Network Evolution

Networks in Cellular Biology

- A. Metabolic Pathways
- B. Regulatory Networks
- C. Signaling Pathways
- D. Protein Interaction Networks - PIN

- E. Other Networks
  - The Internet

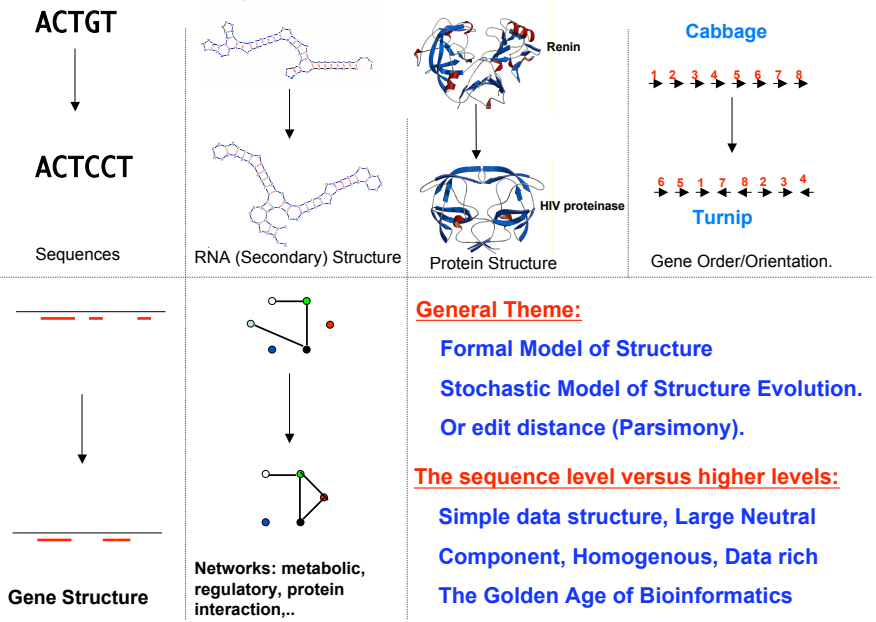
Statistics of Networks

Comparing Networks

- Network Matching
- Stochastic Models of Network

Examples of Comparison and Evolution

Comparative Biology



General Theme:

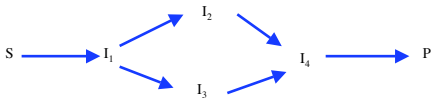
- Formal Model of Structure
- Stochastic Model of Structure Evolution.
- Or edit distance (Parsimony).

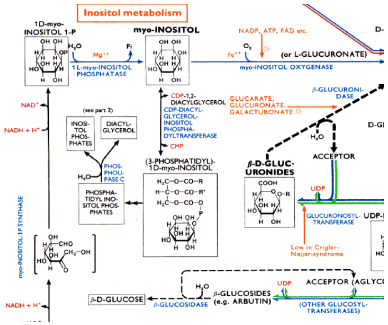
The sequence level versus higher levels:

- Simple data structure, Large Neutral Component, Homogenous, Data rich
- The Golden Age of Bioinformatics

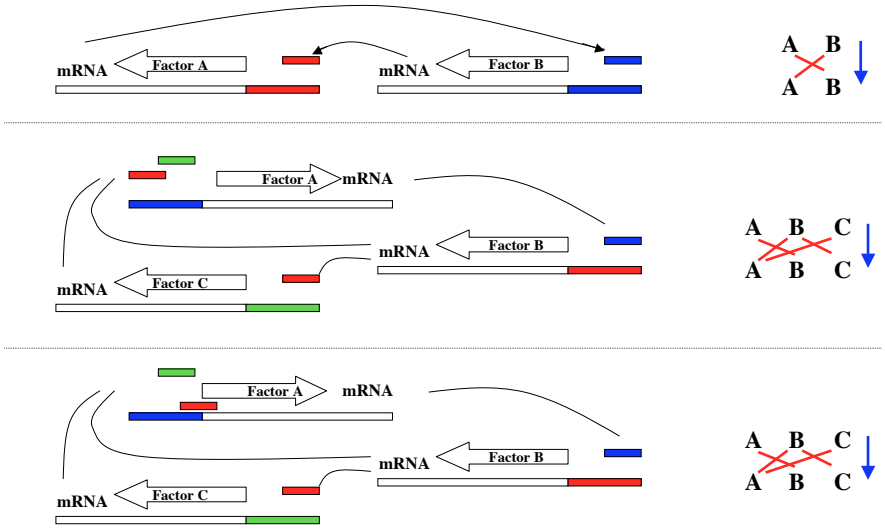
A. Metabolic Pathways

- Flux Analysis
- Metabolic Control Theory
- Biochemical Systems Theory
- Kinetic Modeling



