

# Network models (gene regulation)

Course Bioinformatic Processes 2018/2019; Paulien Hogeweg;  
Theoretical Biology and Bioinformatics Grp Utrecht University

Last time:

CA as modeling tool

generalizations of CA ss multiple model (variants) same/similar behavior  
(over-specification) baseline expectations: “pattern default”

CA and ODE/MAP's as dynamical systems

-alternative simplifications

-common features (types of attractors etc.)

- “almost all cases” ; order parameter - Mean field approximation/assumption

- d

## QUESTIONS?

TODAY

Model of models

- dynamics of mesoscale 'entities' (particles)

*modeling in terms of subsystems (cont)*

Network models

Boolean networks as model for gene regulatory networks:

- multiple attractors (= celltypes)

- domains of attraction, reachability, alternative transients.

- Understanding/interpreting gene knockouts

- Dynamic properties of Encode human regulatory network

# **Comparing/combining modeling approaches**

## **Model entities, model observables, modeltransformations (2) dynamics of mesoscale entities**

---

*Example: Multilevel description of CA  
Mesoscale entities of ECA 54  
(cf Crutchfield et al.)*

ECA 54, the rule under consideration here:

$$\phi(\eta) = \begin{cases} 0, & \eta \in \{111, 110, 011, 000\} \\ 1, & \eta \in \{101, 100, 010, 001\} \end{cases}$$

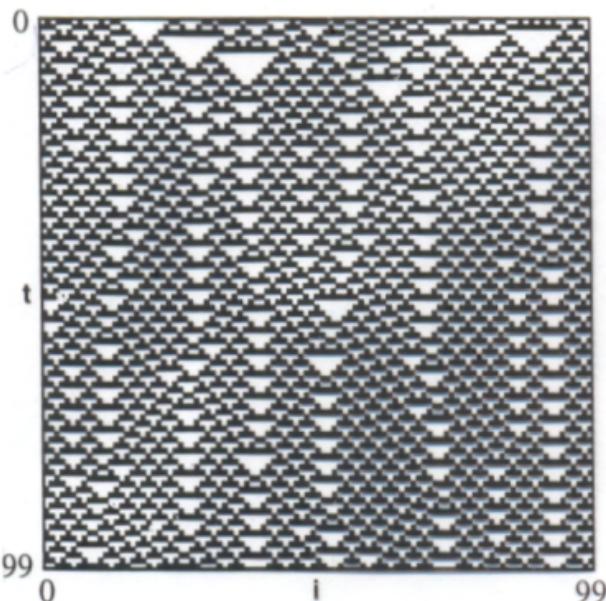


Figure 1. Space-time diagram of ECA 54, starting from an arbitrary initial condition. Boundary conditions are periodic. White squares are cells with value  $s_t^i = 0$ ; black squares are cells with  $s_t^i = 1$ .

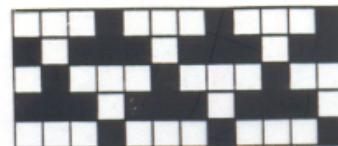


Figure 2. A portion of fig. 1 showing the domain  $\Lambda_{54}$ . Note the spatial phase shift of 2 cells every two iterations.

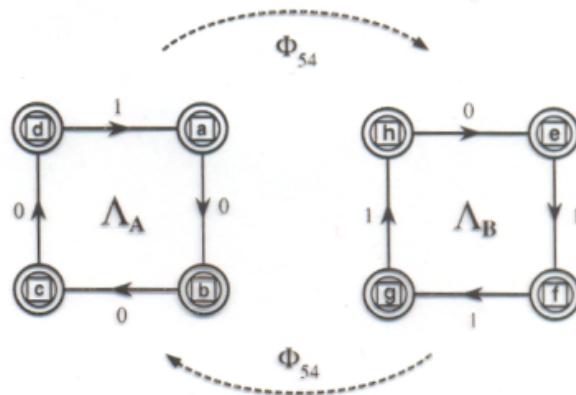


Figure 3. Process graph of ECA 54's domain  $\Lambda_{54}$ . The component on the left is  $\Lambda_A$ ; on the right is  $\Lambda_B$ . As the dotted lines indicate, they are mapped onto one another by the CA ensemble evolution operator,  $\Phi_{54}$ . In this and all following graphs of machines, inscribed circles and squares denote start and accept states, respectively.

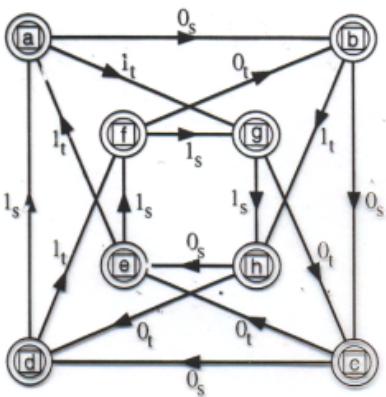


Figure 4. The process graph of  $M_{ST}(\Lambda_{54})$ , the space-time machine for  $\Lambda_{54}$ . States are labelled to correspond to those in the (purely spatial) domain machine  $M(\Lambda_{54})$  in fig. 3.

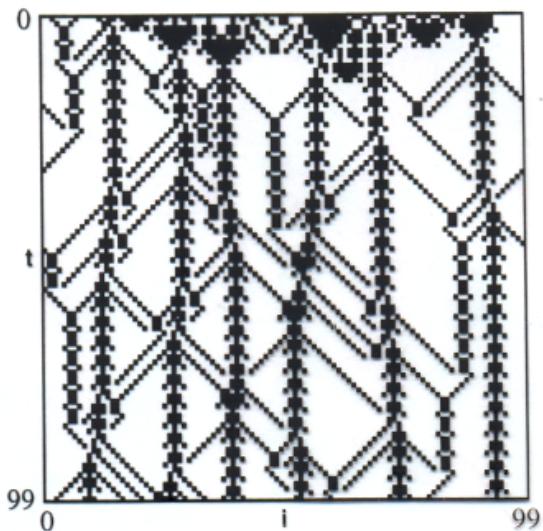


Figure 6. Space-time data of fig. 1, filtered with the domain transducer  $T_{54}^0$  of fig. 5. White cells correspond to sites participating in  $\Lambda_{54}$ ; black cells, to sites with values

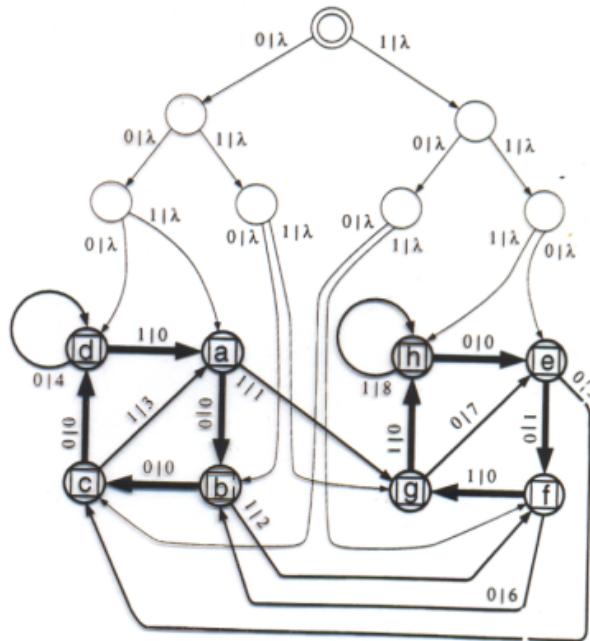


Figure 5. ECA 54's domain filter  $T_{54}^0$ , which maps sites in the domain to 0 and each defect to a unique output in  $\{1, \dots, 8\}$ . Labelled machine states correspond to the domain states of fig. 3.

