

Graduate Computational Genomics

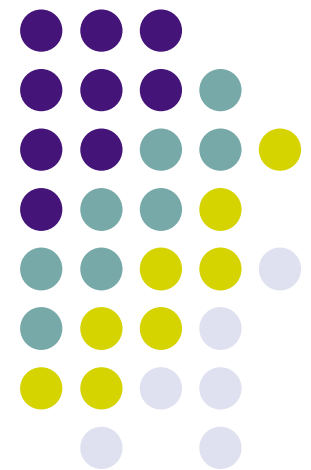
02-710 / 10-810 & MSCBIO2070

Modeling Regulatory Networks

Takis Benos

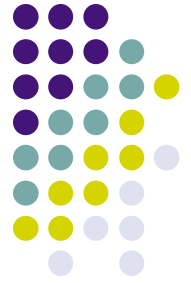
Lecture #24, April 17, 2007

Reading: handouts & papers



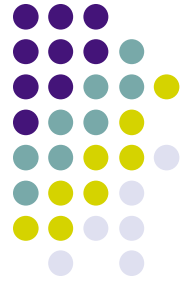


GRNs (cntd)



- GRNs typically model interactions between genes at the transcriptional level
- Philosophy change: from *reductionism* to *globalism*
- Missing
 - Post-translational modifications
 - Protein-protein interactions
 - Cellular (trans)location

GRNs: the data

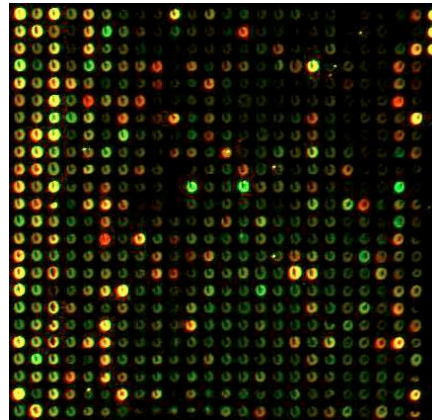


Gene expression



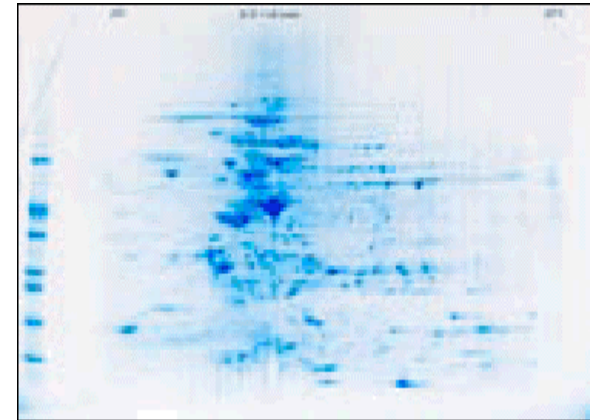
- ✓ Microarrays
- SAGE
- Real-time RT-PCR

