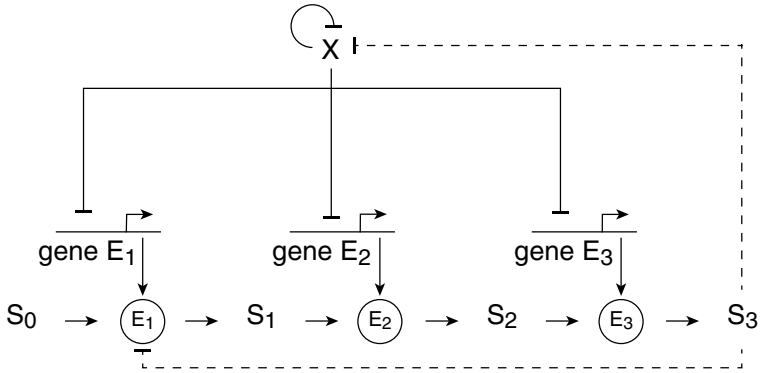


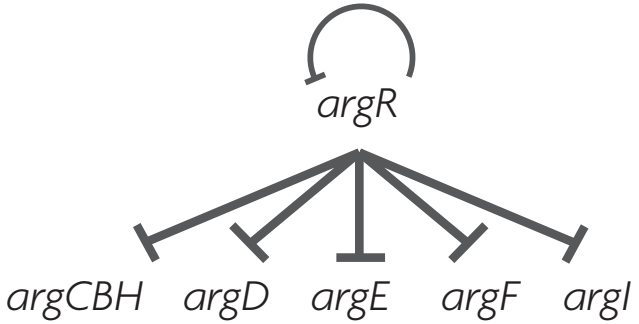
- Independent tuning of delay times (eg. activation thresholds) for different target genes Z_i (eg. promoter sequences) allow sequential expression of target genes in response to rising input X.



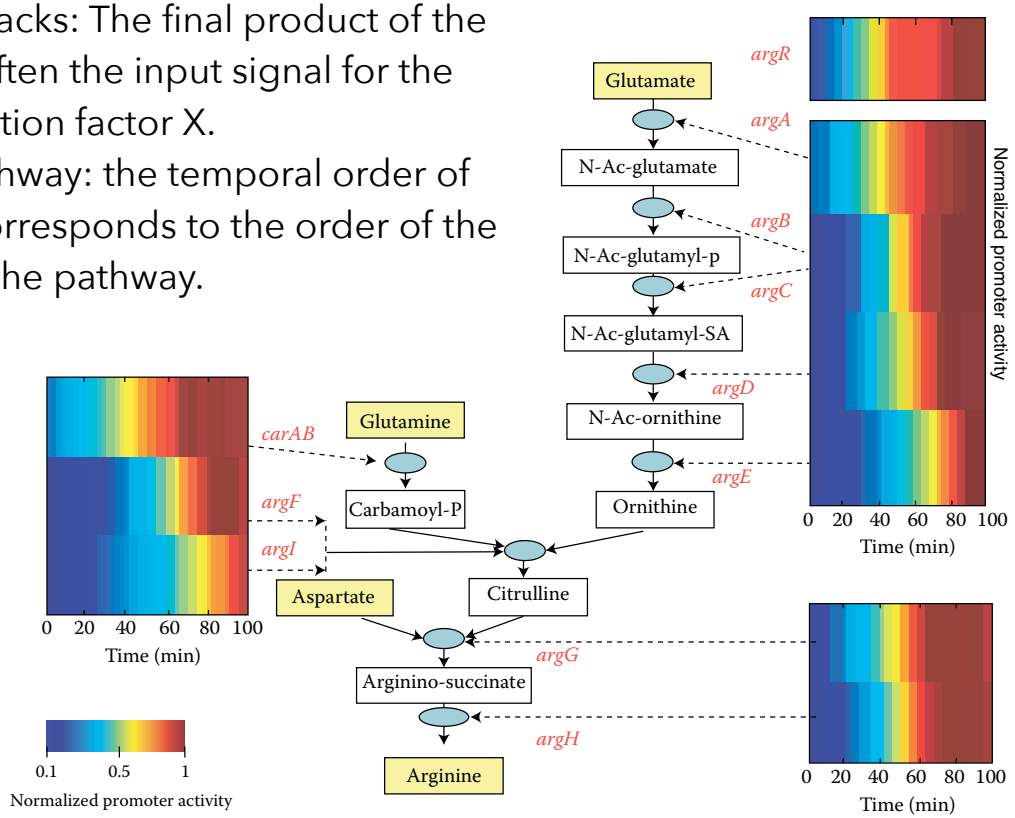
- If X is itself a pulse (eg. as output of I1-FFL) then target genes have a specific temporal programme.
- **The Last IN is the First OUT (LIFO)**

Single input module (SIM): temporal programme





- Other feedbacks: The final product of the pathway is often the input signal for the top transcription factor X.
- Arginine pathway: the temporal order of the genes corresponds to the order of the reactions in the pathway.



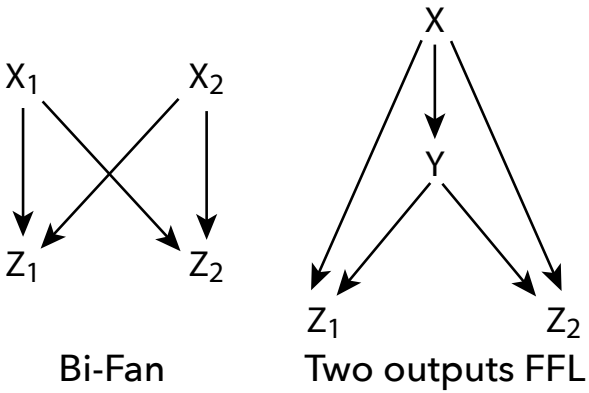
U. Alon. *An introduction to systems biology*. CRC Press (2020)

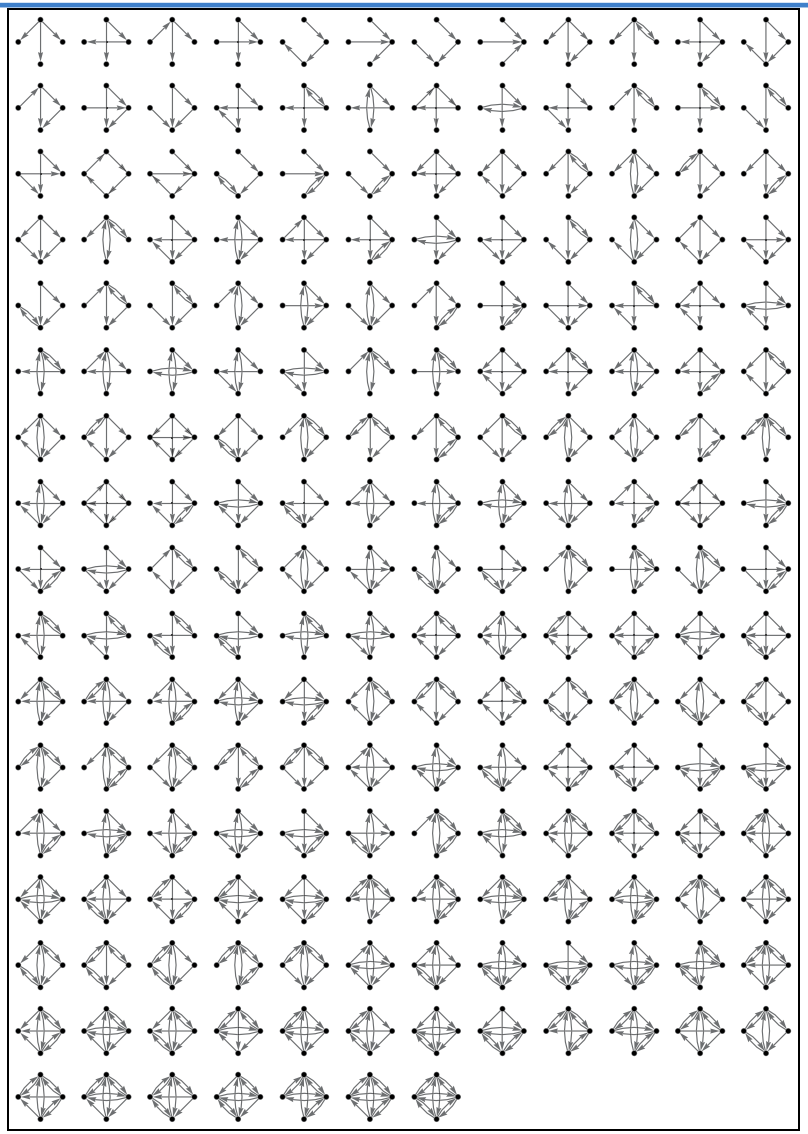
Zaslaver, A. et al. *Nature Genetics,* 36(5):486–491 (2004)

Feedfoward circuits

4-node connected subgraphs

- 199 possible networks
- Only 2 are prominent



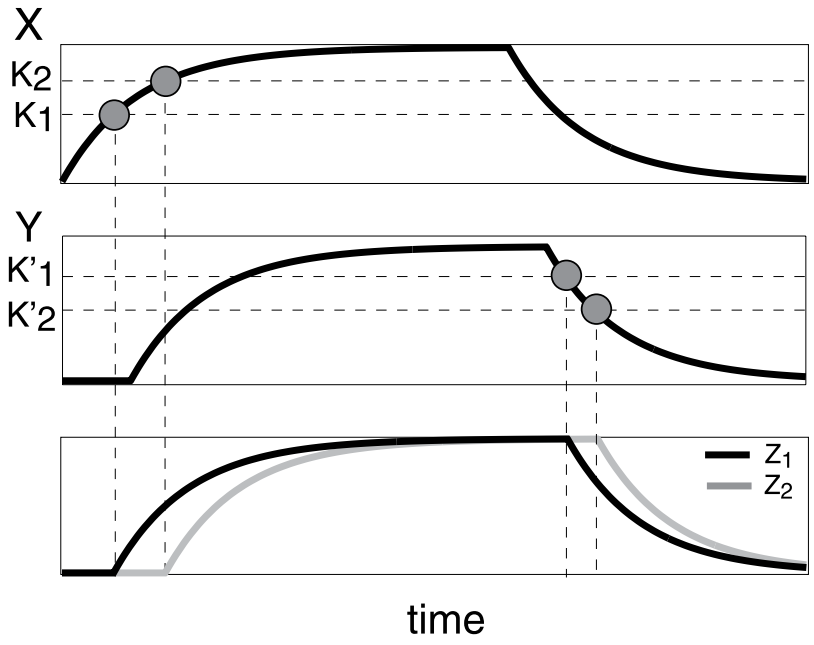


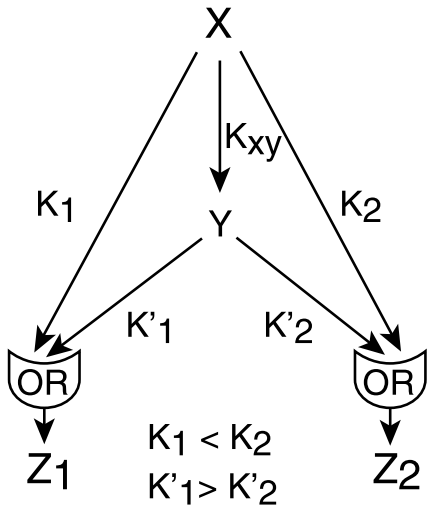
U. Alon. *An introduction to systems biology*. CRC Press (2020)