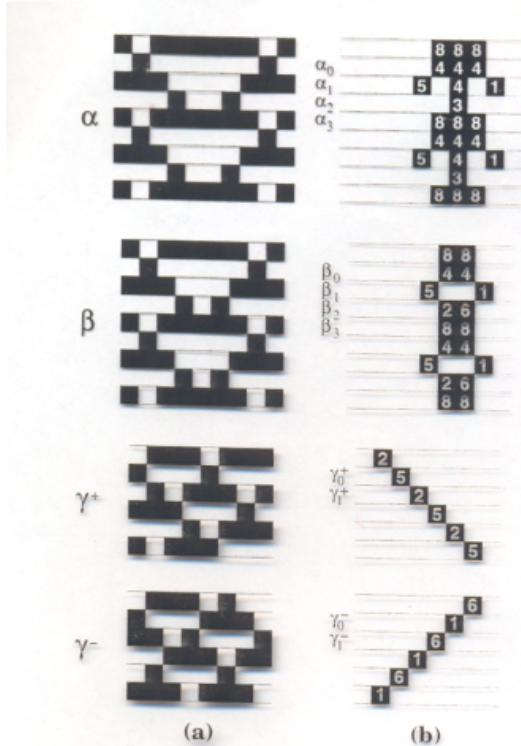


Figure 8. Basic wall structures ("fundamental particles") in space-time patterns of ECA 54. (a) Unfiltered space-time diagrams of the three types of particle  $\alpha$ ,  $\beta$ , and  $\gamma$ , described in the text. (b) Filtered diagrams of the same data, produced by  $T_{54}^0$ . Domain symbols are white cells. All defects are shown in black, with the defect symbol inscribed in white. The temporal phases of the particles, chosen by convention, are printed alongside the filtered strings.

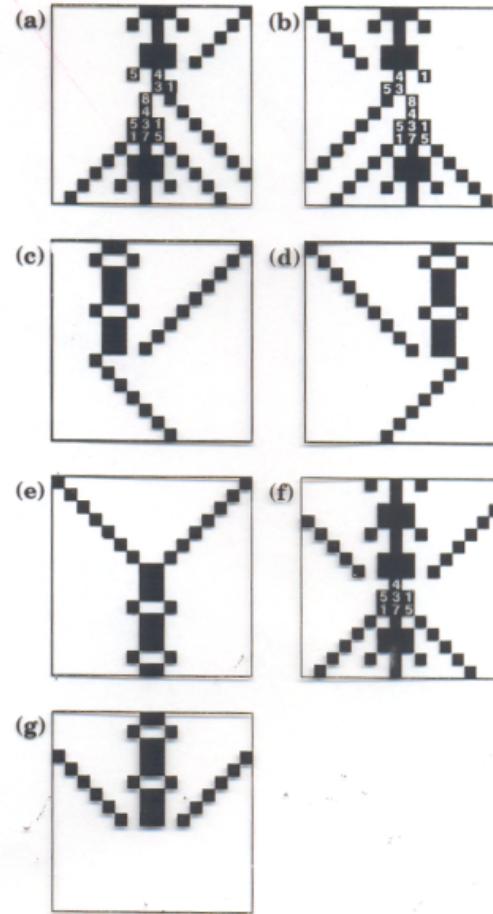


Figure 9. Filtered space-time diagrams of the fundamental interactions among ECA 54's fundamental particles, as listed in table 1. (a)–(e) are two-particle collisions; (f) and (g) are three-particle collisions. Filtering was done with the first version of particle filter  $T_{54}^1$  described in the text. The domain is shown as white, the particles  $\{\alpha, \beta, \gamma^+, \gamma^-\}$  are shown in black. Defects not corresponding to any of the particles are shown in black and have the corresponding  $T_{54}^0$  output symbols inscribed in white.

(a)	$\alpha + \gamma^- \rightarrow \gamma^- + \alpha + 2\gamma^+$
(b)	$\gamma^+ + \alpha \rightarrow 2\gamma^- + \alpha + \gamma^+$
(c)	$\beta + \gamma^- \rightarrow \gamma^+$
(d)	$\gamma^+ + \beta \rightarrow \gamma^-$
(e)	$\gamma^+ + \gamma^- \rightarrow \beta$
(f)	$\gamma^+ + \alpha + \gamma^- \rightarrow \gamma^- + \alpha + \gamma^+$
(g)	$\gamma^+ + \beta + \gamma^- \rightarrow \emptyset$

Table 1. Fundamental interactions among ECA 54's particles. Interactions (a) and (b) induce a spatio-temporal shift in the incident particles, as discussed in the text. Note that the spatial arrangement of input and output particles is respected by the interaction notation.  $\emptyset$  denotes no particles.

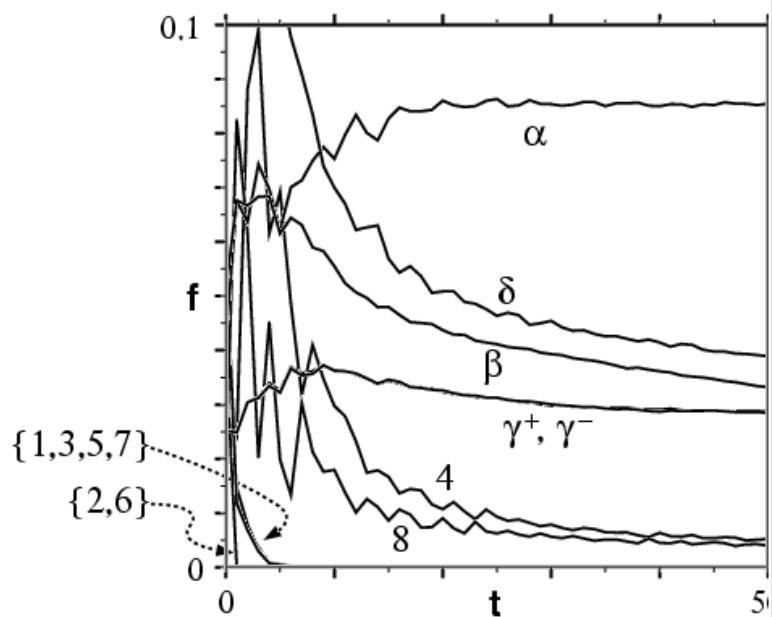
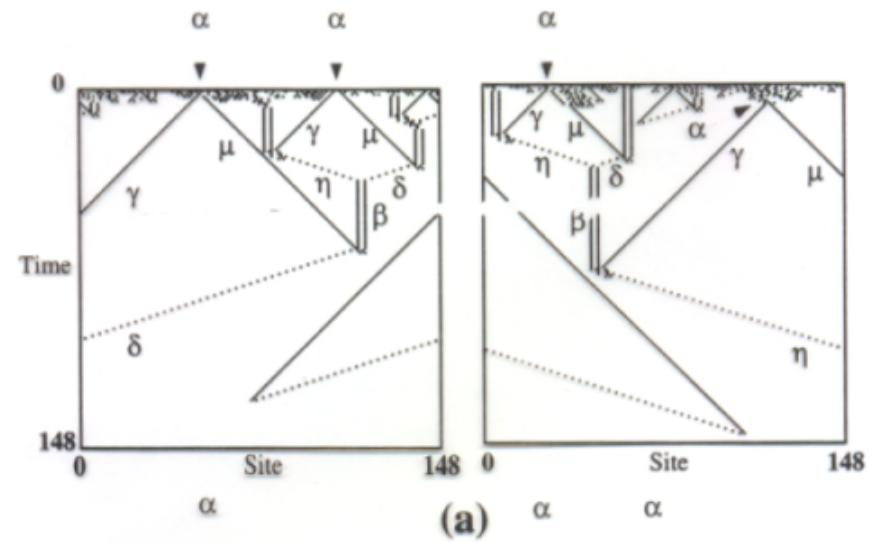
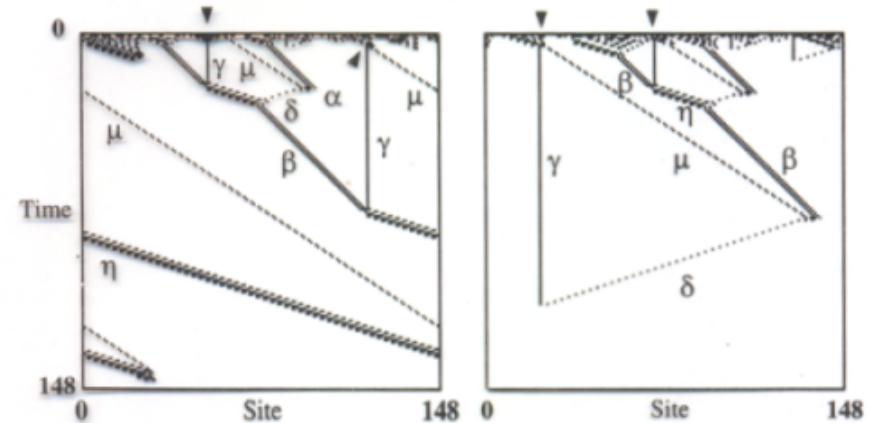


