COMPLEX NETWORKS OF BIOMOLECULAR INTERACTIONS

François Képès

Epigenomics Project

Genopole®, CNRS, University of Évry

Francois.Kepes@epigenomique.genopole.fr

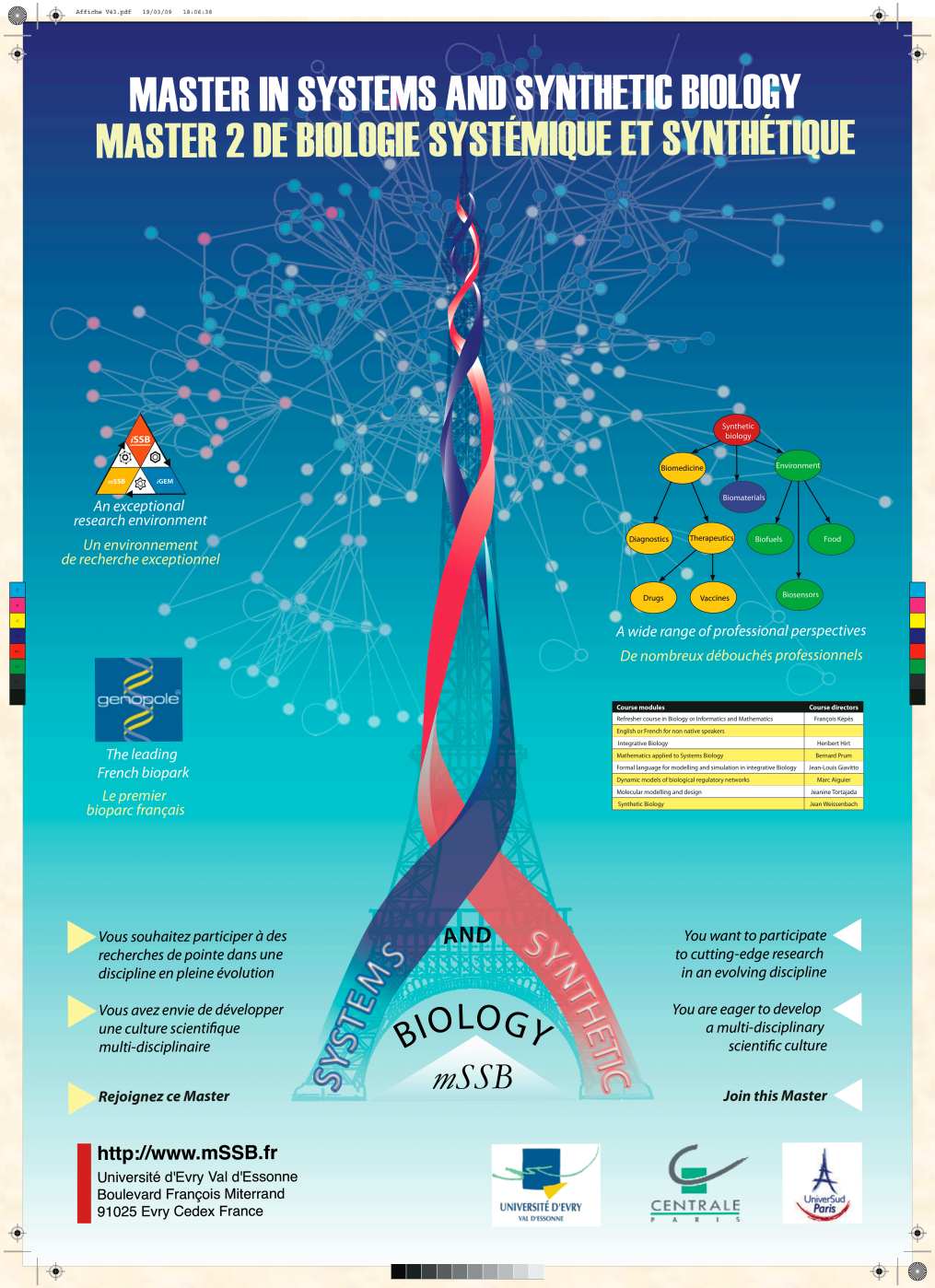
ICTP, June 8th, 2009

master in Systems and Synthetic Biology

www.mSSB.fr

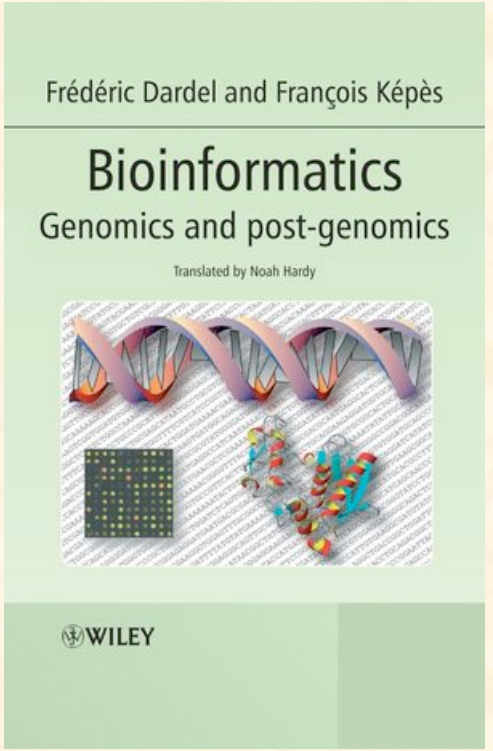
Deadline for application

June 15th



COMPLEX NETWORKS OF BIOMOLECULAR INTERACTIONS

Complex Systems And Interdisciplinary Science - Vol. 3
BIOLOGICAL NETWORKS
edited by François Képès



CONTENT

- **Interacting partners and their interaction**
- **Methods to obtain a network of interactions**
- **Global structure of the network of interactions**
- **Local structure**
- **Dynamics and Modularity**

Interacting partners and their interaction

PROTEIN / PROTEIN

Interactions between proteins

Proteomics

METABOLISM

Interactions between enzymes and metabolites

Metabolomics

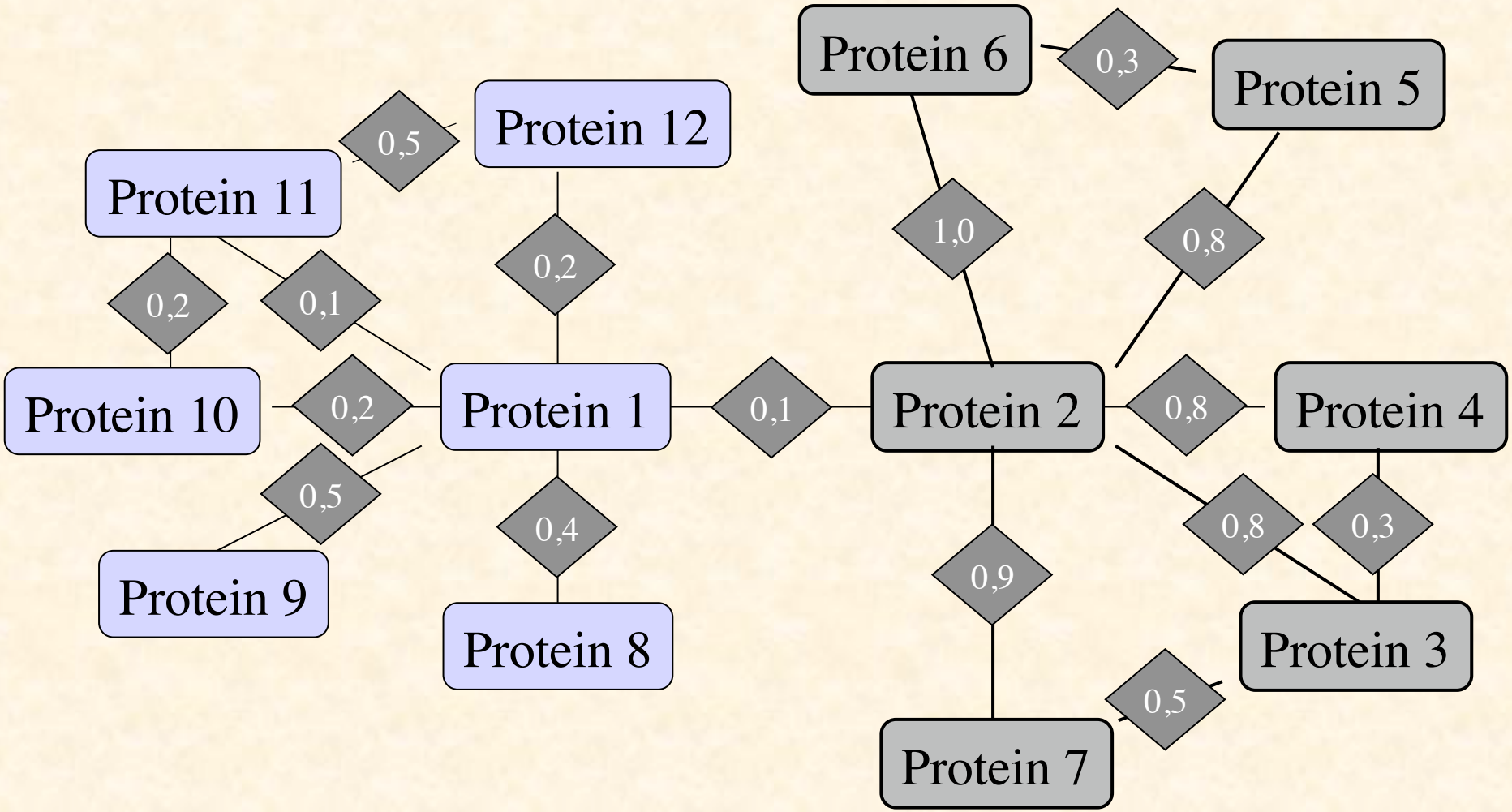
PROTEIN / DNA

Interactions between regulatory proteins
and DNA regulatory regions

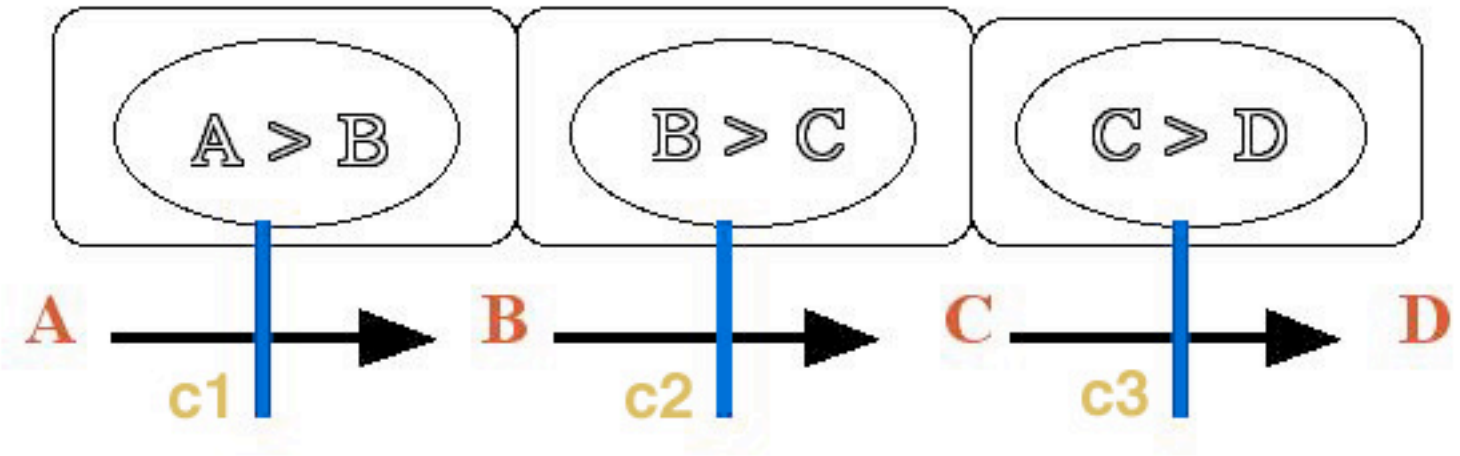
Transcriptomics

OTHERS ...

