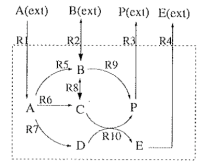# Gene regulatory networks

Lecture 4 of Introduction to Biological Modeling
Oct. 13, 2010

Steve Andrews

Brent lab, Basic Sciences Division, FHCRC

Last week
- metabolic networks
- constraint-based modeling
- pathway analysis
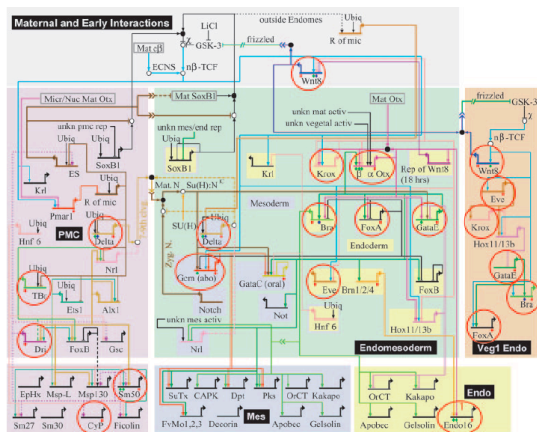- metabolic control analysis

Reading
Milo, Shen-Orr, Itzkovitz, Kashtan, Chklovskii, and Alon "Network motifs: simple building blocks of complex networks" *Science* 298:824, 2002.

Genetic networks

network for sea urchin embryo development

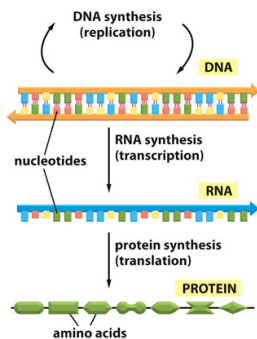Credit: Levine and Davidson, *Proc. Natl. Acad. Sci. USA* 102:4396, 2005.

Genetic regulation
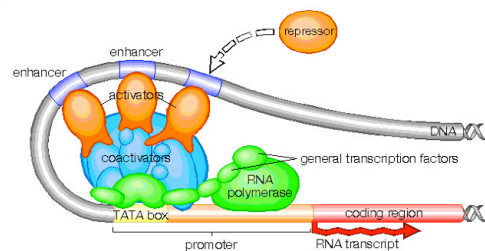
Graphs

Boolean networks

Dynamics

Motifs

Summary

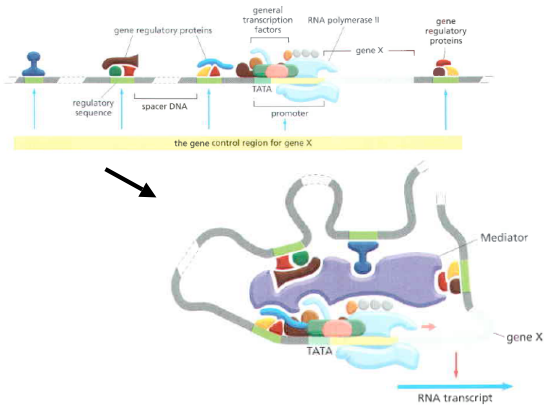## Central dogma of molecular biology

Credit: http://dogma.film.bigbestmovie.com/centraldogmaofmoleculargenetics/

Transcription regulation

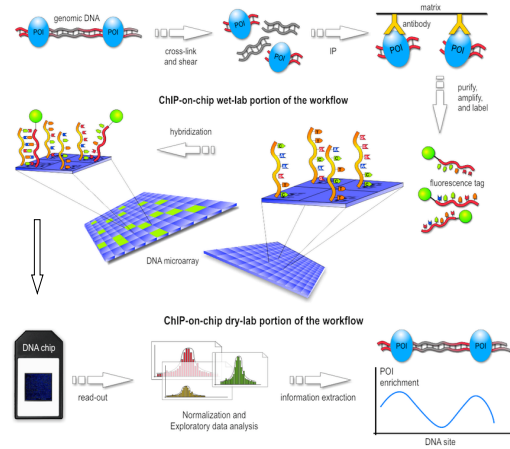Credits: http://www.mun.ca/biology/desmid/brian/BIOL3530/DB_Ch09/DBNDiff.html