Figure 11. Fraction of the CA lattice devoted to the fundamental particles ( $\alpha, \beta, \gamma^+, \gamma^-$ ), to  $\alpha$ - $\gamma$  interactions ( $\delta$ ), and to unrecognized defects  $\{1, \dots, 8\}$  versus time.



(a)



(b)

Figure 8: (a) Version of Fig. 6(a) with the regular domains filtered out, revealing the particles and their interactions. (b) Filtered version of Fig. 6(b).

## conclusions

---

Detection of mesoscale entities

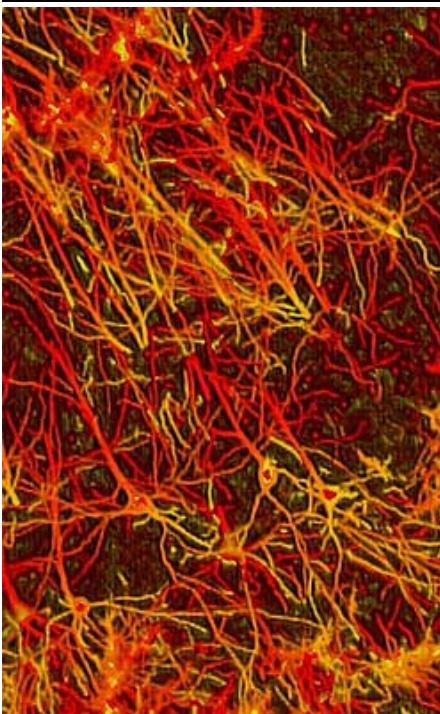
*WITH A DYNAMICS OF THEIR OWN*

Description of system in variable set of  
interacting higher level entities (individual based models)  
("beyond dynamical systems")

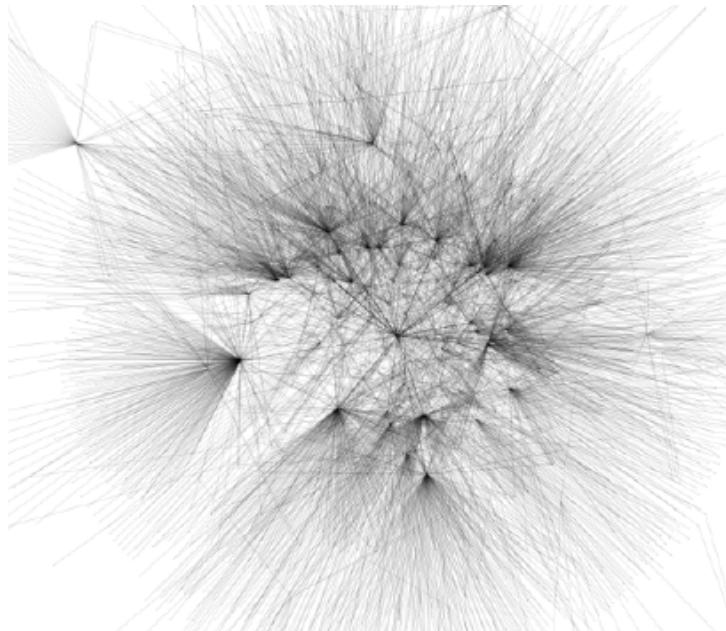
Description in terms of populations of these entities

# dynamical systems: decomposition in many simple systems cont., NETWORKS

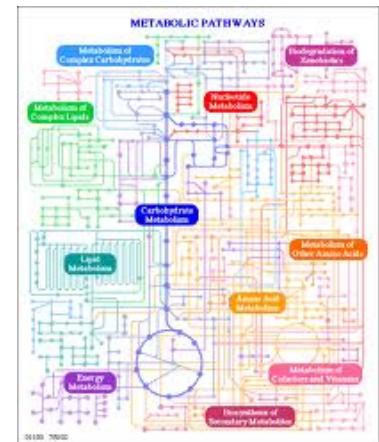
---



Neural net  
connected



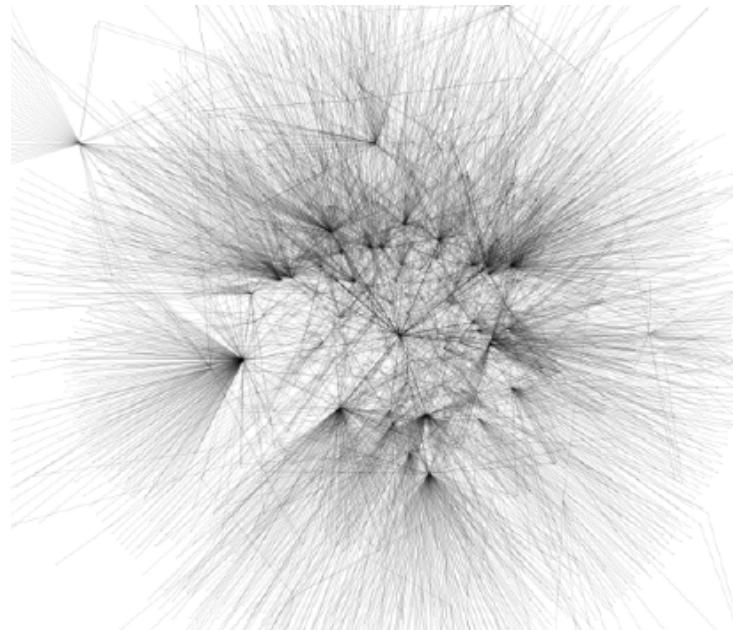
yeast transcription net  
information transfer



Keg metabolic net  
mass conservation  
stoichiometric

# Gene regulation Networks: “full” transcription network of yeast

---



*How does it behave?*

*how special is it?*

*(evolution)*

## Boolean Networks

---

Proposed by S.Kauffman (1969) as model for gene regulation

Like binary CA but  
specific network structure (IO relations)  
specific interaction (not local)  
each node own transition rule  
(Boolean function with k inputs)

