

***Eivind Almaas***  
**Microbial Systems Division**

# Introduction to Biological Networks



UC Davis, May 18<sup>th</sup> 2006



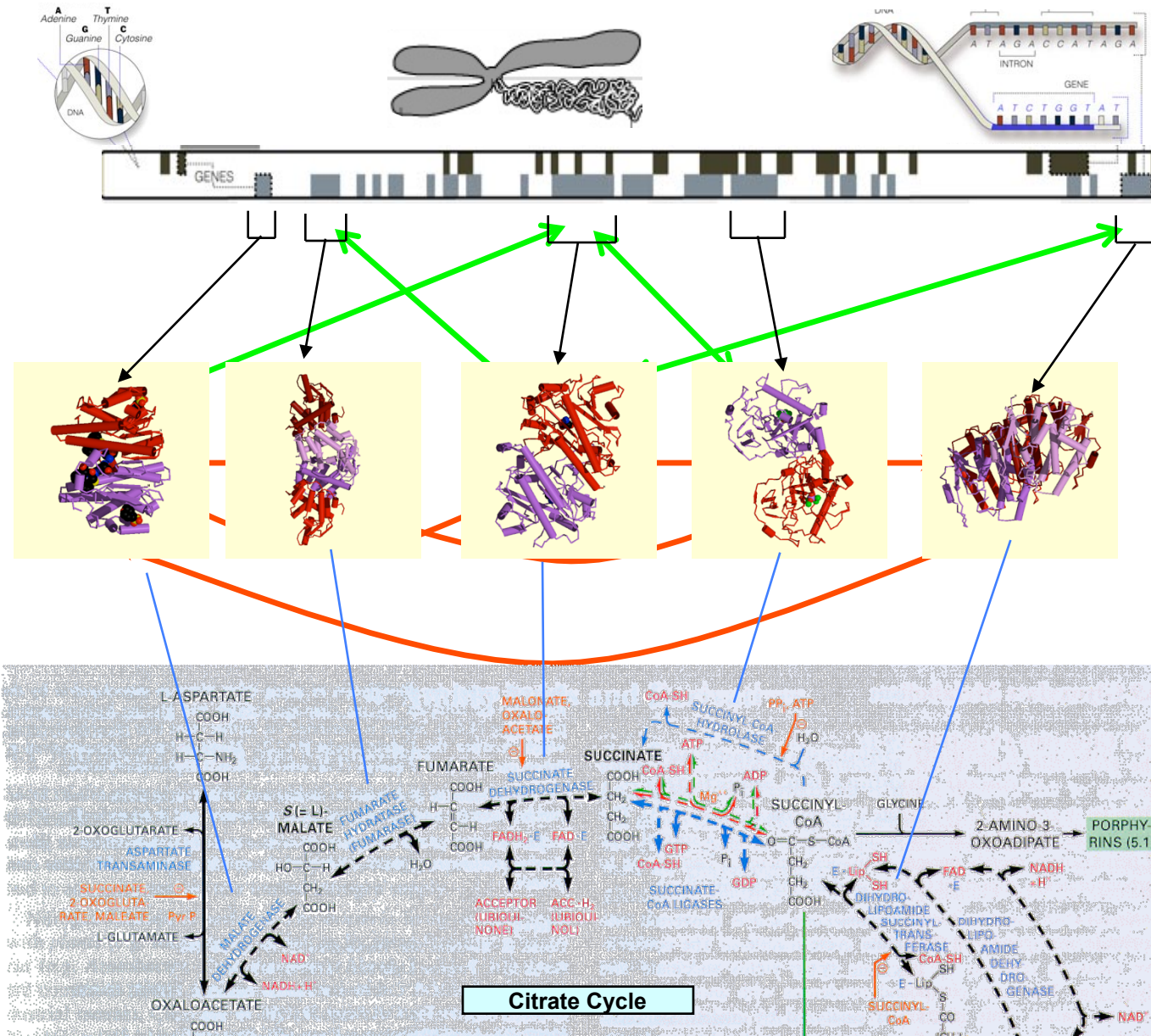
# Biological network examples

---



- Gene-regulation
- Protein interaction
- Metabolism
- Cell signaling
- Cytoskeleton
  
- ...
  