Transcription regulation

ChIP-on-chip

ChIP-on-chip wet-lab portion of the workflow

ChIP-on-chip dry-lab portion of the workflow

ChIP - seq

ACTGGTGACAGGACG
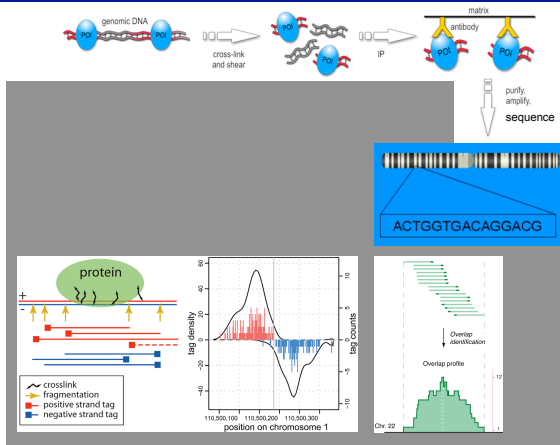
Abstracted gene regulation network

- transcription
- translation
- repression
- activation
- gene
- mRNA
- protein
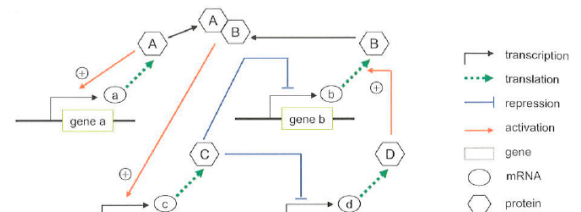
BioTapestry

"an interactive tool for building, visualizing, and simulating genetic regulatory networks"
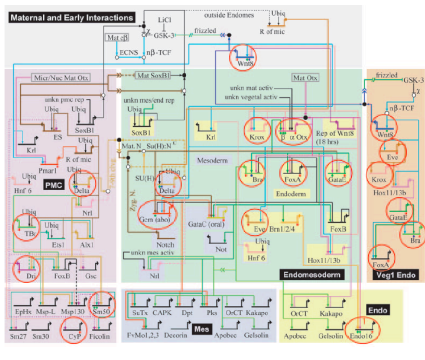
www.biotapestry.org

Hamid Bolouri helped write BioTapestry, and is a Hutch member.

Translation regulation: miRNA

miRNA - microRNA

RISC - RNA-induced silenceing complex

binds to selective mRNAs, to degrade them or slow their translation

Transcription/translation regulation: siRNA
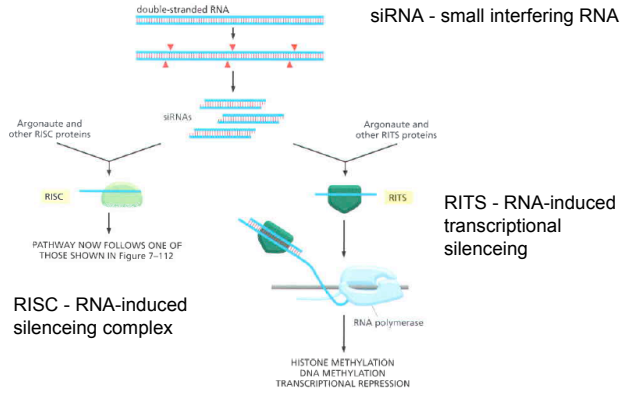
siRNA - small interfering RNA

RITS - RNA-induced transcriptional silenceing

RISC - RNA-induced silenceing complex

13

Gene regulation

regulation

transcriptional activators/repressors

siRNA

miRNA
siRNA
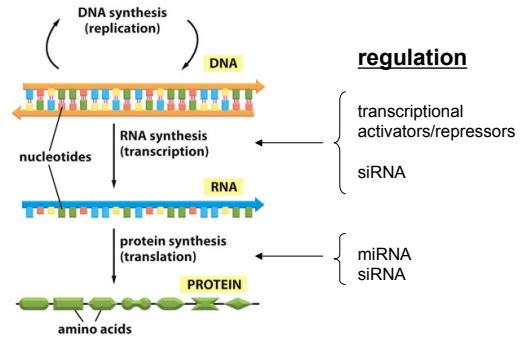
14

Gene regulation *modeling*

Nearly all modeling work is on "classic" transcription regulation by proteins.