**Boolean network : special cases can be mapped into CA  
(homogeneous network structure, “rule-layer” )**

## Multiple attractors

---

**rules**

1	2	3	4	5	6	7	8	9	10
8	6	4	12	5	5	1	7	2	1

**attractors**



basin of attraction

112

cyclength

3

128

4

272

6

512

12

## What kind of behavior do we expect from gene regulation networks?

---

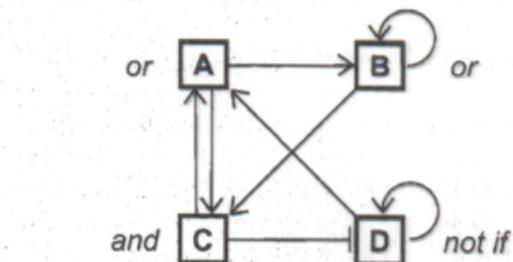
*multiple attractors (cell types)*

*alternative trajectories from A' and A'' to B*

*multiple causes*

*robustness (knockouts)*

**a Network wiring diagram**



Boolean functions:

A: " or "

INPUTS		OUTPUT
C	D	A
0	0	0
0	1	1
1	0	1
1	1	0

B: " or "

INPUTS		OUTPUT
A	B	B
0	0	0
0	1	1
1	0	1
1	1	0

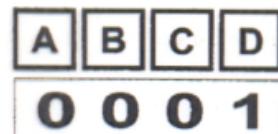
C: " and "

INPUTS		OUTPUT
A	B	C
0	0	0
0	1	0
1	0	0
1	1	1

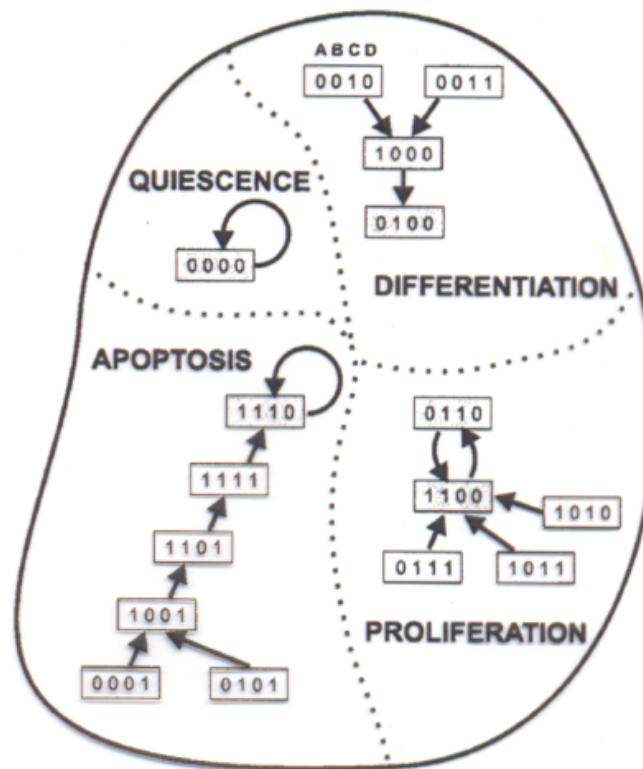
D: " not if "

INPUTS		OUTOUT
C	D	D
0	0	0
0	1	1
1	0	0
1	1	0

**b A network state**

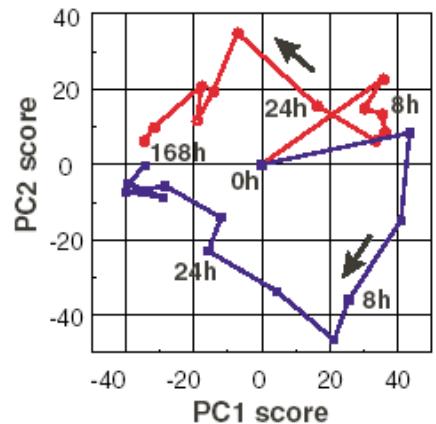
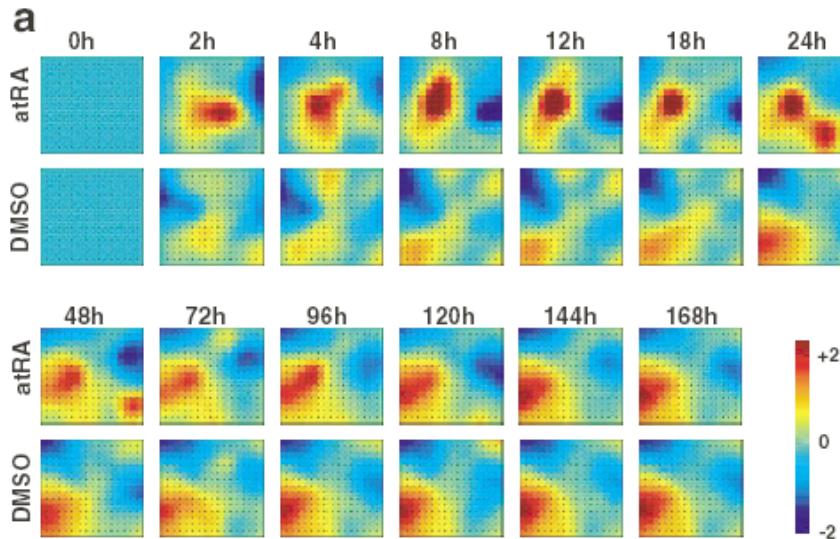


**c Protein activity state space**



# 2 pathways to Neutrophyl differentiation

## Huang et al 2005 (Phys Rev Letters)



gene expression through time

2773 dim state-space,

trajectories in 2D projection

$n^{2773}$  states!