- Neural network
- Lymphatic node system
- Circulatory system

# Cellular networks:



**GENOME**

protein-gene  
interactions

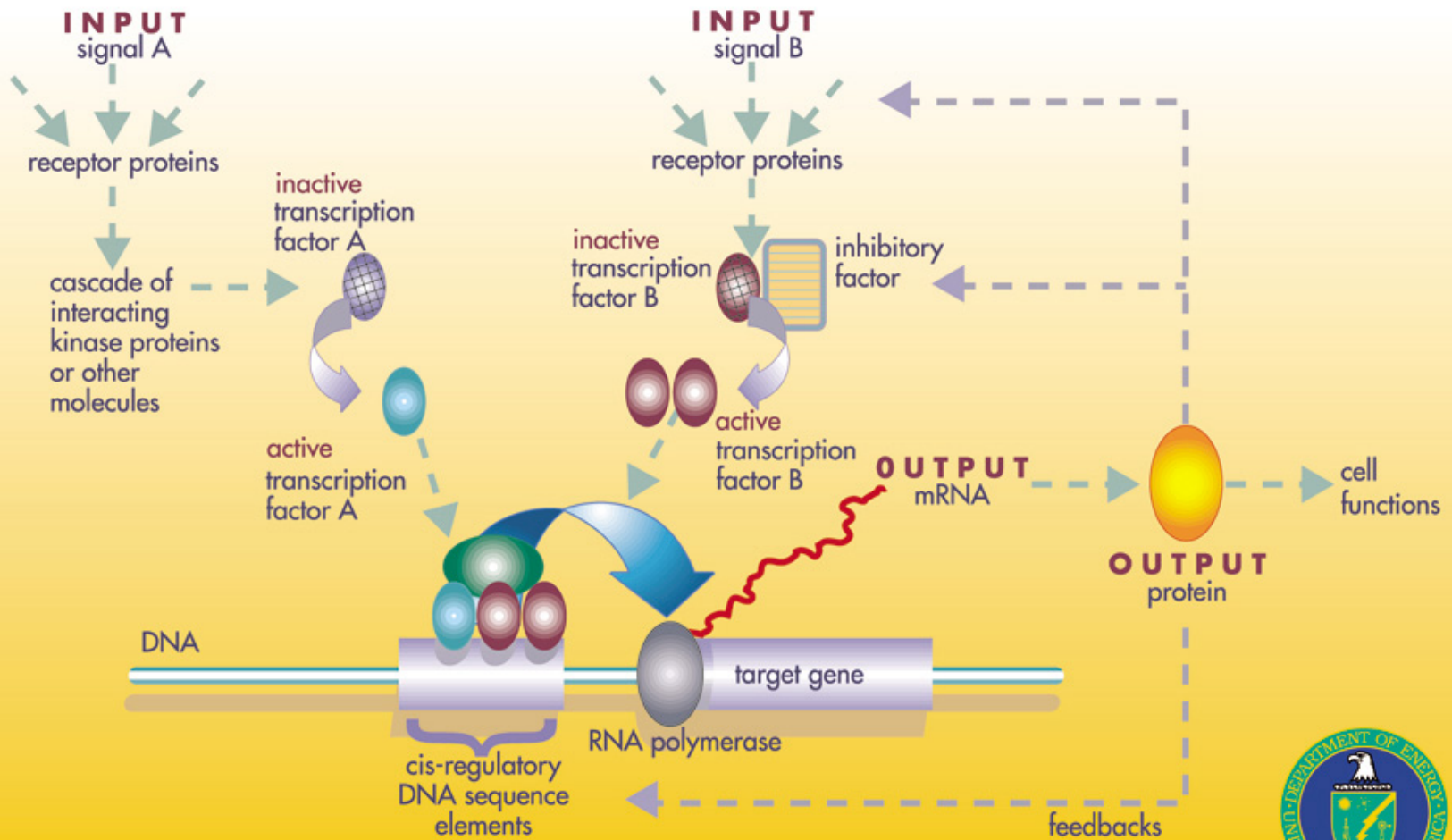
**PROTEOME**

protein-protein  
interactions

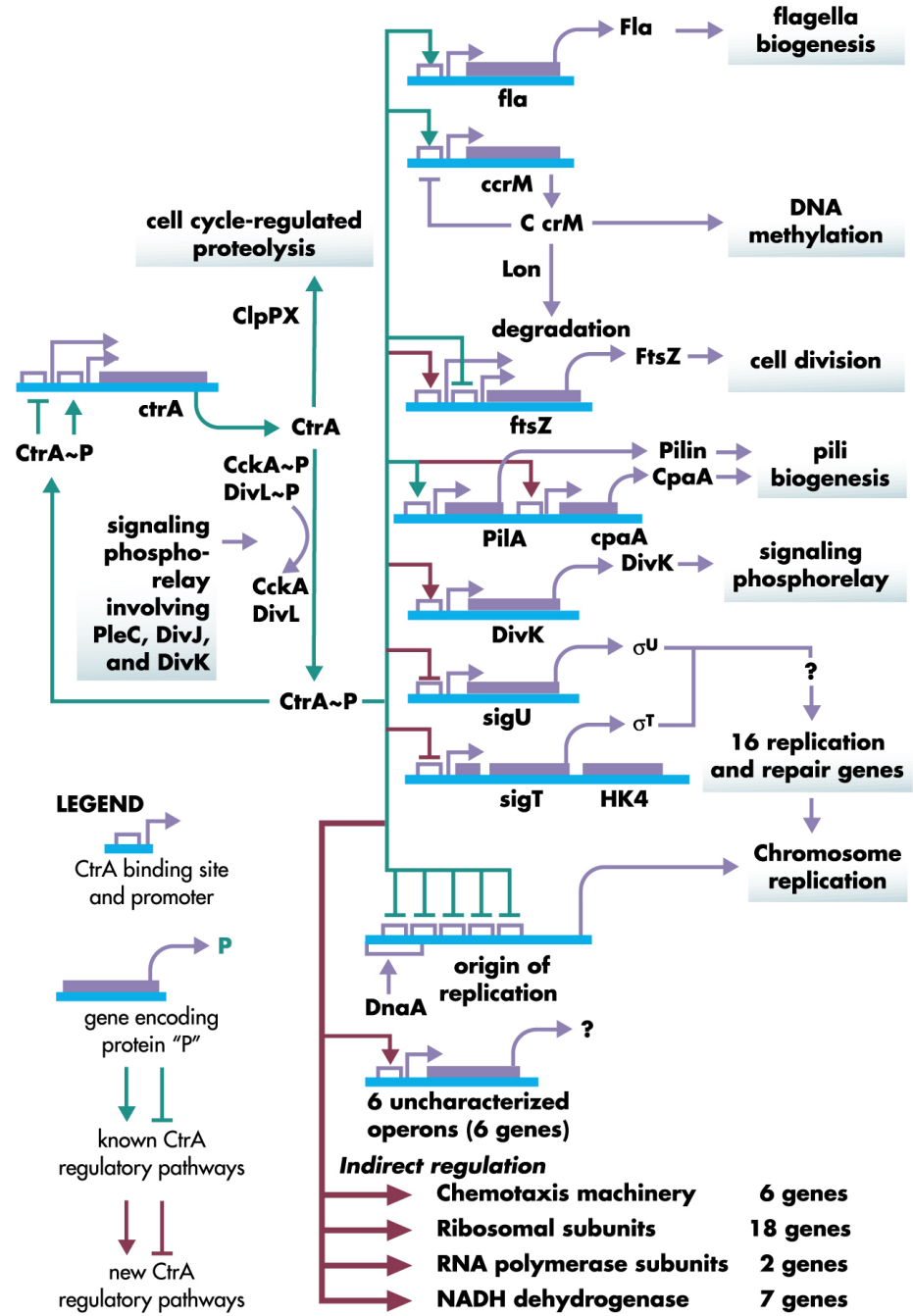
**METABOLISM**

Bio-chemical  
reactions