Two-output coherent FFL

First IN First OUT (FIFO) logic

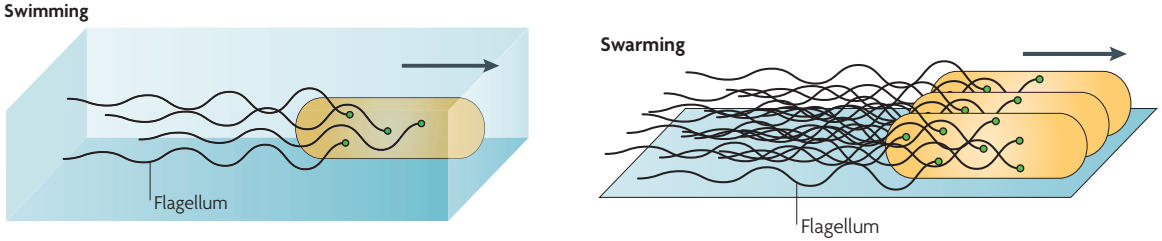
- **OR gates:** the activation order of Z_1 and Z_2 depends on the respective activation thresholds K_1 and K_2 by X.
- Conversely, down regulation order depends on respective activation thresholds K'_1 and K'_2 by Y.



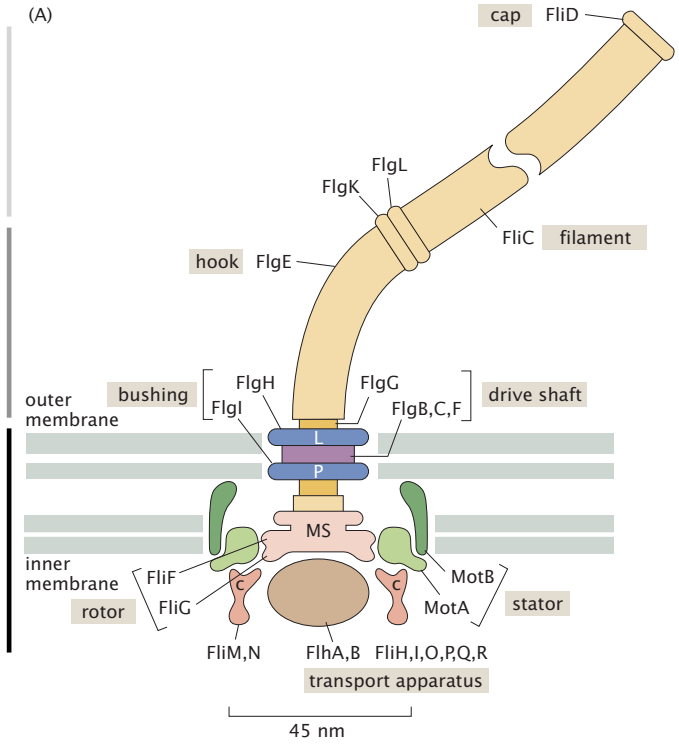


Multi-output FFL

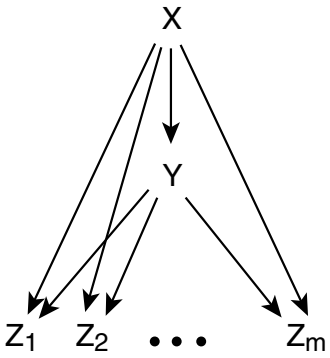
- **Increased viscosity of the environment is associated with a switch to swarm behaviour**
  - –Induction of flagellar genes and increase in number of flagella per cell
  - Flagella cover the entire cell (peritrichous flagella)
  - Peritrichous flagella bundle together when they rotate to increase the effective flagellar stiffness and make force generation more efficient in viscous liquids



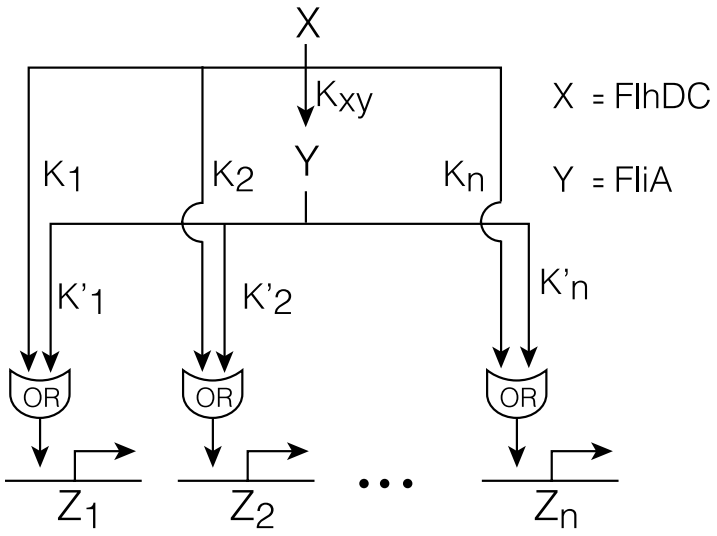
Kearns DB (2010) *Nat Rev Microbiol* 8(9): 634–644. (2010)

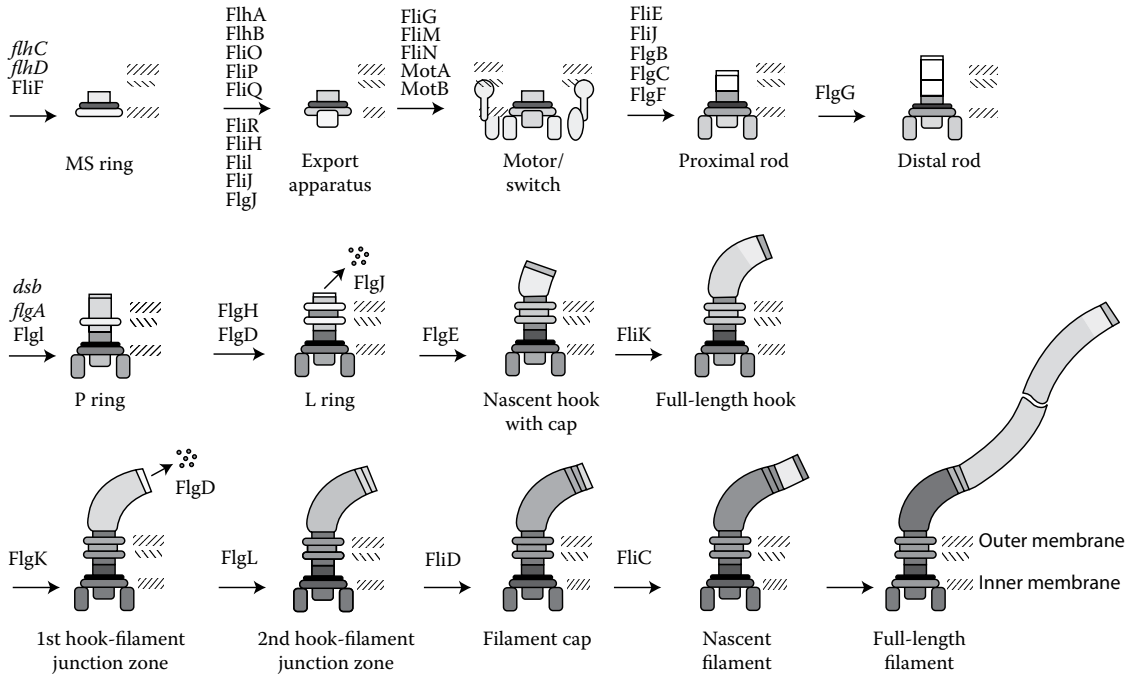


Multi-output coherent FFL



- Hierarchical organisation of activation thresholds K_i and K'_i
- K_i and K'_i are in reverse order





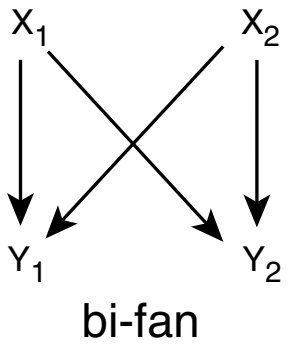
U. Alon. *An introduction to systems biology*. CRC Press (2020)

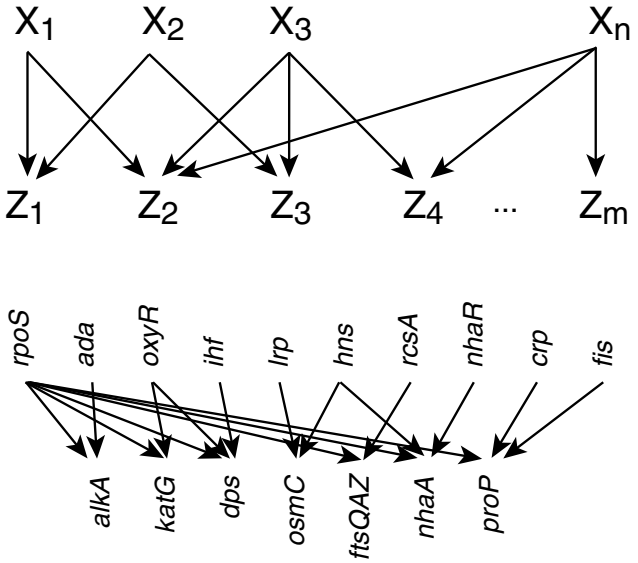
Macnab, R.M. *Annual Review of Microbiology.* 57(1): 77–100 (2003)

Bi-Fans and Dense Overlapping Regulons (DORs)

A combinatorial decision making device

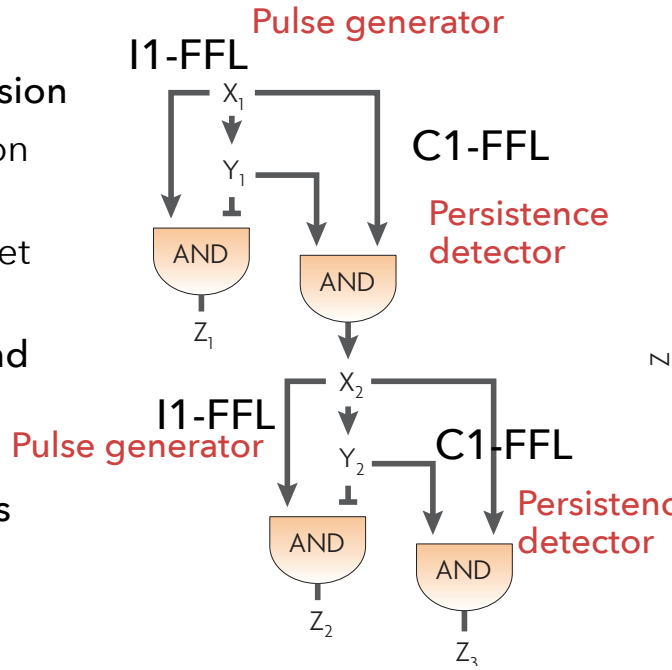
- Multiple inputs (X_i) control multiple outputs (Z_i)
- Denser wiring than in randomised networks
- Multi-dimensional input function characterises input integration at promoters of each output gene