# Robustness: Forcing structures

---

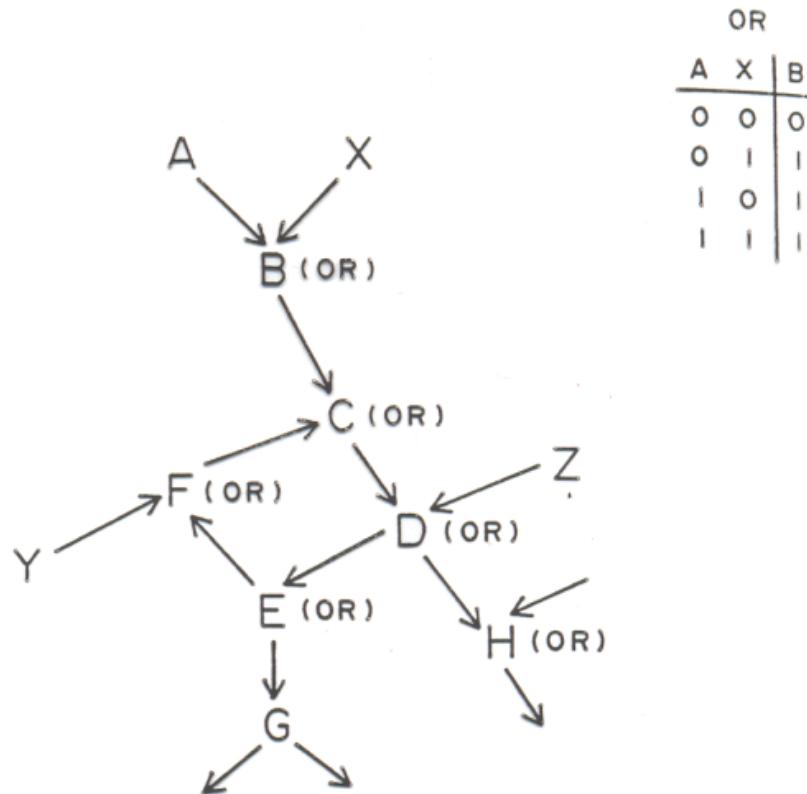


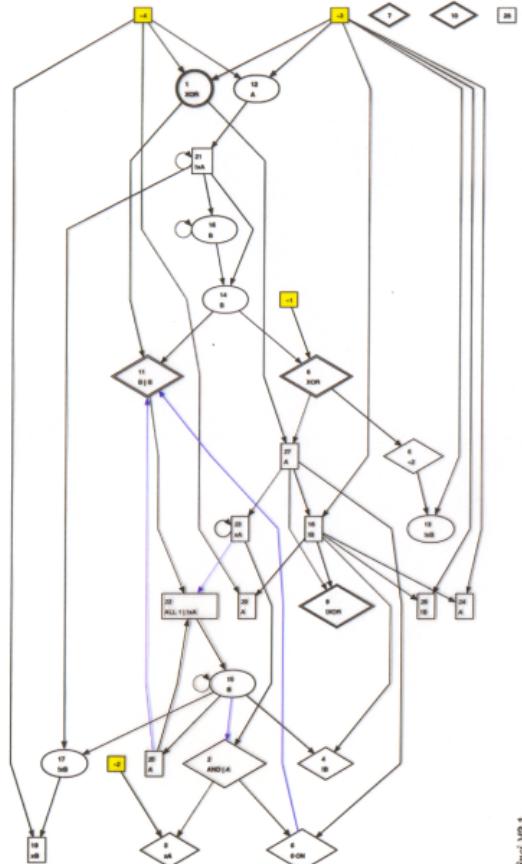
FIGURE 8 Forcing structure among binary elements governed by the Boolean OR function. The forcing "1" value propagates down structure and around forcing loop which eventually is "frozen" into the forced state with "1" values at all elements around the loop. Loop then radiates fixed forced values downstream. From *Origins of Order: Self Organization in Evolution* by S. A. Kauffman. Copyright © 1990 by Oxford University Press, Inc. Reprinted by permission.



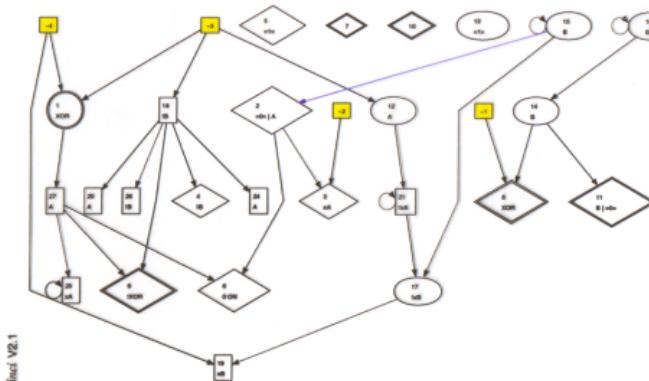
# Evolved gene regulation networks

"existing" vs "functioning"

?? "false positives" ??



53 links



32 links



16 celltypes

18 reg. genes

# Properties of Random Boolean Networks (depending on K)

---

TABLE 1 Properties of Random Boolean Nets for Different Values of  $K^1$

	State Cycle Length	Number of State Cycle Attractors	Homeostatic Stability	Reachability Among Cycles After Perturbation
$K = N$	$0.5 \times 2^{N/2}$	$N/e$	Low	High
$K > 5$ $(B > 1)$	$0.5 \times 2^{BN}$	$\sim N \left[ \frac{\log(\frac{1}{1/2 \pm \alpha})}{2} \right];$ $\alpha = p(K) - 1/2$	Low	High
$K = 1$	Very Long	Very Many	Low	High
$K = 2$	$\sqrt{N}$	$\sqrt{N}$	High	Low

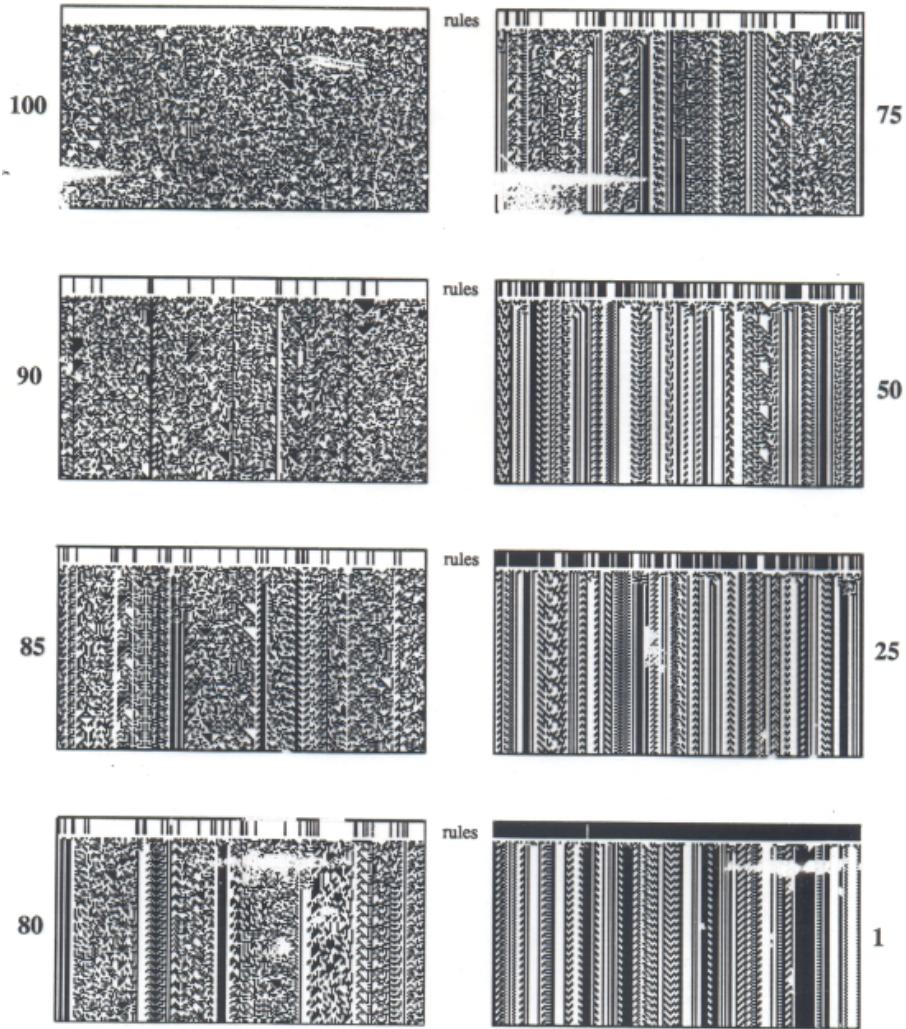
<sup>1</sup> Column 1: state cycle length is median number of states on a state cycle.  
 Column 2: number of state cycle attractors in behavior of one net.  
 ( $\alpha = P_K - 1/2$ , where  $P_K$  is mean internal homogeneity of all Boolean functions on  $K$  inputs; see text.) Column 3: homeostatic stability refers to tendency to return to same state cycle after transient reversal of activity of any one element.  
 Column 4: reachability is number of other state cycles to which net flows from each state cycle after all possible minimal perturbations, due to reversing activity of one element.