ChIP-chip



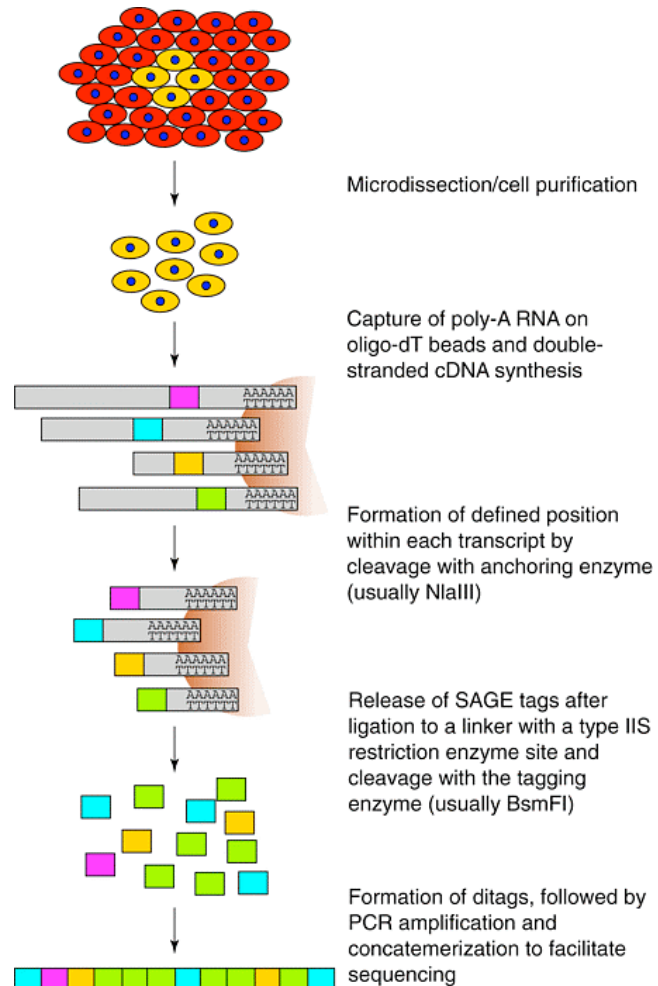
- ✓ Tiling arrays
- Promoter arrays

Proteomics



- 2D gels
- 2D DIGE

SAGE



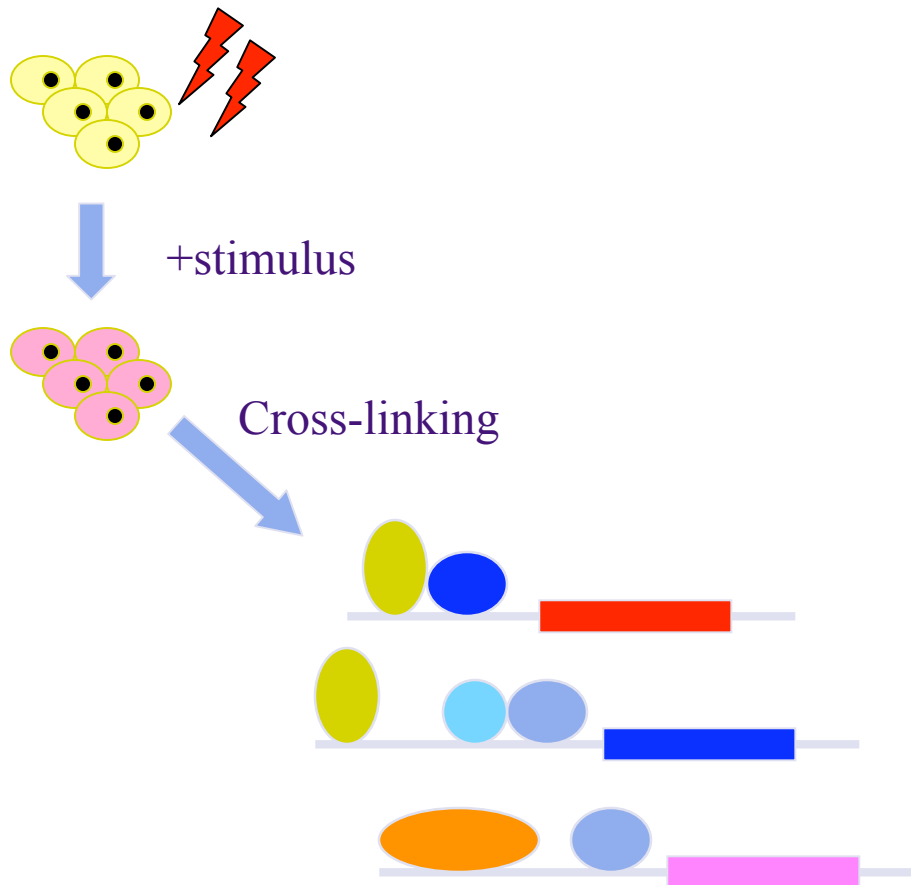
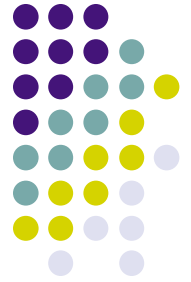
Sequencing and data analysis