Interactions between proteins

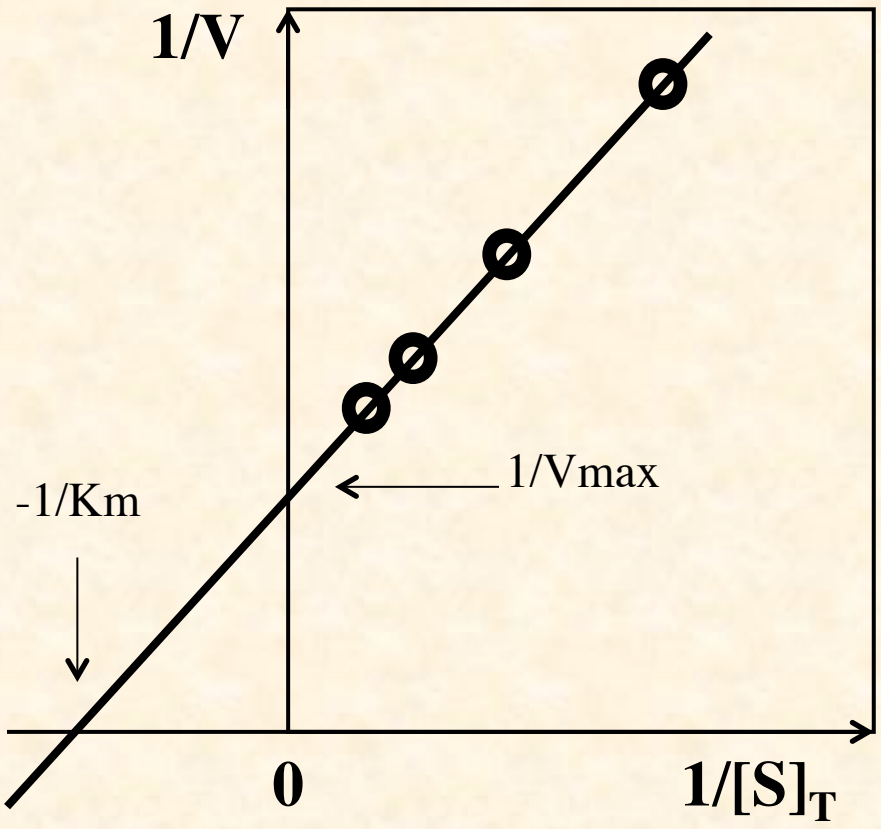


Interactions between enzymes and metabolites



Interactions between enzymes and metabolites

Michaelis-Menten Law

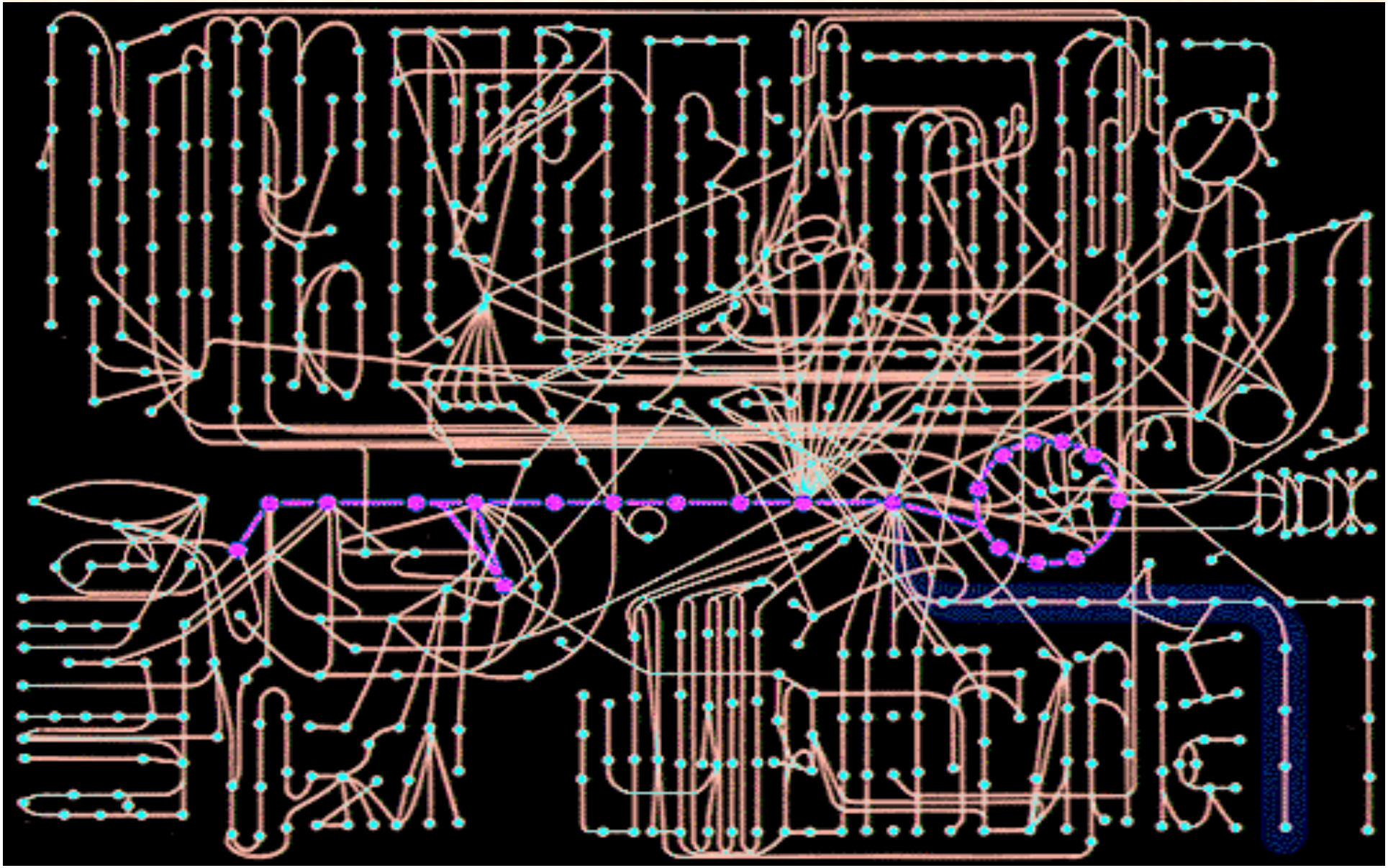


$$1/V = 1/Vmax + Km/Vmax \cdot 1/[S]_T$$

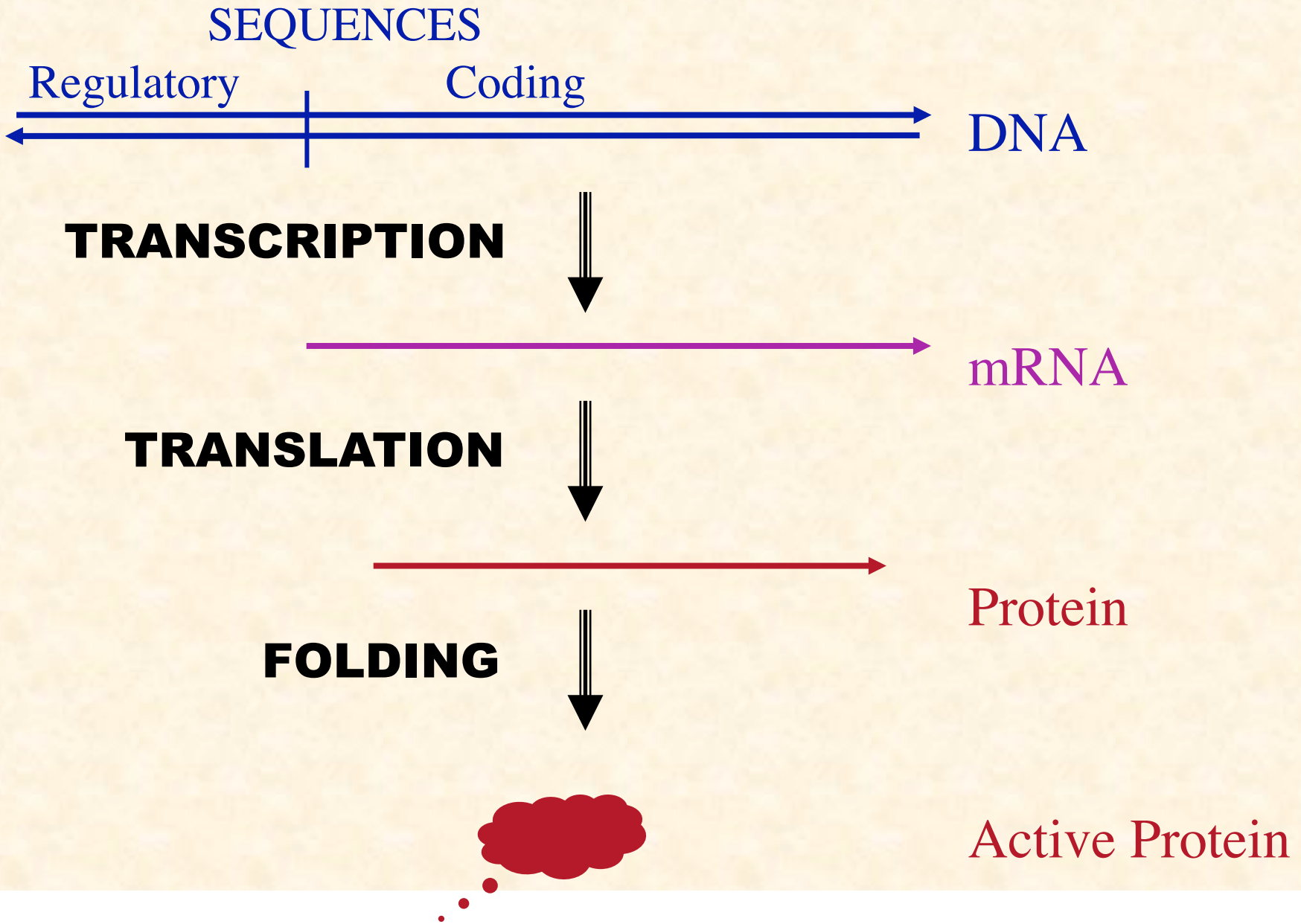
Lineweaver-Burk Equation

Interactions between enzymes and metabolites

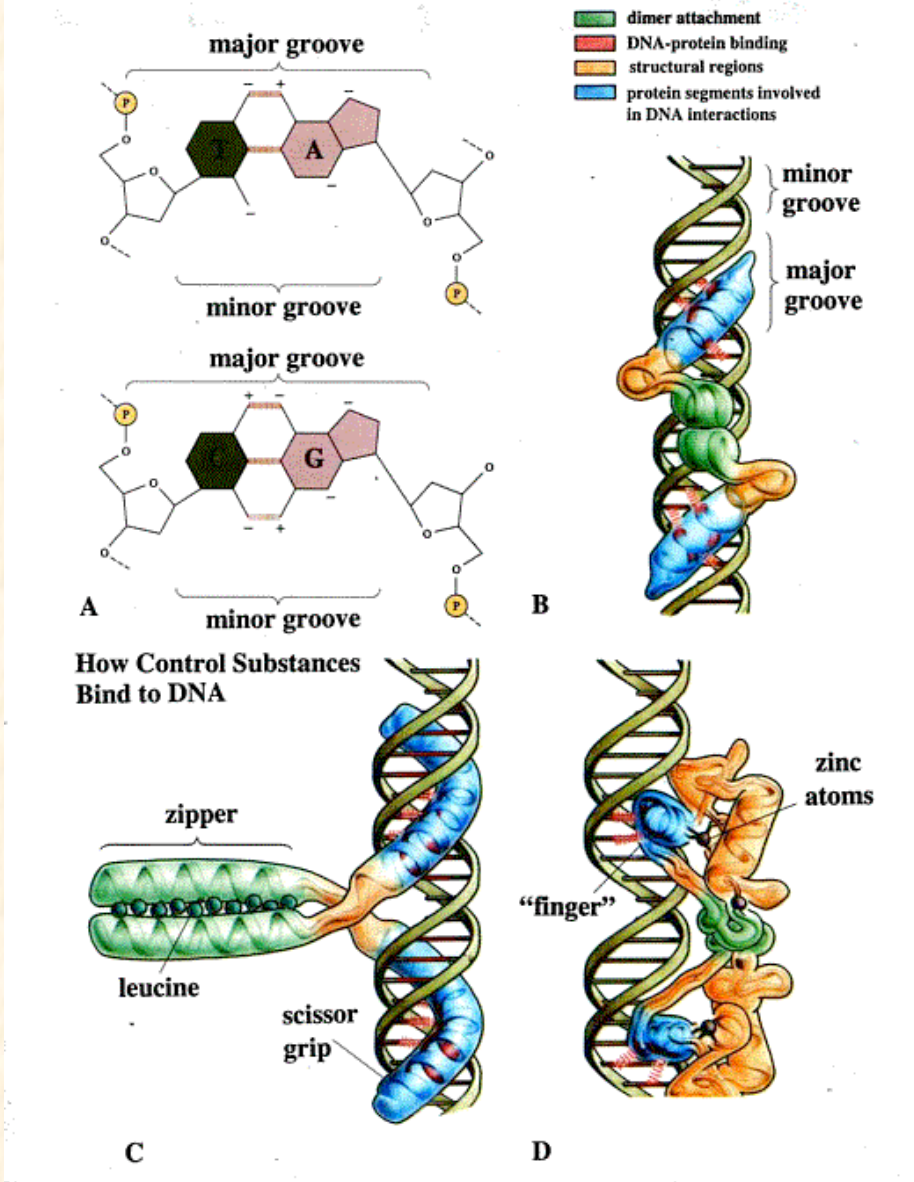
Metabolome



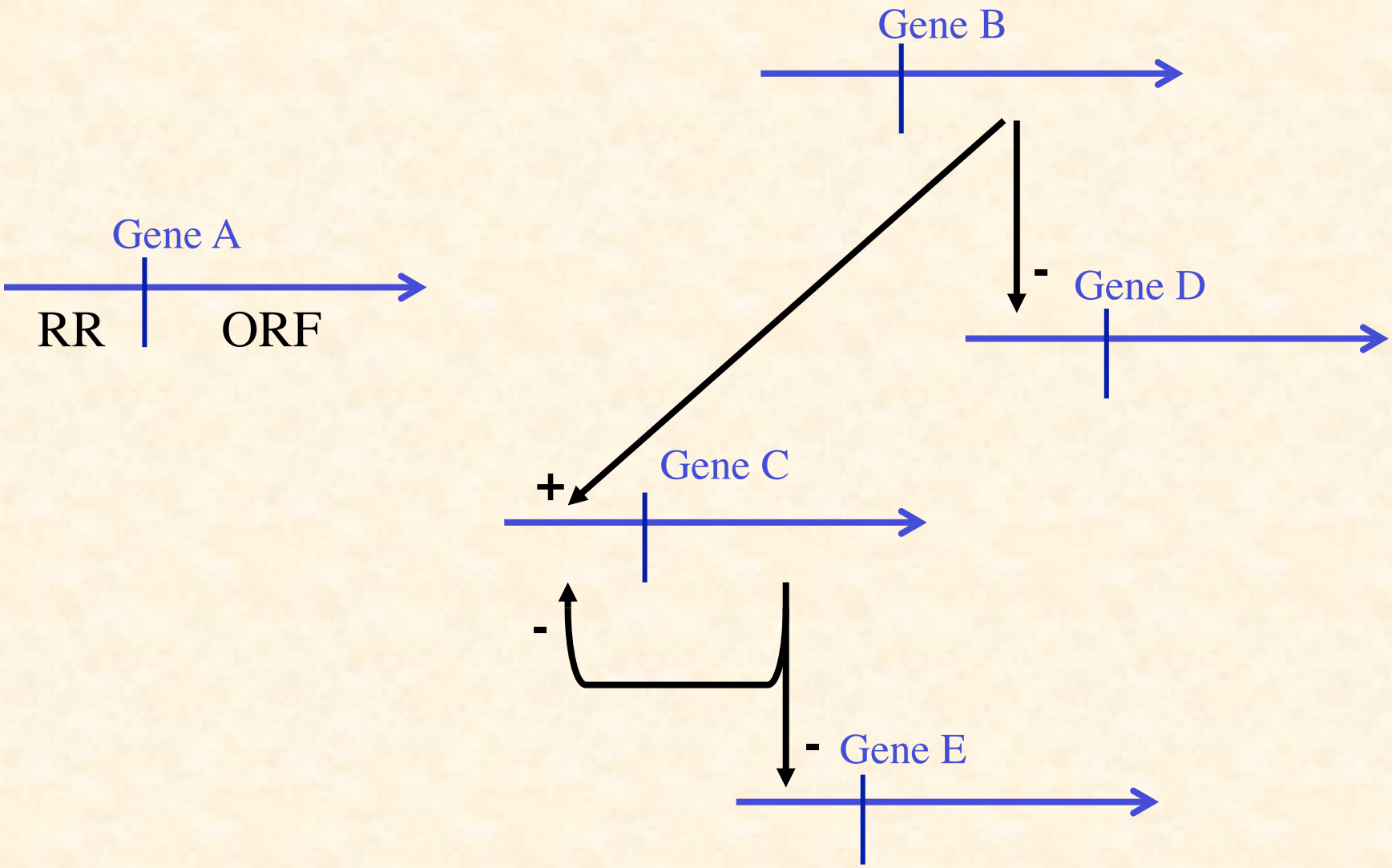
Interactions between regulatory proteins and DNA regulatory regions