15

Genetic regulation

Graphs

Boolean networks

Dynamics
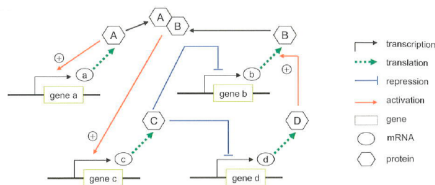
Motifs

Summary

16

Abstraction as a directed graph

Directed graphs

$V = \{a,b,c,d\}$ ← vertices

$E = \{(a,a,+),(a,c,+),(b,c,+),$
$(c,b,-),(c,d,-),(d,b,+)\}$
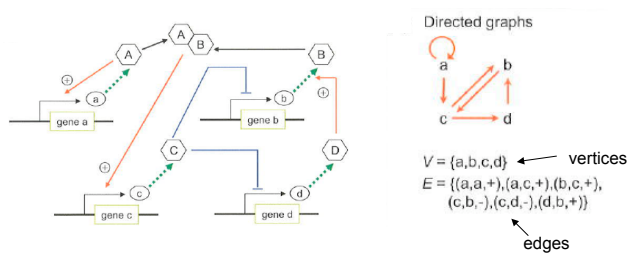
edges

17

Graph basics
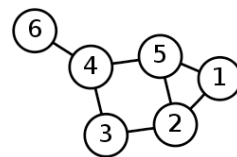
6 nodes
7 edges
cyclic
undirected
diameter = 3 (max pathlength)
degrees, from 1 to 3
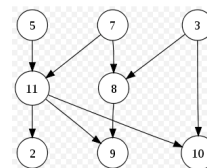
8 nodes
9 edges
acyclic
directed
in-degree from 0 to 2
out-degree from 0 to 3

18

Edge distributions for random graphs

Erdös-Rényi graph

add links randomly
edges ~ <k> ± σ
$P(k)$ is a Poisson distribution

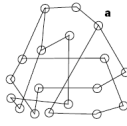*implications:*
• all nodes are roughly equivalent

• mean path ~ log n

examples:
• street maps

Edge distributions for random graphs

Erdös-Rényi graph

add links randomly
edges ~ <k> ± σ
$P(k)$ is a Poisson distribution

implications:
• all nodes are roughly equivalent
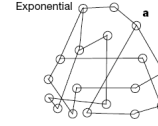
• mean path ~ log n
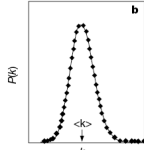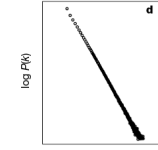
examples:
• street maps

Scale-free graph

power-law distribution of node connectivity
$P(k) \sim <k>^{-\gamma}$

implications:
• lots of "remote" nodes
• a few well-connected hubs
• mean path grows very slowly with n (small-world network)

examples:
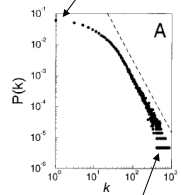• airline routes
• world wide web
• social networks

Scale-free networks

Actor collaborations

lots of actors have worked with few others

a few actors have worked with many others

Yeast protein interaction network
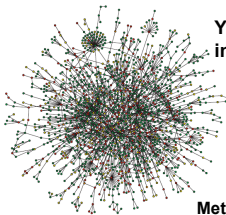
Metabolic networks
C. crescentus　many organisms
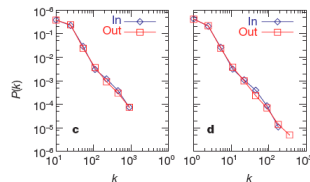
Gene regulation graph topology

