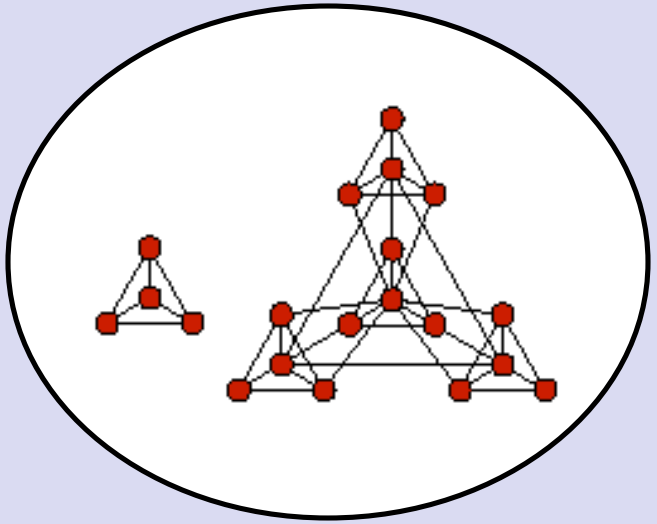
Hierarchical Networks

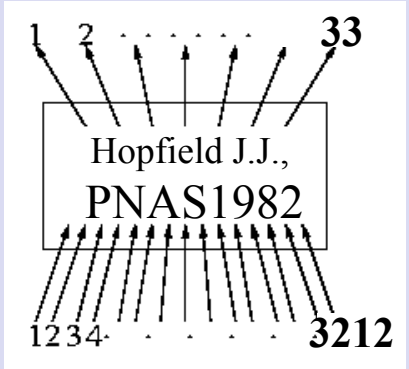
Erzsébet Ravasz
Advisor: **Albert-László Barabási**

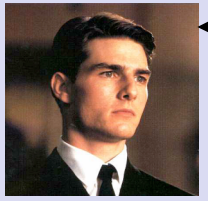


- **Introduction to networks**
- **How to model complex networks?**
- **Clustering and hierarchy**
- **Hierarchical organization of cellular metabolism**
- **The *E. Coli* genetic network**

Introduction to Networks

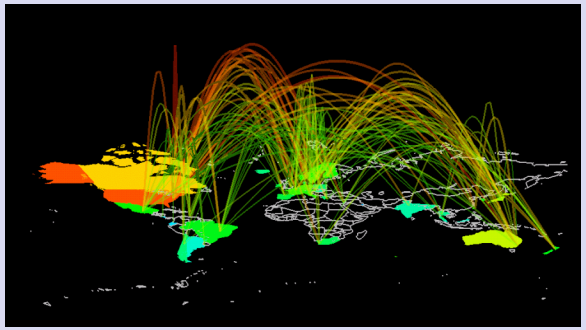
- **Networks in society**
  - Co-authorship
  - Scientific citations
  - Friendships
  - Sexual contacts
  - Movie actor network
  - Business relationships

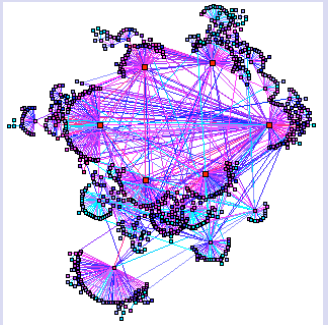




Days of Thunder (1990)
Far and Away (1992)
Eyes Wide Shut (1999)

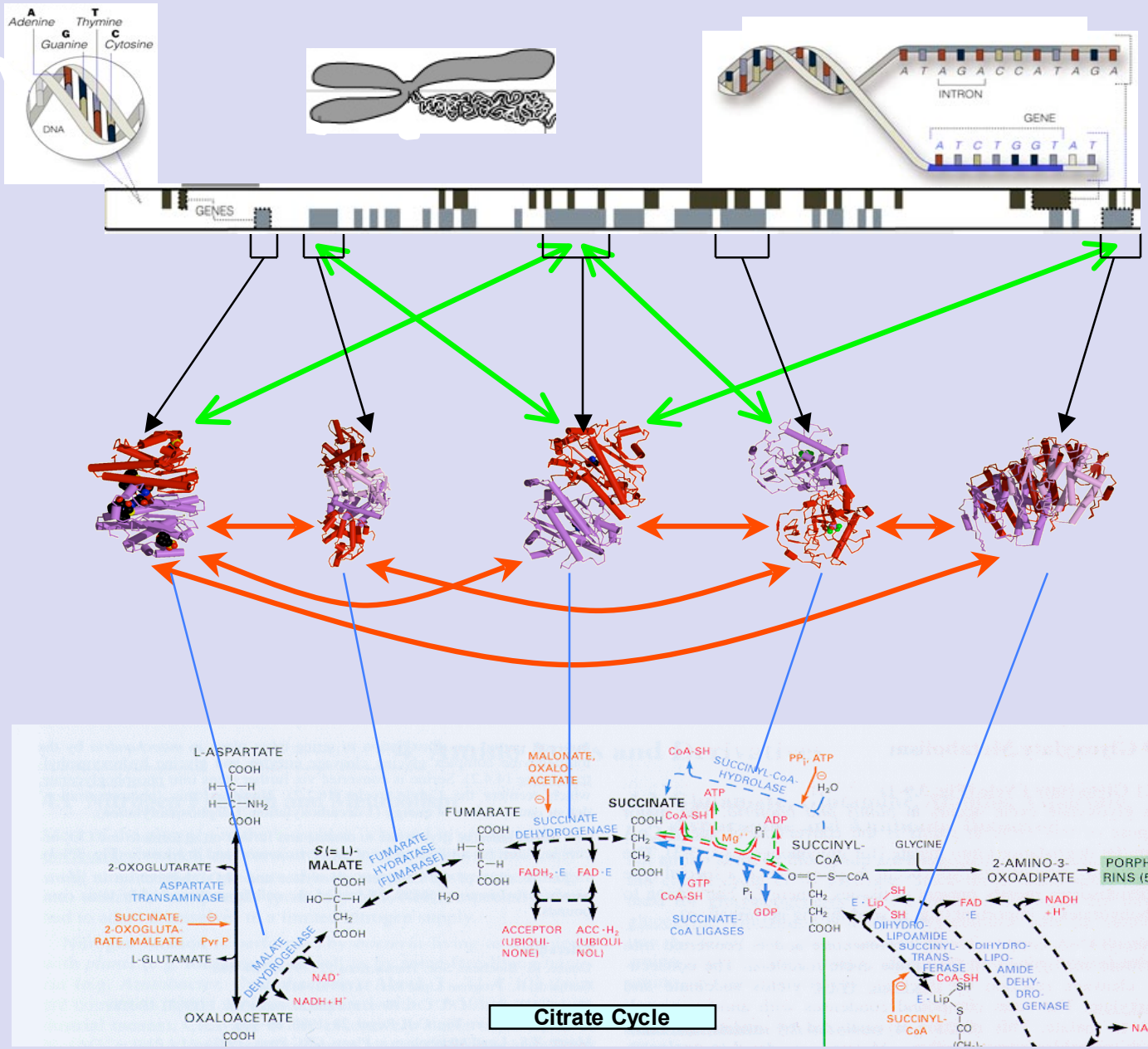
- **Communication networks**
  - Internet
  - World Wide Web