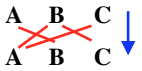


B. Regulatory Networks



Remade from Somogyi & Sniegoski,96. F2

Boolen functions, Wiring Diagrams and Trajectories



Inputs	2	1	1
Rule	4	2	2

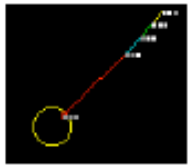
A activates B

B activates C

A is activated by B, inhibited by (B>C)

Point Attractor

A	B	C
1	1	0
1	1	1
0	1	1
0	0	1
0	0	0
0	0	0



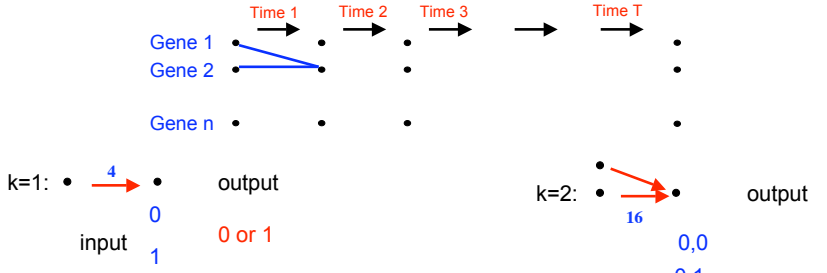
2 State Attractor

A	B	C
1	0	0
0	1	0
1	0	1
0	1	0

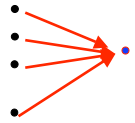
Remade from Somogyi & Sniegoski,96. F4

Boolean Networks

R.Somogyi & CA Sniegoski (1996) Modelling the Complexity of Genetic Networks Complexity 1.6.45-64.



A single function:



2^k

The whole set:

2^{k^k}

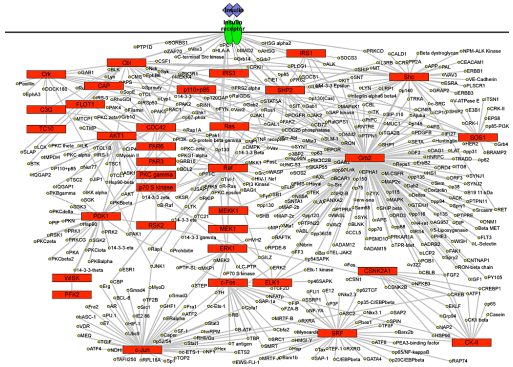
For each gene dependent on i genes: $\binom{k}{i}$ choices of dependent genes. Number of Boolean Rules $\left(\binom{k}{i}2^i\right)^k$

Contradiction: Always turned off (biological meaningless) Tautology: Always turned on (household genes)

C. Signaling Pathways

•Transmits signals from membrane to gene regulation.

•Its function is enigmatic as some of the molecules involved are common to different functions and how cross-interaction is avoided is unknown.



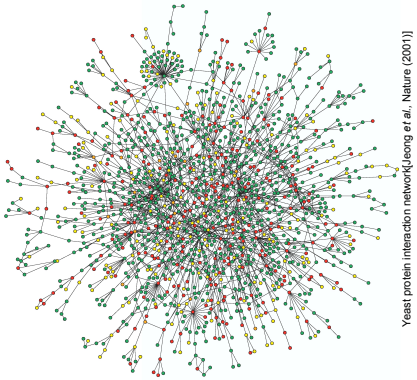
www.hprd.org from Pierre deMeyts

D. Protein Interaction Network

•The sticking together of different protein is measured by mass spectroscopy.

•The nodes will be all known proteins.

•Two nodes are connected if they stick together. This can be indicator of being part of a a functional protein complex, but can also occur for other reasons.