Interactions between regulatory proteins and DNA regulatory regions



Interactions between regulatory proteins and DNA regulatory regions



PROTEIN / PROTEIN

Interactions between proteins

Proteomics

METABOLISM

Interactions between enzymes and metabolites

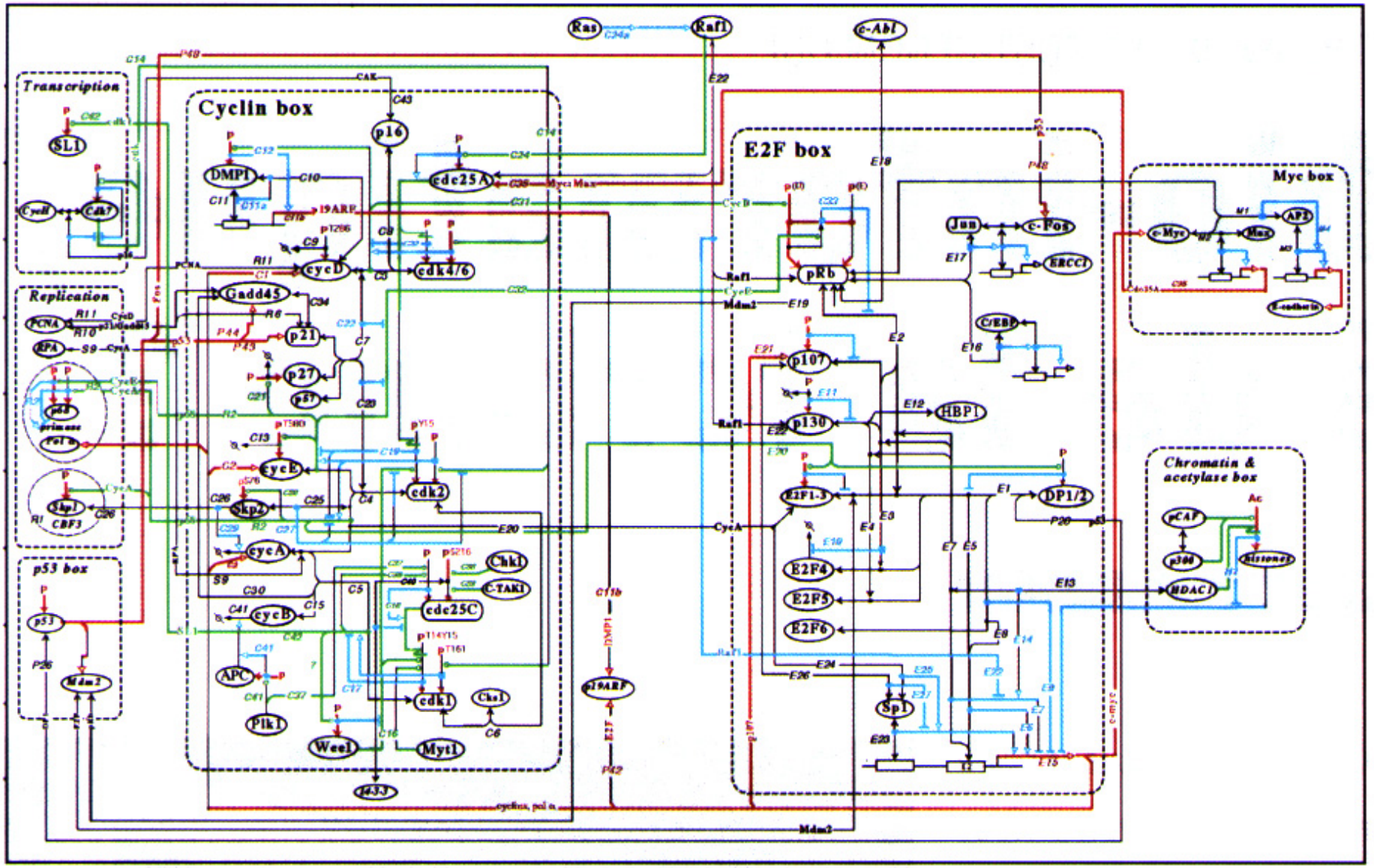
Metabolomics

PROTEIN / DNA

Interactions between regulatory proteins
and DNA regulatory regions

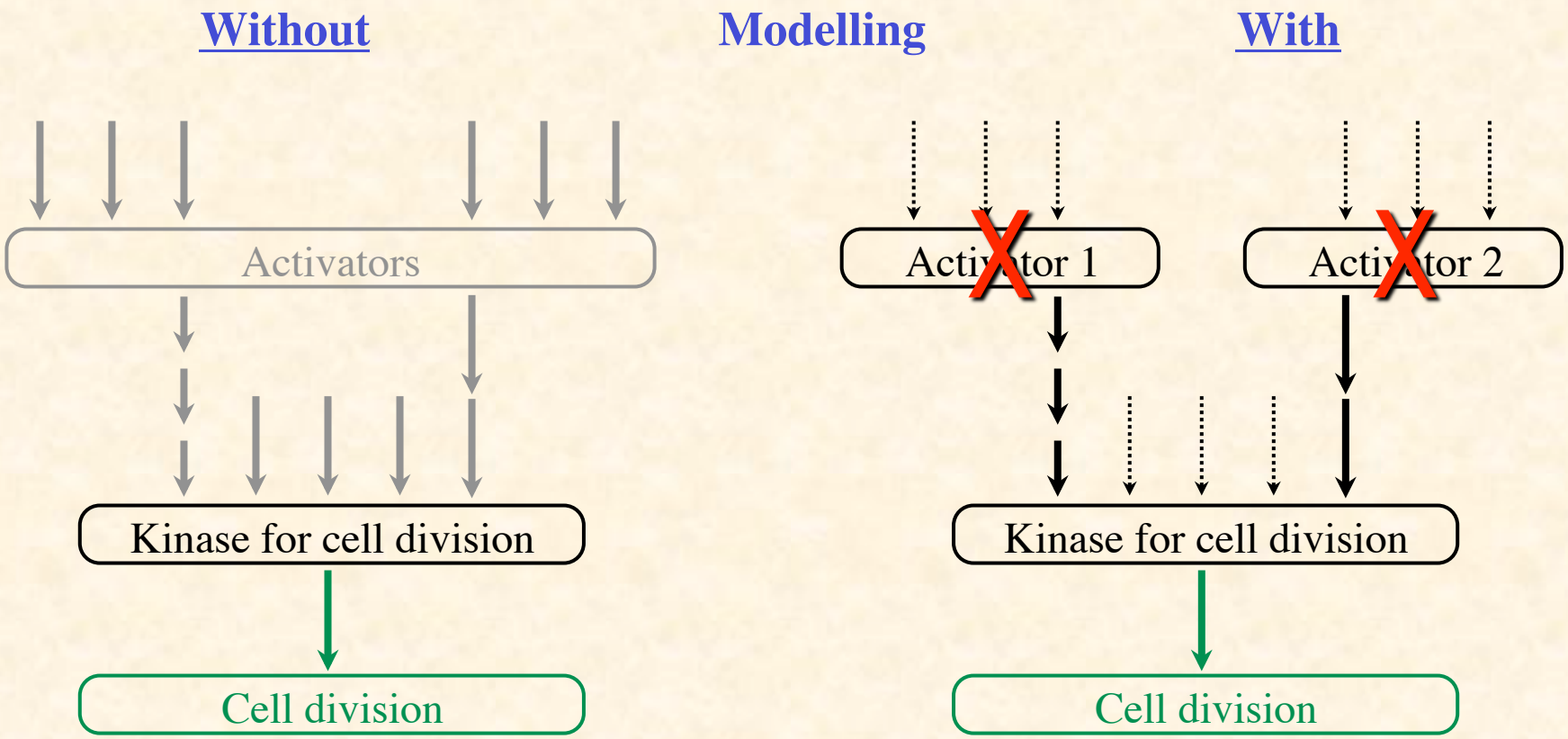
Transcriptomics

A MIXED NETWORK



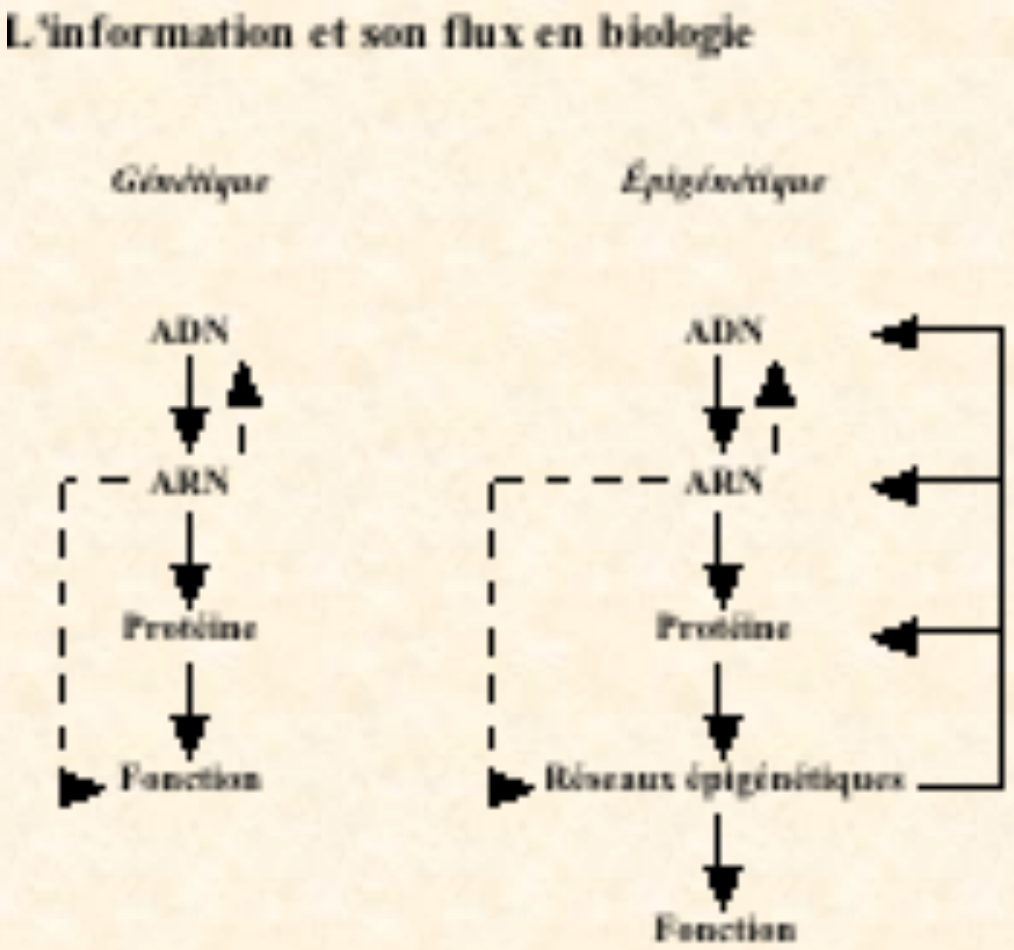
Investment of the Northern American Big Pharma industry for Systems Biology

An example : the discovery of new drugs by systematic screening of a combinatorial chemical bank



Information Flux in Biology

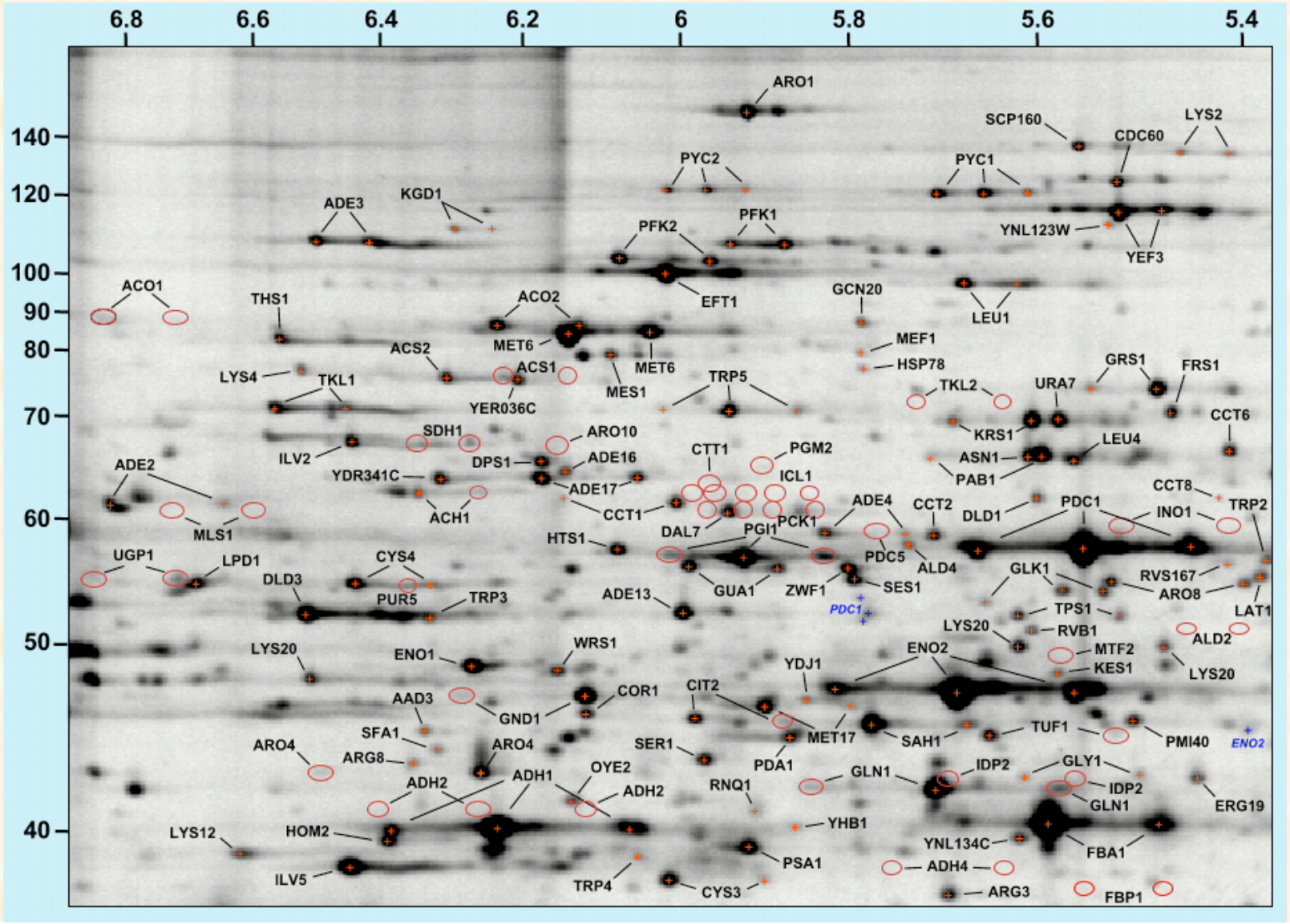
Epigenesis



Methods to obtain a network of interactions (Post-genomics)