**Importance of sampling method: Dependence on K is  
dependence on fraction (non) forcing rules!**

---

# Non forcing rules in 1D CA ( $k=2$ )

---



## conclusion: Boolean Kaufman Networks

---

Important:

Identification of cell state with attractor of gene regulation network

Multiple attractors in simple networks

alternative trajectories to attractor

Domain of attraction: i.e. “robustness”

forcing functions i.e. “robustness”

NOT IMPORTANT (WRONG!) connectivity of 2 “ideal”

## Gene expression data --> Boolean networks

---

*Functional Overlap and Regulatory Links Shape Genetic Interactions between Signalling Pathways*

Sake van Wageningen, Patrick Kemmeren,..... Berend Snel  
and Frank C.P. Holstege Cell Dec 2010

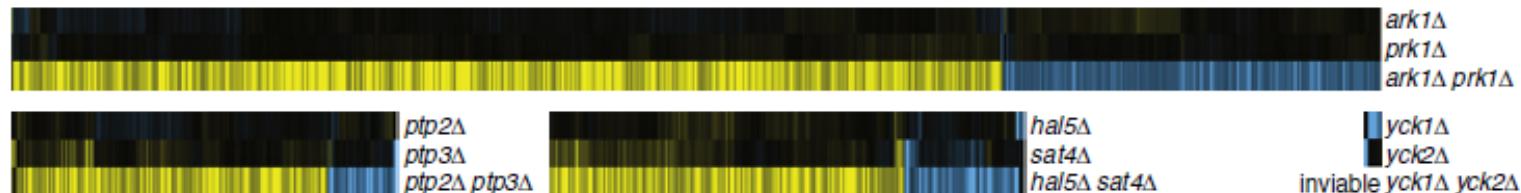
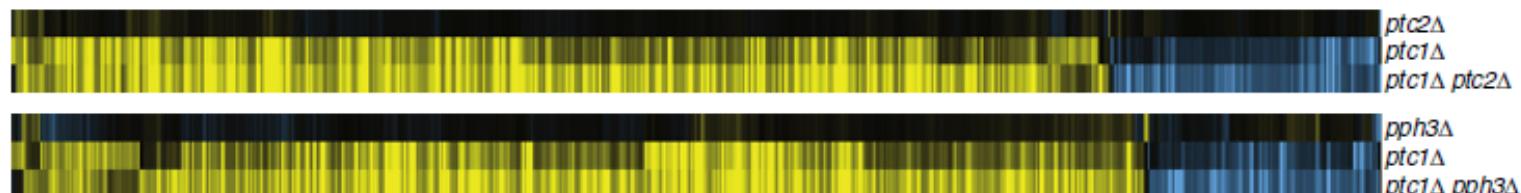
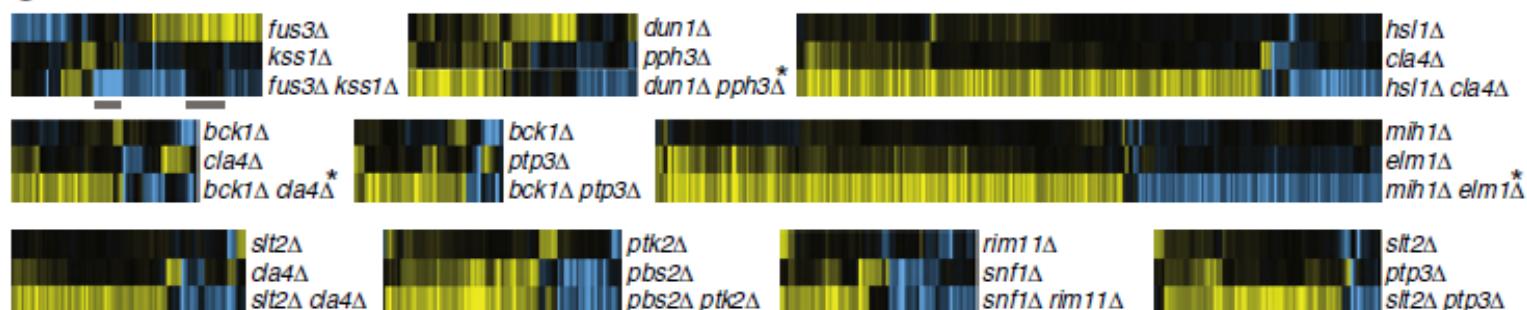
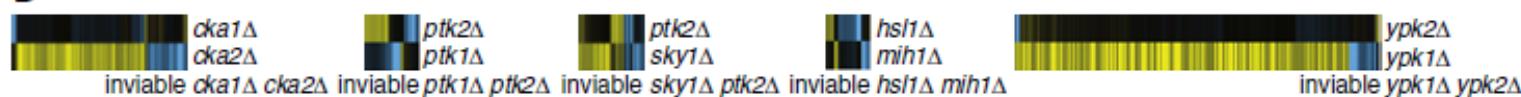
141 kinases, 38 phosphatases in Yeast.

60% single knockouts “no phenotype”  
(== <8 genes different of WT) (single growth condition)

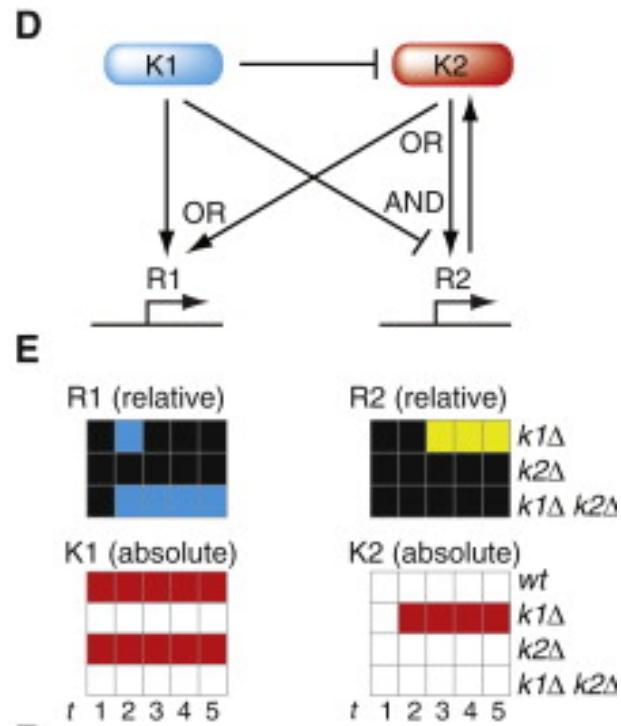
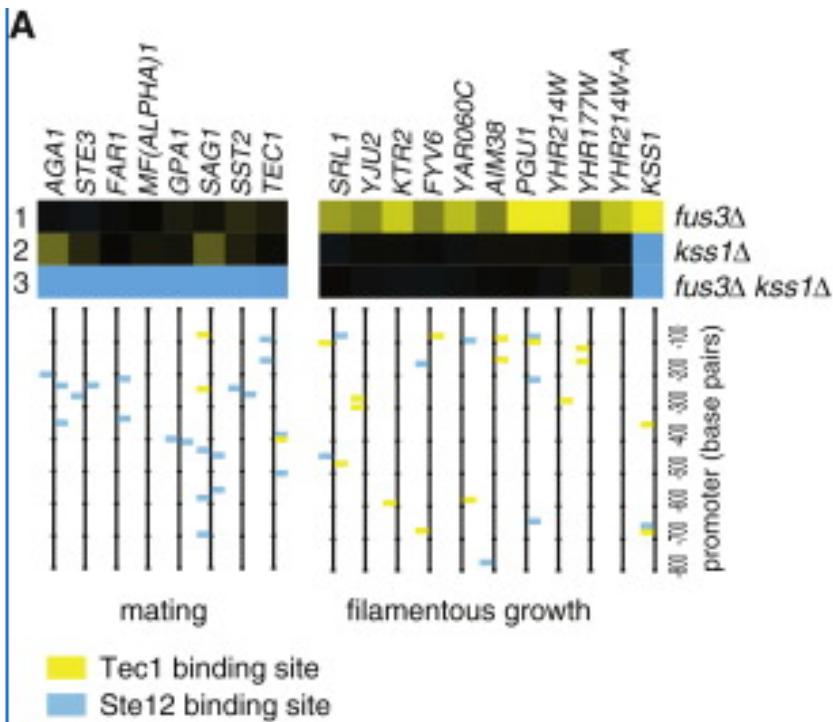
Double knockouts: 21 buffering effects with other kinase/phosphatasse