Serial Analysis of Gene Expression *Velculescu et al., Science (1995)*

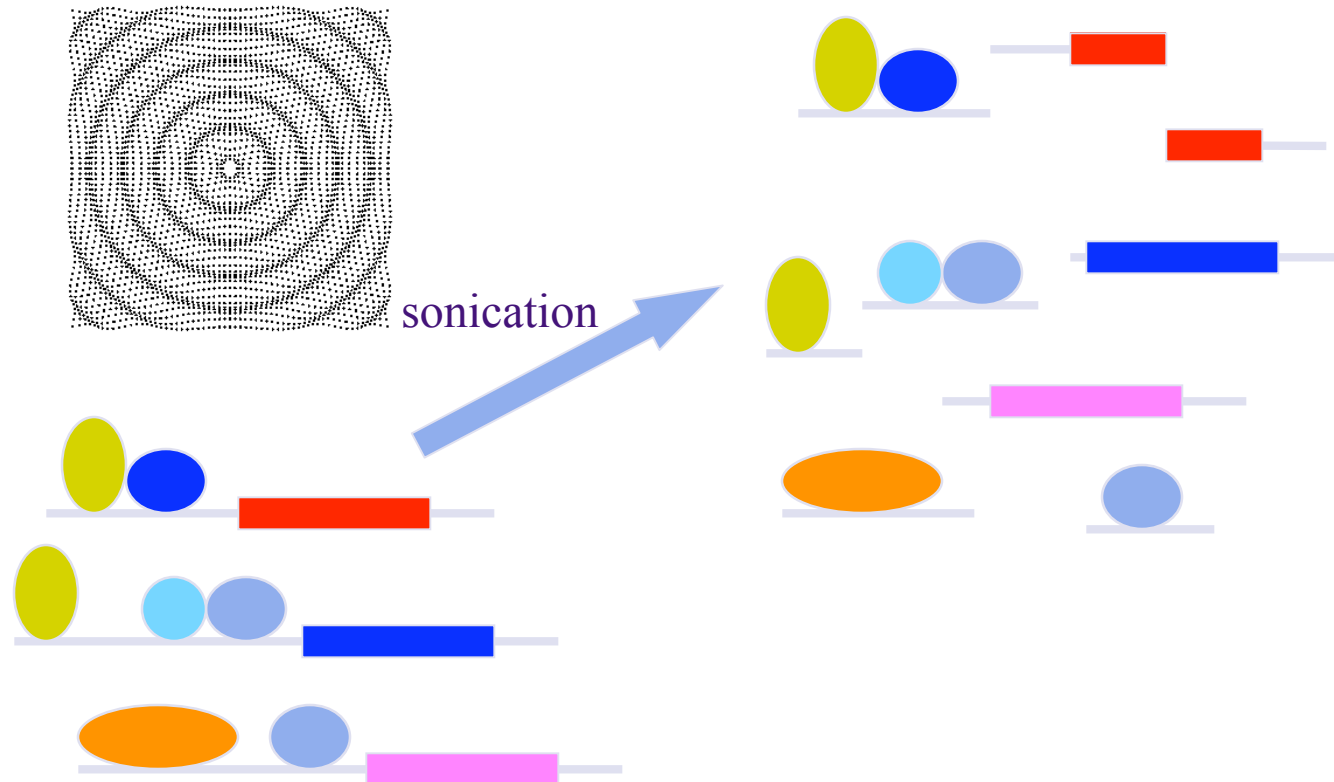
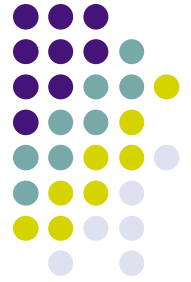
SAGETag	Tag Count (in 100 000 tags)	Absolute abundance
CATGGACGTCCTTAAT	33 TAGS	0.033%
CATGGTGACCTCCTT	63 TAGS	0.063%
CATGTGAAGAGAAGA	22 TAGS	0.022%
CATGAGTGGAGGTGG	9 TAGS	0.009%
NlaIII site		