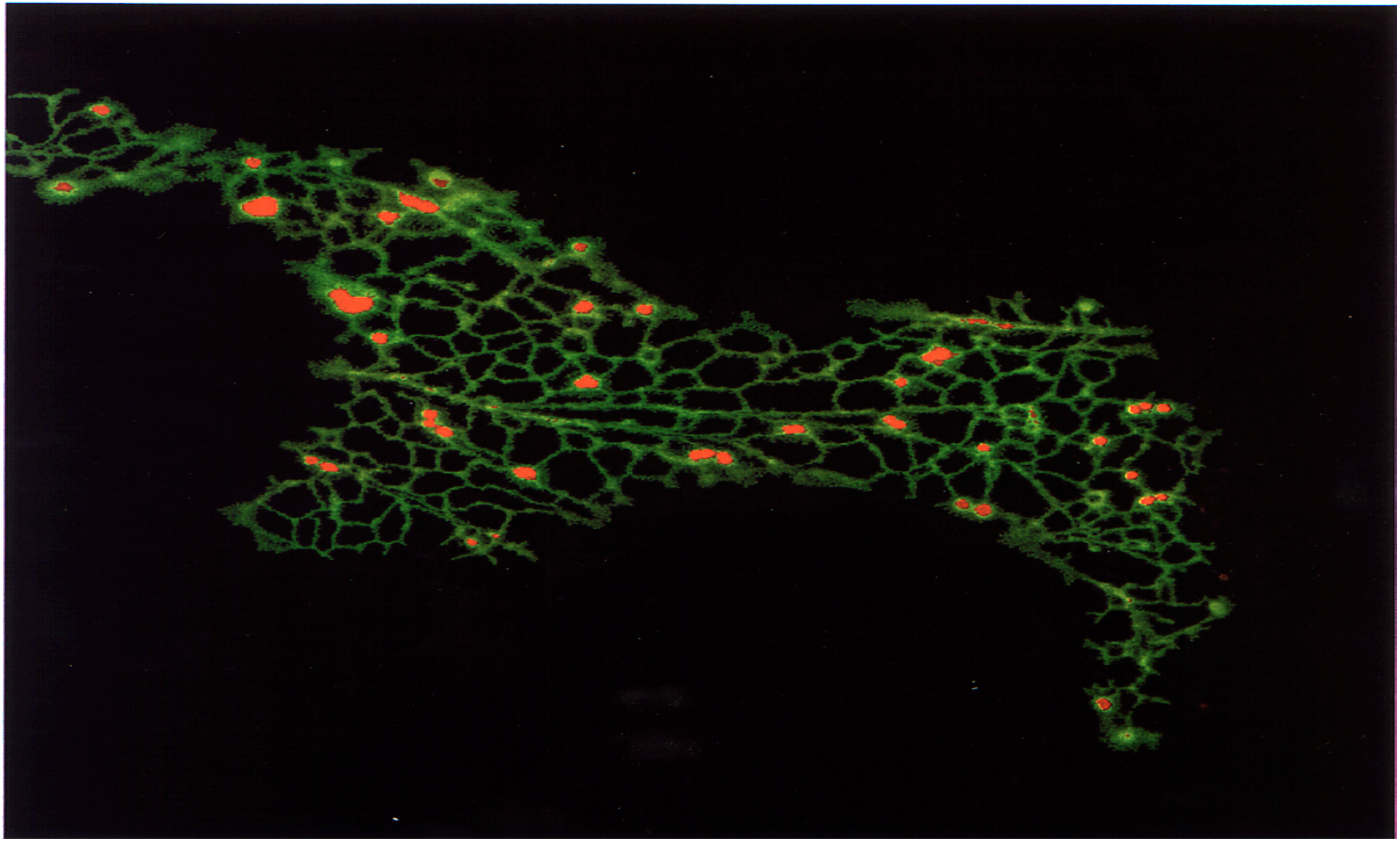
Proteomics

Separation, identification and quantification of proteins



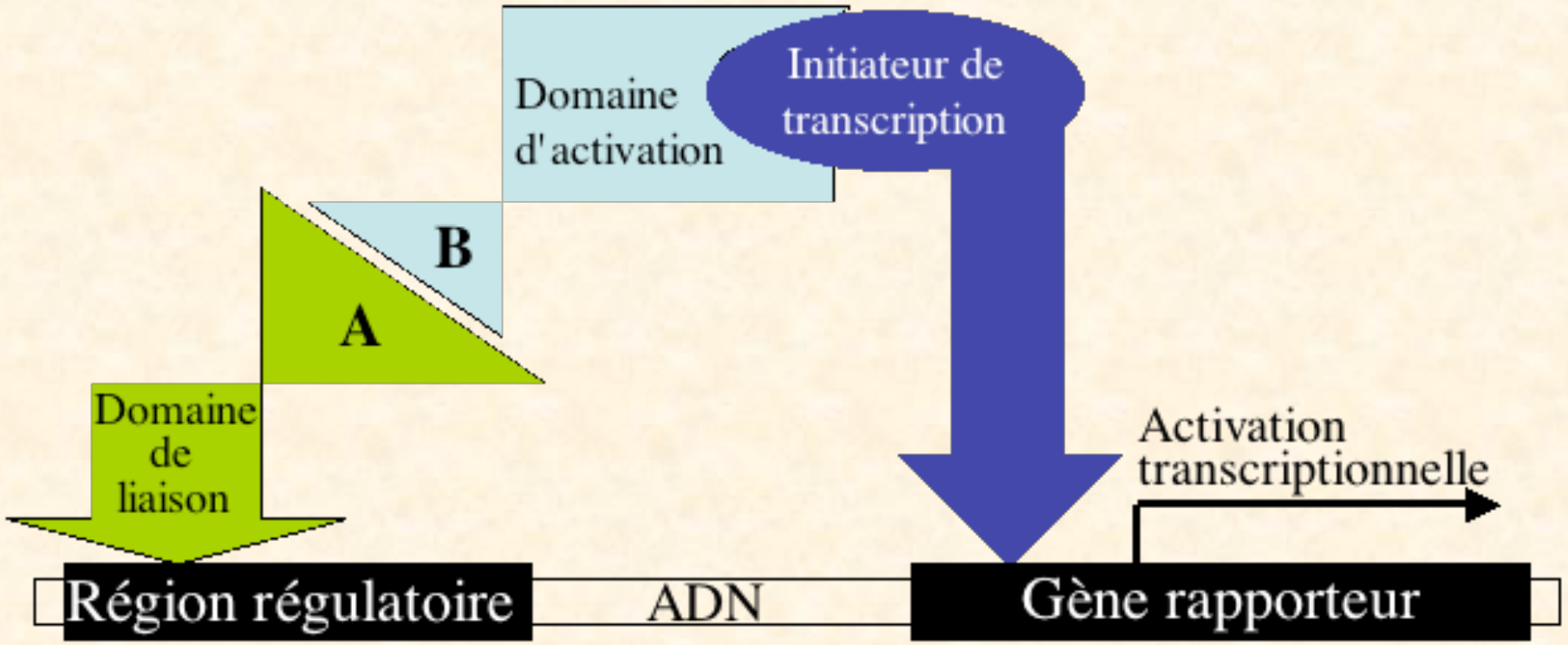
Proteomics

Intracellular Localization



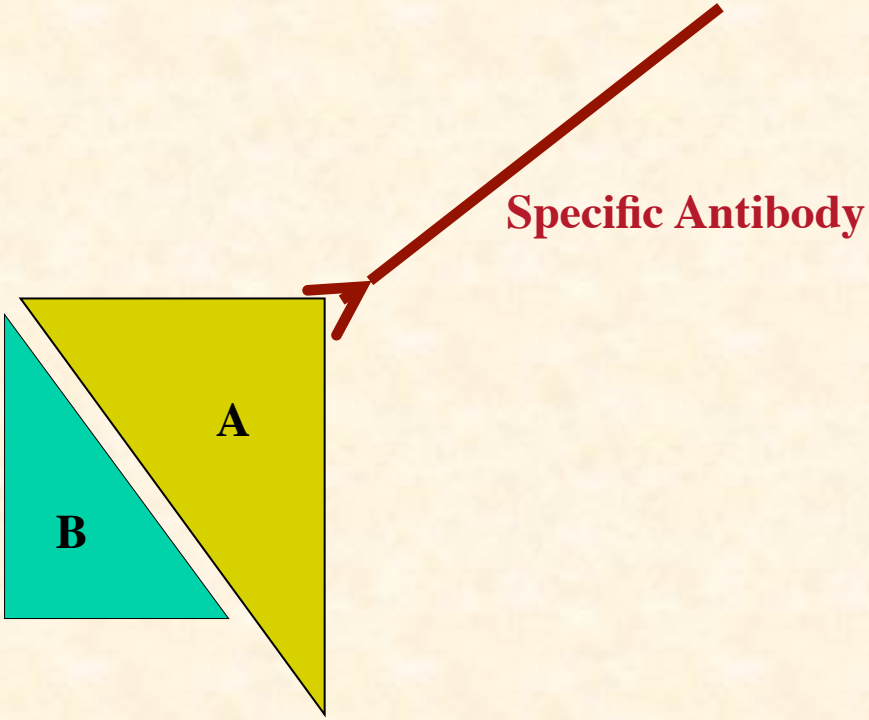
Proteomics

Double-hybrid approaches



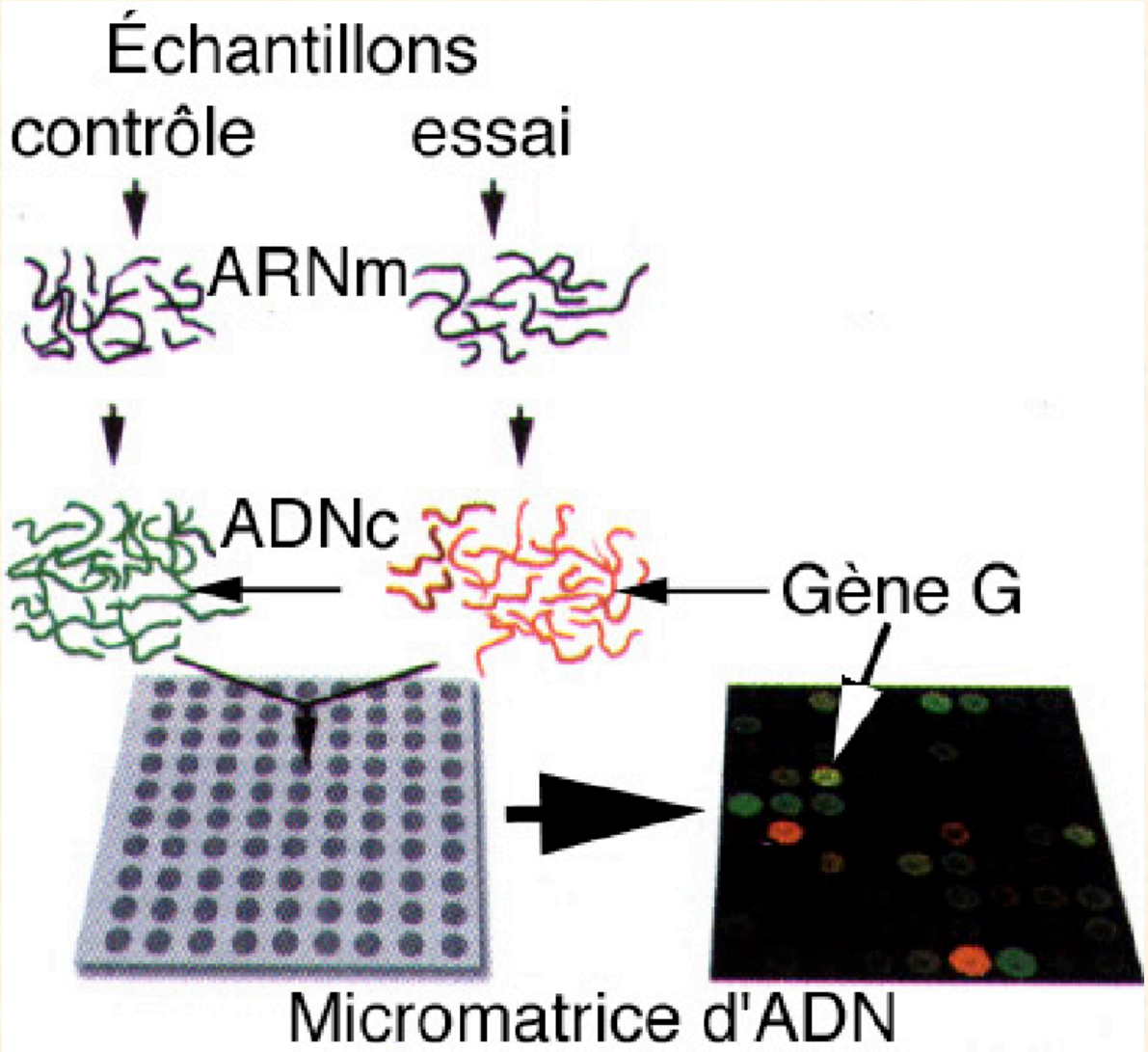
Proteomics

Protein Complexes Assessment



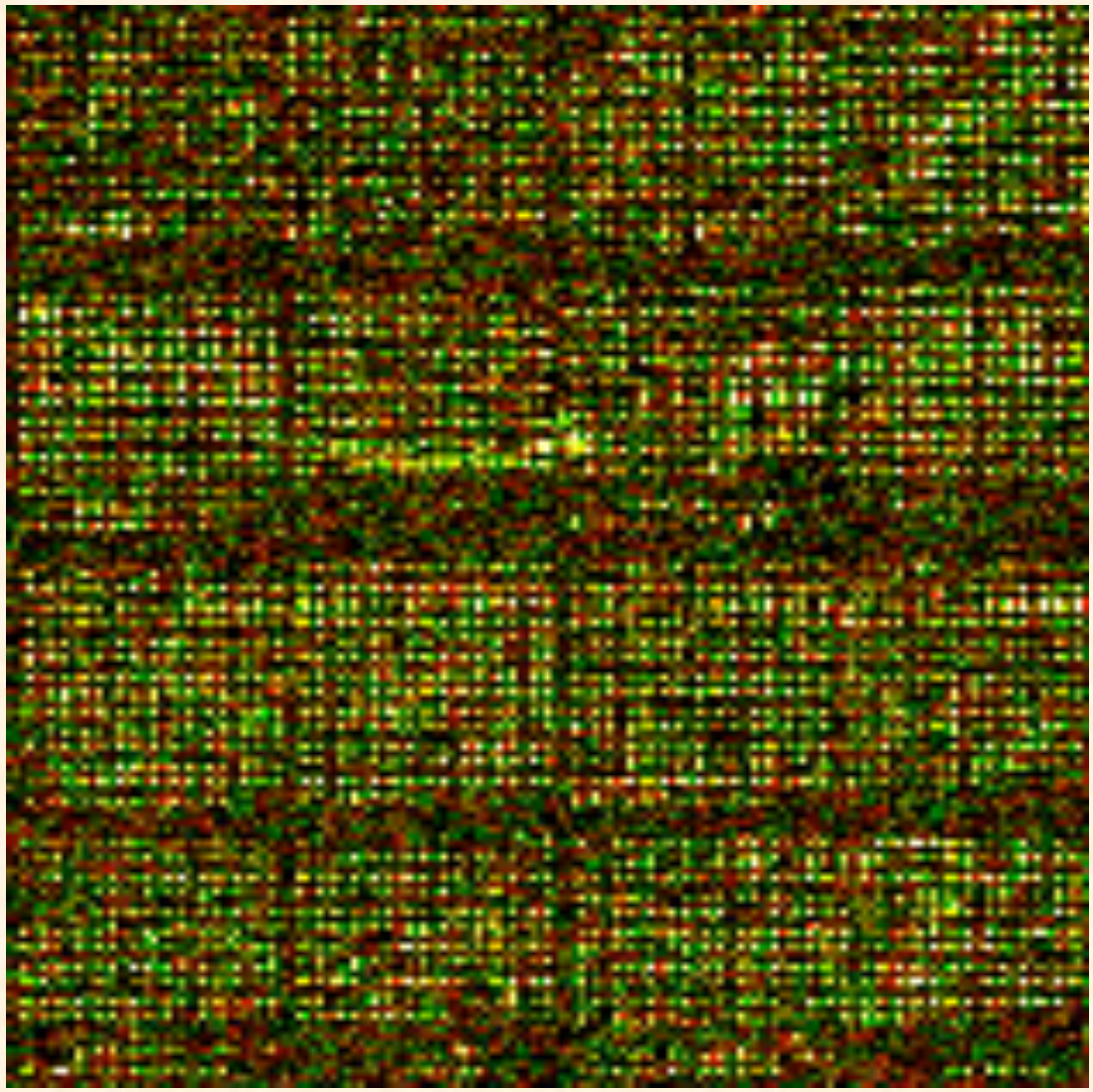
Transcriptomics

Complementary DNA Microarrays



Transcriptomics

Oligonucleotide Biochip



Transcriptomics

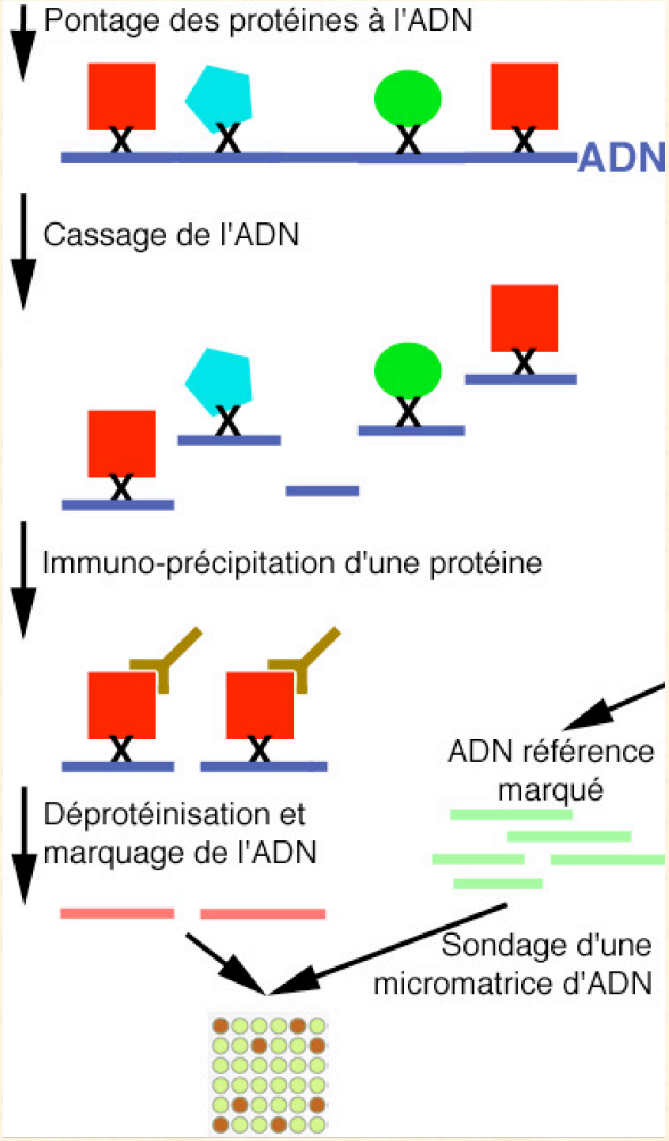
Other techniques :

RT-PCR

SAGE

Transcriptomics

Chromatin Immuno-Precipitation (ChIP)