## A GENE REGULATORY NETWORK



# A CAULOBACTER CELL DIVISION GRN

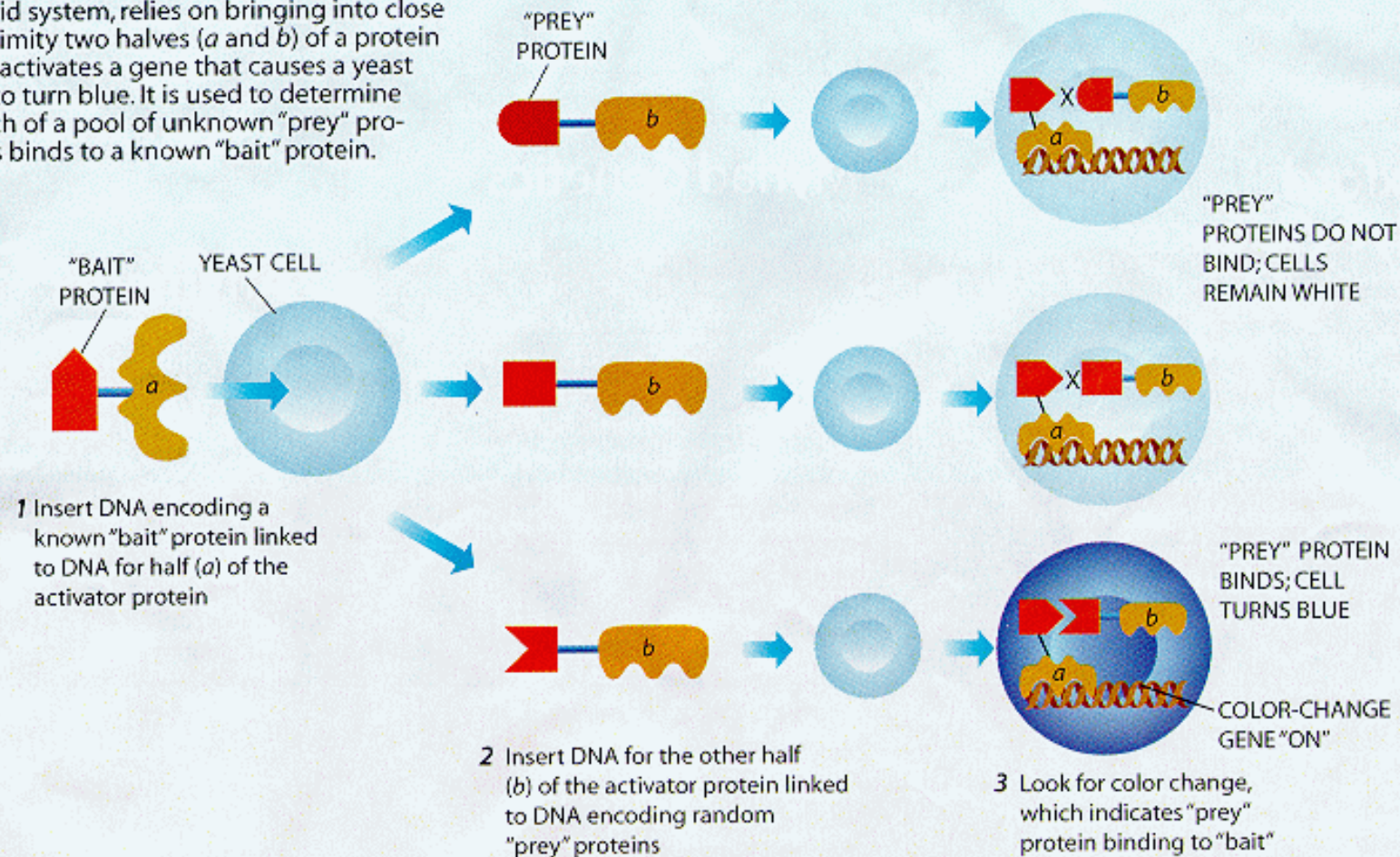


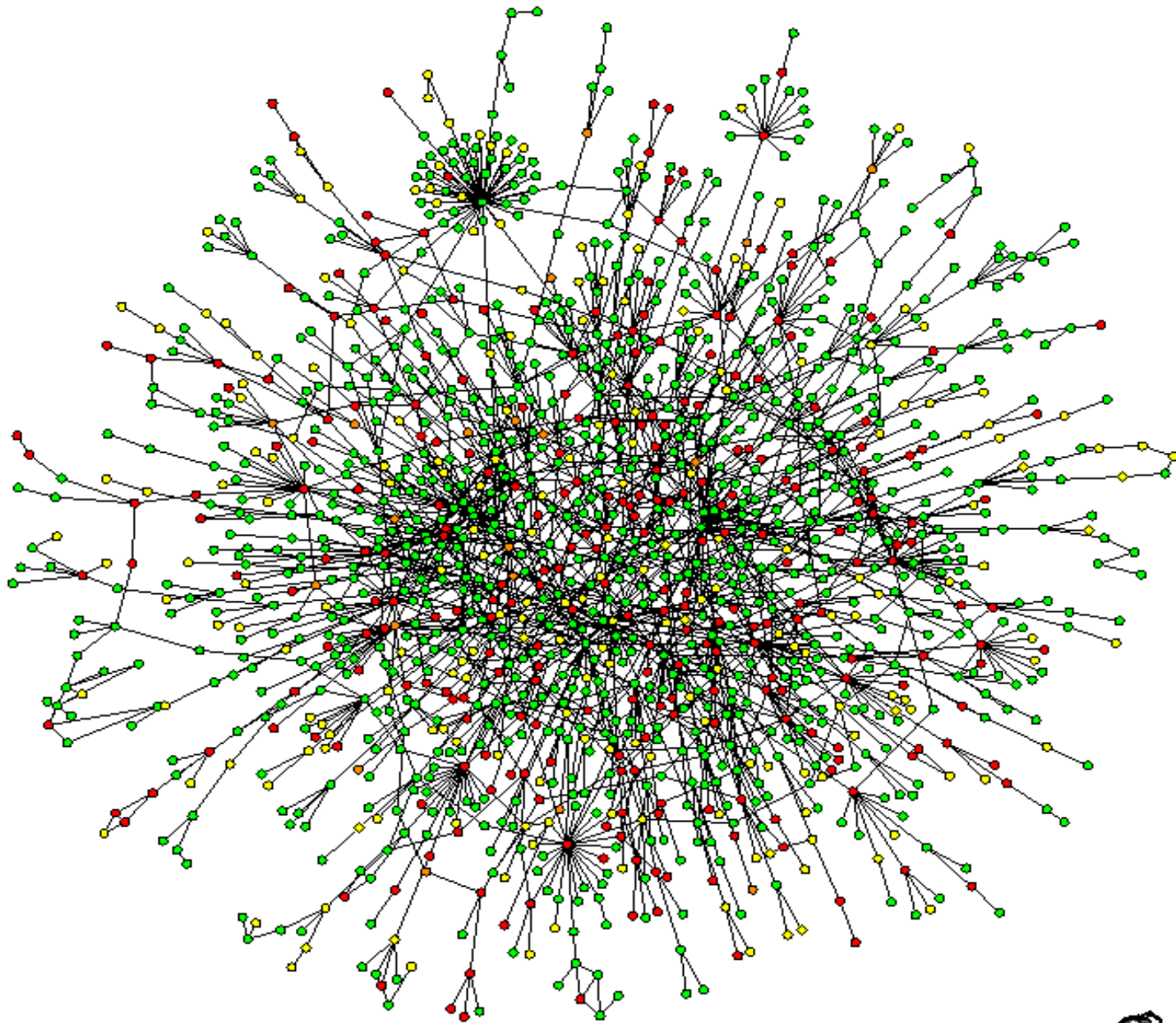
# Protein Interaction Networks (PIN)

# Protein interactions: Yeast two-hybrid method

## Finding Proteins That Interact

One technique, called the yeast two-hybrid system, relies on