trends in Genetics

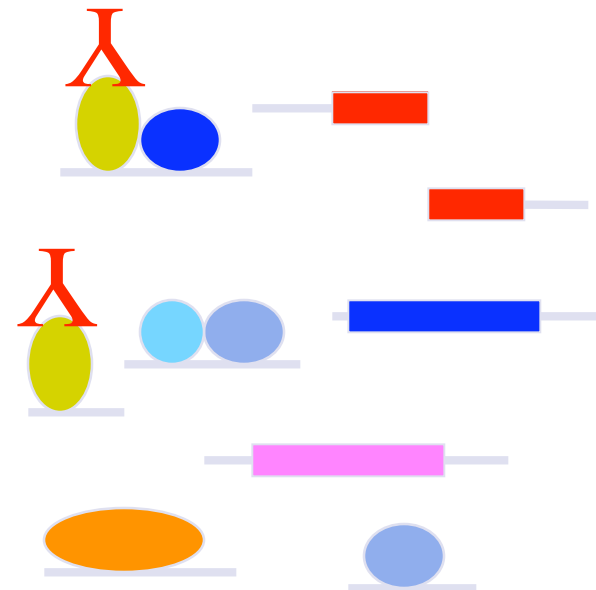
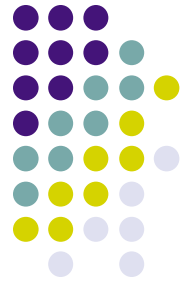
Chromatin immunoprecipitation



Chromatin immunoprecipitation

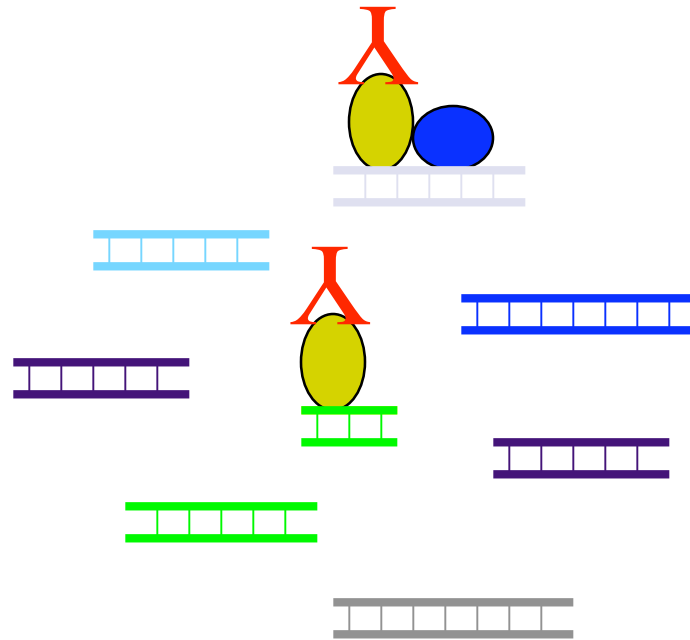
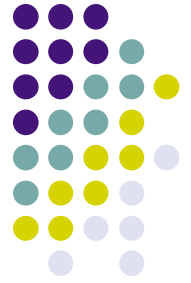


Chromatin immunoprecipitation



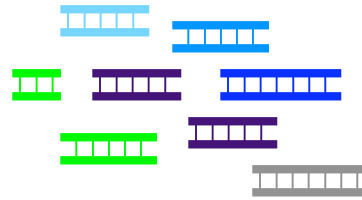
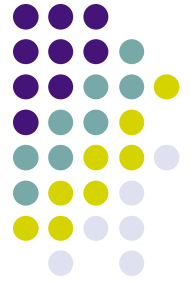
+ Ab

Chromatin immunoprecipitation

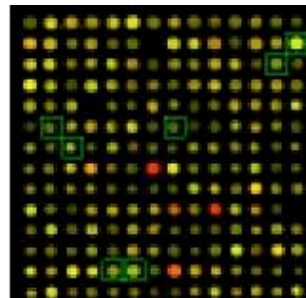


Proteinase K

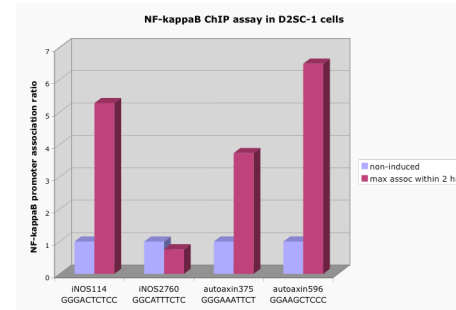
ChIP target verification



LM-PCR (global)