- **And more:**
  - Power grid
  - Airline routes
  - Neuron networks
  - Words linked by synonyms
  - ...

- **Networks in the cell**



GENOME

protein-gene interactions

PROTEOME

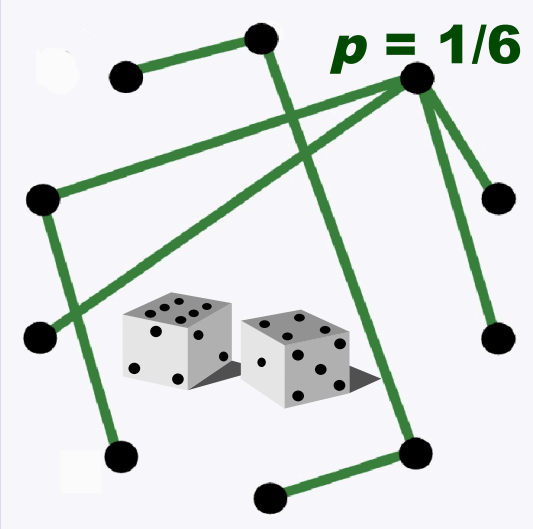
protein-protein interactions

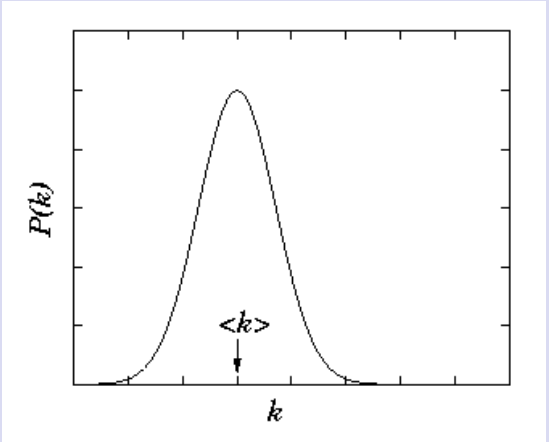
METABOLISM

Bio-chemical reactions

How to model complex networks?

- **The Erdős-Rényi model (1960)**
  - Start with *N* nodes
  - Connect two nodes with probability *p*





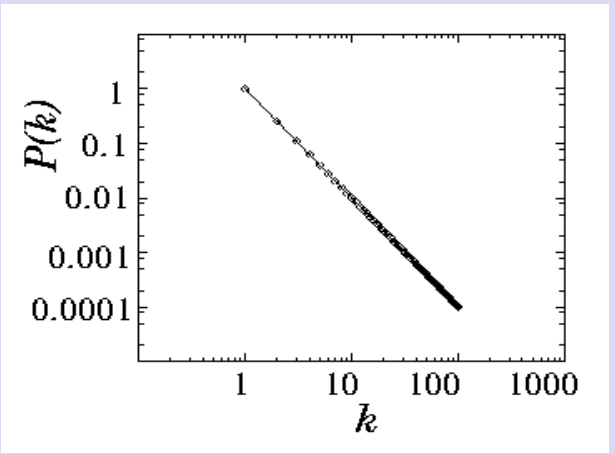
$$P(k) \simeq e^{-pN} \frac{(pN)^k}{k!}$$

  - democratic
  - homogeneous in connectivity

- **Does it work?**
  - No, but ...

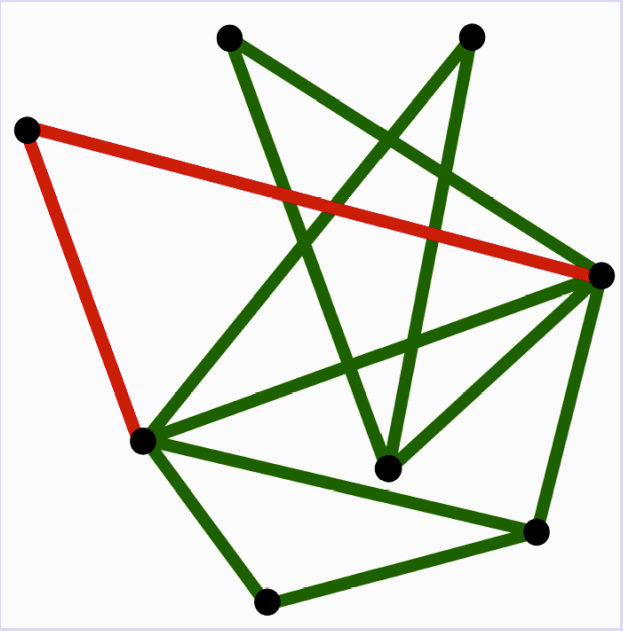
- Co-authorship
- Scientific citations
- Sexual contacts
- Movie actor network

- Internet backbone
- World Wide Web
- Words linked by synonyms
- protein-protein interactions
- Bio-chemical reactions



- Power law connectivity distributions!