bringing into close proximity two halves (*a* and *b*) of a protein that activates a gene that causes a yeast cell to turn blue. It is used to determine which of a pool of unknown "prey" proteins binds to a known "bait" protein.

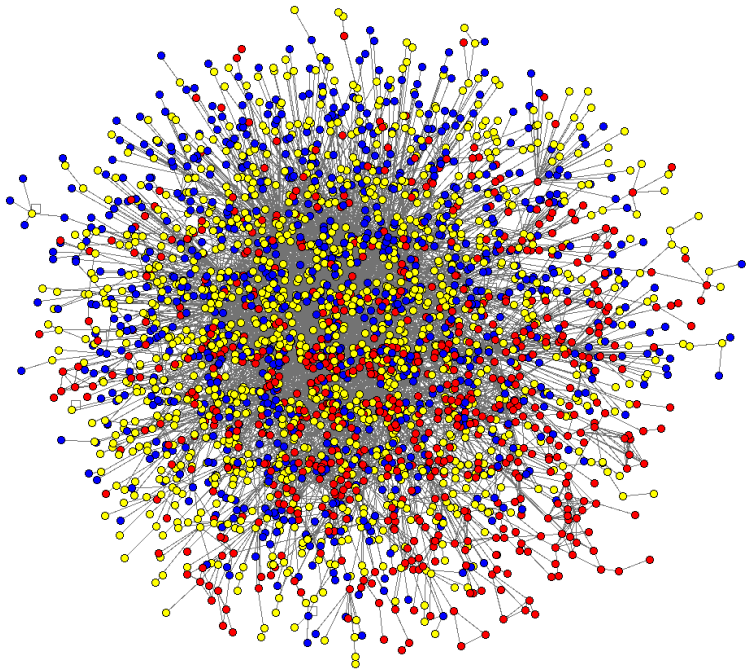
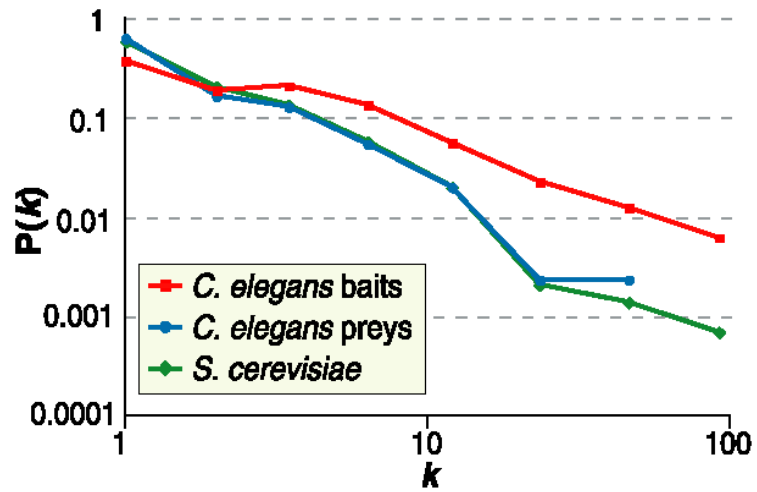