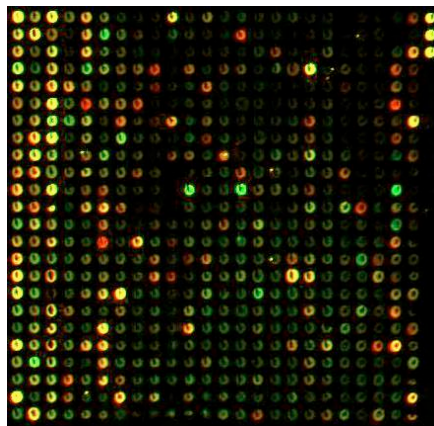
real-time PCR
(gene-specific)



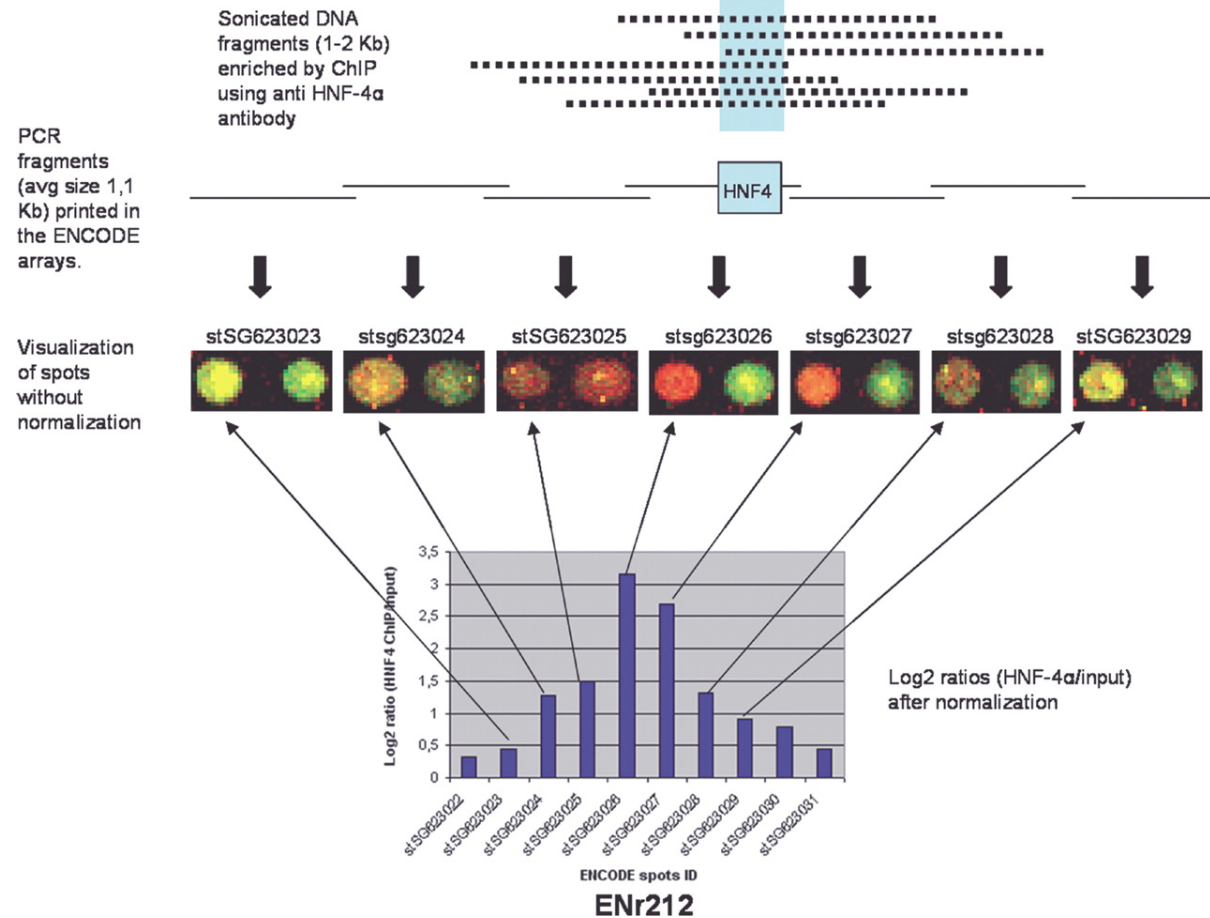
ChIP-chip (cntd)