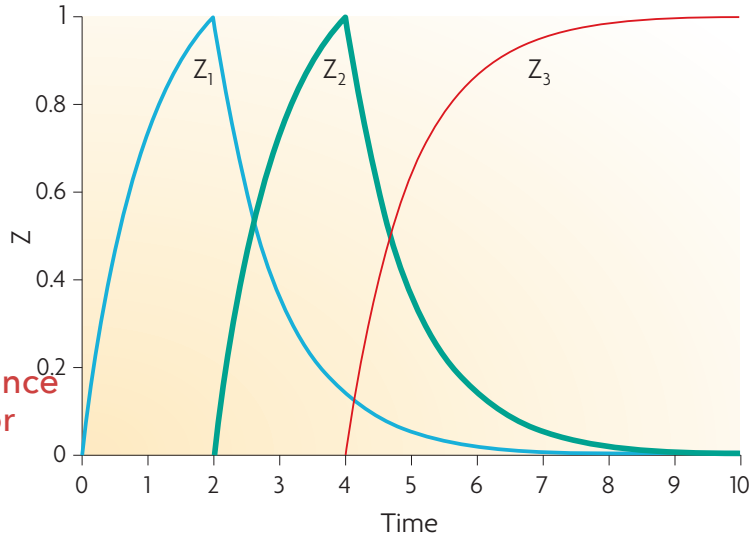




Interlocking FFLs

- 3-wave temporal program of gene expression
- I1-FFL generates Z_1 pulse: X_1 (eg. Starvation signal in *Bacillus subtilis*), triggers rapid Z_1 expression, and repression with a delay (set by Y_1 induction and activation threshold).
- C1-FFL generates activation of X_2 by X_1 and Y_1 with a delay identical to Z_1 repression. It ensures that X_2 (and hence Z_2) is not activated unless X_1, the starvation signal is persistent enough.
- I1-FFL generates Z_2 pulse
- C1-FFL causes sustained activation of Z_3.

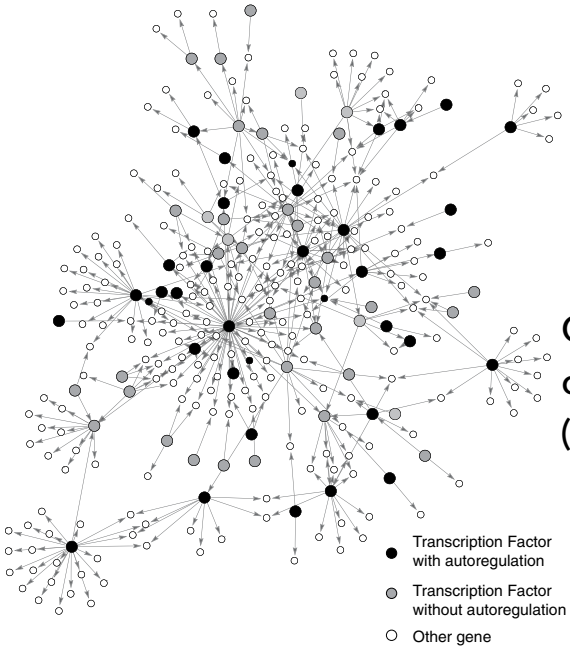




- The FFLs are multi-output FFL: Z_1 - Z_3 represent large group of downstream genes resulting in a 3-wave program of gene expression.
- FFLs assemble into real networks that lend to easy interpretation (understandability), while random association of FFLs do not

Structure of sensory transcriptional network in E. coli

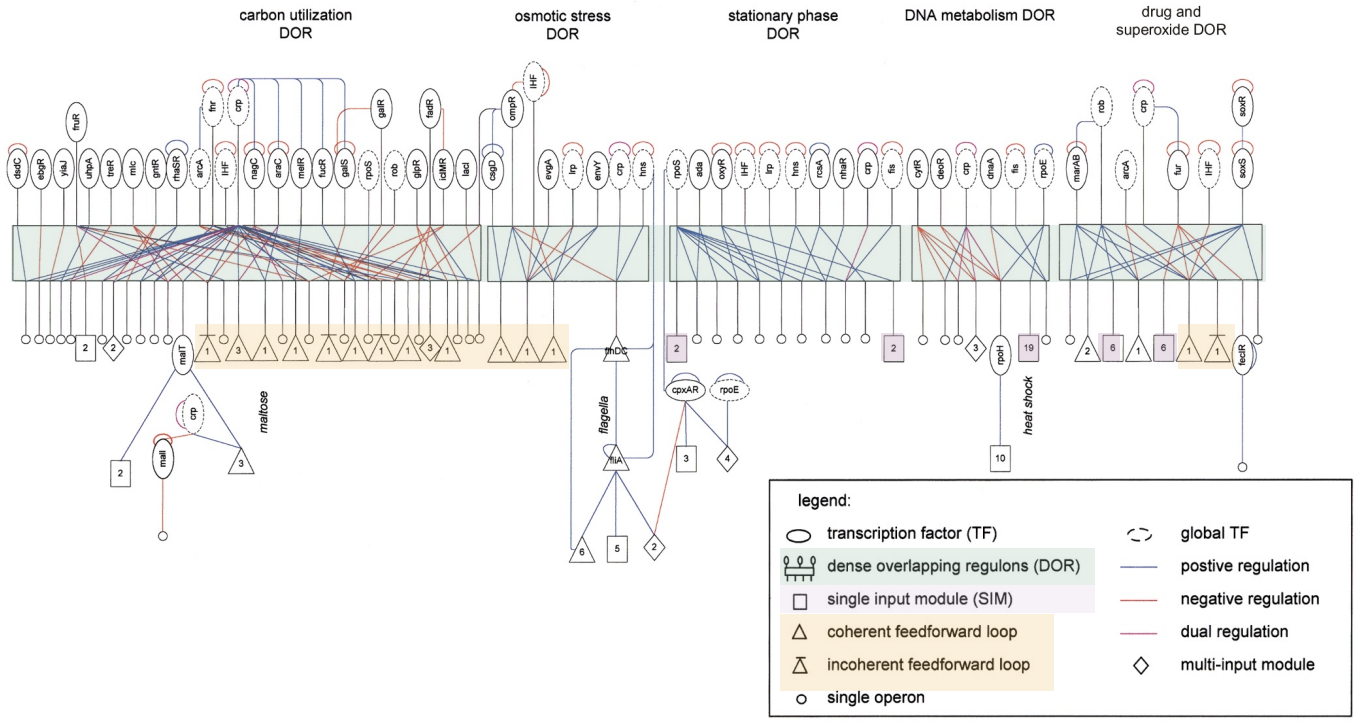
- Graph representation



Combinatorial computation (input function)

- Network motif as substructures

Few core motifs are repeated in Network
FFLs, SIMs, DORs and AR
DORs are in parallel but not in cascade
Core motifs generate temporal programs or response to fluctuating environments



S.S. Shen-Orr et al U. Alon *Nature Genetics* 31:64-68 (2002)

Network motifs in transcriptional sensory systems

- core property of *reversible* responses to environment inputs
- this is very common in Bacteria and Yeast.

1. AR: Auto-regulation – negative feedback loop
   - Response accelerator and Robustness

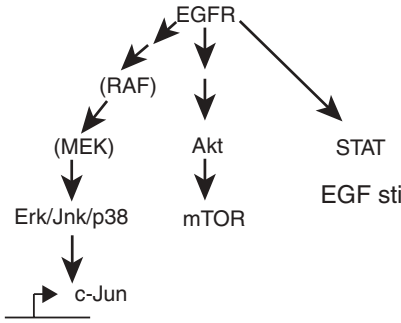
2. FFLs: feedfoward loops
   - Persistance detector (C1-FFL)
   - Response accelerator (I1-FFL)
   - Pulse generator (I1-FFL)
   - temporal waves (interlocking FFLs): FIFO

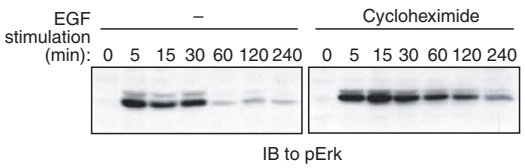
3. SIMs: single input modules
   - Temporal programmes (LIFO)

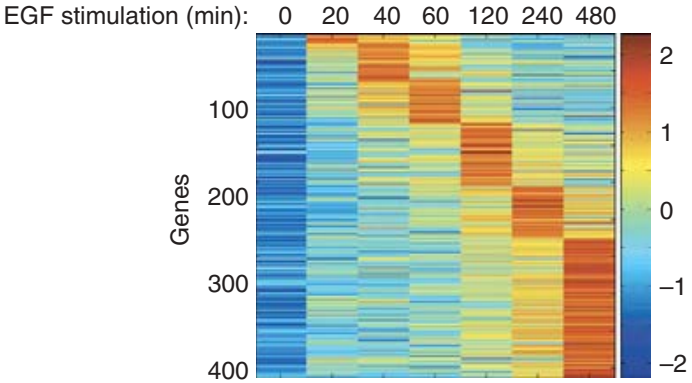
4. DORs: dense overlapping regulons
   - Combinatorial control

Sensory systems in Eukaryotes

- Response to signalling pathways during development and in adults

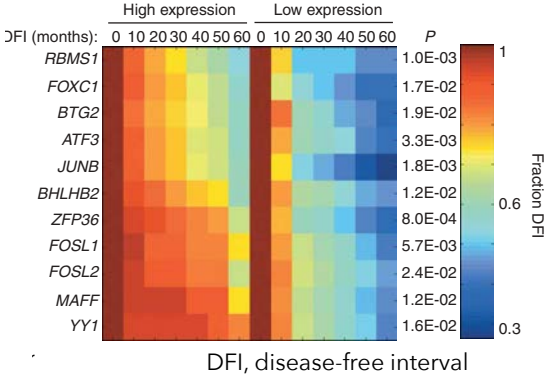






- The EGF signalling pathway is attenuated following sustained activation.
- Negative feedback exerted on ERK (MAPK) activation. Requires translation (gene regulation)
- Immediate early (IEG), delayed early (DEG) and late genes (LG) are activated in a wave

- In human tumours, transcription repressor DEGs are often downregulated.
- Correlation between expression of repressors and survival of individuals.