Transcriptomics

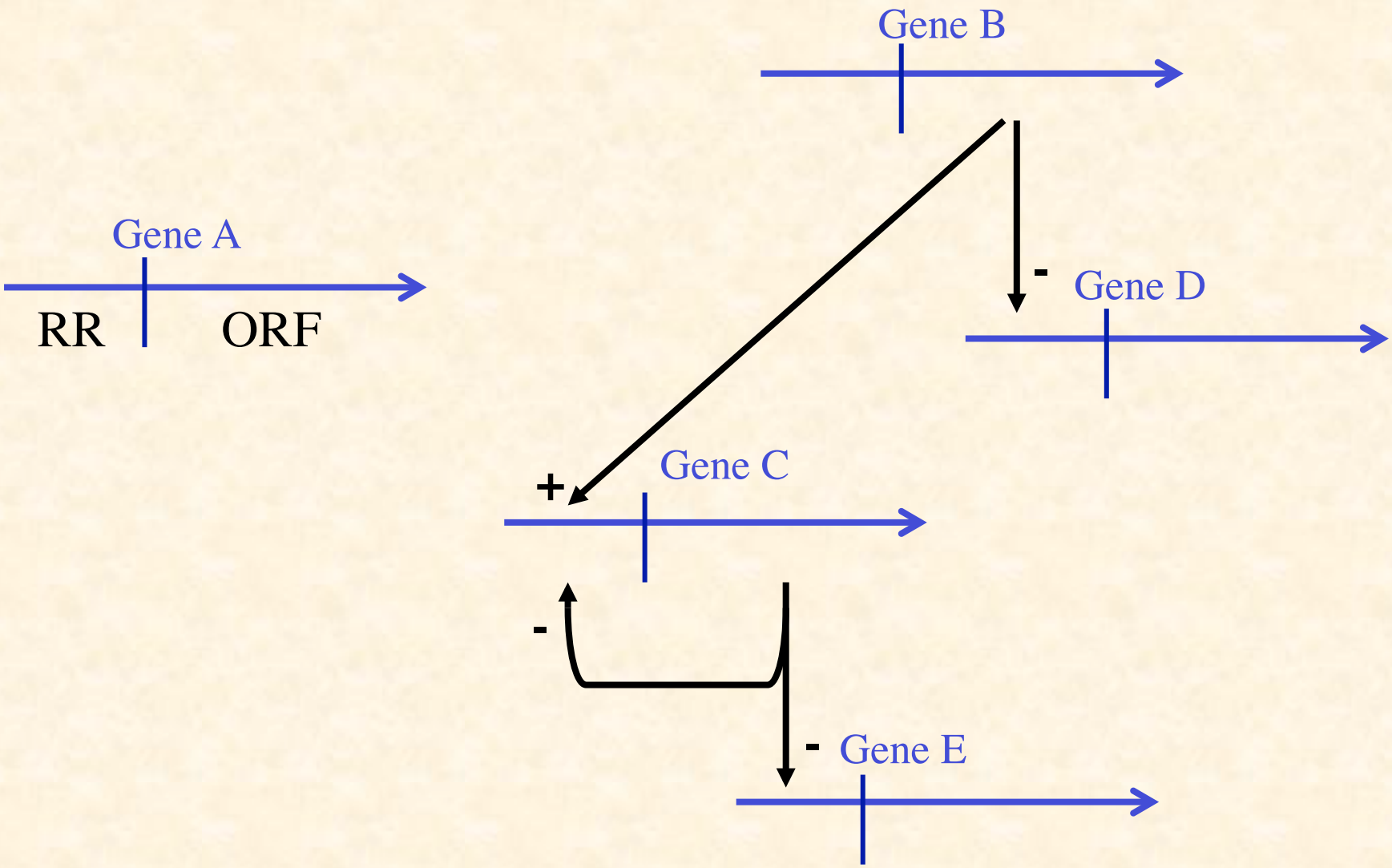
Bioinformatics

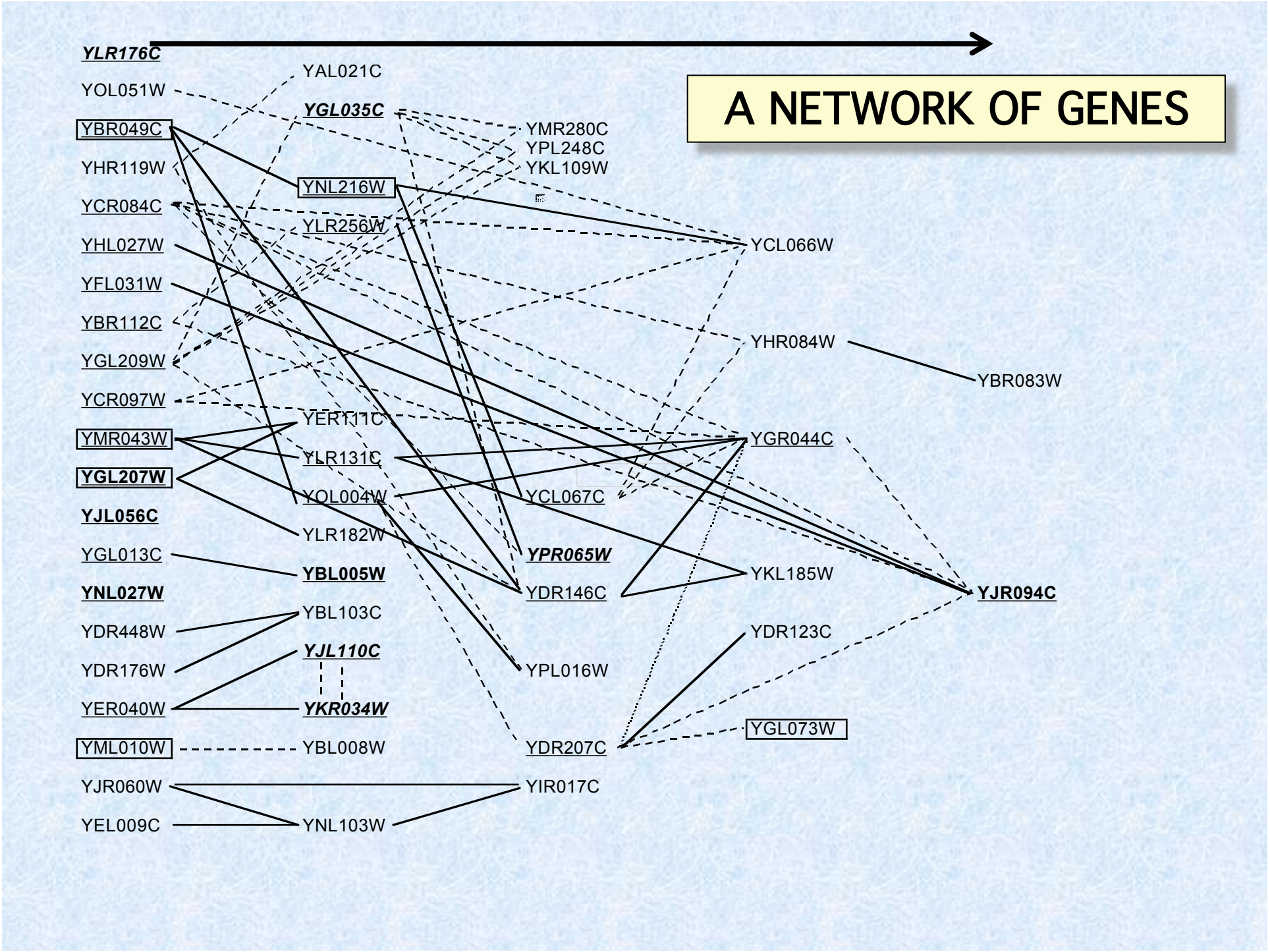
Information content expressed in bits

Global Structure of the Network of Interactions

SEVERAL GENES

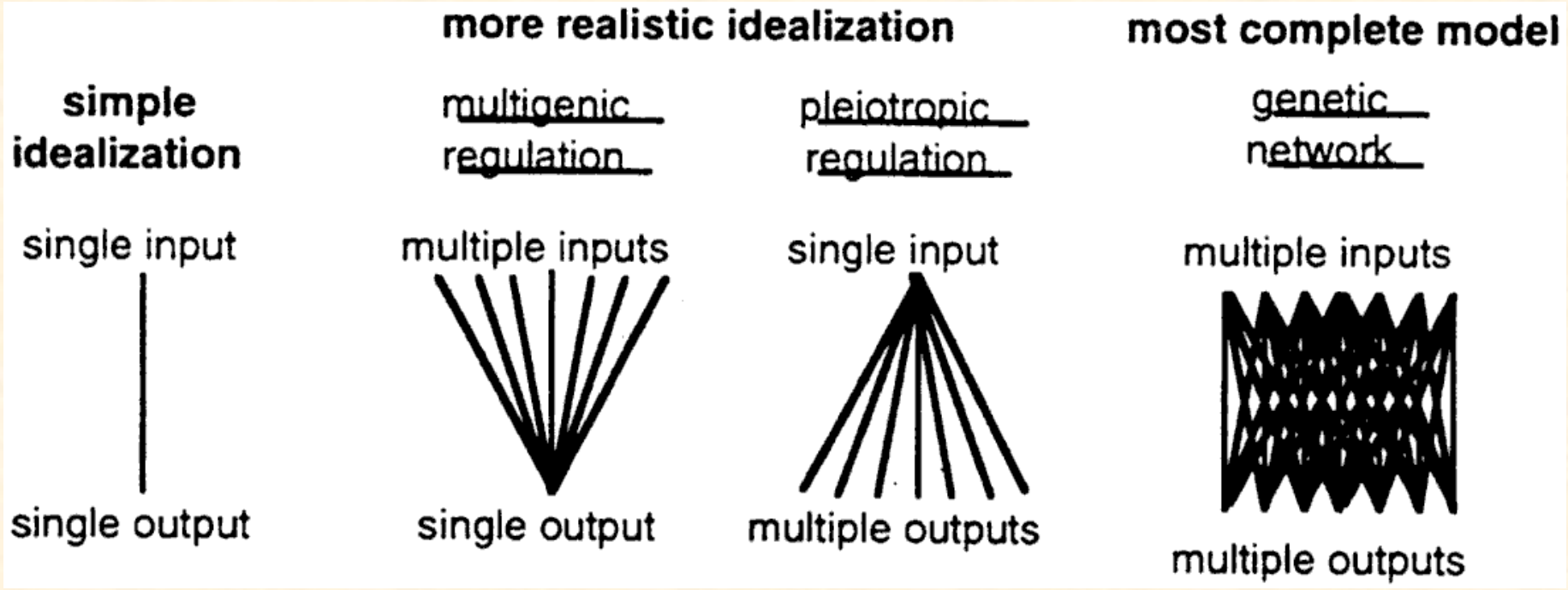
Interactions between regulatory proteins and DNA regulatory regions




A NETWORK OF GENES
YLR176C
YAL021C
YOL051W
YGL035C
YBR049C
YMR280C
YPL248C
YKL109W
YHR119W
YNL216W
YCR084C
YLR256W
YCL066W
YHL027W
YFL031W
YHR084W
YBR083W
YBR112C
YGL209W
YCR097W
YER111C
YMR043W
YGR044C
YLR131C
YGL207W
YCL067C
YOL004W
YJL056C
YLR182W
YPR065W
YGL013C
YBL005W
YKL185W
YNL027W
YDR146C
YJR094C
YBL103C
YDR448W
YDR123C
YJL110C
YDR176W
YPL016W
YER040W
YKR034W
YML010W
YBL008W
YGL073W
YDR207C
YJR060W
YIR017C
YEL009C
YNL103W

MODELING A GENETIC NETWORK

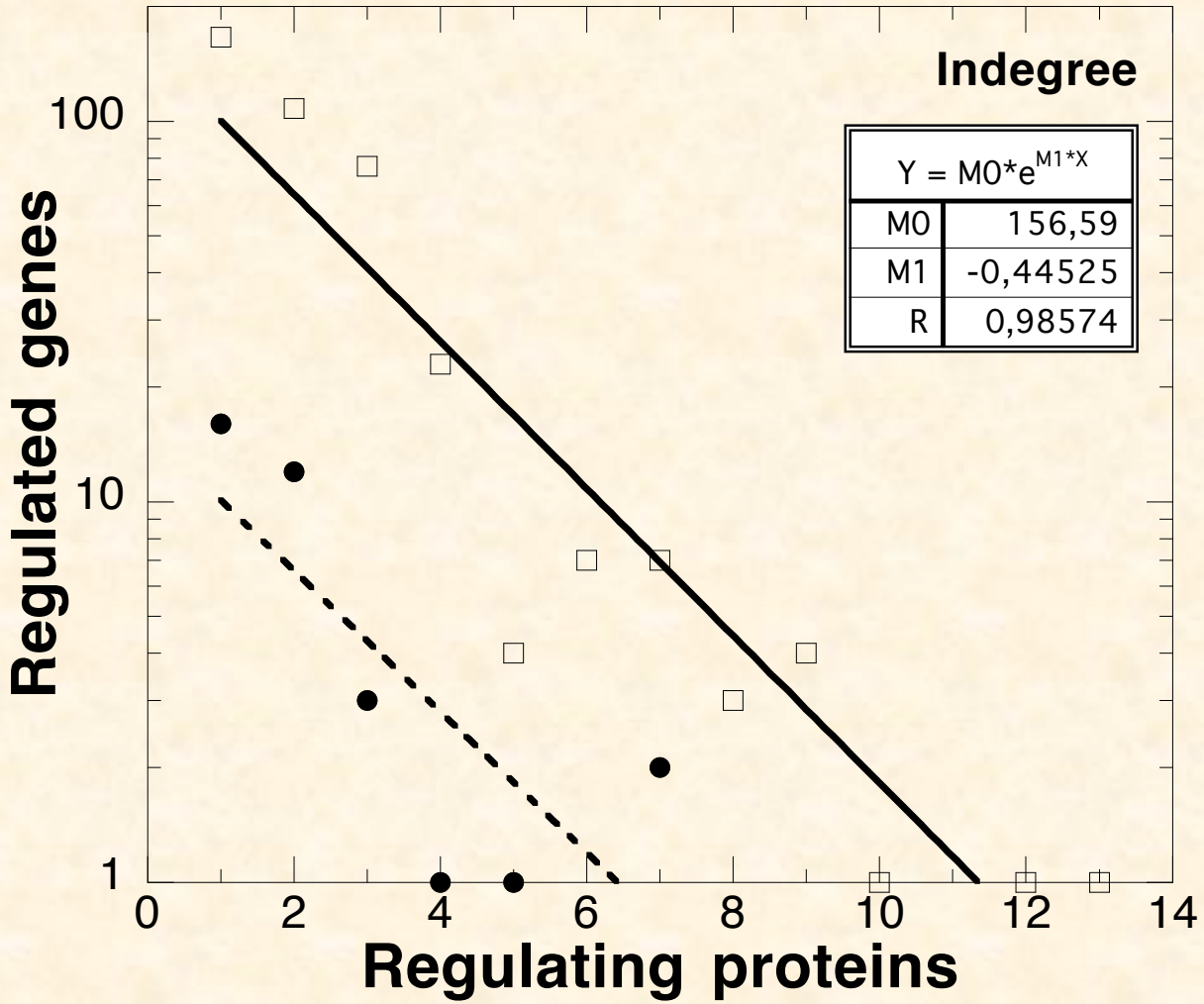
Pleiotropic or multigenic regulation ?