ChIP-chip



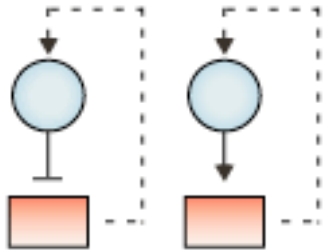
✓ Tiling arrays



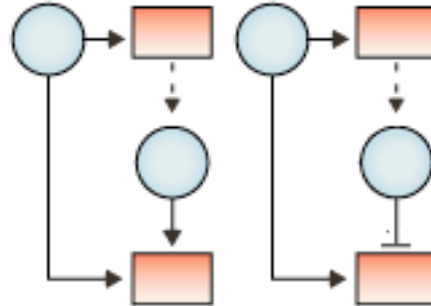
Network components



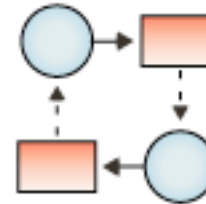
Autoregulatory loops



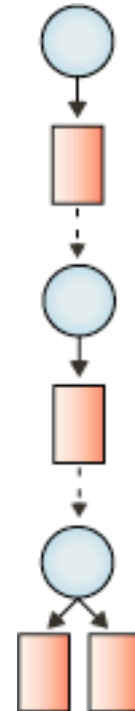
Feed-forward interactions



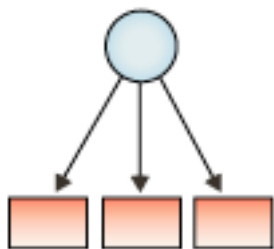
Multi-component loops



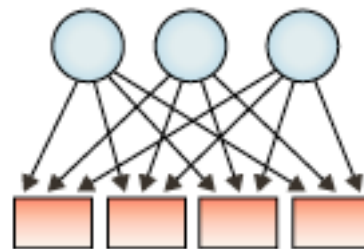
Regulator chains

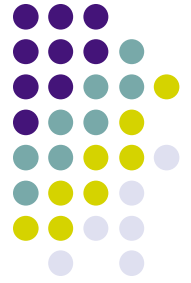


Single layer



Multiple layer





Methods for modeling GRNs

Ways to model GRNs:

- Boolean networks
 - Weight matrices
 - Graphical models
 - Differential equations
-
- Curse of dimensionality
 - Number of parameters increase exponentially to the number of nodes (yeast: ~140 TFs; human: ~2,000 TFs)
 - Solution: constrain the model

Boolean networks



- Genes are represented as nodes in a *directed graph*.
- Connection between genes (nodes) A and B exist iff there is a causal relationship between them.
- Genes (nodes) can be either ON or OFF.
- Time is considered to be discrete.

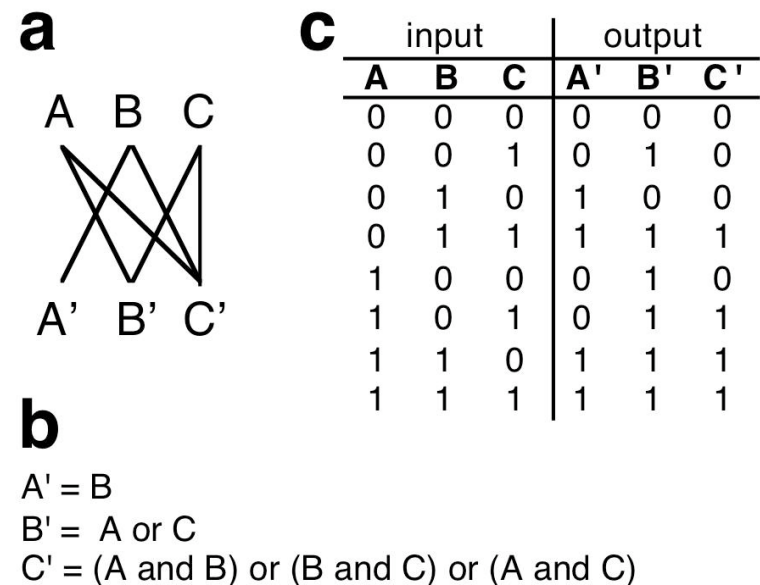
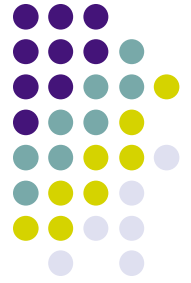