P. Uetz et. al. Nature 403, 601 (2000)  
H. Jeong et. al. Nature 411, 41 (2001)

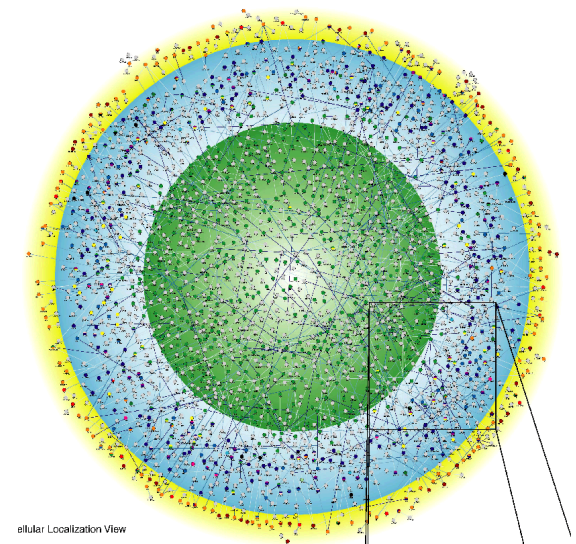
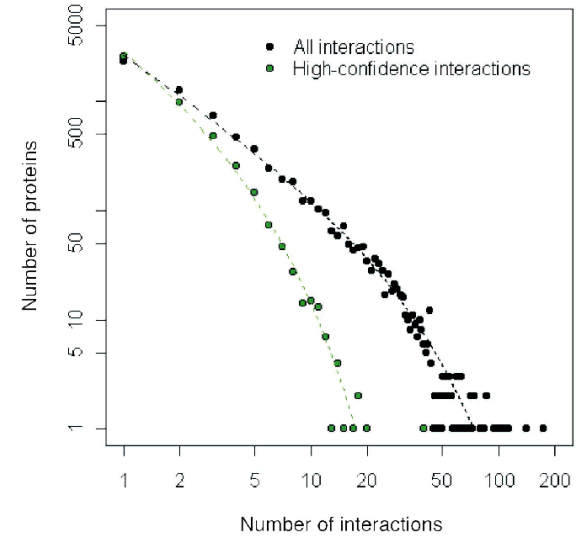


## *C. Elegans*



Li et al, Science **303**, 540 (2003)

## *Drosophila M.*



Giot et al, Science **302**, 1727 (2003)

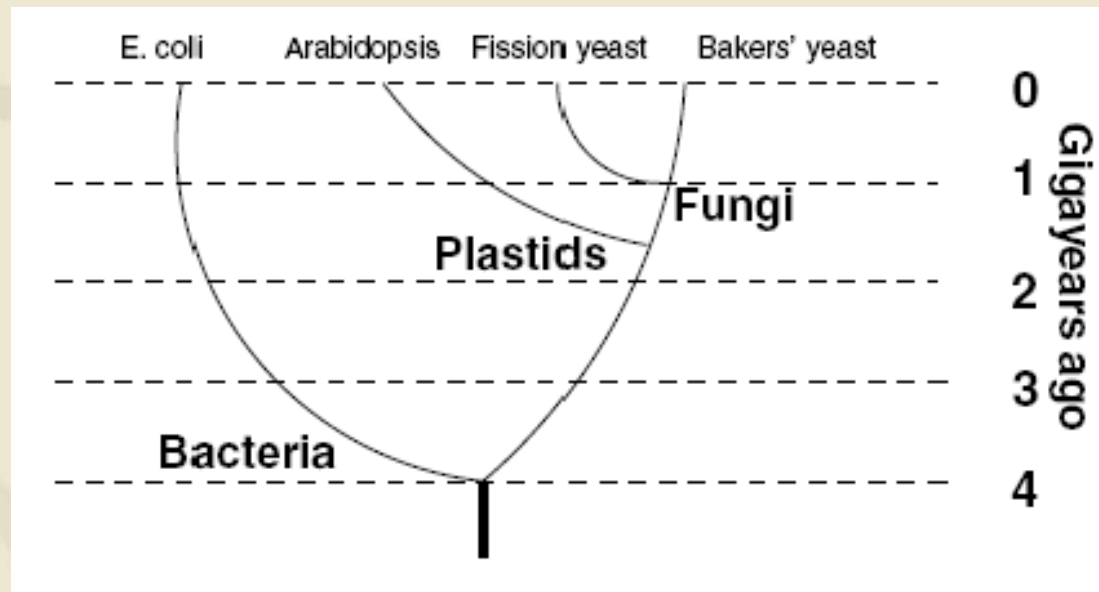
# PINs are scale-free...



Protein interaction networks are scale-free.

- is this because of preferential attachment?
- another mechanism?
- how can we determine the cause?