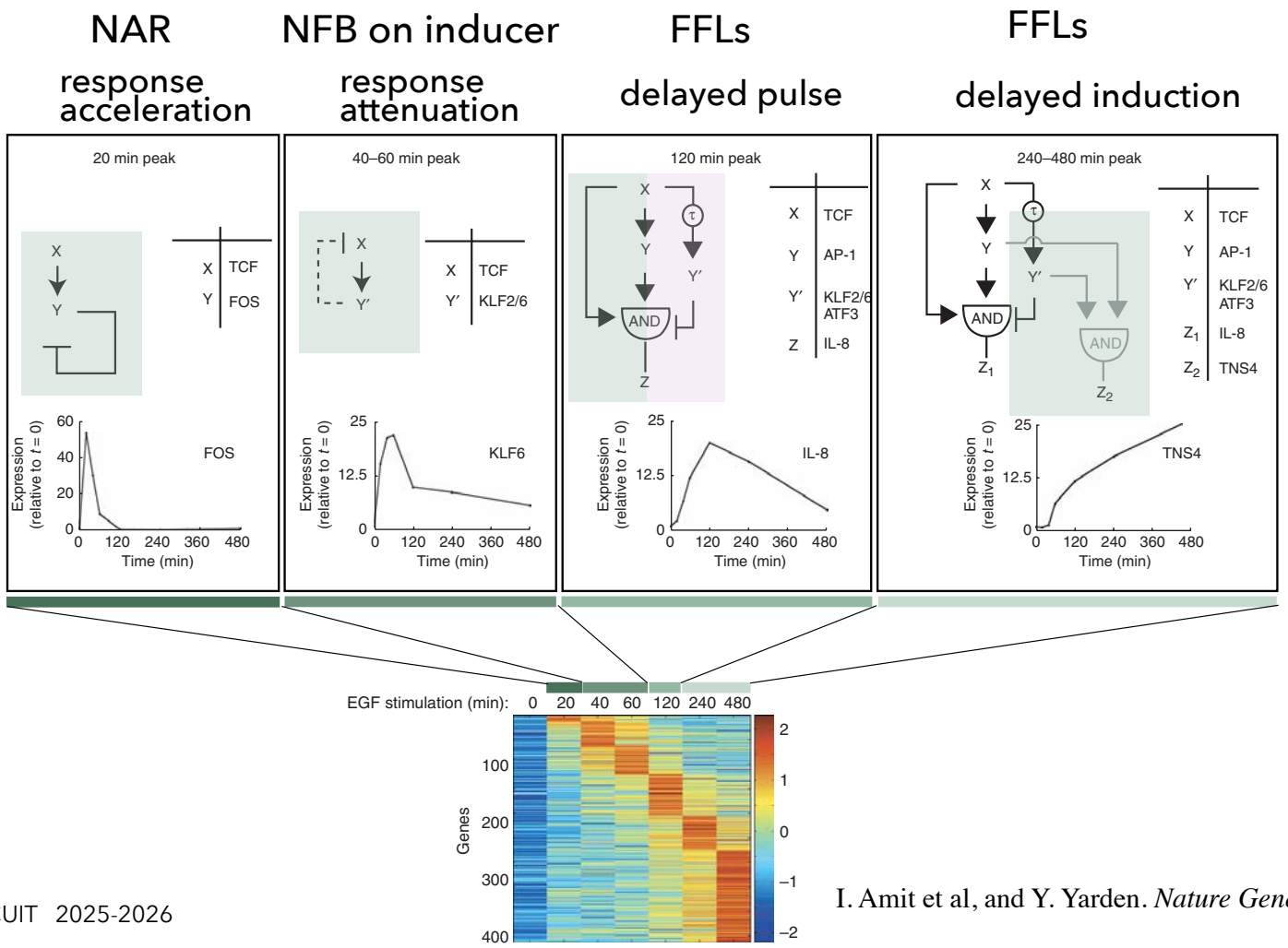


Fraction of individuals with no evidence of disease (prostate cancer)

Ido Amit et al, and Y. Yarden. *Nature Genetics*, 39:503-512 (2007)

Sensory systems in Eukaryotes

- Response to signalling pathways during development and in adults



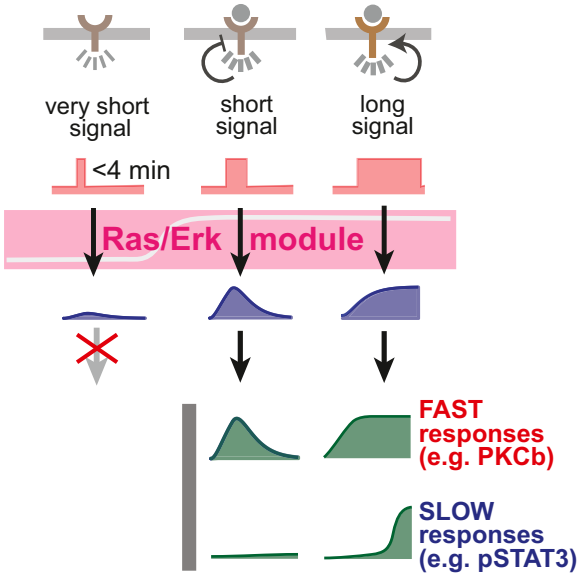
I. Amit et al, and Y. Yarden. *Nature Genetics*, 39:503-512 (2007)

Temporal information decoding using network motifs

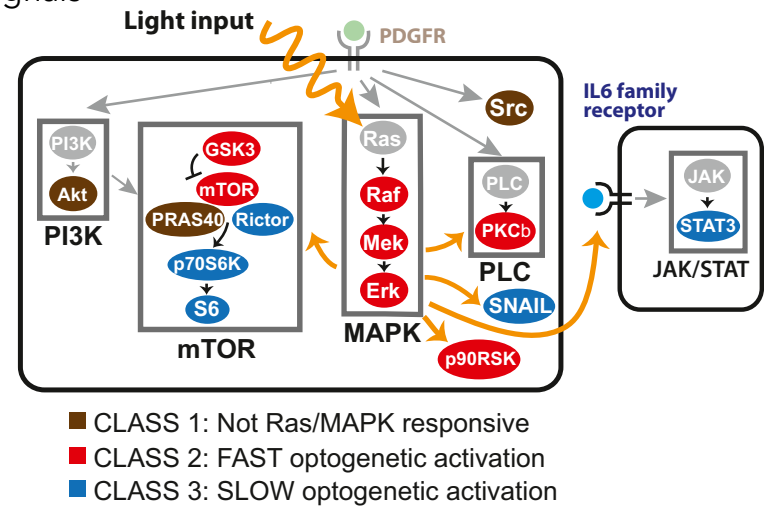
- **Short and long signals can have different cellular outcomes**

- Use optogenetics to perturb the dynamics of Ras signalling
- Precision sensing at the single cell level: Each cell is capable of singular and stable response over hours.
- ERK signalling is a **high bandwidth low pass filter**.
- **Differential modular decoding downstream of Ras/ERK:**
  - *Fast module* faithfully transmit Ras dynamics
  - *Slow module* is a persistence detector that only conveys long lasting signals

Encoding of signal *duration*



Modular decoding



Toettcher JE, Weiner OD, Lim WA. *Cell* 155:1422–34 (2013)

Mangan, S., and Alon, U. Structure and function of the feed-forward loop network motif. *PNAS* 100, 11980–11985 (2003)

Network motifs

Sensory system networks

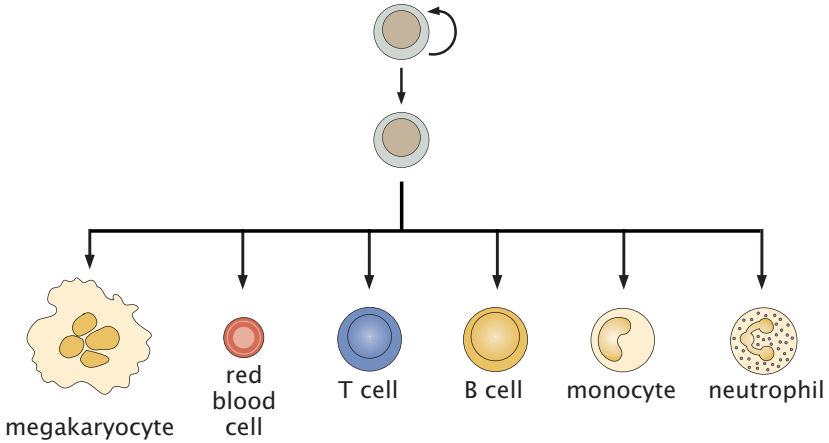
- 1-2 nodes networks:
  - Simple regulation and autoregulation: response accelerator
- 3 nodes networks: Feedforward loops: Persistence detector, pulse generator, response accelerator
- Programmes (temporal, combinatorial)
  - Single input modules (SIMs)
  - Feedforward circuits (multi FFLs)
  - Dense overlapping regulons (DORs)

Developmental networks:

- Positive autoregulation: slowed response time/noise filtering
- Bistable networks: memory

Developmental Networks

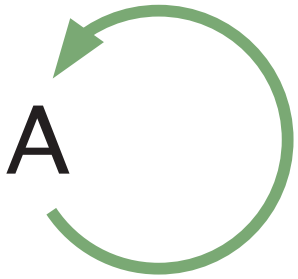
- How to induce irreversible cellular responses (eg. cell fates) ?
- How to induce responses as a function of variation in input signal?