Fig. 1 A simple Boolean network. a) Wiring diagram. b) Logical (Boolean) rules. c) Complete state transition table defining network. The input column corresponds to the state at time= t , the output column (elements marked by prime) corresponds to the state at time= $t+1$.



Boolean networks: REVEAL

REVerse Engineering ALgorithm:

Liang, Fuhrman & Somogyi, *PSB* (1998) 3: 18-29

- They use mutual information to infer the wiring of the network (associations between genes)

$$H(X) = - \sum_x p(x) \cdot \log(p(x))$$

$$H(X,Y) = - \sum_{x,y} p(x,y) \cdot \log(p(x,y))$$

$$M(X,Y) = H(X) + H(Y) - H(X,Y)$$

- “*X exactly determines A*” when: $M(A^t, X^{t-1}) = H(A^t)$
- In each step, i , they calculate the $k=i$ links

REVEAL (cntd)



Input entropies

H(A)	1.00
H(B)	1.00
H(C)	1.00
H(A,B)	2.00
H(B,C)	2.00
H(A,C)	2.00
H(A,B,C)	3.00

$$H(X) = - \sum p(x) \log p(x)$$

$$H(X,Y) = - \sum p(x,y) \log p(x,y)$$

$$M(X,Y) = H(X) + H(Y) - H(X,Y)$$

$$M(X,[Y,Z]) = H(X) + H(Y,Z) - H(X,Y,Z)$$

Determination of inputs for element A

H(A')	1.00		
H(A',A)	2.00	M(A',A) 0.00	M(A',A) / H(A') 0.00
H(A',B)	1.00	M(A',B) 1.00	M(A',B) / H(A') 1.00
H(A',C)	2.00	M(A',C) 0.00	M(A',C) / H(A') 0.00

①

Rule table for A

rule no. 2	
input	
B	A'
0	0
1	1

Determination of inputs for element B

H(B')	0.81		
H(B',A)	1.50	M(B',A) 0.31	M(B',A) / H(B') 0.38
H(B',B)	1.81	M(B',B) 0.00	M(B',B) / H(B') 0.00
H(B',C)	1.50	M(B',C) 0.31	M(B',C) / H(B') 0.38
H(B',[A,B])	2.50	M(B',[A,B]) 0.31	M(B',[A,B]) / H(B') 0.38
H(B',[B,C])	2.50	M(B',[B,C]) 0.31	M(B',[B,C]) / H(B') 0.38
H(B',[A,C])	2.00	M(B',[A,C]) 0.81	M(B',[A,C]) / H(B') 1.00

③

Rule table for B

rule no. 14	
input	
A	C
0	0
0	1
1	0
1	1

Determination of inputs for element C

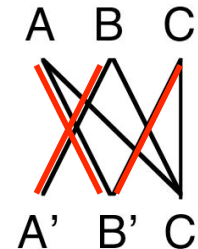
H(C')	1.00		
H(C',A)	1.81	M(C',A) 0.19	M(C',A) / H(C') 0.19
H(C',B)	1.81	M(C',B) 0.19	M(C',B) / H(C') 0.19
H(C',C)	1.81	M(C',C) 0.19	M(C',C) / H(C') 0.19
H(C',[A,B])	2.50	M(C',[A,B]) 0.50	M(C',[A,B]) / H(C') 0.50
H(C',[B,C])	2.50	M(C',[B,C]) 0.50	M(C',[B,C]) / H(C') 0.50
H(C',[A,C])	2.50	M(C',[A,C]) 0.50	M(C',[A,C]) / H(C') 0.50
H(C',[A,B,C])	3.00	M(C',[A,B,C]) 1.00	M(C',[A,B,C]) / H(C') 1.00