# Comparison of proteins through evolution



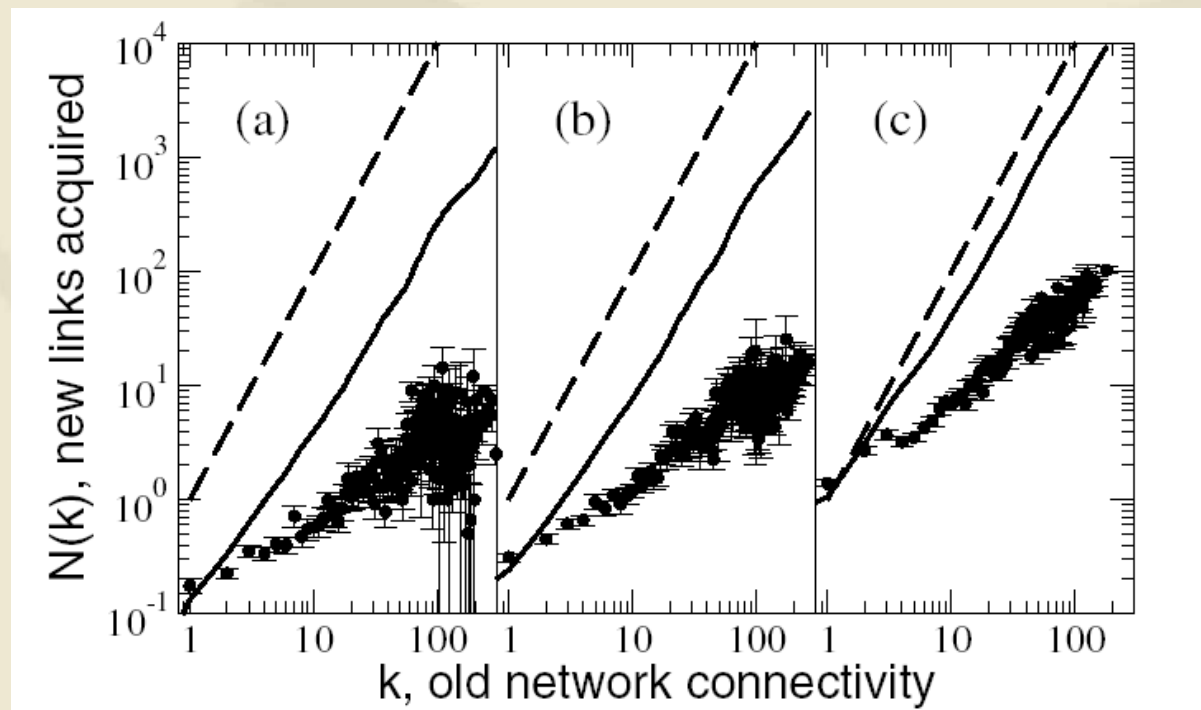
Use Protein-Protein BLAST (Basic Local Alignment Search Tool)  
-check each yeast protein against whole organism dataset  
-identify significant matches (if any)

# Preferential Attachment!



$$\frac{\partial k_i}{\partial t} \propto \Pi(k_i) \sim \frac{\Delta k_i}{\Delta t} \quad \text{For given } \Delta t: \Delta k \propto \Pi(k)$$

**k vs.  $\Delta k$  : linear increase in the # of links**

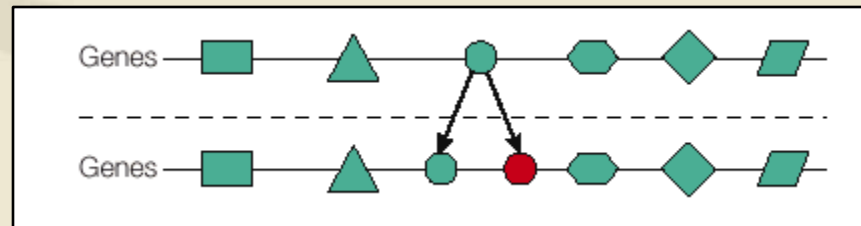


*S. Cerevisiae* PIN: proteins classified into 4 age groups

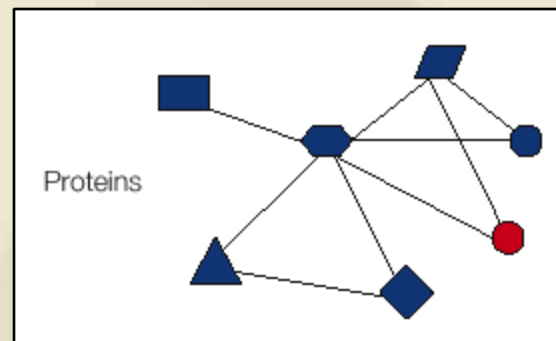
# SF topology from: duplication & diversification



Copying DNA: when mistake (gene duplication) happens



Effect on network:



**Proteins with more interactions are more likely to get a new link:**

$$\Pi(k) \sim k$$

**preferential attachment**

Wagner (2001); Vazquez *et al.* 2003; Sole *et al.* 2001; Rzhetsky & Gomez (2001); Qian *et al.* (2001); Bhan *et al.* (2002).

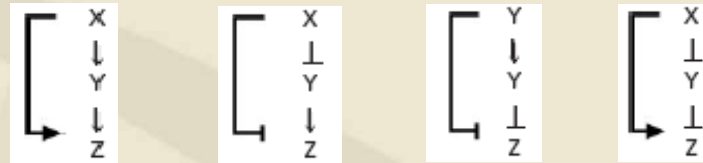
**How can we dissect the PIN?**

# Network motifs



Definition: A motif is a recurrent network module

Examples:



- Can think of networks as constructed by combining these “basic” building blocks
- Do these motifs have special properties?

R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, U. Alon  
“Network Motifs: Simple Building Blocks of Complex Networks”, Science 2002

# PIN motifs and evolution



#	Motifs	Number of Yeast Motifs	Natural Conservation Rate	Random Conservation Rate	Conservation Ratio
1		8,692	12.74 %	4.58 %	2.78
2		158,912	4.64 %	1.10 %	4.34
3		2,310	15.19 %	1.07 %	13.82
4		3,580,496	0.66 %	0.28 %	2.41
5		1,621,883	2.50 %	0.19 %	12.83
6		10,546	8.73 %	0.18 %	47.23
7		112,492	5.73 %	0.14 %	41.23
8		7,662	14.71 %	0.14 %	103.39
9		739	13.67 %	0.28 %	77.69
10		17,118	5.03 %	0.09 %	134.68
11		199	12.27 %	0.00 %	inf

**Table 3SM: The evolutionary conservation of non-complex motifs.** The third column denotes the number of motifs obtained by counting all subgraphs of two to five nodes found in the yeast protein interaction network of 3,143 proteins and 8702 interactions. This network has been cleaned from interactions which originate from annotated complexes of the Swissprot database. Thus, only 31 proteins but 564 interactions were removed from the original network. As set of orthologs we choose 678 proteins that have an ortholog in each of the five studied higher eukaryotes, and identified all motifs for which each component belongs to this evolutionary conserved protein subset. The natural conservation rate shows the fraction of the original yeast motifs that are evolutionary fully conserved, i.e., each of their protein components belong to the 678 orthologs of the list. The random conservation rate denotes the fraction of motifs which are found to be fully conserved for a random ortholog distribution. The last column denotes the ratio between the natural and the random conservation ratios, indicating that all motifs are highly conserved.