Yeast protein interaction network(Jeong *et al.*, Nature (2001))

E. Other Networks

More Sub-Cellular

- Alternative Splicing Graph

Cellular

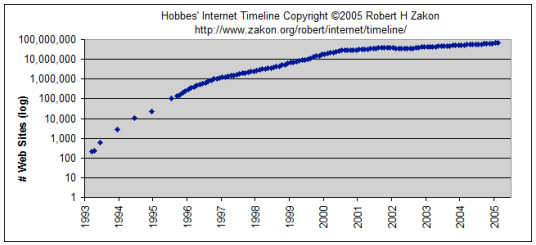
- Neural Networks
- Immunological Networks

Above the Cell

- Disease Networks
- Genealogical Networks

Non-biological Networks

- Social Networks
- The Internet
- Collaboration Networks
- Semantic Networks
- Publications and references



Network Description and Statistics I

Barabasi & Oltvai, 2004

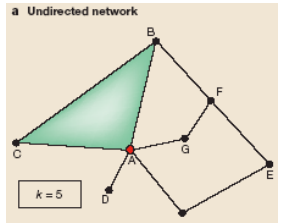
- Degree
- Shortest Path $Dist(i,j)$
- Mean Path Length
- Diameter: $Max_{i,j}\{Dist(i,j)\}$
- Clustering Coefficient - $C_I=2T_I/n_I(n_I-1)$

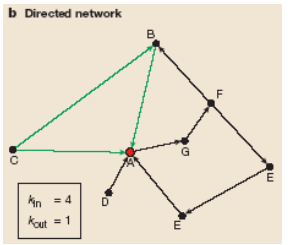
  $C_A=2/20$
- Degree Distribution - P(k)
- Scale Free Networks $P(k)\sim k^{-\gamma}$ $\gamma>2$
- Hubs: multiply connected nodes

  The lower γ, the more hubs.

Small World Property:

Graph connected and path lengths small



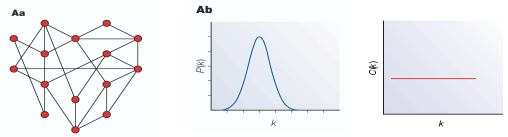


Remade from Barabasi, 2004

Network Description and Statistics II

Barabasi & Oltvai, 2004

A. Random Networks [Erdos and Rényi (1959, 1960)]



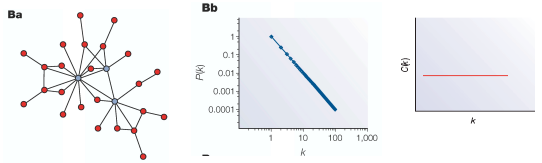
$$P(k) = \frac{e^{-\bar{k}}\bar{k}^{k}}{k!}$$

Mean path length ~ ln(k)

Phase transition:

Connected if: $p \geq \ln(k)/k$

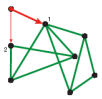
B. Scale Free [Price,1965 & Barabasi,1999]



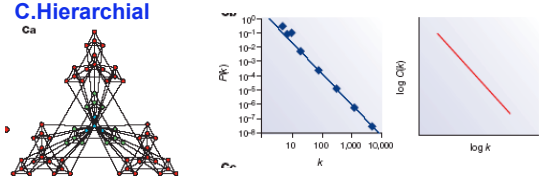
$P(k) \sim k^{-\gamma},\ k >> 1,\ 2 < \gamma$

Mean path length ~ lnln(k)

Preferential attachment. Add proportionally to connectedness



C.Hierarchial

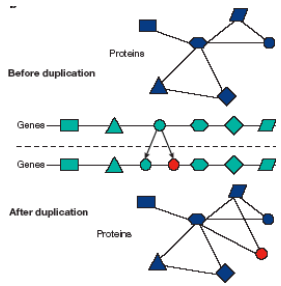


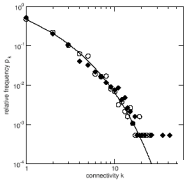
Copy smaller graphs and let them keep their connections.

Network Evolution

Barabasi & Oltvai, 2004 & Berg et al. ,2004

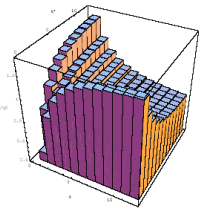
- **A gene duplicates**
- **Inherits it connections**
- **The connections can change**





Berg et al. ,2004

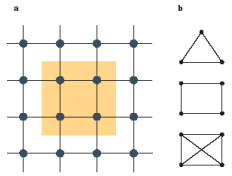
- **Gene duplication slow ~10^{-9}/year**
- **Connection evolution fast ~10^{-6}/year**
- **Observed networks can be modeled as if node number was fixed.**

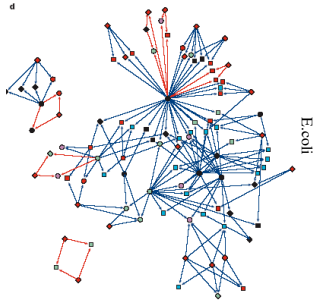


Network Alignment & Motifs

- Global Network Matching
- Network integration
- Network Search
- Motifs

Barabasi & Oltvai, 2004





A Model for Network Inference I

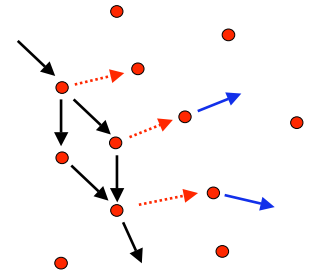
- A given set of metabolites:
- A given set of possible reactions - arrows not shown.
- A core metabolism: →
- A set of present reactions - M black and red arrows