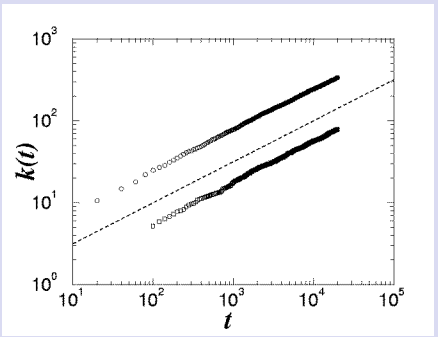
- **The scale-free model**
  - Real networks grow continuously.
  - Highly connected nodes are more likely to receive new links.

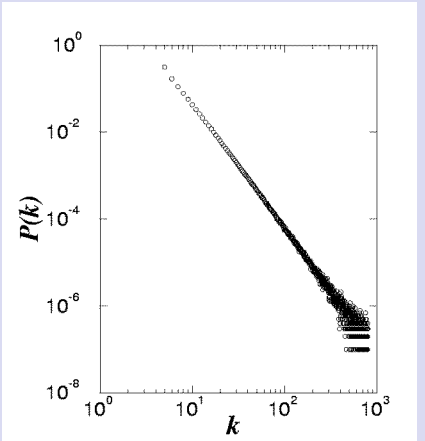


- GROWTH
  - At each time-step a new node is added with *m* links.
- PREFERENTIAL ATTACHMENT
  - The probability that an existing node receives a link from the new one is proportional to it's connectivity.

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j}$$

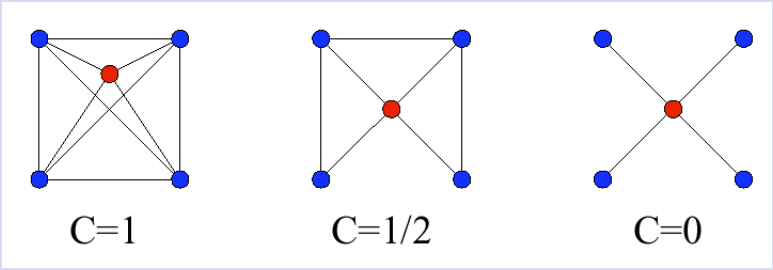
- Mean field solution
  - $k_i(t) = m\left(\frac{t}{t_i}\right)^{0.5}$
  - $P(k) = \frac{2m^2}{k^3}$



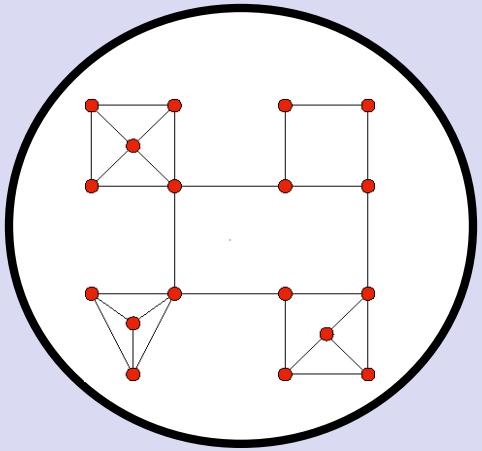


Clustering and hierarchy

- **Clustering coefficient**
  - Defined as the probability that a node's neighbors are linked.
  - Scale-free model $C \sim N^{-0.75}$
  - Real networks have high average clustering coefficients!