INDEGREE DISTRIBUTION

EXPONENTIAL
Average 2.3

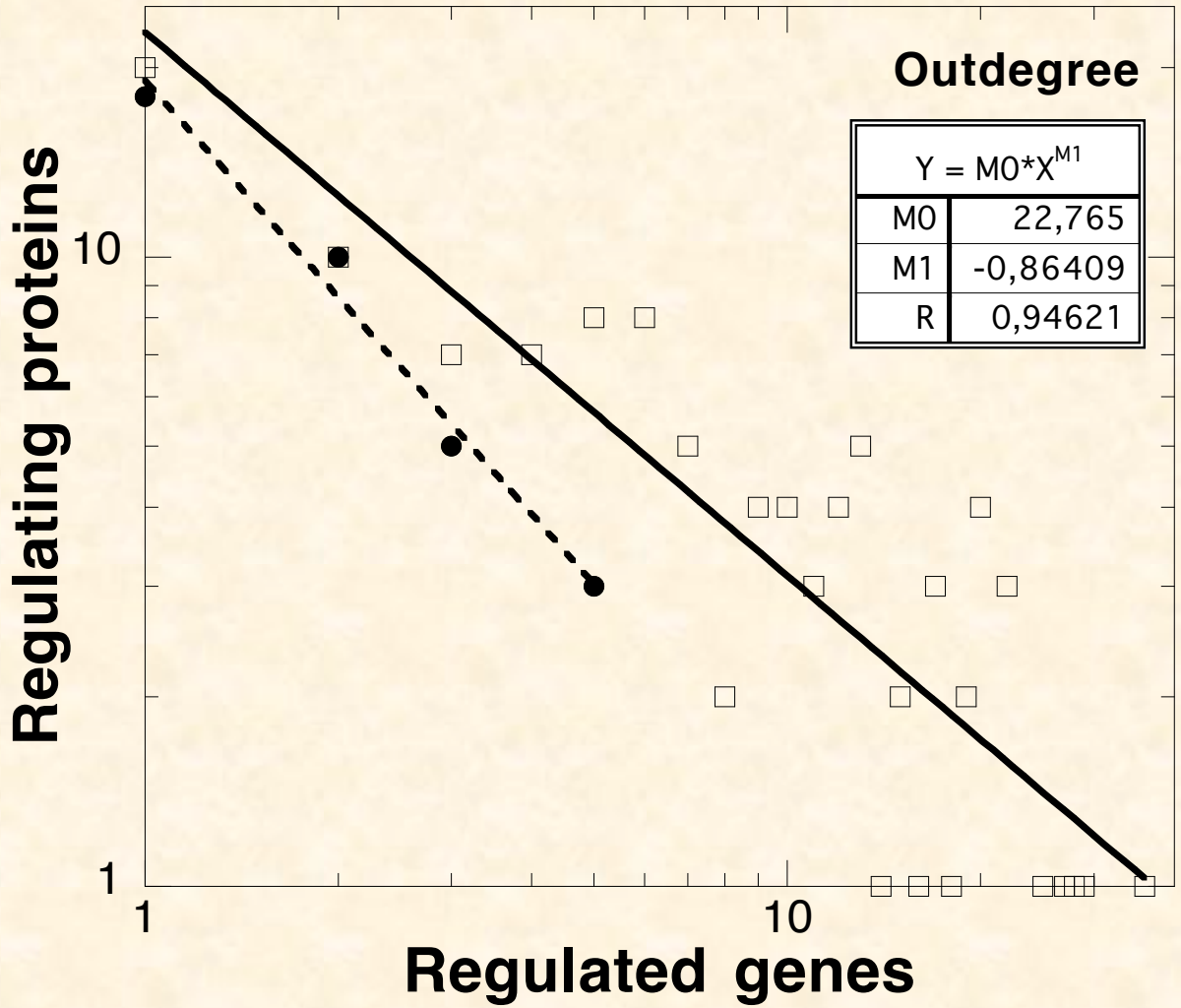
Global Topology



OUTDEGREE DISTRIBUTION

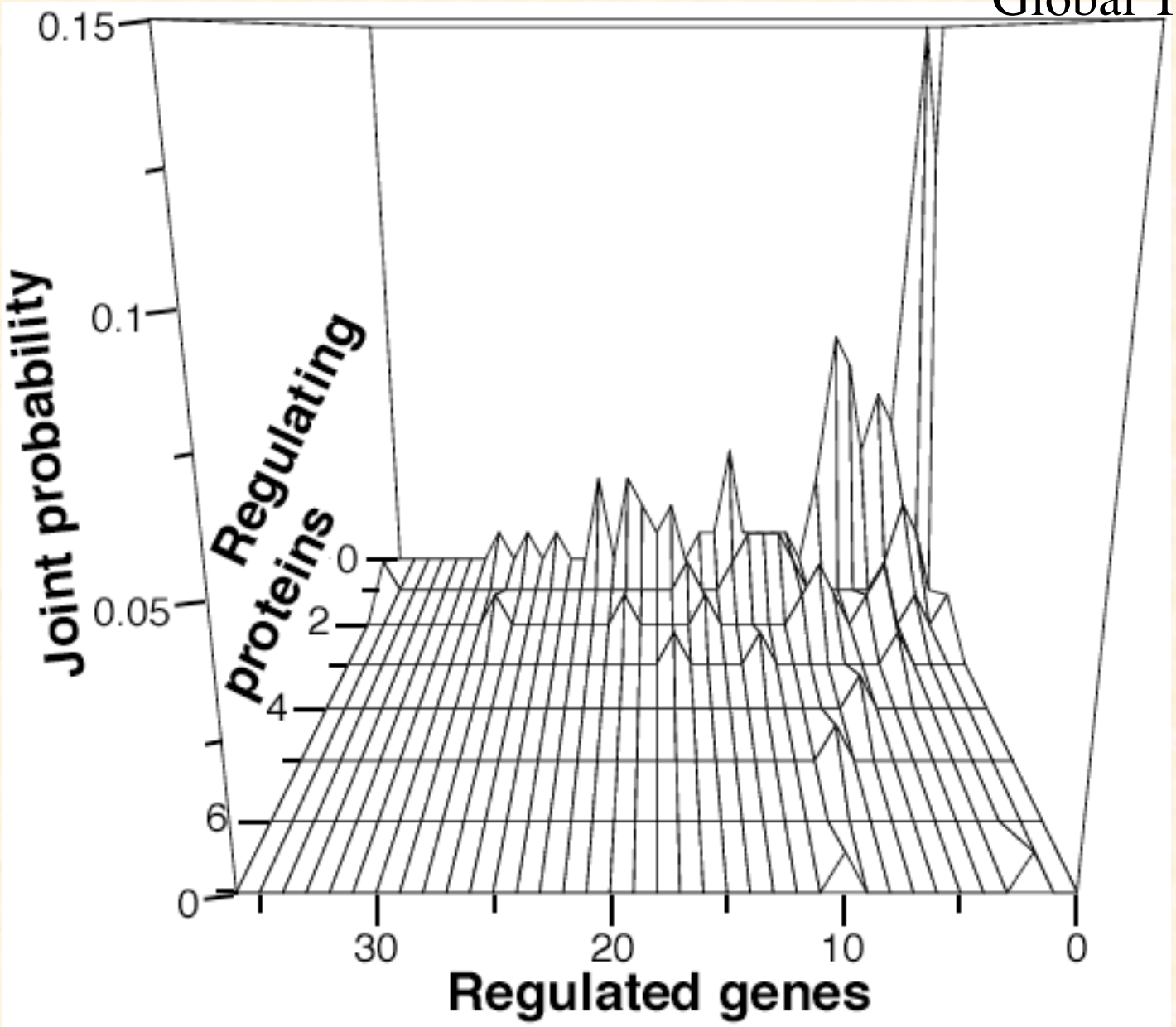
POWER
Average 8.3

Global Topology



JOINT IN / OUT DISTRIBUTION

Global Topology

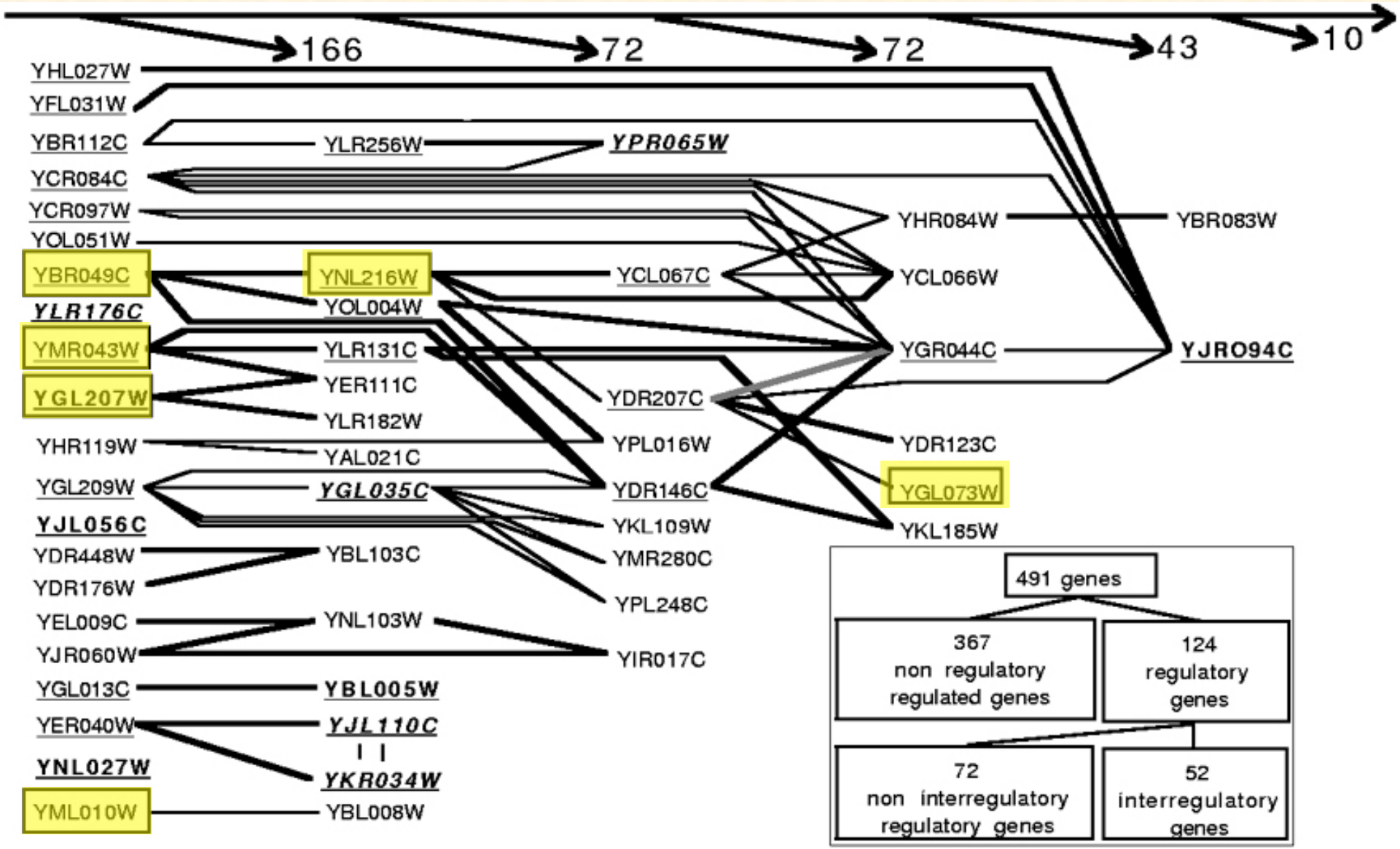


ESSENTIALITY AND DESCENT SIZE

6 essential inter-regulatory genes

ORF	*Direct descent*	*Overall descent*
YBR049C	18	86
YMR043W	29	52
YNL216W	9	49
YGL207W	11	22
YGL073W	12	12
YML010W	1	2

ESSENTIAL INTER-REGULATORY GENES

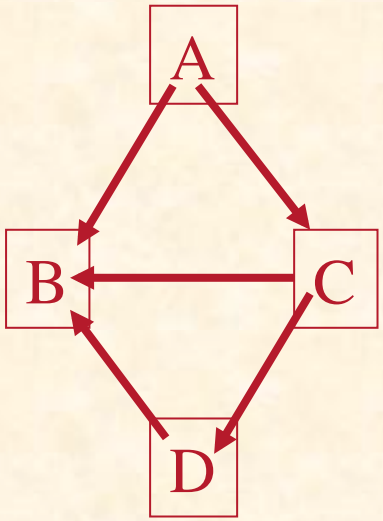


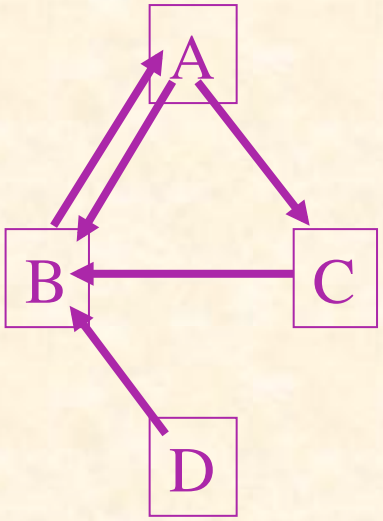
Lethal gene inactivation

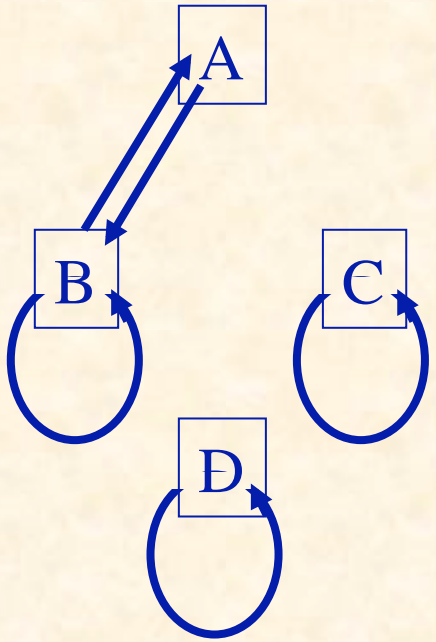
Local Structure of the Network of Interactions

LOCAL APPORTIONMENT OF EDGES PER VERTEX

Local topology







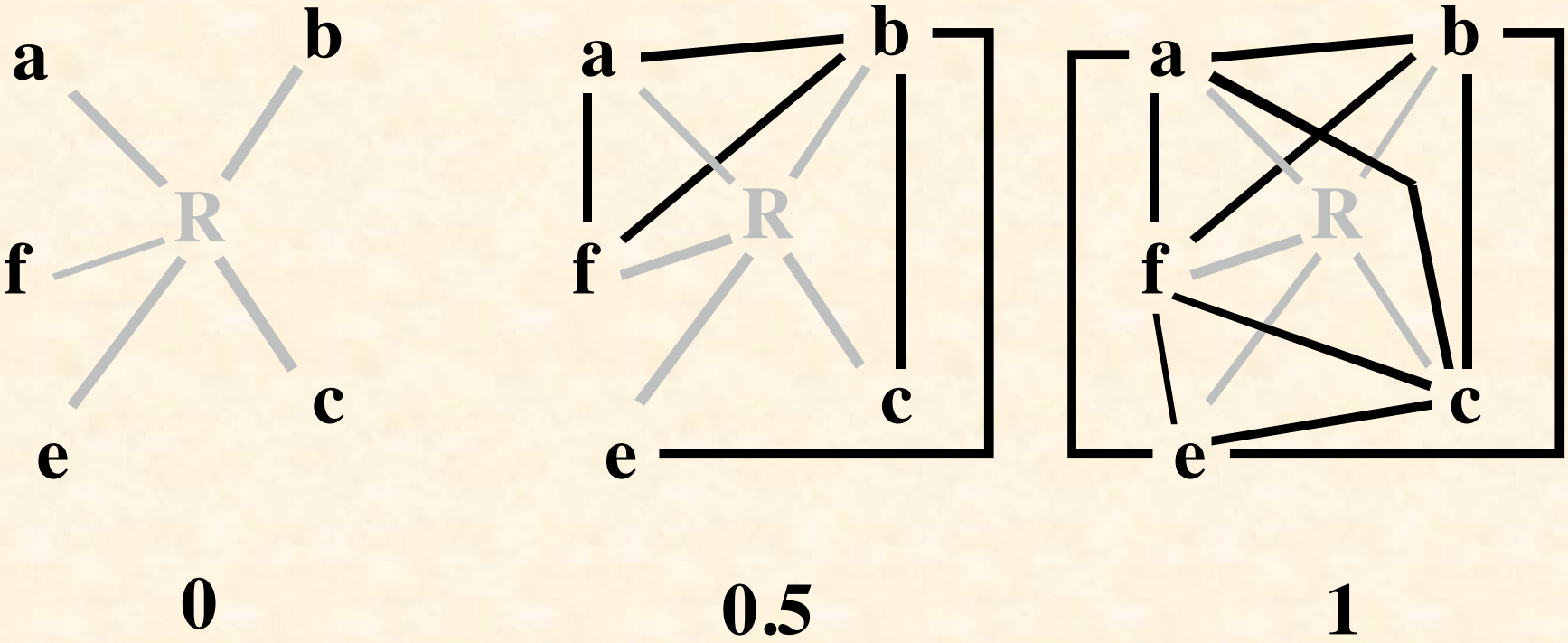
UNIFORM

SMALL WORLD

HIGHLY CLUSTERED
FRAGMENTED

CLIQUISHNESS (clustering coefficient)

Local topology



Dynamics and Modularity

MOTIFS

Homogeneous networks:

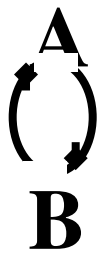
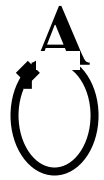
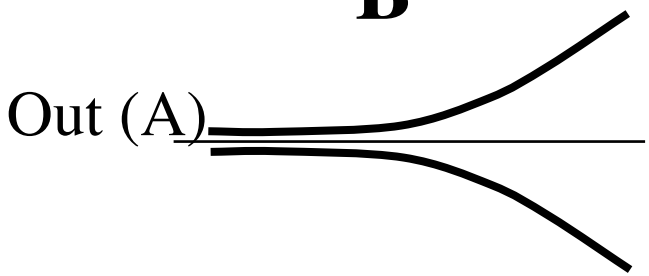
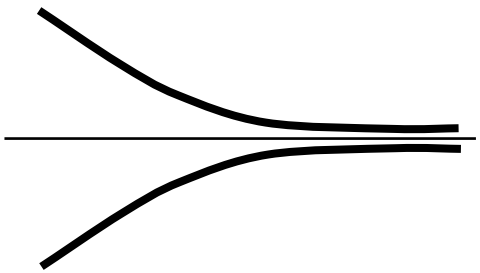
Why partition at all ?

- Modularity
- Dynamical explanation
- Biological relevance
- Compositionnality

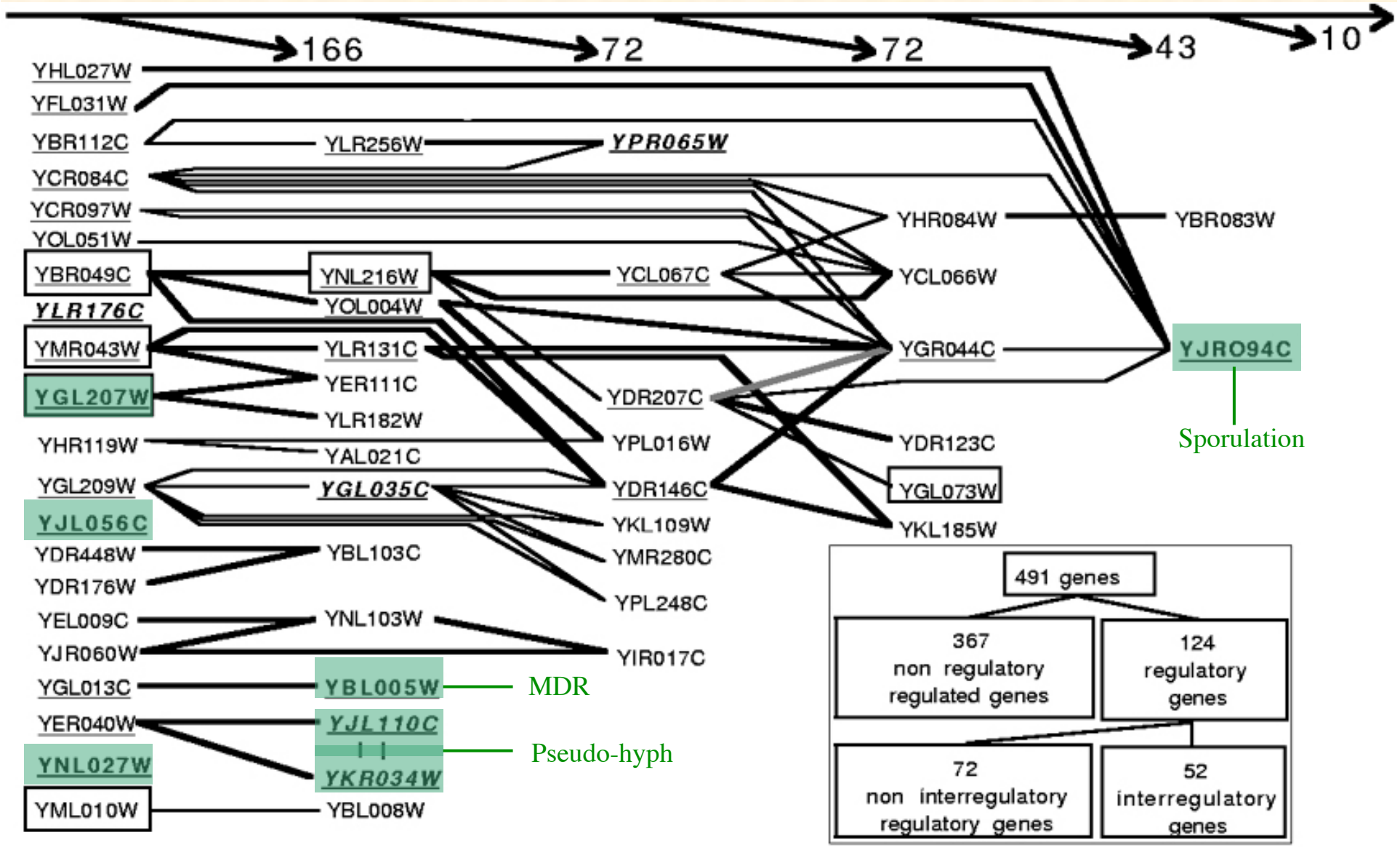
MOTIFS

Topology | Physical properties | Qualitative dynamics | Numerical simulation | *In vivo* simulation

MOTIFS

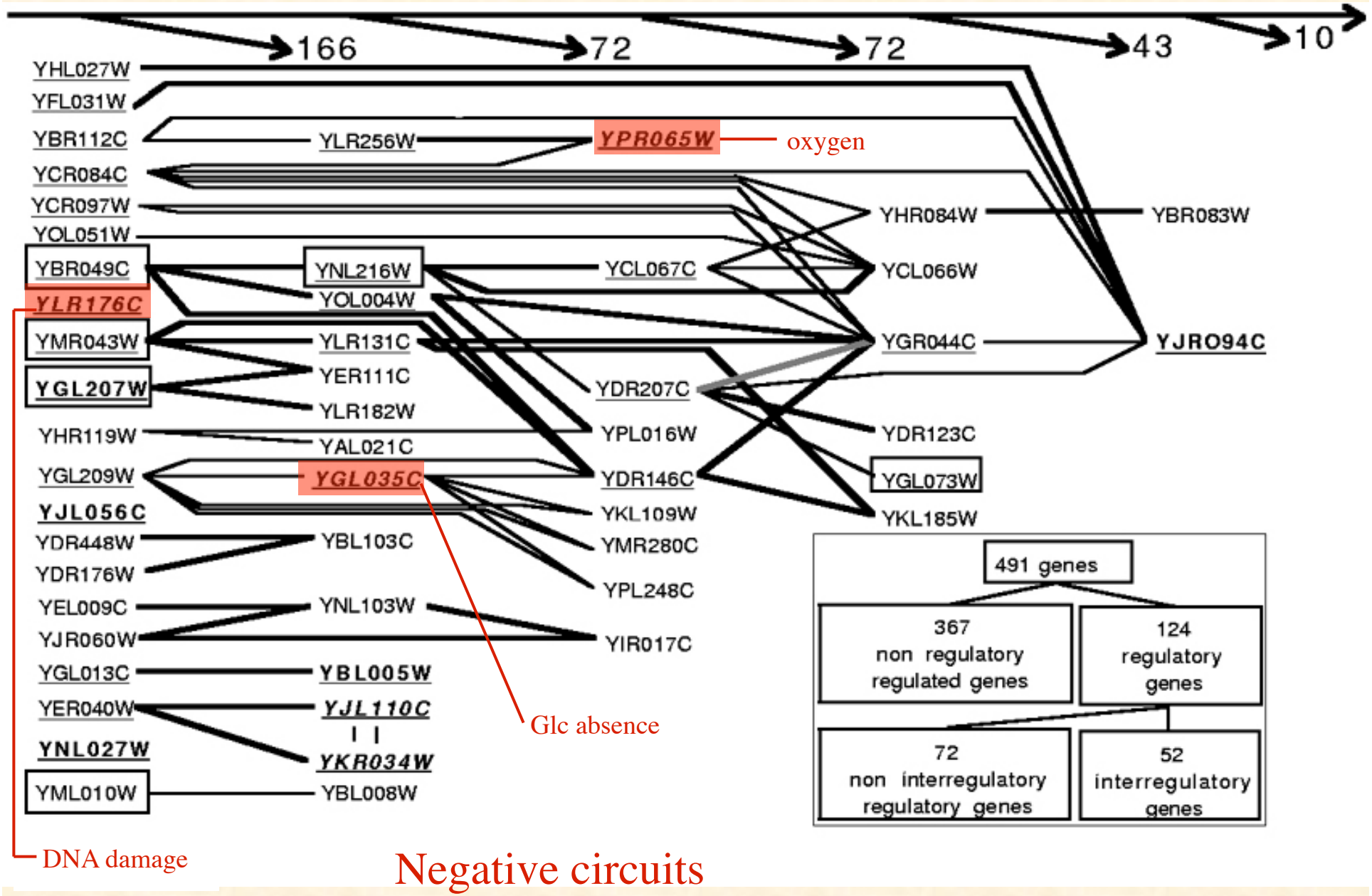
	MODULES	FEEDBACK CIRCUITS	
	SIGN	POSITIVE	NEGATIVE
# negative interactions		Even	Odd
Dynamic property		Multistationarity	Homeostasis
Biological property		Differentiation	Stable regulation
Topology			
Qualitative dynamics (*in vivo* and numerical simulations carried out)			