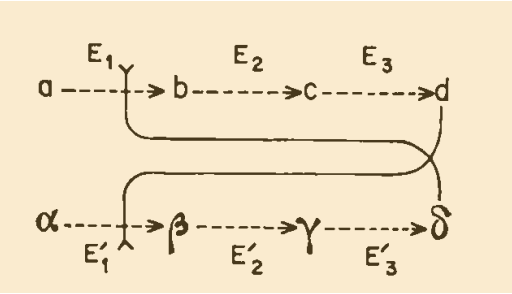
- Bistable networks: Memory

Genetic Switch

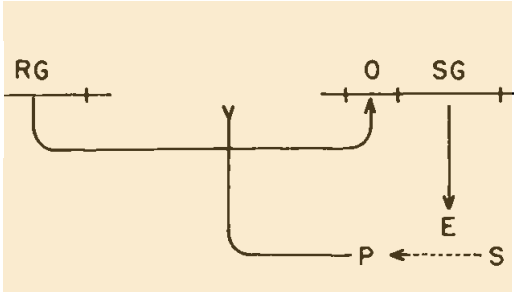
Bistable networks and memory



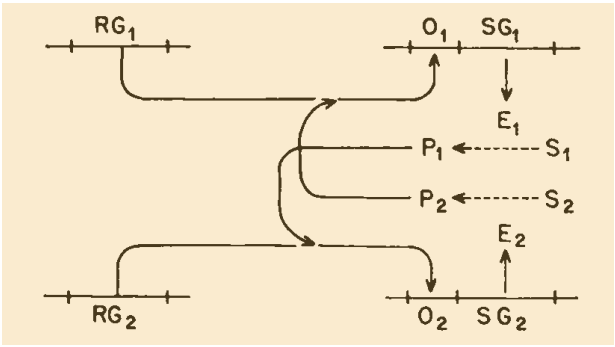
Bistable networks and memory



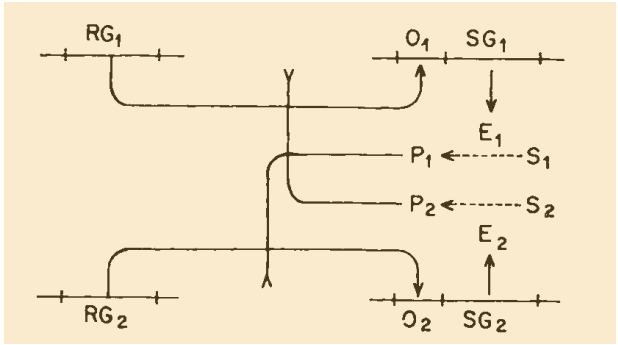
- Network with cross inhibitory feedback



- Inducible system positive feedback circuit (via double inhibition)

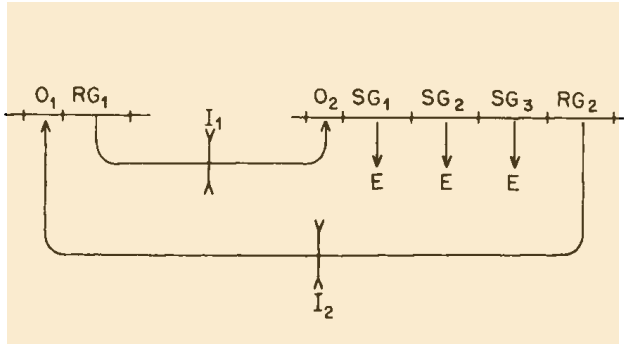


- Network with cross inhibition



- Network with co-activation

« Let us study a certain number of theoretical model systems in which we shall use only the controlling elements known to exist in bacteria, interconnected however in an arbitrary manner. »

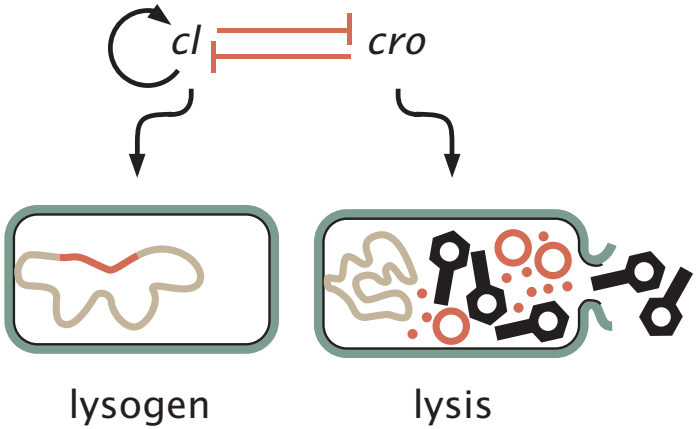


- Network with double negative feedback, ie. positive feedback

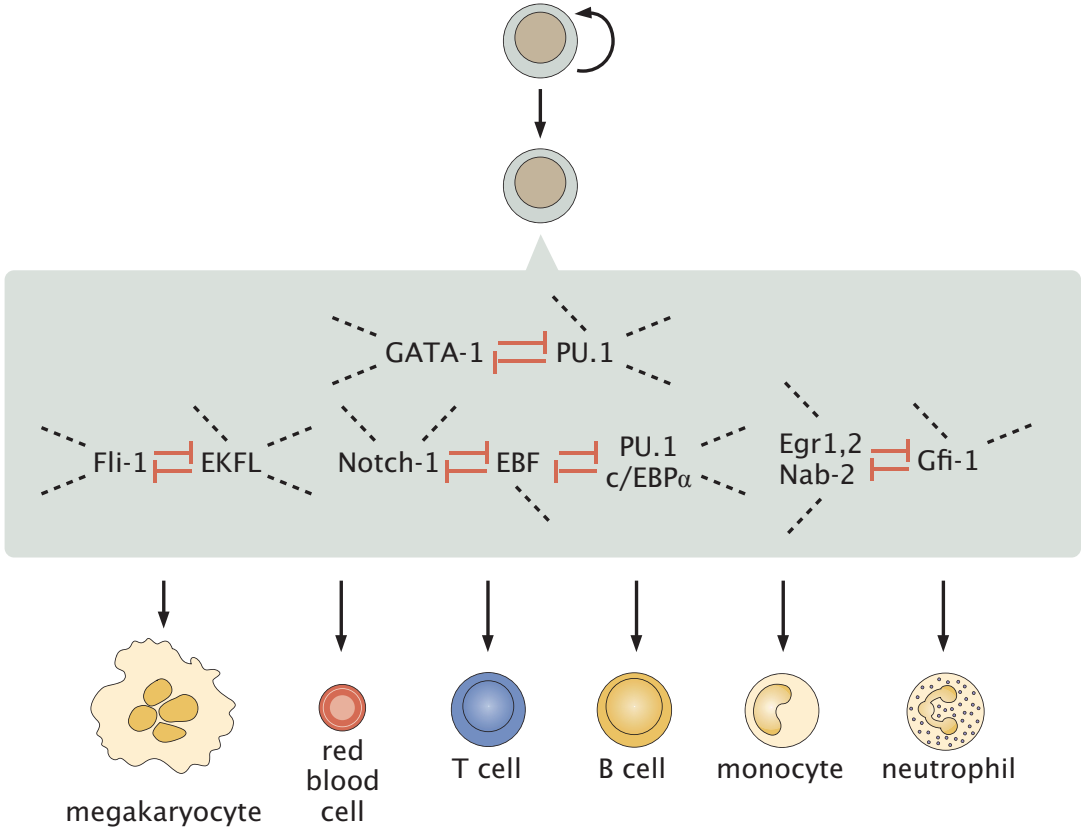
Monod, J. & Jacob, F. *Cold Spring Harb. Symp. Quant. Biol.* 26, 389–401 (1961).

Bistable networks and memory

- Phage lambda in *E.coli*



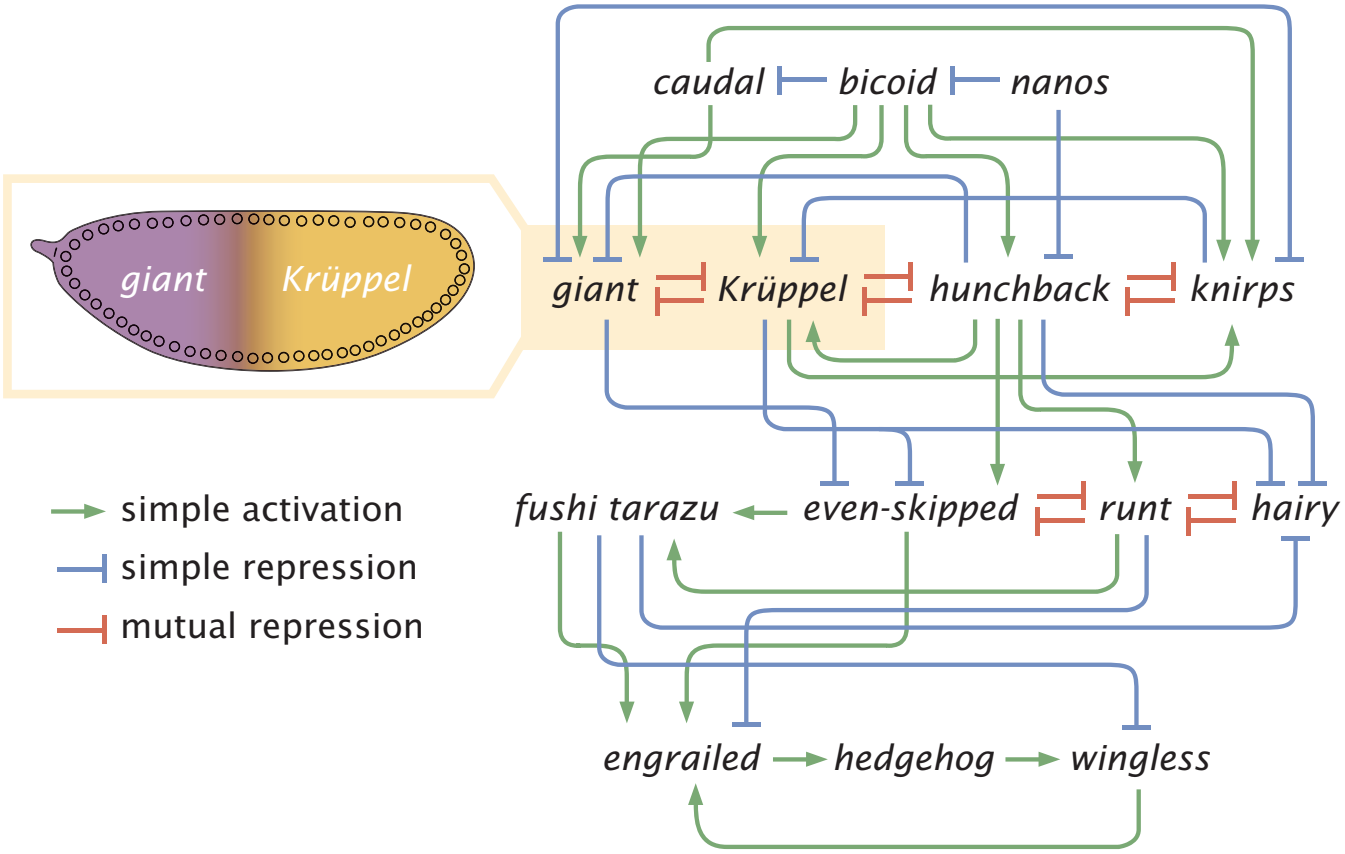
- Haematopoietic lignage



Hernan G. Garcia and Rob Phillips, *Physical genomics*

Bistable networks and memory

- Spatial patterning in *Drosophila* embryos

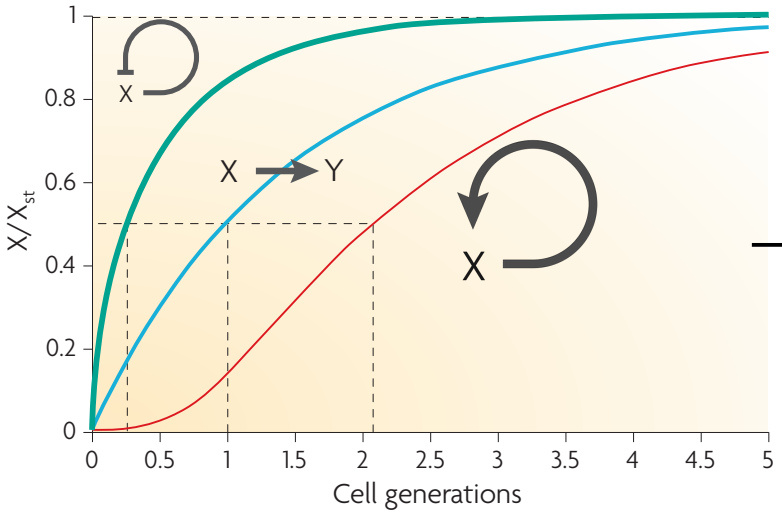


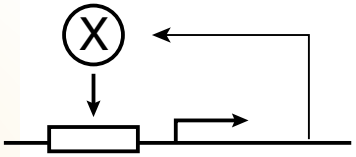
Hernan G. Garcia and Rob Phillips, *Physical genomics*