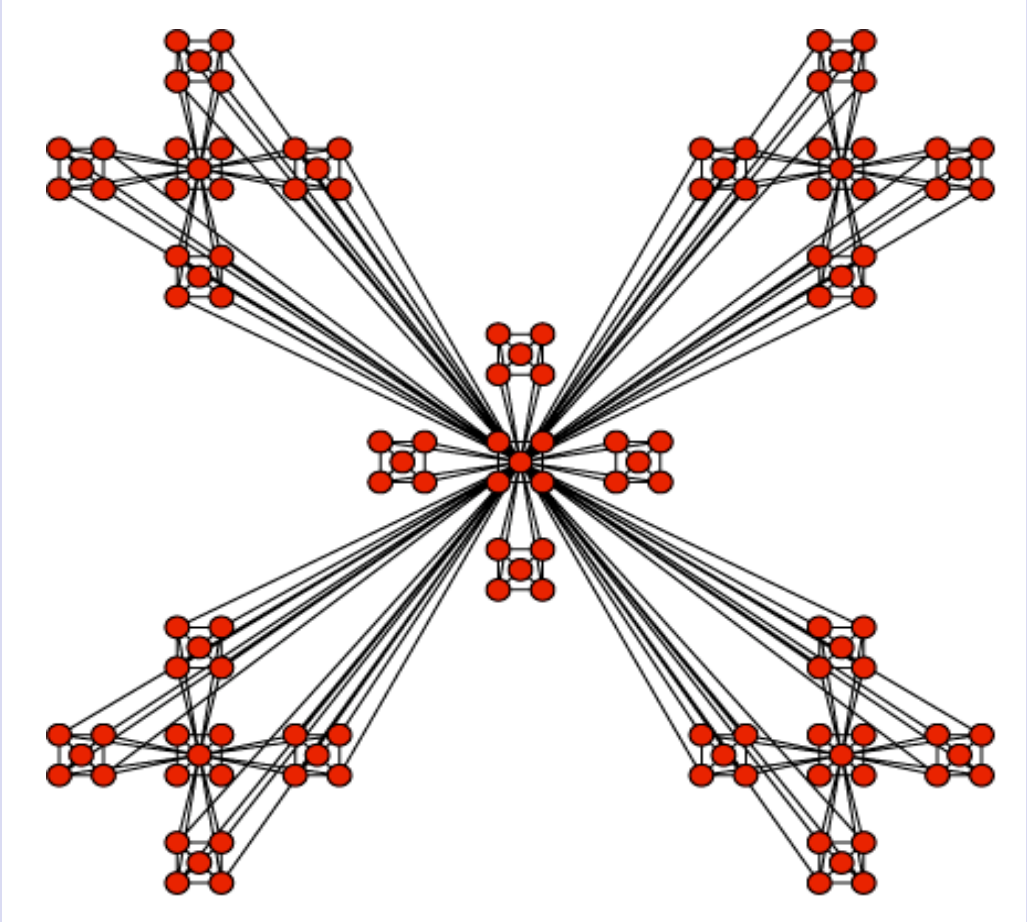
- **Why?**
  - *Modular* organization



- Protein complexes
- WWW communities
- Scientific groups
- Research areas
- Circles of friends

- **Modular and scale-free?**
  - ❖ Inspired by the *E. Coli* metabolic network
  - Hierarchical modularity

- **Hierarchical scale-free model**



- **Power law degree distribution**

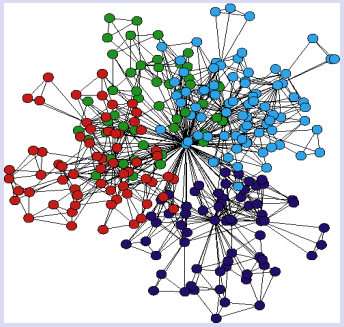
$$\gamma = 1 + \ln 5 / \ln 4$$

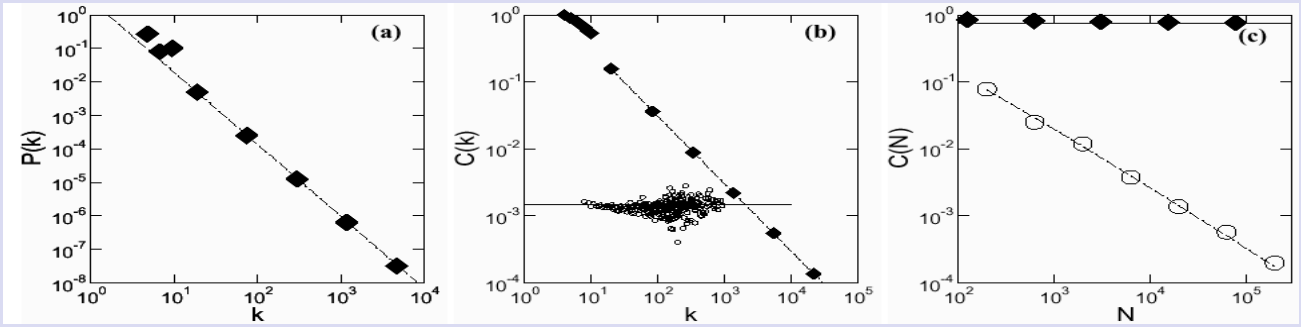
- **Power law scaling of *C(k)***

$$C(k) \sim k^{-1}$$

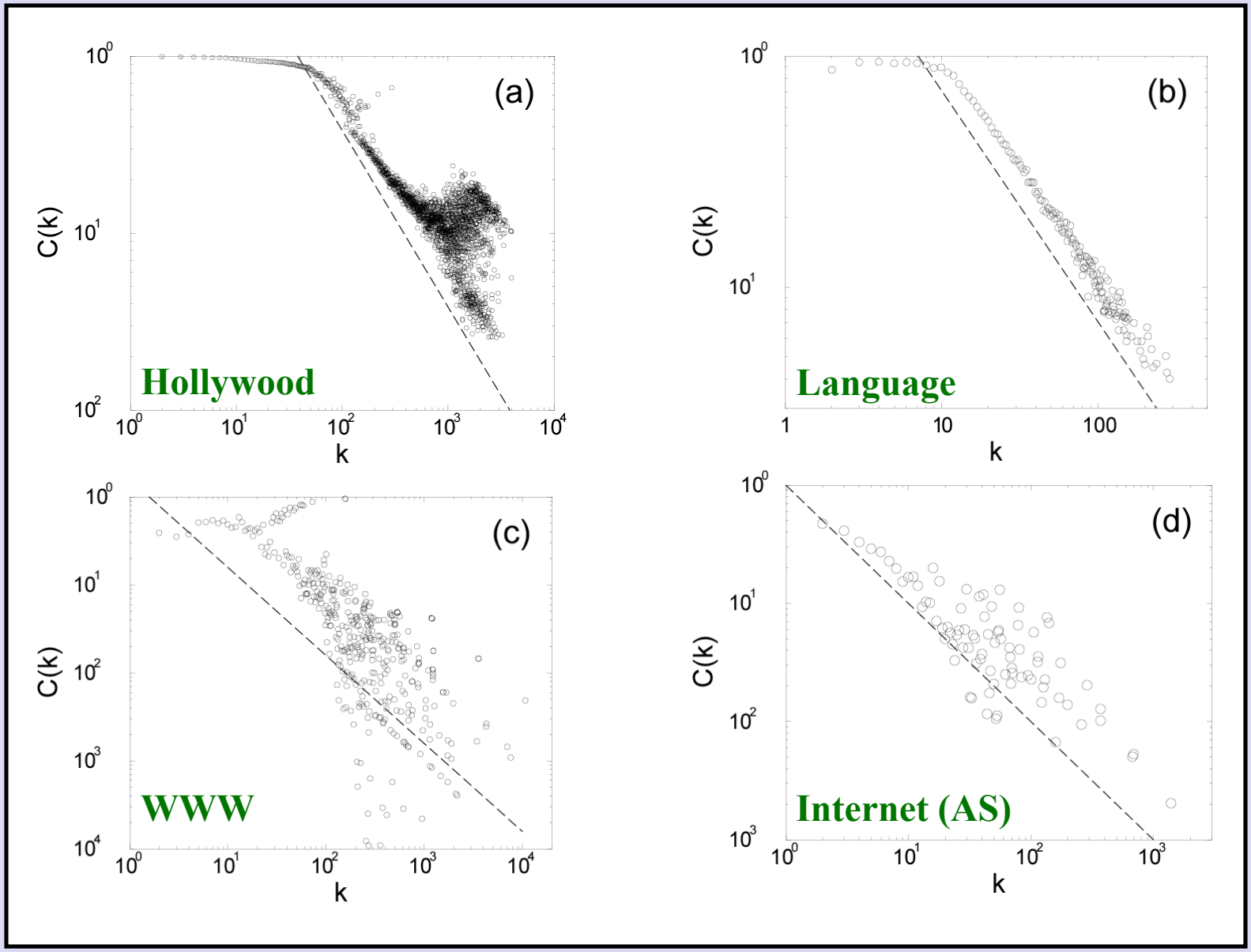
- **Size independent average clustering**