FEEDBACK CIRCUITS AMONG INTERREGULATORY GENES



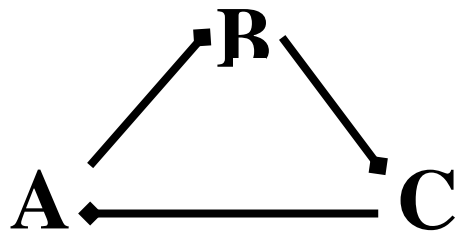
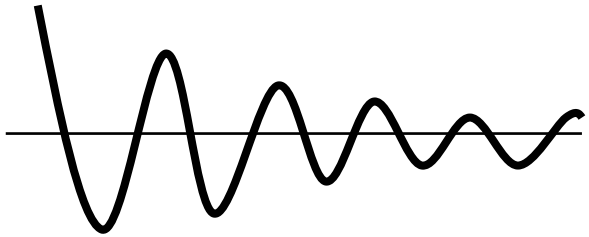
Positive circuits

FEEDBACK CIRCUITS AMONG INTERREGULATORY GENES



Negative circuits

MOTIFS

MODULES	FEEDBACK CIRCUITS
TYPE	NEGATIVE
# negative interactions	Odd
Dynamic property	Oscillator
Biological property	Stable regulation or oscillation
Topology	
Qualitative dynamics (*in vivo* and numerical simulations carried out)	

MOTIFS

MODULES **TRIANGLES (Feed-forward loops)**

TYPE COHERENT INCOHERENT

Dynamic property

Filter out pulses
Respond to persistent stimulations
Rapidly shutdown

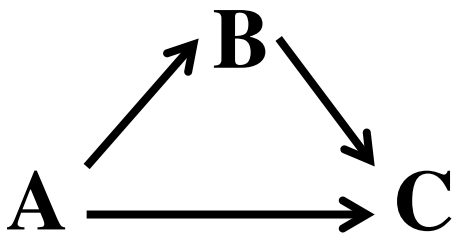
Initially reacts strongly
Later comes back
to intermediate levels

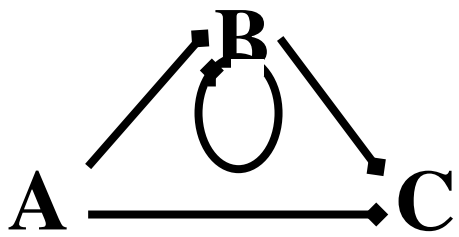
Biological property

Decide from fluctuating signal

Easily reverse

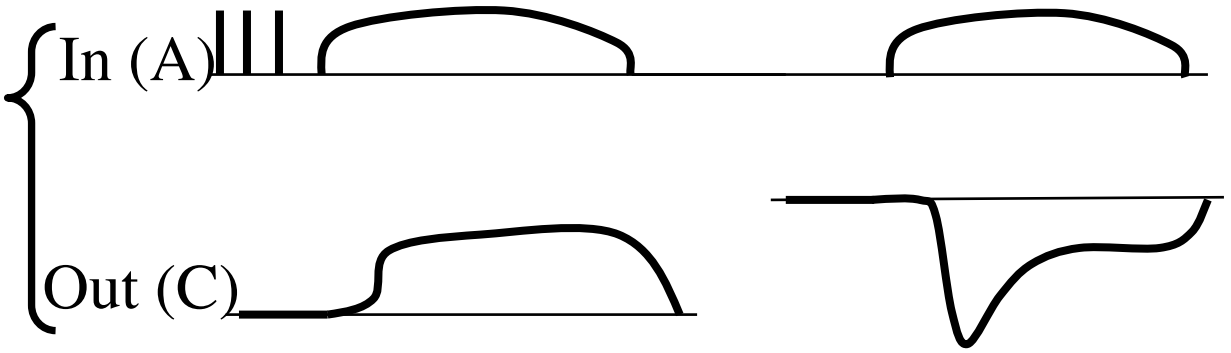
Topology



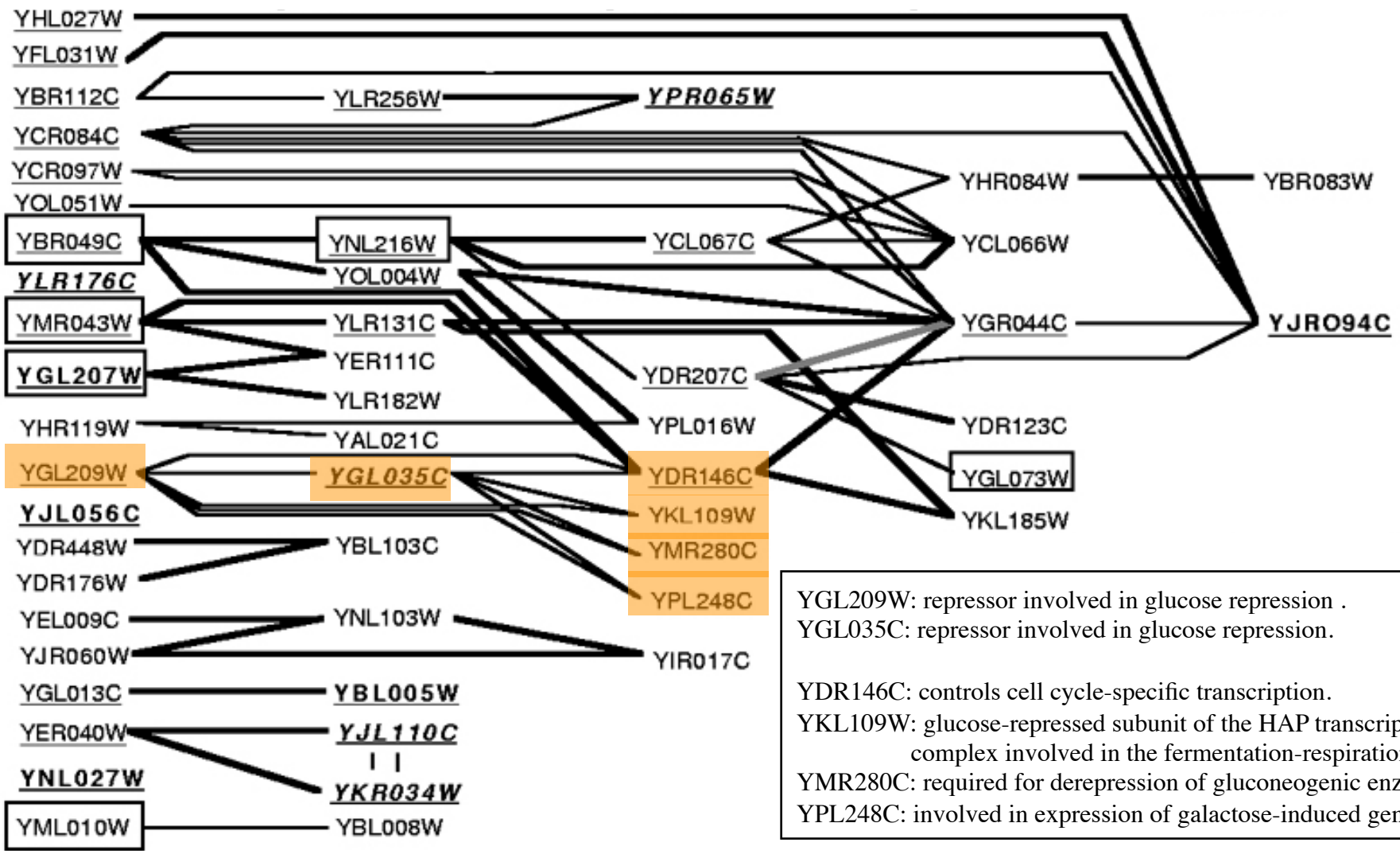


Qualitative dynamics

(Coherent triangle was numerically simulated)



Incoherent Triangles



YGL209W: repressor involved in glucose repression .
YGL035C: repressor involved in glucose repression.

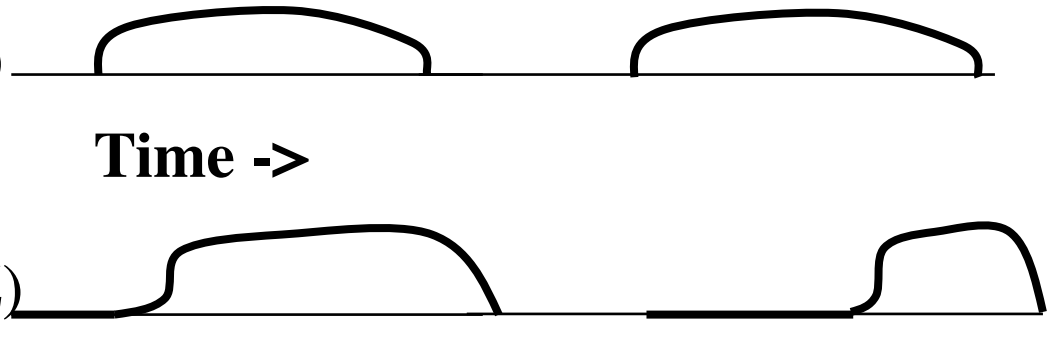
YDR146C: controls cell cycle-specific transcription.
YKL109W: glucose-repressed subunit of the HAP transcriptional complex involved in the fermentation-respiration shift.
YMR280C: required for derepression of gluconeogenic enzymes.
YPL248C: involved in expression of galactose-induced genes.

MOTIFS

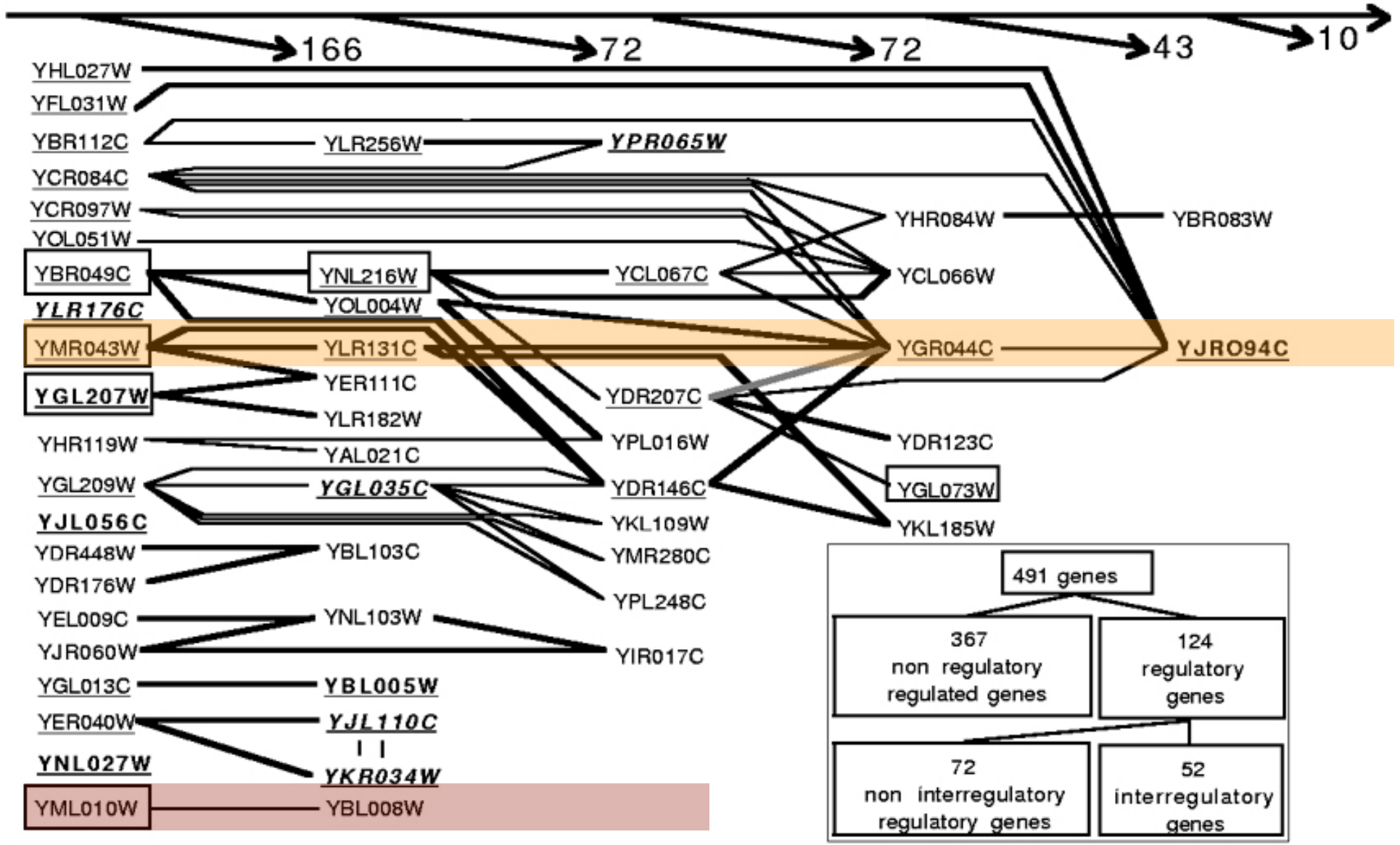
	MODULES	CASCADES (Linear set of regulations)	
TYPE		SHORT	LONG
Dynamic property		Rapidly shoots up	Long lag before it shoots up
Biological property		Fast response in microbe	Time counting in multicellular
Topology		A → Z	A → B → … → Z

Qualitative dynamics

(Not simulated yet)



Cascades


166
72
72
43
10
YHL027W
YFL031W
YBR112C
YLR256W
YPR065W
YCR084C
YCR097W
YOL051W
YBR049C
YNL216W
YCL067C
YHR084W
YBR083W
YCL066W
YLR176C
YOL004W
YMR043W
YLR131C
YGR044C
YJRO94C
YGL207W
YER111C
YDR207C
YLR182W
YHR119W
YPL016W
YDR123C
YAL021C
YGL209W
YGL035C
YDR146C
YGL073W
YJL056C
YKL109W
YKL185W
YDR448W
YBL103C
YMR280C
YDR176W
YPL248C
YEL009C
YNL103W
YJR060W
YIR017C
YGL013C
YBL005W
YER040W
YJL110C
YNL027W
YKR034W
YML010W
YBL008W
491 genes
367
non regulatory
regulated genes
124
regulatory
genes
72
non interregulatory
regulatory genes
52
interregulatory
genes

MOTIFS

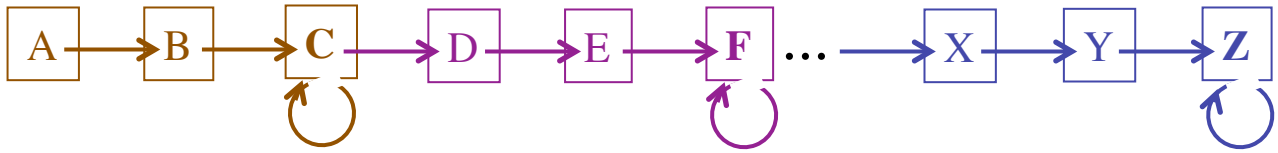
MODULES **Combination of long cascade and positive circuits**

TYPE — LOCK-ON

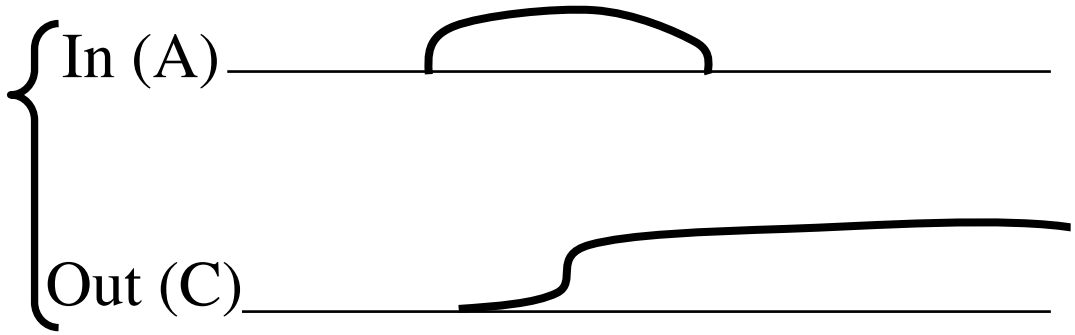
Dynamic property — Ratchet

Biological property — Succession of time lags and differentiation events in multicellular development