Transcription factors per gene
exponential distribution
most genes have 1 or 2 TFs
very rare for a gene to have > 6 TFs
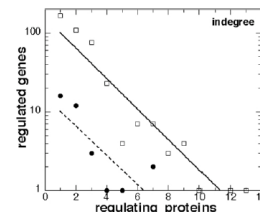
Genes per transcription factor
power-law distribution
most TFs bind to few genes
some TFs bind to lots of genes
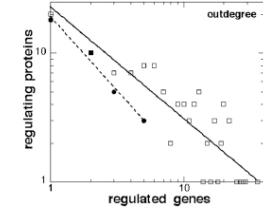
note linear vs. log scales

data are from yeast

Modularity

Modules are largely independent regions of networks
- arise from hierarchical networks
- clusters that are largely separate from each other

scale-free　　hierarchical
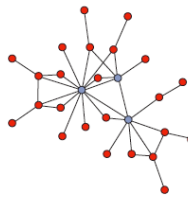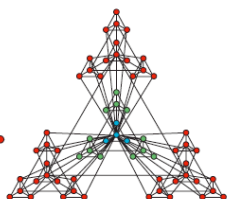
Yeast genetic regulation is modular

All Factors

DNA / RNA / protein biosynthesis

Cell cycle

Metabolism

B. subtilis genetic regulation

- different module activities in different conditions
- gene network topologies differ between endogenous and exogenous

endogenous - internally initiated events (cell cycle, sporulation)
exogenous - externally initialted events (diauxic shift, DNA damage, stress response)

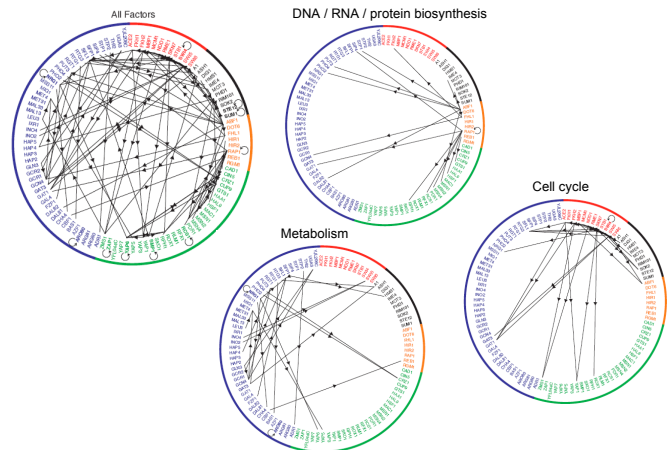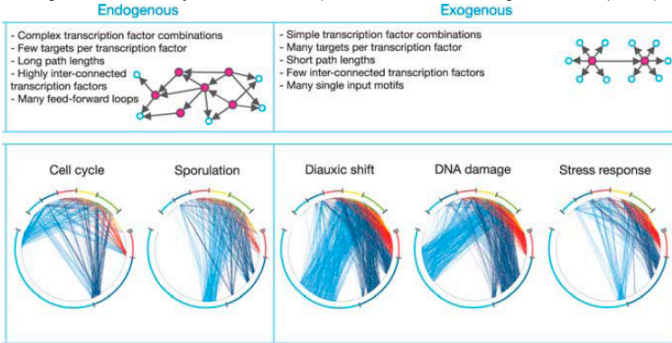| Endogenous	Exogenous
- Complex transcription factor combinations - Few targets per transcription factor - Long path lengths - Highly inter-connected transcription factors - Many feed-forward loops	- Simple transcription factor combinations - Many targets per transcription factor - Short path lengths - Few inter-connected transcription factors - Many single input motifs

Cell cycle   Sporulation   Diauxic shift   DNA damage   Stress response

Credit: Luscombe et al. *Nature* 431:308, 2004.

Boolean networks

Boolean network

$a(t+1) = a(t)$
$b(t+1) = (\text{not } c(t)) \text{ and } d(t)$
$c(t+1) = a(t) \text{ and } b(t)$
$d(t+1) = \text{not } c(t)$