Protein BLAST against:  
 A. thaliana  
 C. elegans  
 D. melanogaster  
 M. musculus  
 H. sapiens

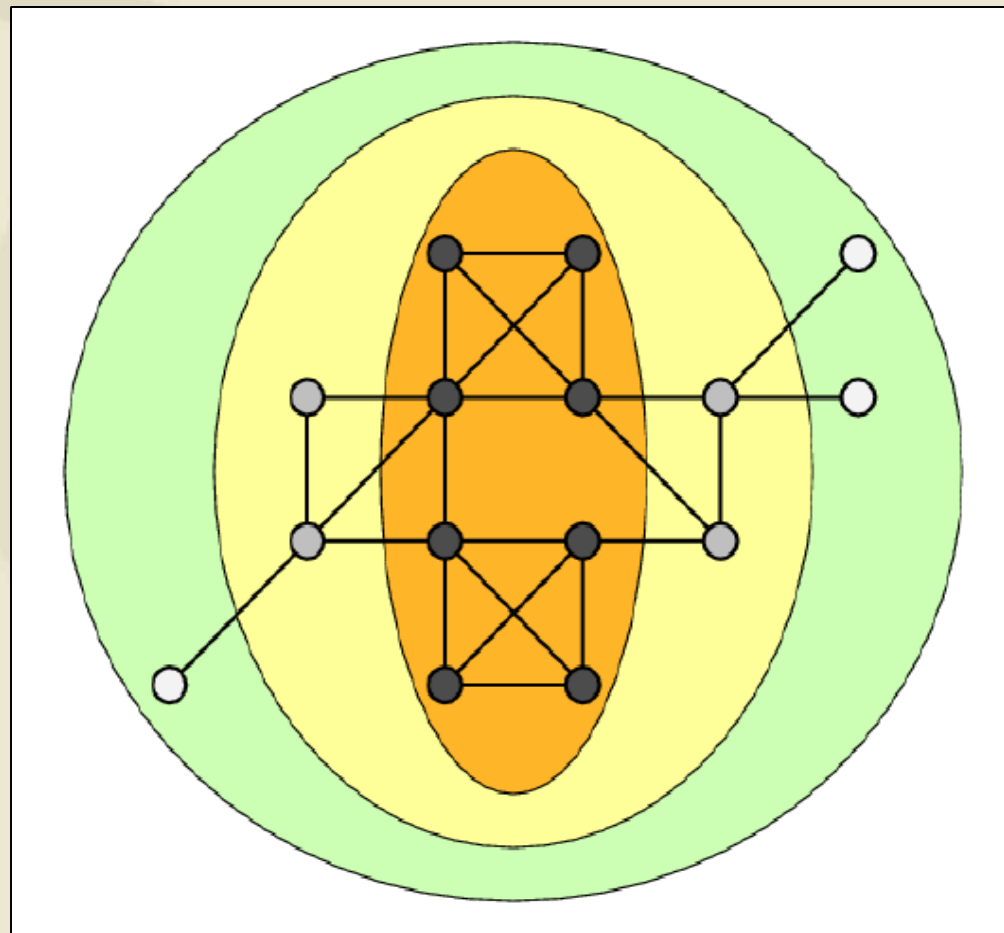
S. Wuchty, Z.N. Oltvai,  
 A.-L. Barabasi, 2003.