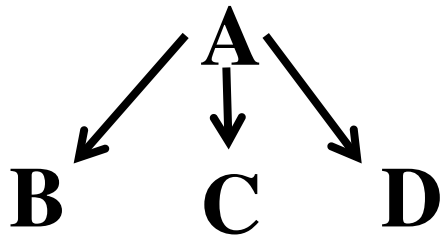
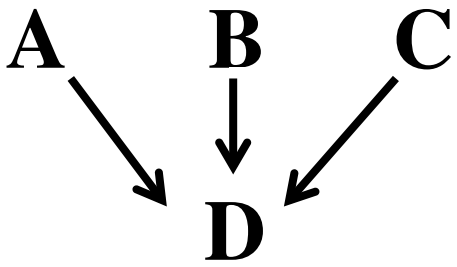
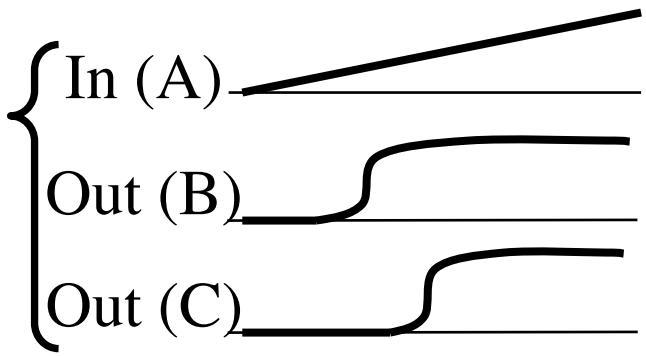
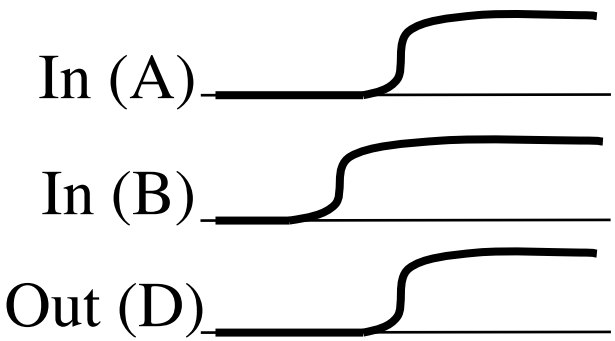
Topology



Qualitative dynamics for a single-colored series (Not simulated yet)



MOTIFS

MODULES	SIM / SOM	
TYPE	Single-Input Module	Single-Output Module
Dynamic property	Ordered temporal response	' AND ' logical gate
Biological property	Sequential firing based on differential thresholds	Co-regulation
Topology		
Qualitative dynamics (Verified *in vivo*)		

MOTIFS OF MOTIFS

Is it possible to spot combinations of motifs ?

What could be said about their representation ?

What could be said about their topology ?

What could be said about their dynamics ?

Some of the present challenges

1) Recompose motifs in a useful way

2) Top-down partition into 'functional' modules

3) Go into more global dynamics

4) Analyze mixed networks

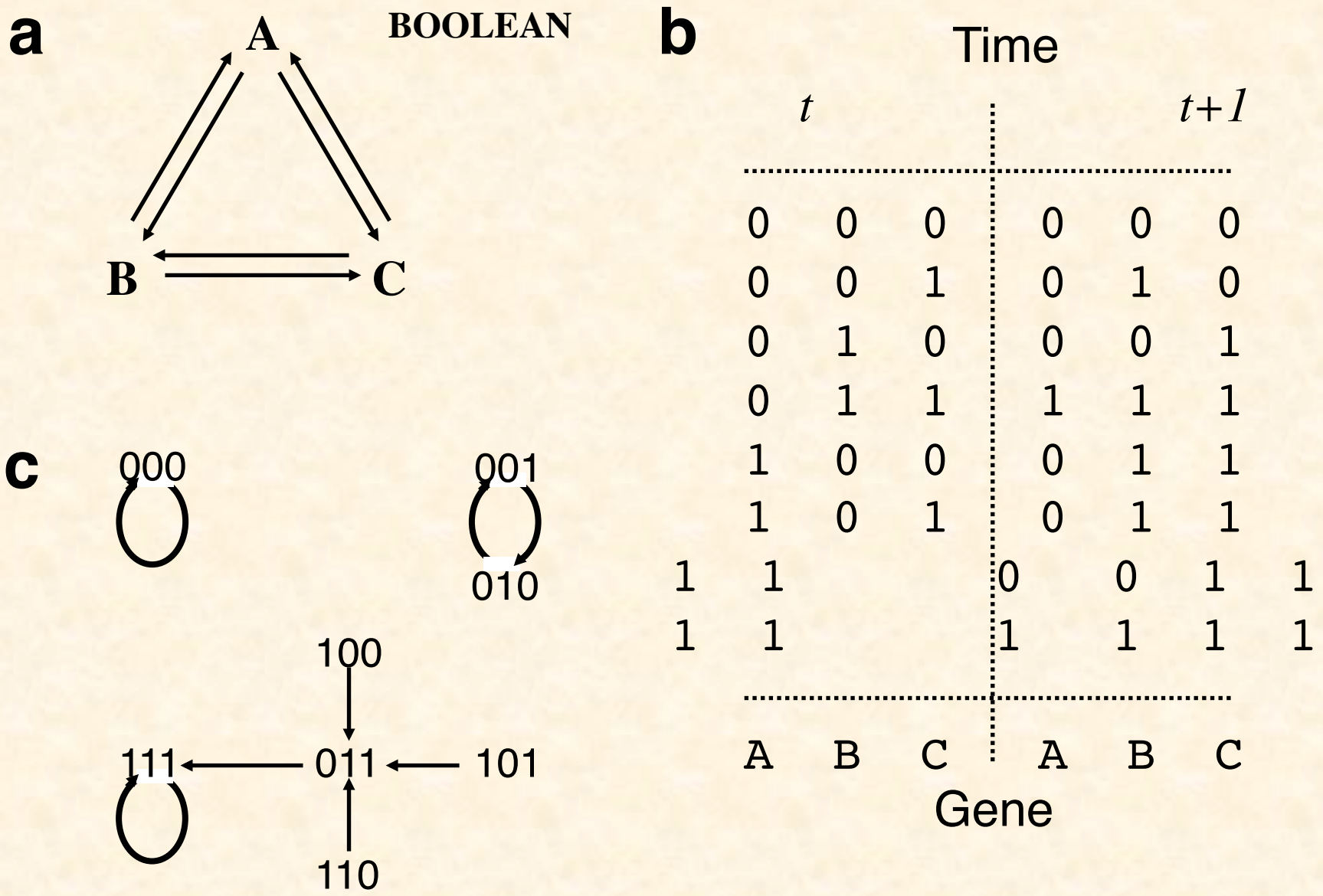
5) Realistically model evolution of networks

6) Unfold topologies in geometrical cellular space

7) Control

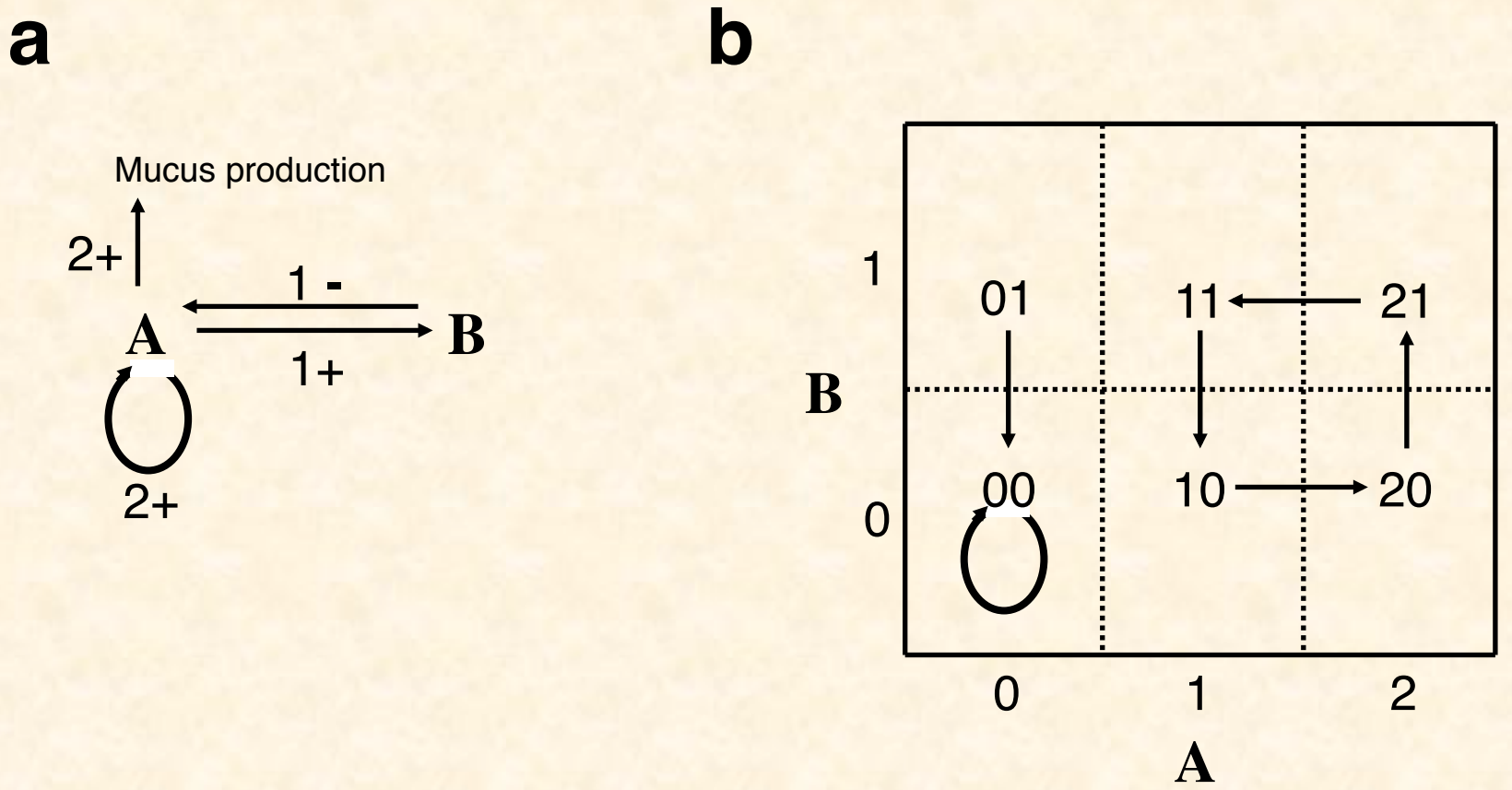
8) Exploit hybrid formalisms

Mathematical formalisms



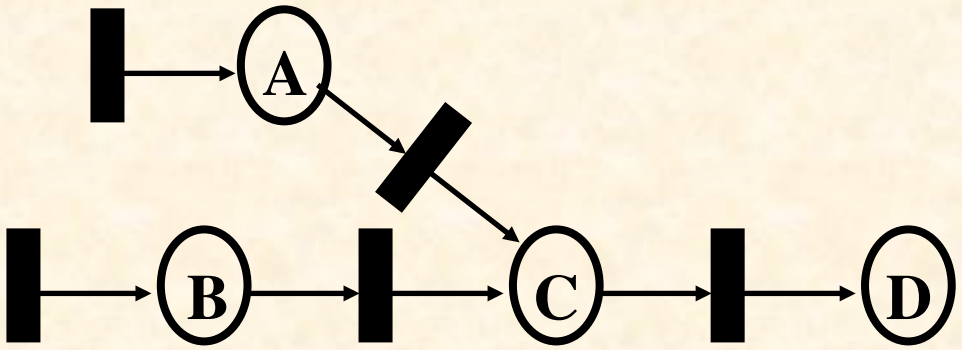
A small boolean network. a) The wiring diagram in a boolean network with three genes (A, B and C), each an input to the other two. **b)** The boolean rules for the diagram shown in a), assuming that gene A represents an *AND* gate, while genes B and C each represent an *OR* gate. **c)** The state transition graph of the boolean network depicted in a) and b). Each triplet of digits correspond to a state for genes A-C, from left to right.

RENÉ THOMAS



A small logical network "à la René Thomas". a) The regulatory interactions for mucus production in the opportunistic pathogen *Pseudomonas aeruginosa*. Two genes, encoding an activator (A) and an inhibitor (B) of mucus production are considered. Each edge in the graph is labeled with the rank number of the threshold, followed by the sign of its regulatory influence (-, inhibition; +, activation). Given parameters (not shown here), a dynamics may be deduced. **b)** The asynchronous state graph. This graph is one among several graphs that would fulfill the constraints based on biological knowledge or hypotheses. A can take any value among {0, 1, 2}, and B among {0, 1}. Thresholds are represented by dashed lines, and transitions by arrows. The graph shows two steady states, one for A = 0, and one cycle: 11 -> 10 -> 20 -> 21 -> 11.

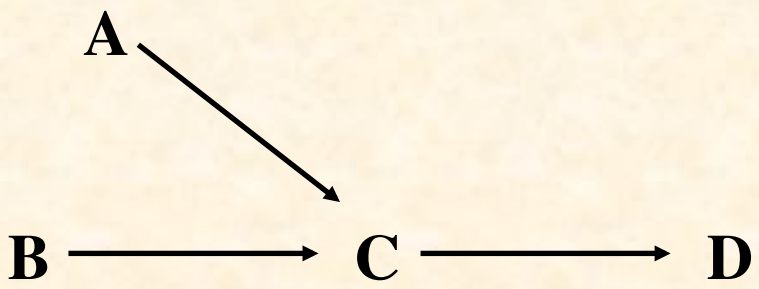
PETRI NET



A Petri net representation of a set of regulatory interactions. Circles denote places identified by a letter, black rectangles are transitions, and arrows are arcs. Only discrete elements are shown.

BAYESIAN

a



b

$p(X_A)$

$p(X_B)$

$p(X_C \mid X_A, X_B)$

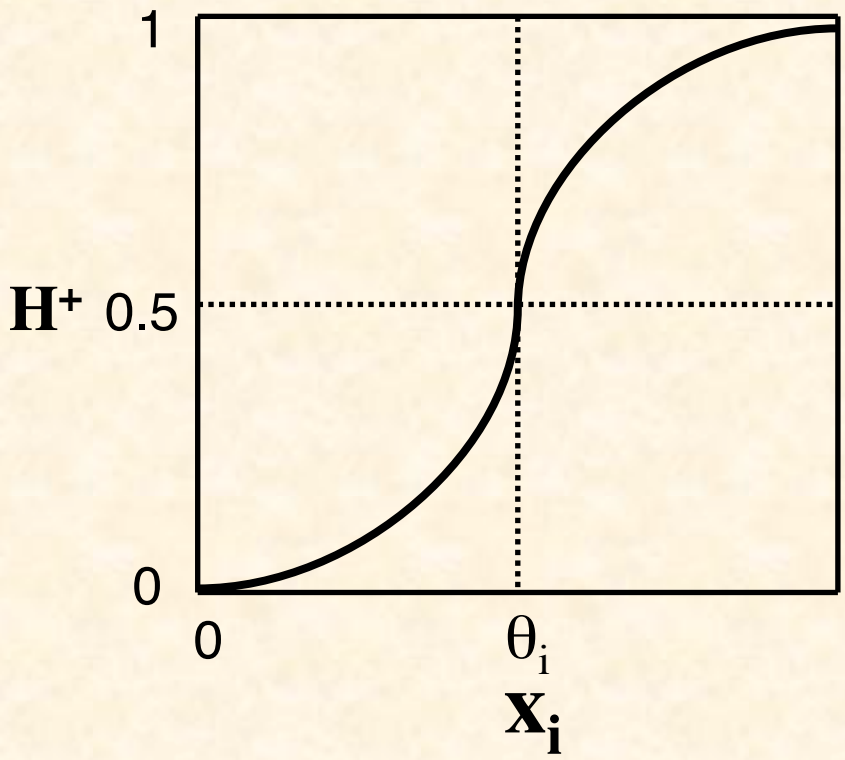
$p(X_D \mid X_C)$

c

$$p(X) = p(X_A)\ p(X_B)\ p(X_C \mid X_A, X_B)\ p(X_D \mid X_C)$$

A bayesian network. a) In this directed acyclic graph, a vertex corresponds to a molecular entity such as a gene or a protein, and holds a random variable representing the gene expression level or the protein concentration. **b)** A conditional probability distribution is defined for the variable of each vertex, given the variables of its direct inputs in the directed graph. **c)** A joint probability distribution is finally defined from all conditional distributions.

Ordinary Differential Equations



Hill activation function H^+. θ_i is the threshold of the regulatory influence of the i^{th} molecule onto its targets, for which $H = 0.5$. $S > 0$ is the steepness parameter. For $S > 1$, Hill curves show a sigmoidal shape, as shown on this graph. As S increases, so does the sigmoidicity of the curve.

For which purpose do we need non-biologists in Biology ?

a) To hold the pipettes

b) To cope with the massive amount of data

c) To cope with the lack of data

d) No opinion