⑤

Rule table for C

rule no. 170	
input	
A	B
0	0
0	1
1	0
1	1

a



b

A' = B

B' = A or C

C' = (A and B) or (B and C) or (A and C)

c

input			output		
A	B	C	A'	B'	C'
0	0	0	0	0	0
0	0	1	0	1	0
0	1	0	1	0	0
0	1	1	1	1	1
1	0	0	0	1	0
1	0	1	0	1	1
1	1	0	1	1	1
1	1	1	1	1	1

REVEAL: performance



- In principle, it requires computation of the probabilities of the 2^N state transition pairs for each of the (k of N) possible wirings.
- Solution: do a sampling in the input patterns
- Should use the same sample to calculate all quantities
- How many sampled input patterns?

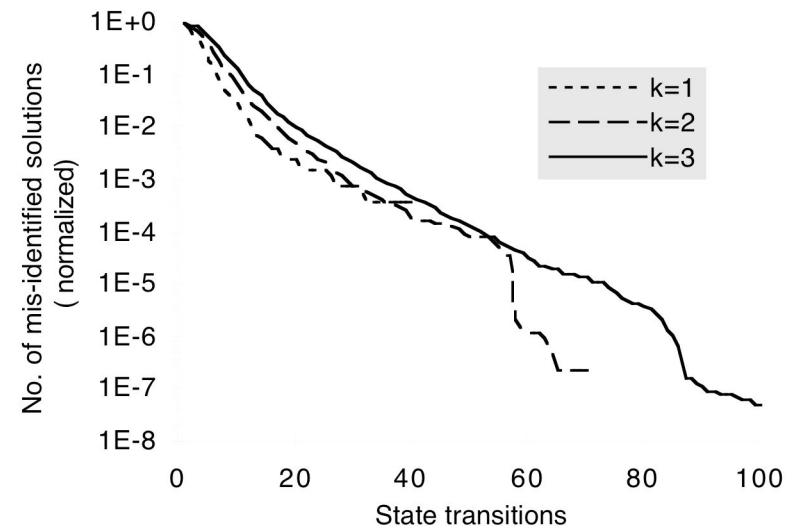
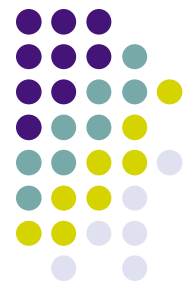
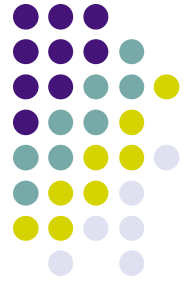


Fig. 6 Reduction of mis-identified network wiring solutions. The number of erroneous wirings identified by the M-analysis (normalized) versus the number of state transition pairs used for *effective* k value $k=1,2,3$. The data was obtained by averaging over 50 random wirings for a network with 50 elements. Note that a correct solution is always found; this is subtracted from the plotted number of solutions.

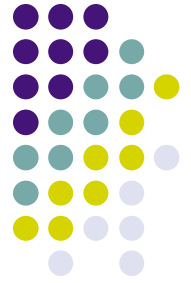




For some reading

- Baohong Zhang *et al.*, “*Computational identification of microRNAs and their targets*” (2006) *Computational Biology and Chemistry* **30**: 395-407
and references therein...
- *rna22*: Miranda *et al.*, “*A pattern-based method for the identification of microRNA binding sites and their corresponding heteroduplexes*” (2006) *Cell* **126**: 1203-1217

Acknowledgements



Some of the slides used in this lecture are adapted or modified slides from lectures of:

- Serafim Batzoglou, Stanford University
- David Corcoran, University of Pittsburgh

Pictures from the following:

- <http://wmd.weigelworld.org/bin/mirnatools.pl?page=6>