Boolean logic

AND   OR   NOT

$x \wedge y$ $x \vee y$ $\neg x$

output is 1 if:
  x AND y are 1   x OR y are 1   x is NOT 1

	y 0	1
x 0	0	0
1	0	1

	y 0	1
x 0	0	1
1	1	1
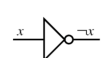
x	output
0	1
1	0

Transition table

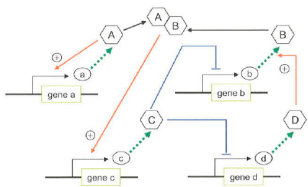state names are "on" genes

Boolean network

$a(t+1) = a(t)$
$b(t+1) = (\text{not } c(t)) \text{ and } d(t)$
$c(t+1) = a(t) \text{ and } b(t)$
$d(t+1) = \text{not } c(t)$

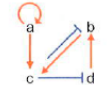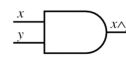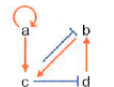| before	after
0	→ d
d	→ bd
c	→ 0
cd	→ 0
b	→ d
bd	→ bd
bc	→ 0
bcd	→ 0
a	→ ad
ad	→ abd
ac	→ a
acd	→ a
ab	→ acd
abd	→ abcd
abc	→ ac
abcd	→ ac

Discrete state space

state names are "on" genes

Boolean network

$a(t+1) = a(t)$
$b(t+1) = (\text{not } c(t)) \text{ and } d(t)$
$c(t+1) = a(t) \text{ and } b(t)$
$d(t+1) = \text{not } c(t)$

before	after
0	→ a
d	→ bd
c	→ 0
cd	→ 0
b	→ d
bd	→ bd
bc	→ 0
bcd	→ 0
a	→ ad
ad	→ abd
ac	→ a
acd	→ a
ab	→ acd
abd	→ abcd
abc	→ ac
abcd	→ ac

state space

c   cd   bc   bcd   ab   abc
b   0   acd   ac
d   a
bd   ad   abcd
abd
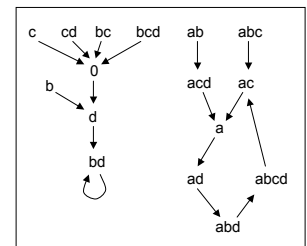
Is gene regulation Boolean?

E. coli Lac operon combines cAMP and IPTG signals

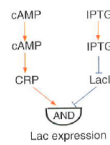### Boolean behavior

x OR y   x AND y

wild-type   **mutant 1**   **mutant 2**

(a)   (b)   (c)
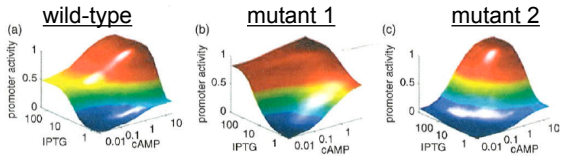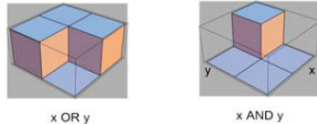
• gene regulation is sort of Boolean, but not sharp
• gene regulation is plastic with respect to mutations

Credits: Klipp *et al. Systems Biology*, 2009; Mayo et al. *PLoS Biology* 4:e45, 2006.

31

Transcription factor reactions

RNA polymerase
Activator CRP
Repressor LacI

ODE model of transcription and translation

nucleotides, DNA → mRNA → degrade

amino acids → protein → degrade
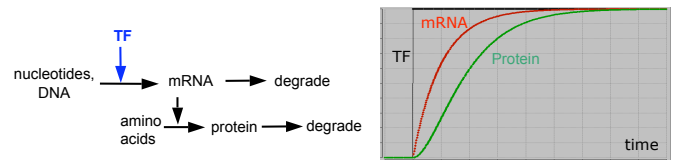
Y = fractional saturation of TF binding to DNA

$$\frac{d[\mathrm{mRNA}]}{dt} = k_t Y - k_{dm}[\mathrm{mRNA}]$$ k_t is the maximal rate of transcription

$$\frac{d[\mathrm{P}]}{dt} = k_s[\mathrm{mRNA}] - k_{dp}[\mathrm{P}]$$ k_s is the protein synthesis rate/mRNA concentration unit

At steady-state

$$[\mathrm{P}]_{ss} = \frac{k_s k_t}{k_{dp} k_{dm}} Y$$ **steady-state [protein] ~ TF binding**