Positive Autoregulation (PAR)

- PAR slows down the response time

PAR is common in developmental networks
Suitable in developmental processes that take a long time (hours)
It helps filter noisy input via integration

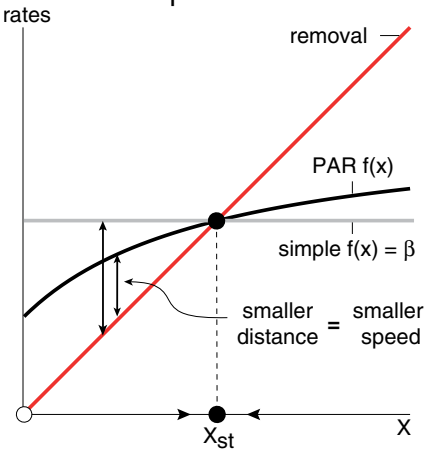




- Equation: $dX/dt = f(x) - \alpha X$

$f(X)$ is an increasing function of X (and not a constant as in simple regulation), reflecting the autoregulation: As X increases, so does the rate of production of X due to the positive feedback.

- Rate analysis shows that the approach to the final steady state X_{st} is slower than in simple regulation.

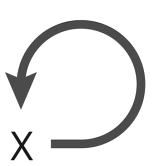


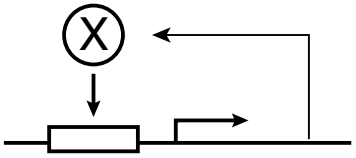
Positive Autoregulation (PAR)

- PAR can ensure sensitivity and noise filtering

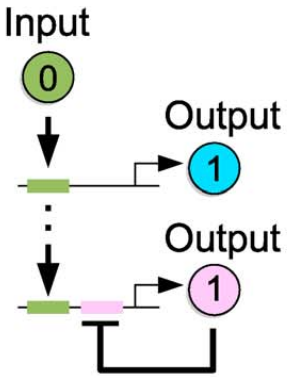
Negative Feedback increases sensitivity (gain) but amplifies noise

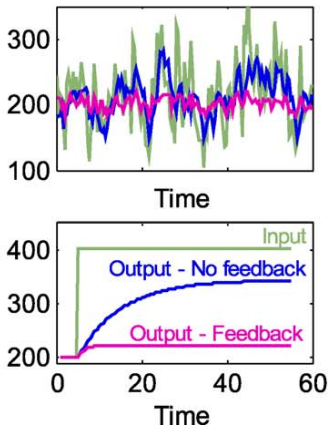
Positive Feedback increases sensitivity AND can buffer noise, by allowing longer averaging over delayed response





$$\frac{dn_1}{dt} = \beta_1 \frac{n_0^{h_0}}{1+n_0^{h_0}} \frac{1}{1+n_1^{h_1}} - \frac{n_1}{\tau_1},$$





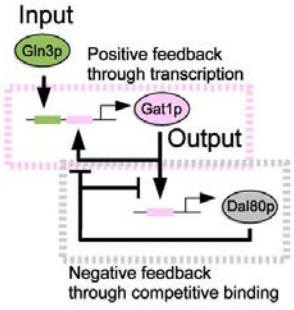
sensitivity/gain
(relative change in output following a change in the input)

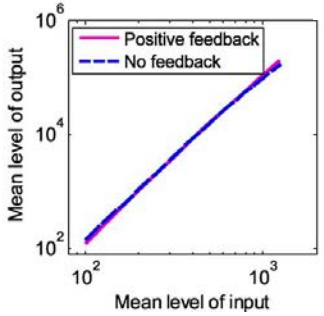
$$s = \frac{\langle n_0 \rangle}{\langle n_1 \rangle} \frac{d\langle n_1 \rangle}{d\langle n_0 \rangle} = \frac{d\ln\langle n_1 \rangle}{d\ln\langle n_0 \rangle}$$

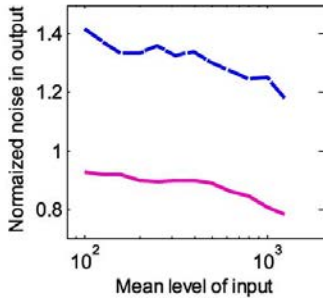
noise amplification
(ratio of output and input noise)

$$\bar{\eta} = \frac{\eta_1}{\eta_0} = \frac{std(n_1)/\langle n_1 \rangle}{std(n_0)/\langle n_0 \rangle}$$

Gat1 self-activation = positive feedback → noise buffering via time averaging.







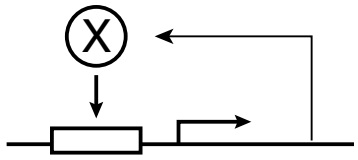
Dal80 repression = fast negative feedback → stabilization to prevent runaway or bistability.

G. Hornung and N. Barkai *Plos Comp. Biol.* 4, 55-61 *(2008)*

Positive Autoregulation (PAR) and bistability

- PAR can cause bistability when the autoregulation function shows cooperativity.

$$dX/dt = f(x) - \alpha X$$

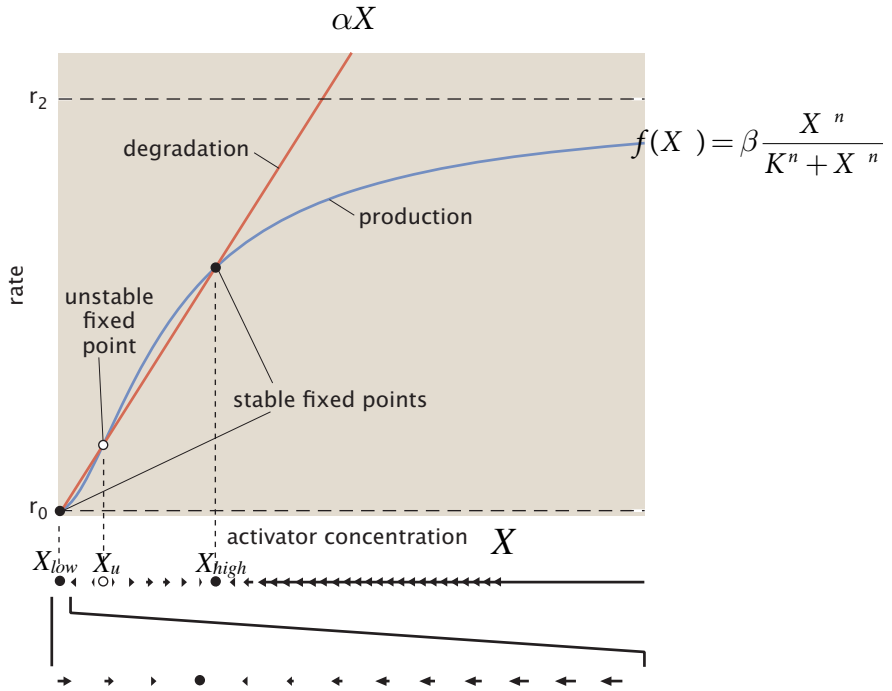


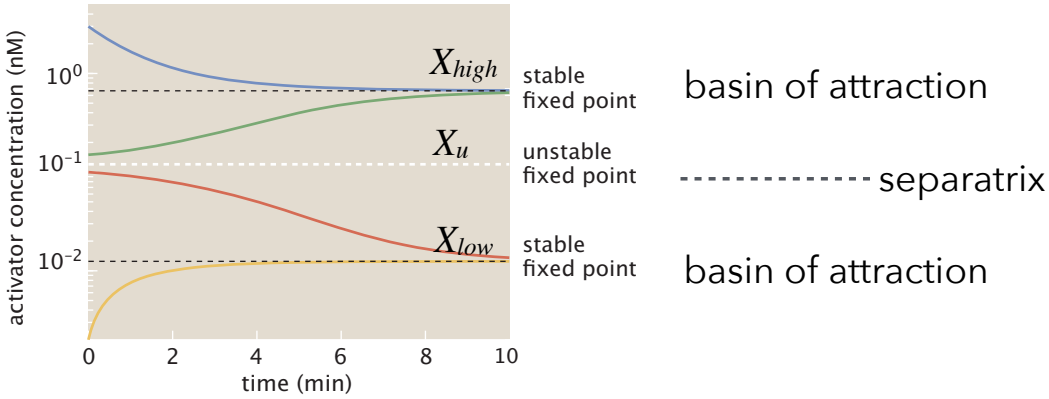
$f(X)$ has a sigmoid shape, eg. Hill function:

$$f(X) = \beta \frac{X^n}{K^n + X^n}$$

Rate analysis reveals 3 fixed points, 2 are stable (X_{low} , X_{high}) and 1 is unstable (X_u).

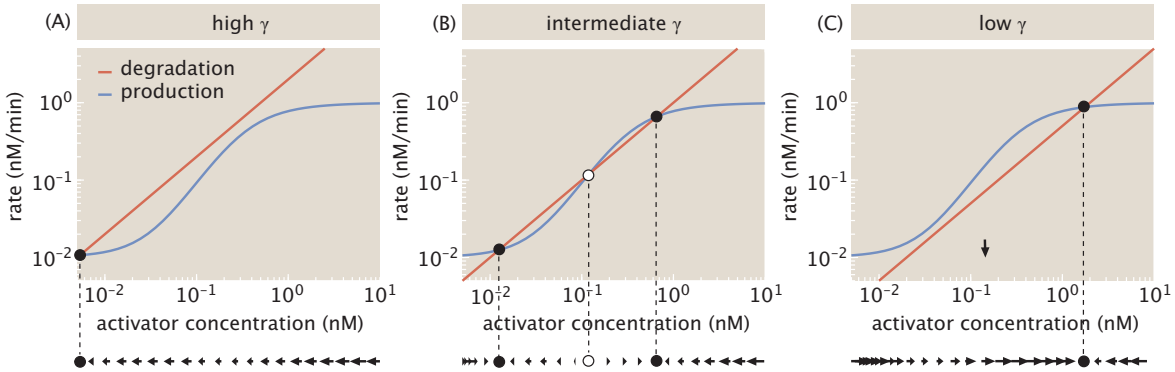
Evolution of X as a function of the initial conditions with respect to X_{low}, X_u and X_{high}