$$C \simeq 0.743$$



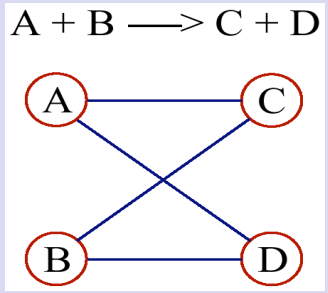


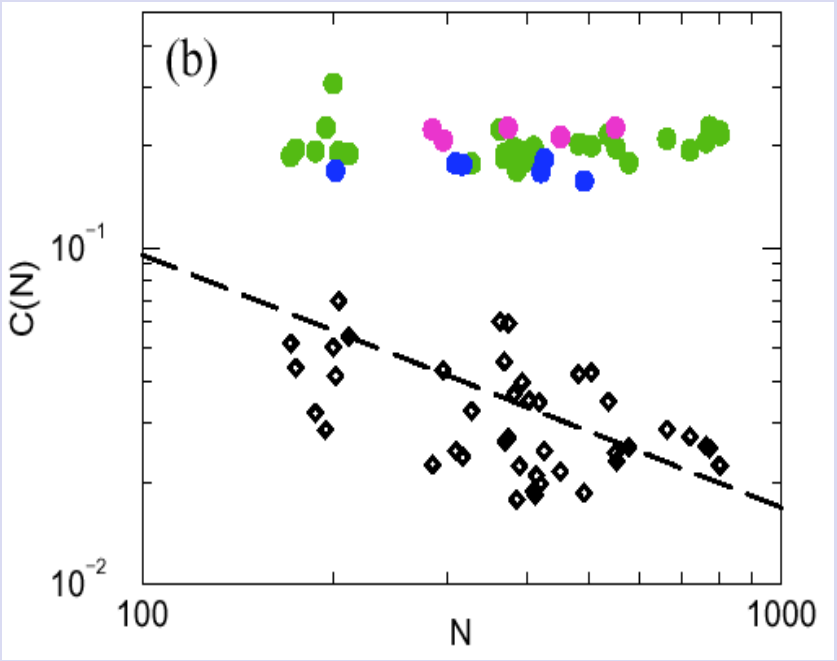
- **Hierarchy captured in real networks**