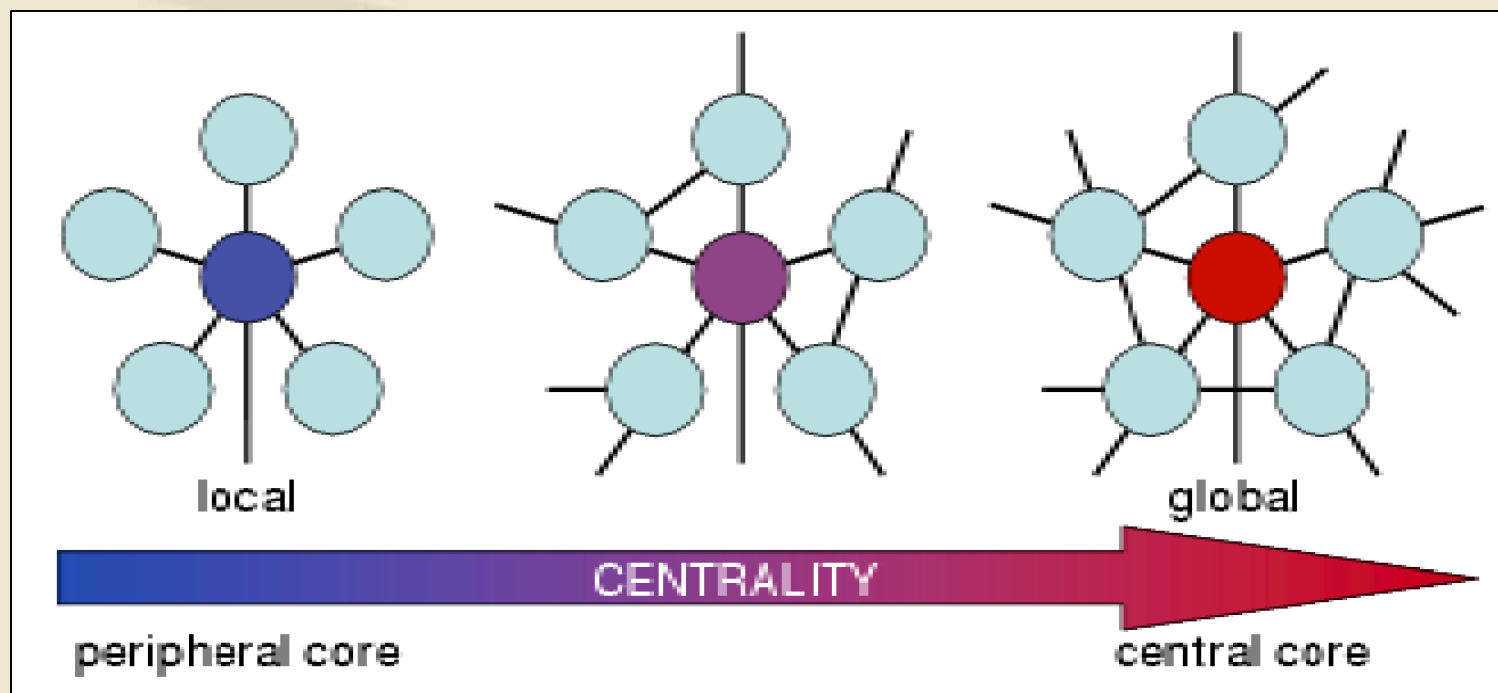
# Network peeling

Core decomposition method:

- the  $k$ -core consists of *all* nodes with degree  $\geq k$ .
- recursively remove nodes with degree  $< k$ .



# Local vs. global centrality



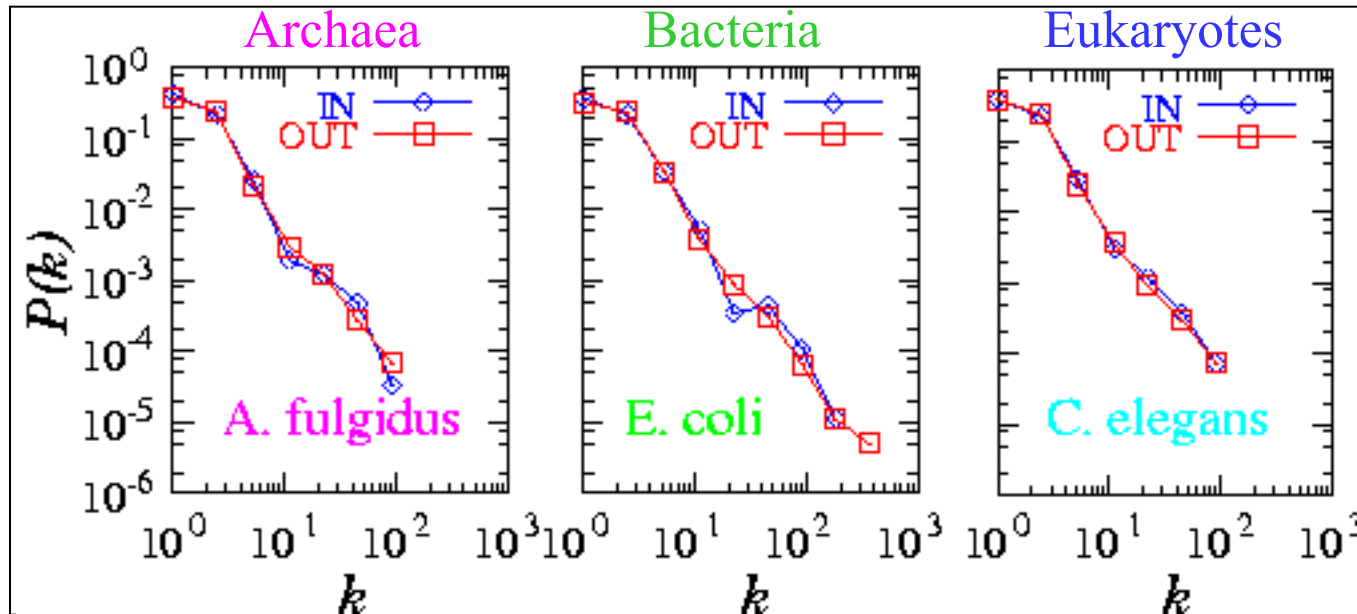
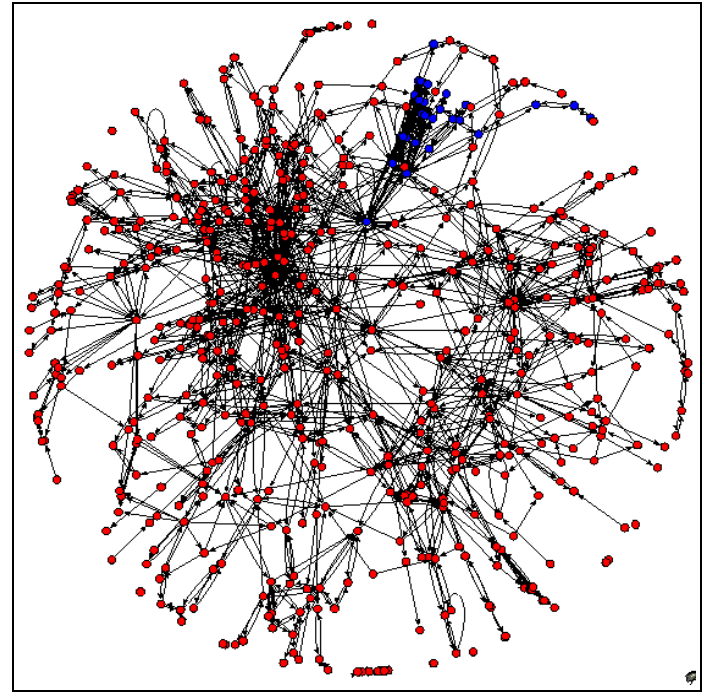
# Metabolic Networks



# Metabolic Networks:

**Nodes:** chemicals (substrates)

**Links:** chem. reaction



100+  
organisms, all  
domains of life  
are **scale-free**  
networks.

H. Jeong, B. Tombor, R. Albert, Z.N. Oltvai, and A.L. Barabasi, *Nature* **407**, 651 (2000).