Transcriptional activation by one factor

Activator (A)

transcription → mRNA → degradation

synthesis → P → degradation
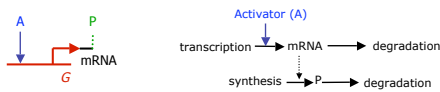
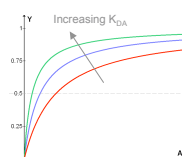Y = fractional saturation of A binding to DNA
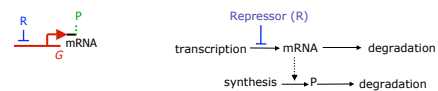K = binding constant of A to DNA

$$K = \frac{[A:DNA]}{[A][DNA]}$$

$$Y = \frac{[A:DNA]}{[DNA] + [A:DNA]}$$
$$= \frac{K[A][DNA]}{[DNA] + K[A][DNA]}$$
$$= \frac{K[A]}{1 + K[A]}$$ ← Hill equation

Increasing K_{DA}

$$Y = \frac{K[A]^n}{1 + K[A]^n}$$ ← Hill equation with cooperativity n

---

Transcriptional repression

Repressor (R)
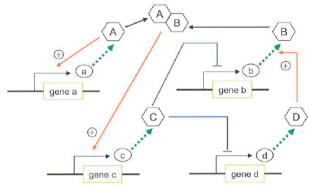
transcription ⊣ mRNA → degradation

synthesis → P → degradation

$$Y_R = \frac{K_R[R]}{1 + K_R[R]}$$ (fraction of DNA occupied by R)

$$1 - Y_R = \frac{1}{1 + K_R[R]}$$

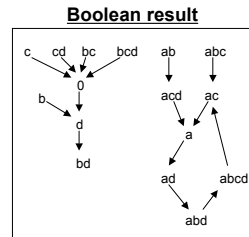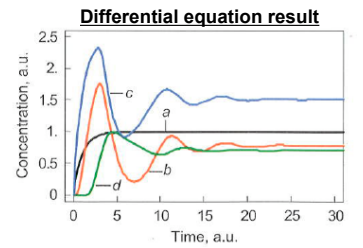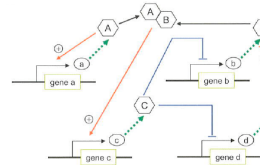$$\frac{d[\mathrm{mRNA}]}{dt} = k_t(1 - Y) - k_{dm}[\mathrm{mRNA}]$$

Example system dynamics

$$\frac{da}{dt} = v_a - k_a a$$

$$\frac{db}{dt} = V_b \frac{K_b d^{n_d}}{1 + K_b d^{n_d}} \cdot \frac{1}{1 + K_{ic} c^{n_c}} - k_b b$$

$$\frac{dc}{dt} = V_c \frac{K_c (ab)^{n_{ab}}}{1 + K_c (ab)^{n_{ab}}} - k_c c$$

$$\frac{dd}{dt} = V_d \frac{1}{1 + K_{ic} c^{n_c}} - k_d d$$

k_a, k_b, k_c, k_d – degradation rate constants
v_a – rate of gene a expression
V_b, V_c, V_d – maximal rates of b, c, d expression
K_b, K_c, K_{ic} – binding constants
n_d, n_{ab}, n_c – Hill coefficients

Credit: Klipp et al. *Systems Biology*, 2009.

Example system dynamics

Differential equation result

With these parameters,
Boolean result is a poor approximation

Boolean result

Credit: Klipp et al. *Systems Biology*, 2009.

Genetic regulation
Graphs
Boolean networks
Dynamics
Motifs
Summary

Motifs

Motifs are highly represented sub-graphs

real network

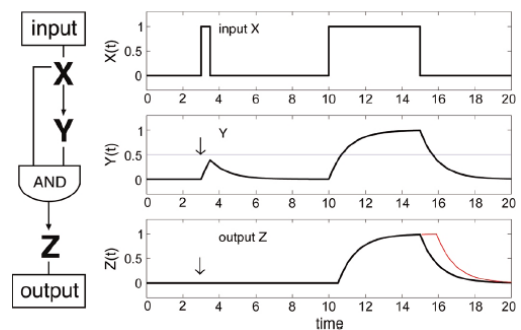randomized networks

motif:

Credit: Milo et al., *Science* 298:824, 2002..
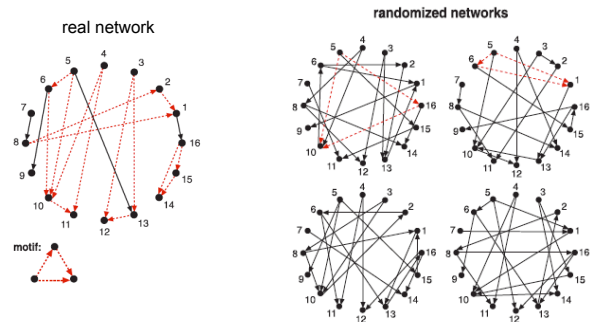
Transcriptional motifs

Motifs in *E. coli* transcriptional network

Credit: Shen-Orr et al., *Nat. Genetics* 31:64, 2002.
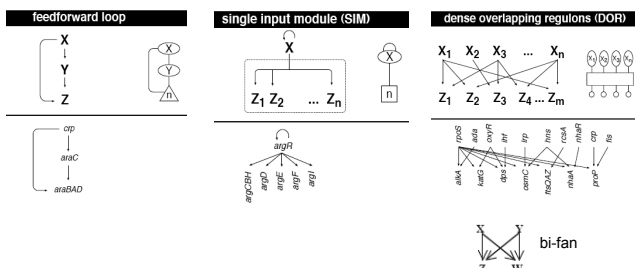
Feed-forward loop motif

filters out brief inputs – noise reduction

Credit: Shen-Orr et al., *Nat. Genetics* 31:64, 2002.

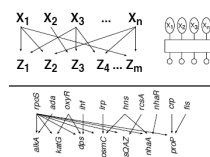## Single-input module motif

enables one TF to control a cascade of processes
(e.g. flagellar assembly)
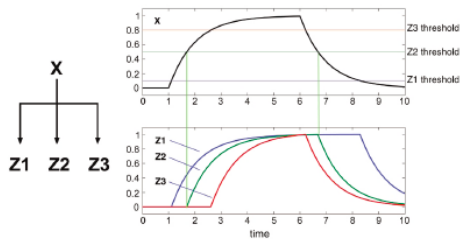
Dense overlapping regulon motif
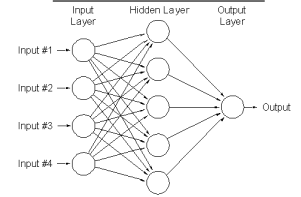
Shen-Orr *et al.* speculate that DORs
perform core computational tasks

Genetic regulation

Graphs

Boolean networks

Dynamics

Motifs

Summary

Summary

genetic regulation
 • transcription factors
 • miRNA, siRNA
graphs
 • random, scale-free
 • modular
Boolean networks
 • simple
 • discrete state space
dynamics
 • Hill functions for TF binding
motifs
 • feed-forward loop
 • single-input
 • dense overlapping regulon

Two books
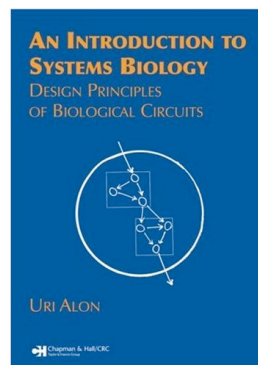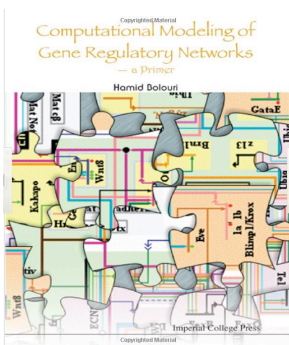
Homework

Next week's class is on stochasticity and robustness

(Back in Pelton)

<u>Read</u>
Rao, Wolf, Arkin, "Control, exploitation and tolerance of
intracellular noise" *Nature* 420:231, 2002.

(Arkin, Ross, and McAdams, "Stochastic kinetic analysis of
developmental pathway bifurcation in phage l-infected *Escherichia coli*
cells" *Genetics* 149:1633-1648, 1998.)