Restriction R:

A metabolism must define a connected graph

M + R defines

1. a set of deletable (dashed) edges D(M): ⇢
2. and a set of addable edges A(M): →



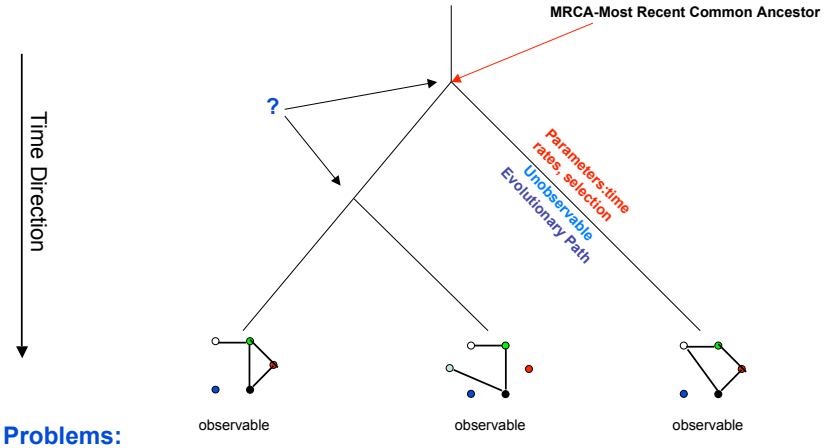
Let μ be the rate of deletion

λ the rate of insertion

Then

$$\frac{dP(M)}{dt} = \lambda \sum_{M' \in D(M)} P(M') + \mu \sum_{M'' \in A(M)} P(M'')$$
$$- P(M)[\lambda |D(M)| + \mu |A(M)|]$$

A Model for Network Inference II



3 Problems:

i. Test all possible relationships.

ii. Examine unknown internal states.

iii. Explore unknown paths between states at nodes.

Recommended Literature

A.Cornish-Bowden (1995) Fundamentals of Enzyme Kinetics Portland Press

David Fell (1997) Understanding the Control of Metabolism. Portland Press.

Gottschalk (1987) Bacterial Metabolism (2nd edition) Springer

R. Heinrich & S.Schuster (1996) The Regulation of Cellular Systems. Chapman and Hall.

Gerhard Michal (ed.) (1999) Biochemical Pathways. Wiley

Savageau, M.(1976.) Biochemical Systems Theory. Addison-Wesley.

Stephanopoulos, G. et al. (1999) Metabolic Engineering. Academic Press.

Dandekar, T. et al. (1999) Pathway Alignment: application to the comparative analysis of glycolytic enzymes. J. Biochem. 343.115-124.

JS Edwards et al (2001) In silico predictions of E.coli metabolic capabilities are consistent with experimental data. Nature Biotechnoology 19.Feb. 125-130.

Karp, P (2001) Pathway Databases: A Case Study in Computational Symbolic Theories. Science 293.2040-

Schuster, S et al. (1999) Detection of elementary flux modes in biochemical networks. TIBTech vol 17.53-59.

Schilling, C., D.Letscher and B.O.Palsson. (2000) J. Theor.Biol.203.229-248. "Theory for the Systemic Definition of Metabolic Pathways from a Pathway-Oriented Perspective."

Schilling, C and B.O.Palsson. (2000) J. Theor.Biol.203.249-283. "Assessment of the Metabolic Capabilities of Haemophilus influenzae Rd. through a Genome-scale Pathway Analysis."

Schuster, S et al. (1999) Detection of elementary flux modes in biochemical networks. TIBTech vol 17.53-59.

P.D'haeseleer, Liang & Somogyi (2000) Genetic network inference: from co expression clustering to reverse engineering. Bioinformatics 16.8.707-726

T.Akutsu, Miyano & Kuhara (2000) Inferring qualitative relations in genetic networks and metabolic pathways. Bioinformatics 16.8.727-734.

Liang & Somogyi (1998) Genetic network inference: from co-expression clustering to reverse engineering. PSB

T.Akutsu, Miyano & Kuhara (1999) Identification of genetic networks from a small number of gene expression patterns under the boolean network model. PSB 4.17-28