## **double knockout expression profiles**

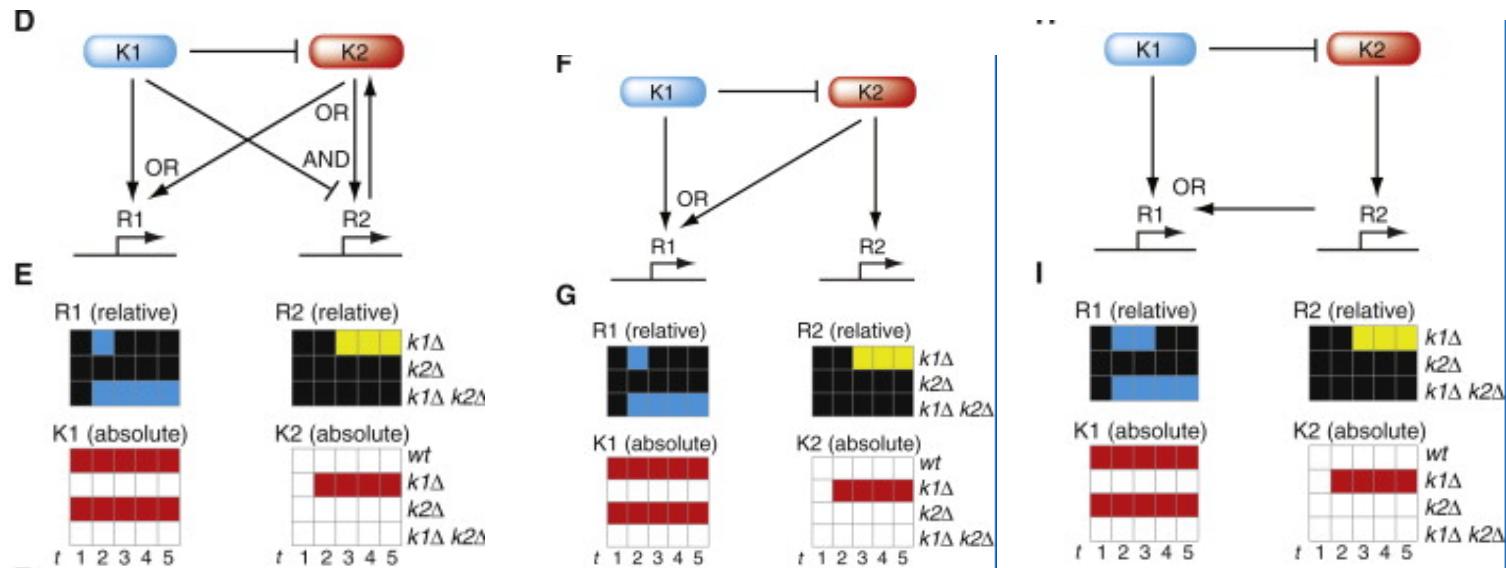
---

**A****B****C****D**

# example of mixed epistasis filamentous growth vs mating



## 2 simpler networks with same effect (complexer network most similar to exp. inferred network)



Many networks (max 2 inputs per node) with same effect!

---

Edges	Models	Correct models	Root models
2	0	0	0
3	352	0	0
4	4,960	2	2
5	32,896	6	1
6	129,280	24	8
7	294,912	28	7
8	331,776	46	10