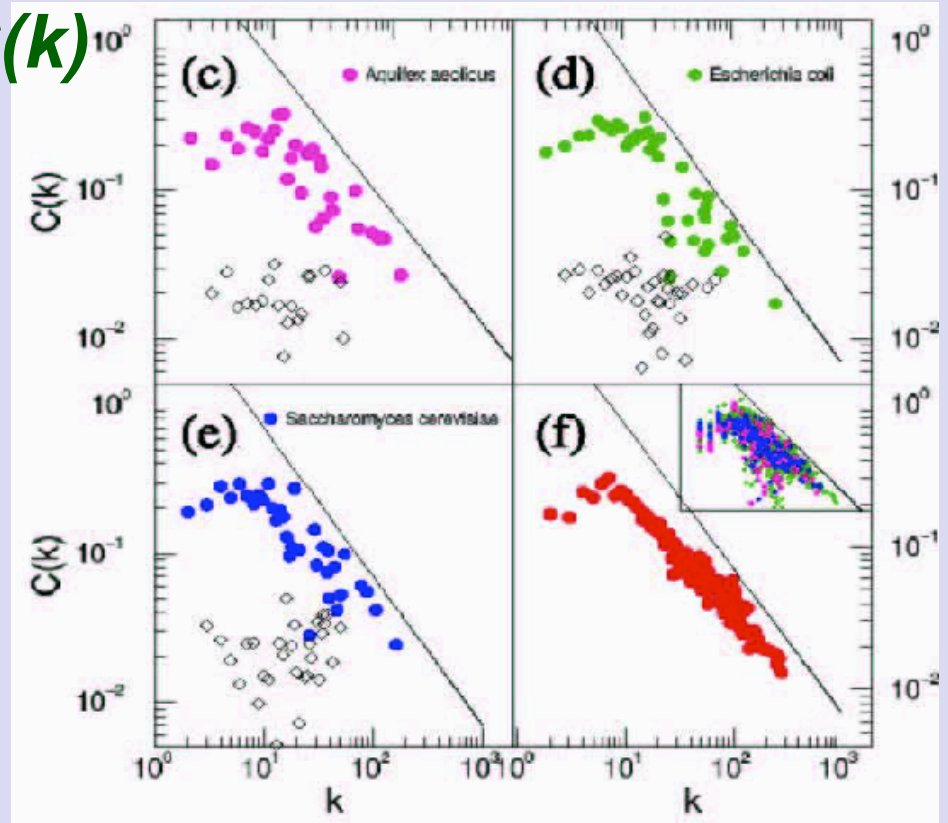


Hierarchical organization of cellular metabolism

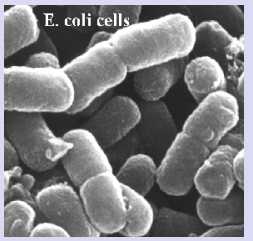
- Graph representation of metabolism
- Clustering coefficient is size independent
- Power-law scaling of *C(k)*



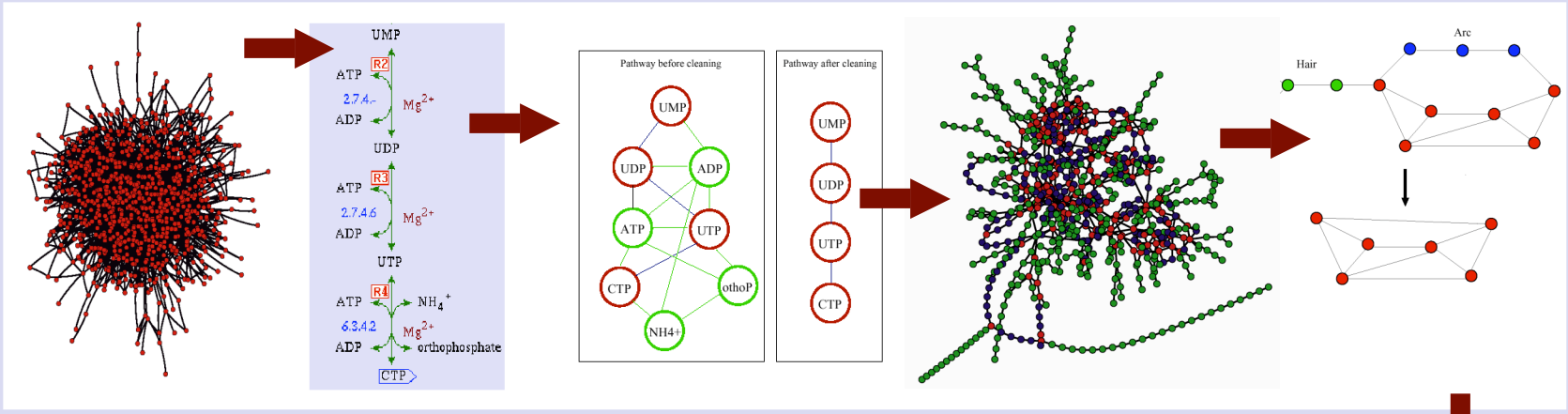




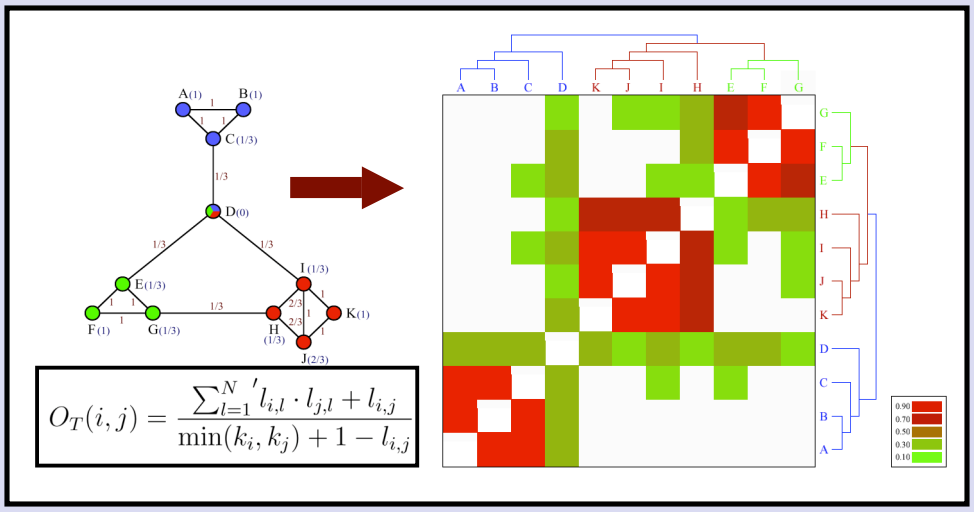
• Modules in the *E. Coli* metabolic network