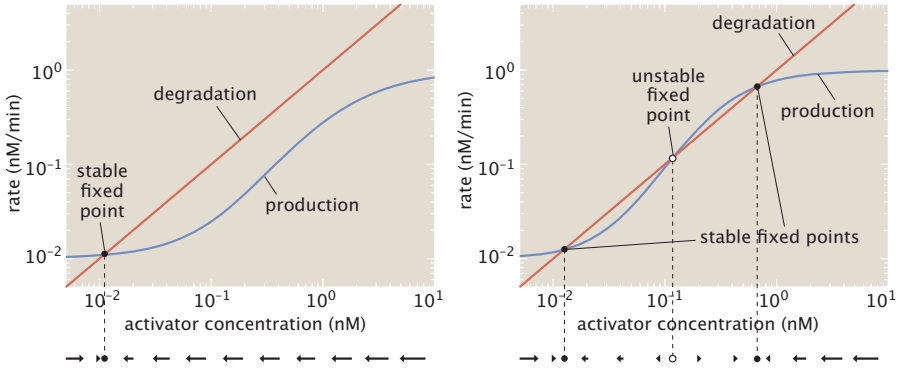


Positive Autoregulation (PAR) and bistability

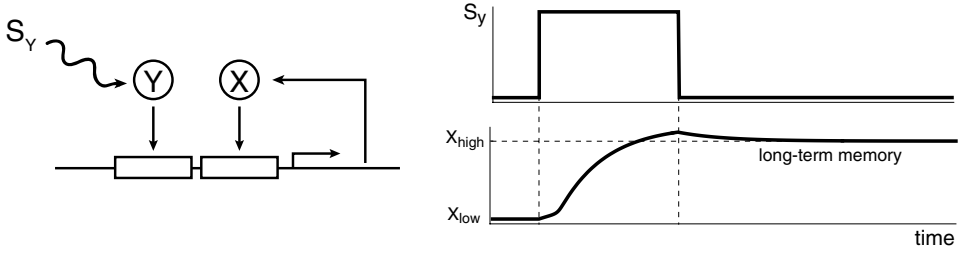
- Bistability depends on the degradation rate



- Bistability depends on the cooperativity

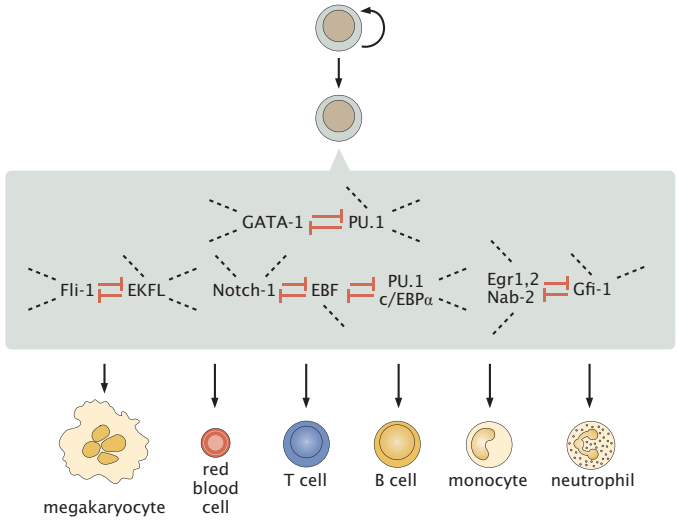


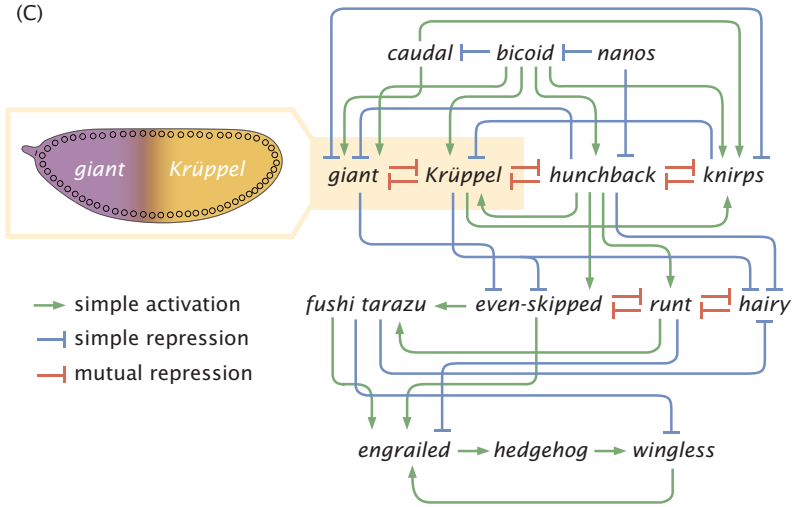
- Bistability imparts memory:



Two-node positive feedbacks

- Frequent in developmental networks
- Double positive: Lock on system Y ⇆ X X and Y *both* ON or OFF
- Double negative: Toggle switch Y ⊣⊢ X X ON and Y OFF or vice versa





Double negative FB: Toggle switch

Y ⊣ X

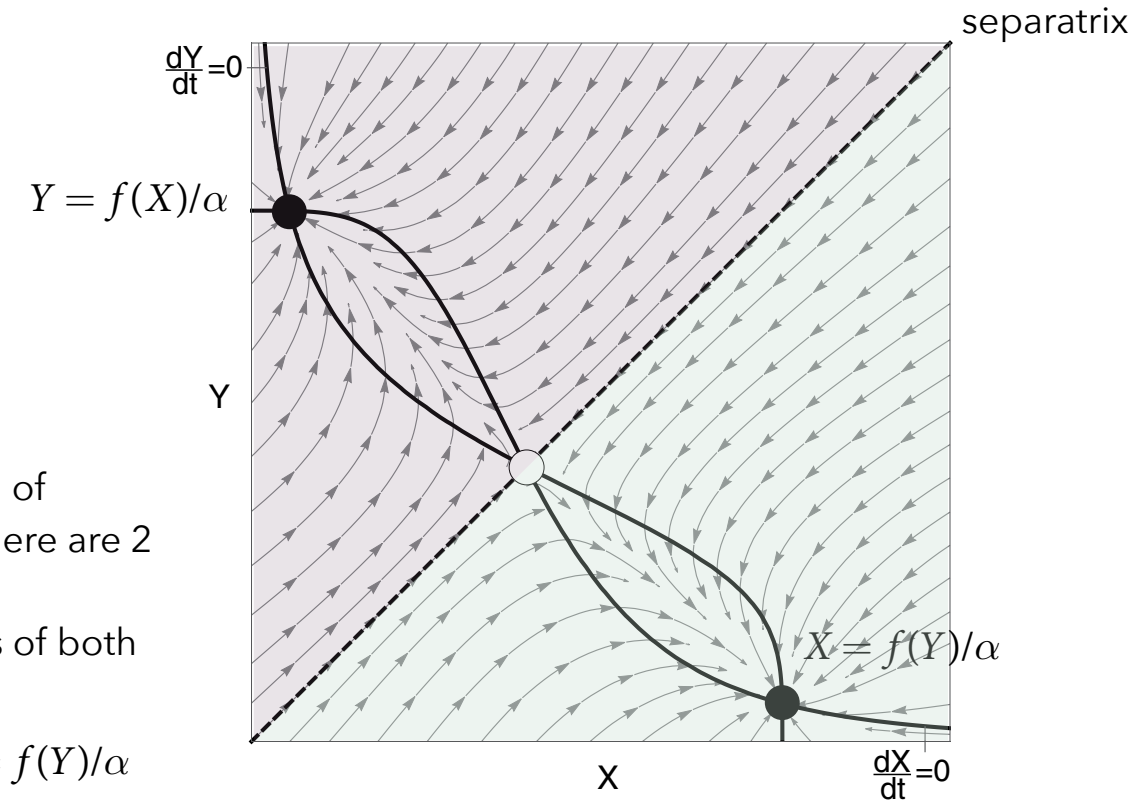
$$\frac{dX}{dt} = f(Y) - \alpha X$$

$$\frac{dY}{dt} = f(X) - \alpha Y$$

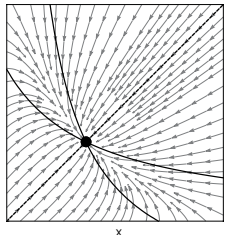
Where $f(X)$ is a decreasing Hill function $n>1$.

Null clines are curves in X, Y plane, where the rate of change of one protein (eg. X) is zero ($dX/dt = 0$). There are 2 null clines for X and Y.
Existence of fixed points at points of intersections of both null clines.

Null clines are Hill functions: $Y = f(X)/\alpha$, $X = f(Y)/\alpha$



If f is not a Hill function (or any other function with sigmoid shape, eg. logistic) the system is monostable

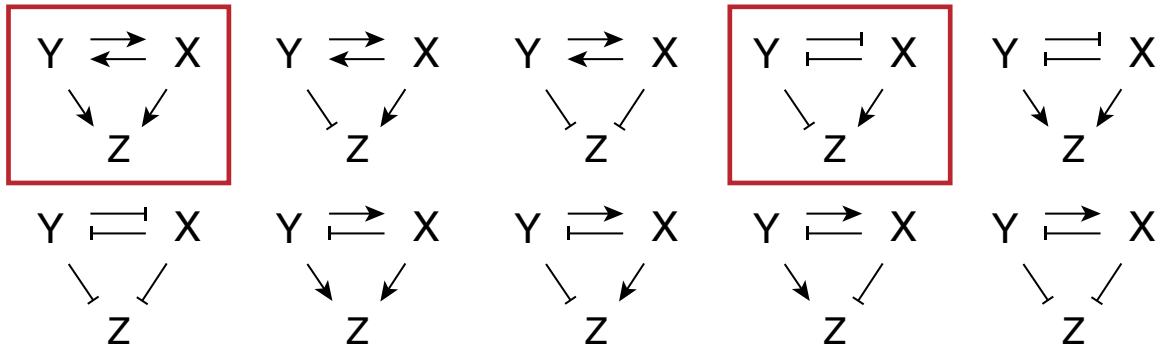


Y ⇄ X Y ⊣ X

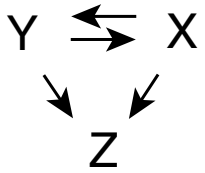
PAR on X and Y can turn a double FB without cooperatively into a bistable network