# all buffering pairs: Many non-homologs!; many mixed

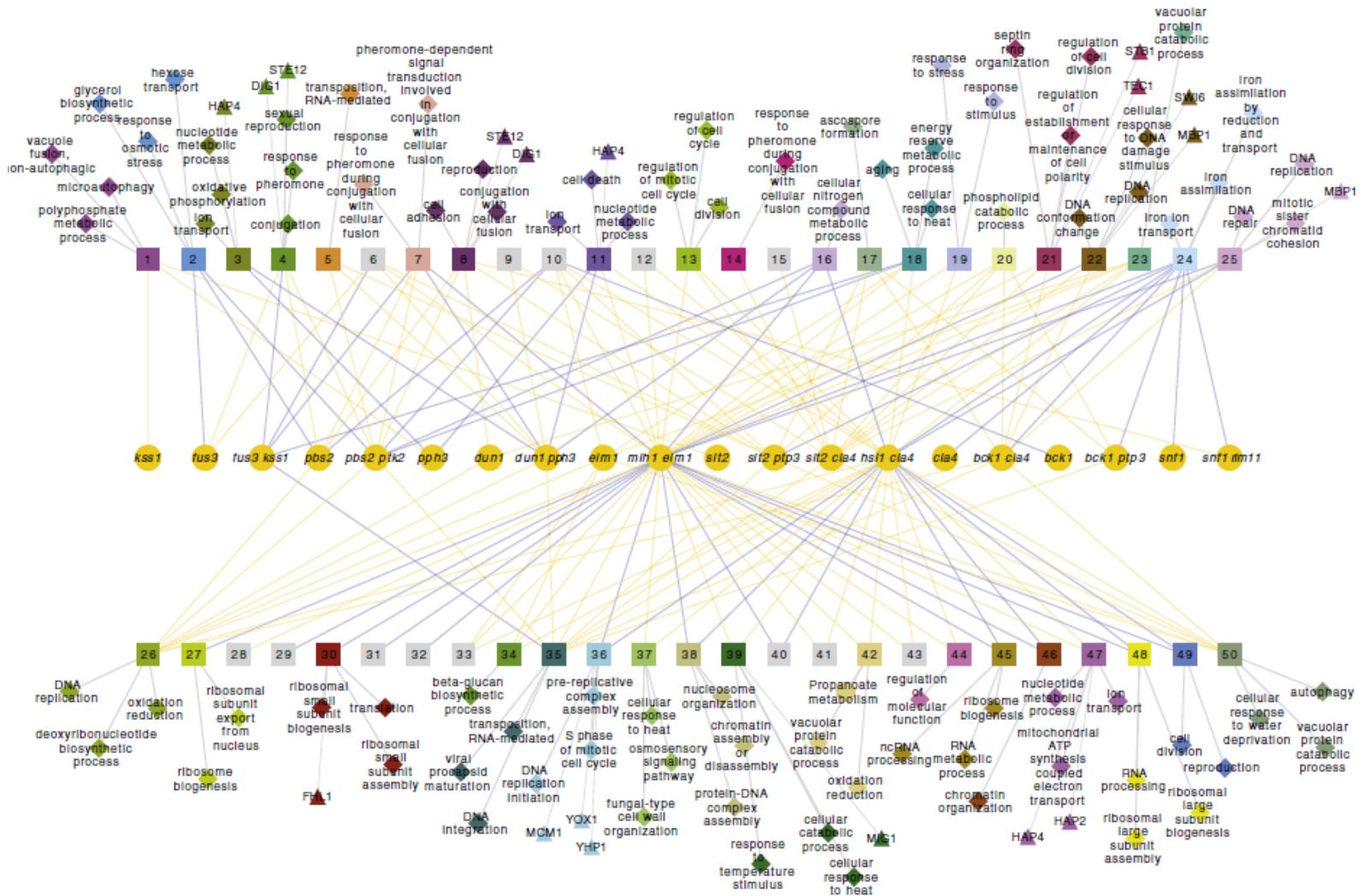
---

**Table 1. Buffering Relationships between Kinases and Phosphatases**

Gene 1	Gene 2	Type	Duplication	Time (Years Ago)	Buffering Relationship
<i>HAL5</i>	<i>SAT4</i>	kk	old	600 M – 2 G	complete redundancy
<i>ARK1</i>	<i>PRK1</i>	kk	whole genome	125 M	complete redundancy
<i>PTP2</i>	<i>PTP3</i>	pp	recent	125 M – 600 M	complete redundancy
<i>YCK1</i>	<i>YCK2</i>	kk	whole genome	125 M	complete redundancy <sup>a</sup>
<i>PTC1</i>	<i>PTC2</i>	pp	old	600 M – 2 G	quantitative redundancy
<i>PTC1</i>	<i>PPH3</i>	pp	not homologous		quantitative redundancy
<i>PBS2</i>	<i>PTK2</i>	kk	ancient	>2G	mixed epistasis
<i>CLA4</i>	<i>SLT2</i>	kk	ancient	>2G	mixed epistasis
<i>CLA4</i>	<i>HSL1</i>	kk	ancient	>2G	mixed epistasis
<i>SNF1</i>	<i>RIM11</i>	kk	ancient	>2G	mixed epistasis
<i>BCK1</i>	<i>PTP3</i>	kp	not homologous		mixed epistasis
<i>SLT2</i>	<i>PTP3</i>	kp	not homologous		mixed epistasis
<i>FUS3<sup>b</sup></i>	<i>KSS1</i>	kk	recent	125 M – 600 M	mixed epistasis
<i>ELM1</i>	<i>MIH1</i>	kp	not homologous		mixed epistasis <sup>a</sup>
<i>CLA4</i>	<i>BCK1</i>	kk	ancient	>2G	mixed epistasis <sup>a</sup>
<i>DUN1</i>	<i>PPH3</i>	kp	not homologous		mixed epistasis <sup>a</sup>
<i>CKA2</i>	<i>CKA1</i>	kk	recent	125 M – 600 M	not classified <sup>a</sup>
<i>YPK1<sup>b</sup></i>	<i>YPK2</i>	kk	whole genome	125 M	not classified <sup>a</sup>
<i>PTK1</i>	<i>PTK2</i>	kk	whole genome	125 M	not classified <sup>a</sup>
<i>HSL1</i>	<i>MIH1</i>	kp	not homologous		not classified <sup>a</sup>
<i>SKY1</i>	<i>PTK2</i>	kk	ancient	>2G	not classified <sup>a</sup>

**regulatory network via mixed response networks**

---



**Figure 6. Multiprocess Control through Signaling Components with Mixed Epistasis**

# Boolean networks

## Boolean functions vs threshold functions

---

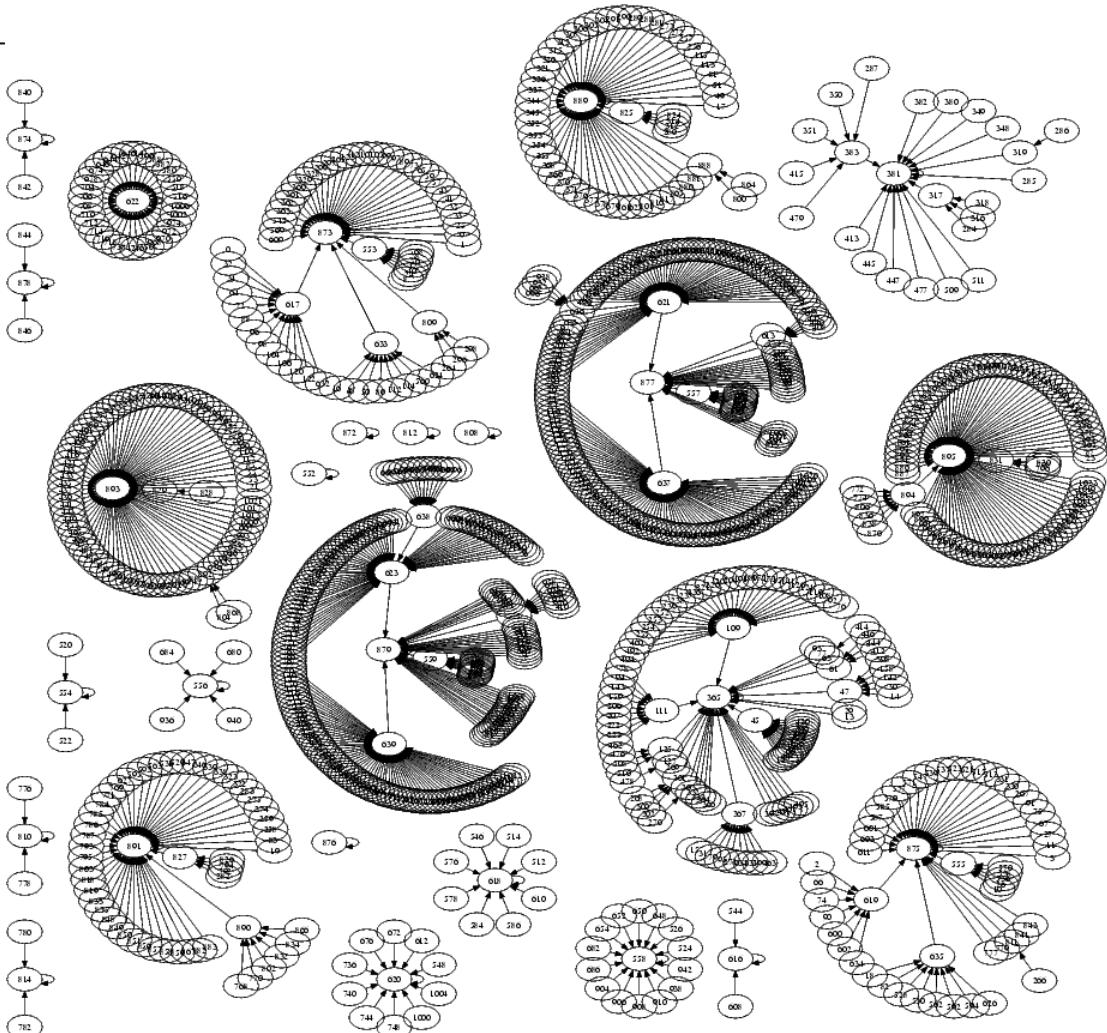
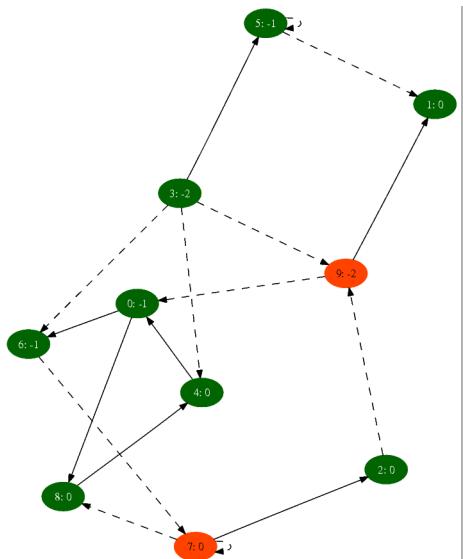
$$V_j(t+1) = \begin{cases} 1, & \text{if } \sum_k T_{jk} V_k(t) + I_j > 0 \\ 0, & \text{otherwise} \end{cases}$$

*how to make a XOR?*

# simple random network (threshold dynamics)

$$V_j(t+1) = \begin{cases} 1, & \text{if } \sum_k T_{jk} V_k(t) + \\ & \text{otherwise} \end{cases}$$

(Anton Crombach)  
MULTIPLE  
ATTRACTORS



ONLY 10 nodes (=genes)!

STATESPACE