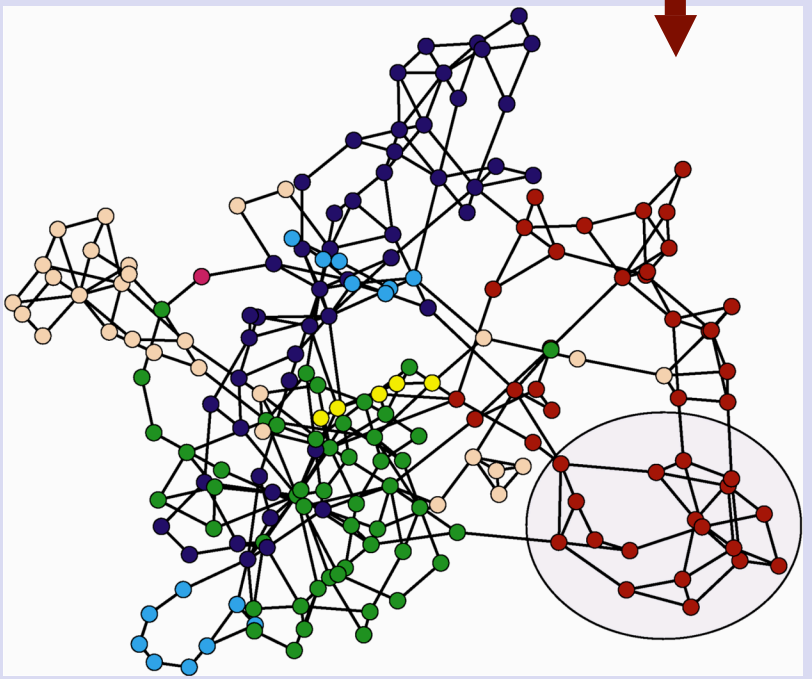
➢ Reduced graph representation



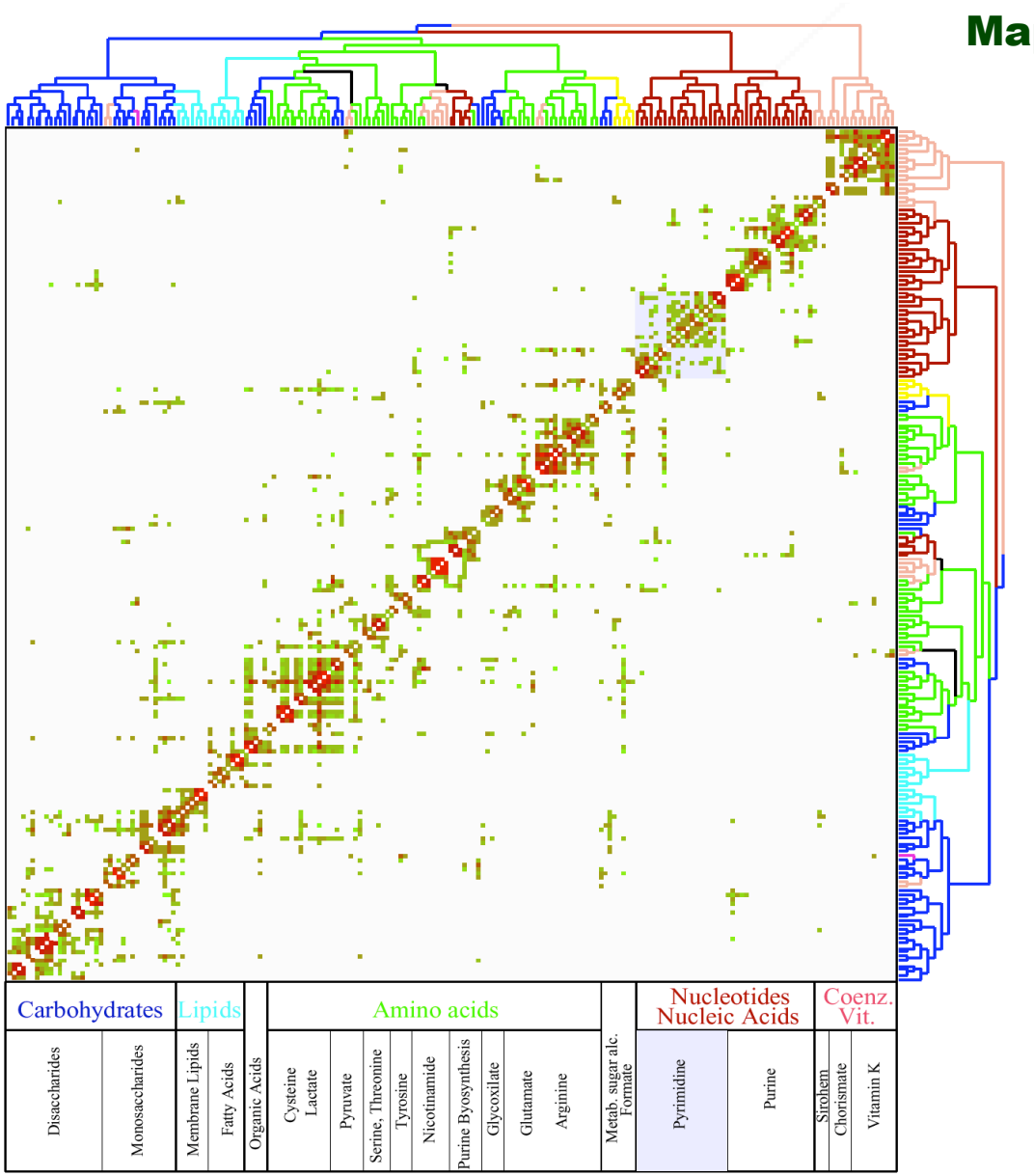
➢ Clustering



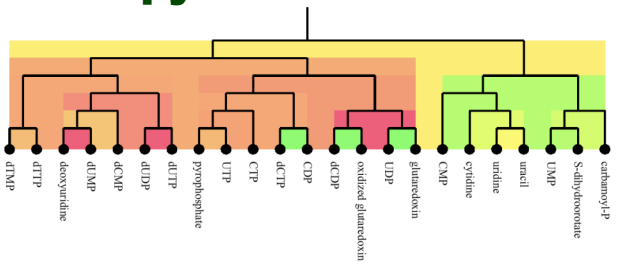
$$O_T(i,j) = \frac{\sum_{l=1}^{N} {}' l_{i,l} \cdot l_{j,l} + l_{i,j}}{\min(k_i, k_j) + 1 - l_{i,j}}$$