Three-node feedbacks

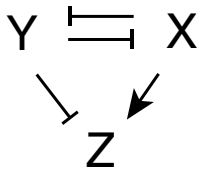
Regulating Feedbacks



Double-positive feedback loop



Double-negative feedback loop

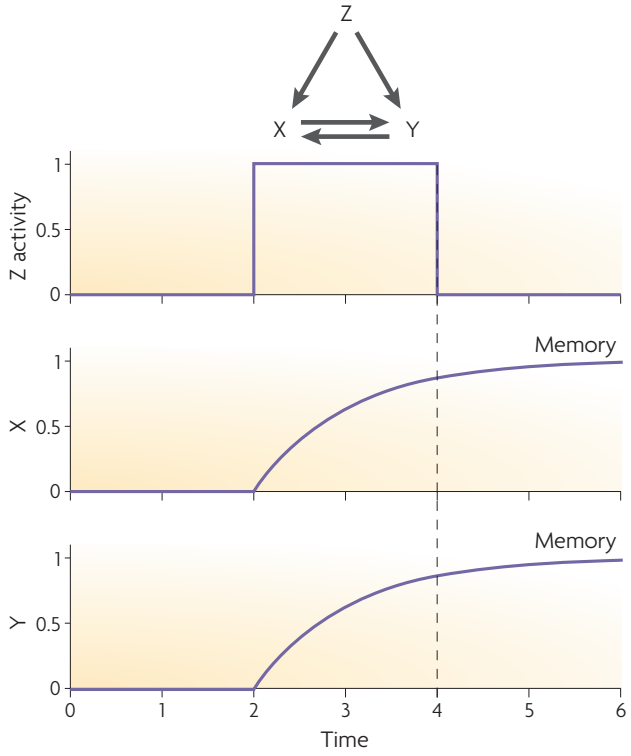


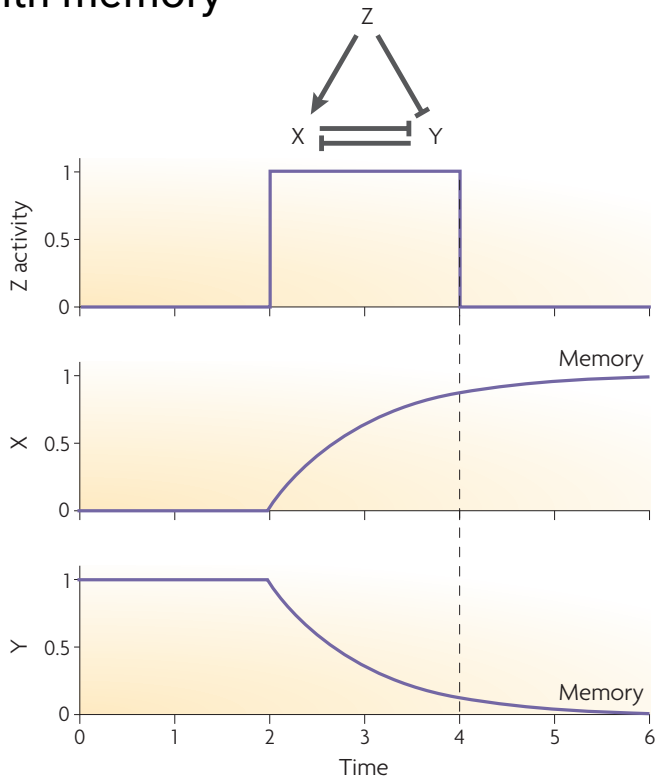
Coherent networks: the two arms, such as X->Z and X->Y->Z have same sign.

Three-node feedbacks

Common motifs of Regulated Feedbacks

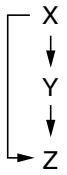
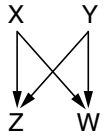
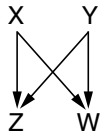
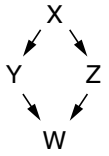
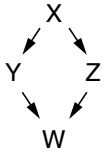
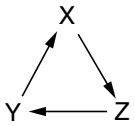
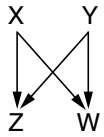
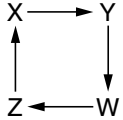
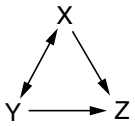
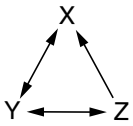
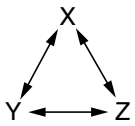
- Bistable networks with memory





Beyond biochemical and regulatory networks

- Different implementation of similar fundamental logic designs

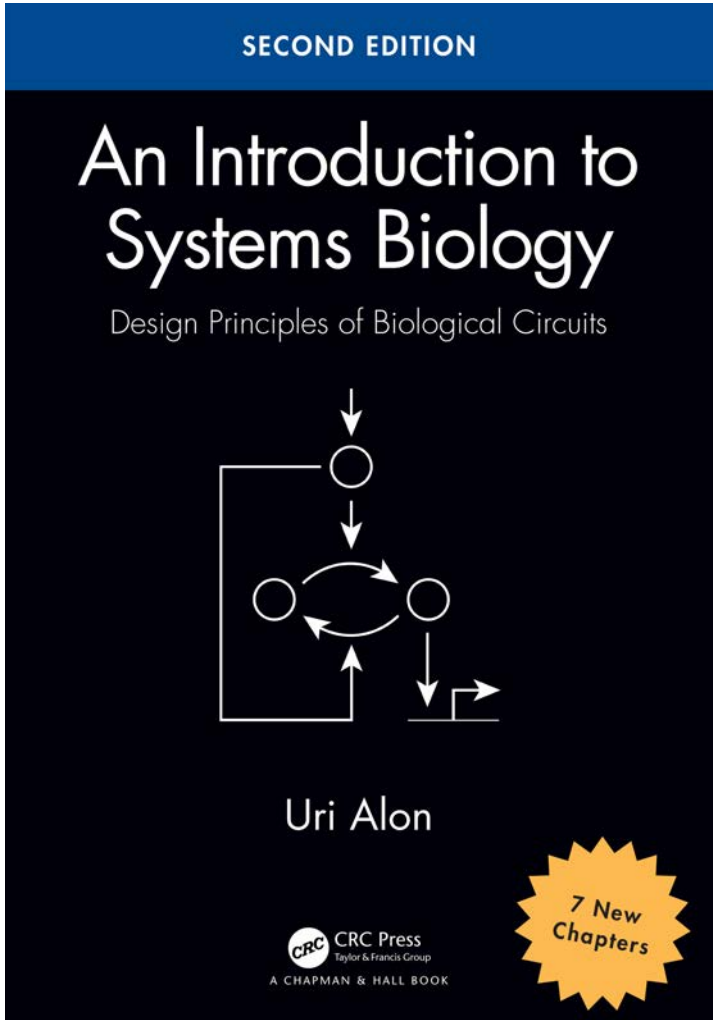
Transcription networks	 Feedforward loop	 Bi-fan	
Neurons	 Feedforward loop	 Bi-fan	 Diamond
Food webs	 Cascade	 Diamond	
Electronic circuits (fractional multipliers)	 Three-node feedback loop	 Bi-fan	 Four-node feedback loop
Social Networks	 Regulating mutual	 Semi-Clique	 Clique

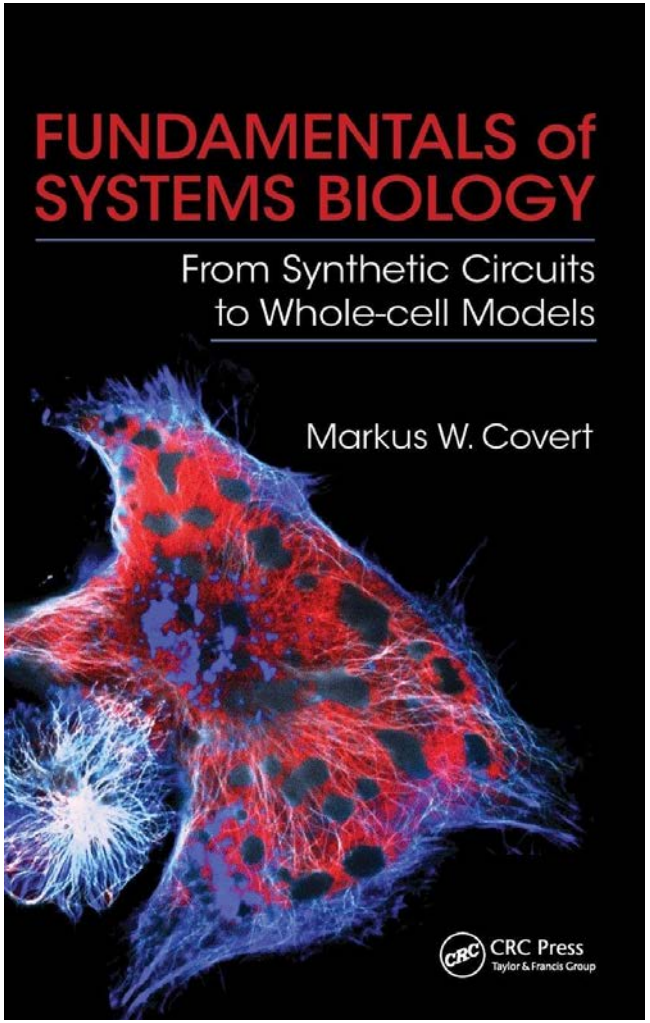
R. Milo et al and U. Alon. *Science*, 298: 824-827 (2002)

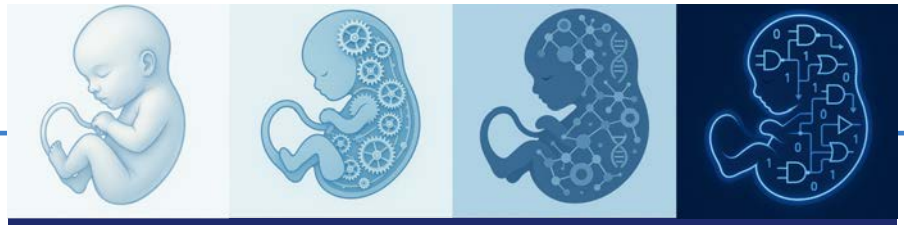
Conclusions

- Deciphering the logic of biochemical and gene networks is possible
- The existence of network motifs reflects evolutionary constraints on the search for algorithms that fulfil a computational task.
- Examples of computational modules:
  - Response accelerator: NAR
  - Persistence detector: C-FFL
  - Noise filtering: PAR
  - Pulse generator and Temporal waves: I-FFL and interlocked FFLs
  - Robustness: NAR
  - Some algorithms combine tasks: ie. PAR increases sensitivity but not noise.

Books recommendations







Thomas LECUIT, chaire Dynamiques du vivant

Qu'est-ce que l'information biologique ? (II)

COURS : 20 novembre > 18 décembre 2025

COURS

Le jeudi de 10 h à 12 h
Amphithéâtre Guillaume Budé

Jeudi 20 Novembre 2025
Introduction :
approche computationelle du vivant

Jeudi 27 Novembre 2025
Complexité et information
au cours du développement

Jeudi 4 Décembre 2025
Vision logique des flux d'information

Jeudi 11 Décembre 2025
Vision dynamique des flux d'information

Jeudi 18 Décembre 2025
Apprentissage non neuronal
dans un système biologique

COLLOQUE

De 9h à 18h
Amphithéâtre Maurice Halbwachs

Vendredi 26 juin 2026
Information flow and computation in living systems

Les cours et colloques
sont gratuits, en accès libre,
sans inscription préalable.

Image générée par une I.A. © T. Lecuit.

COLLÈGE DE FRANCE 1530

Thomas Römer, administrateur du Collège de France
11, place Marcelin-Berthelot, 75005 Paris
www.college-de-france.fr

Année académique 2025/2026