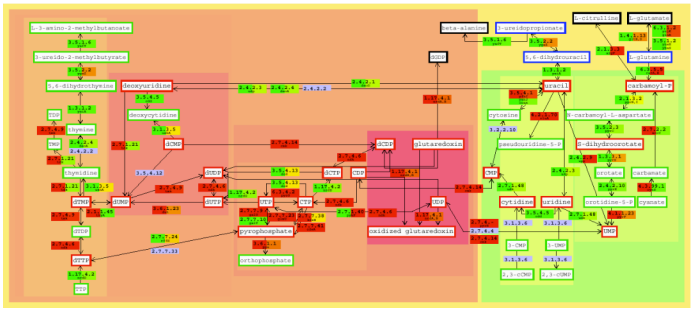


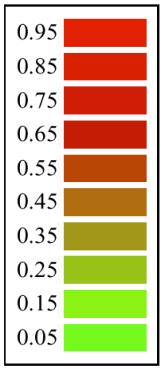
➢ Clusters and their biological function



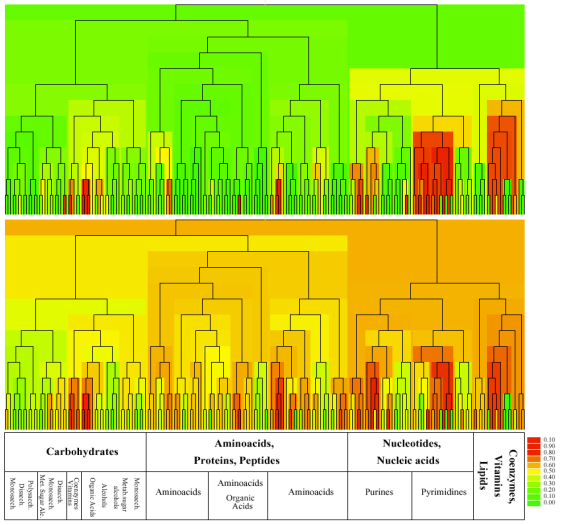
Map of the pyrimidine metabolism







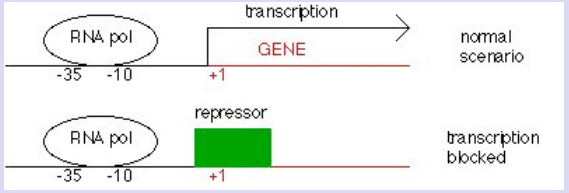
Cluster lethality

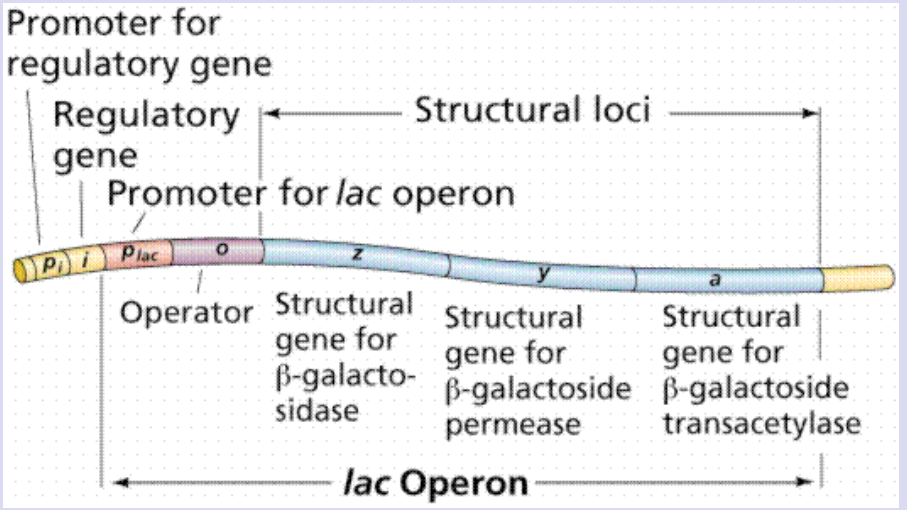


The *E. Coli* genetic network

- **Transcription and it's regulation**
  - ❖ Promoter regions and operons
  - ❖ Transcriptional regulation
    - ▶ Activation

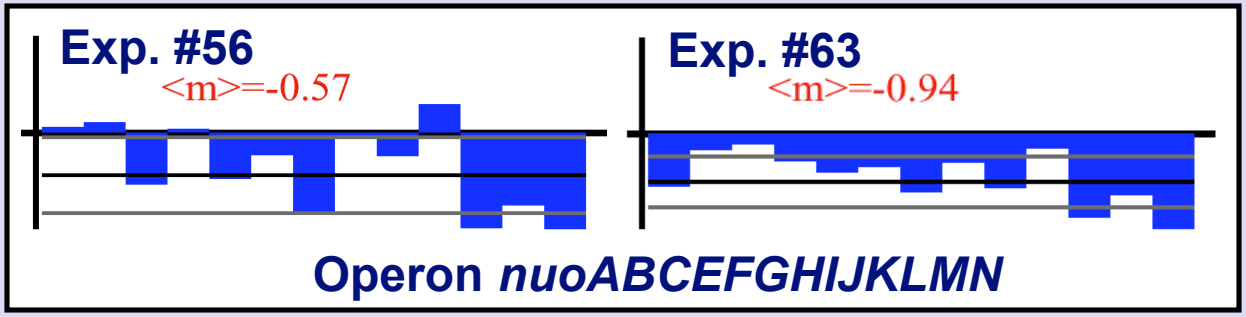
    - ▶ Repression



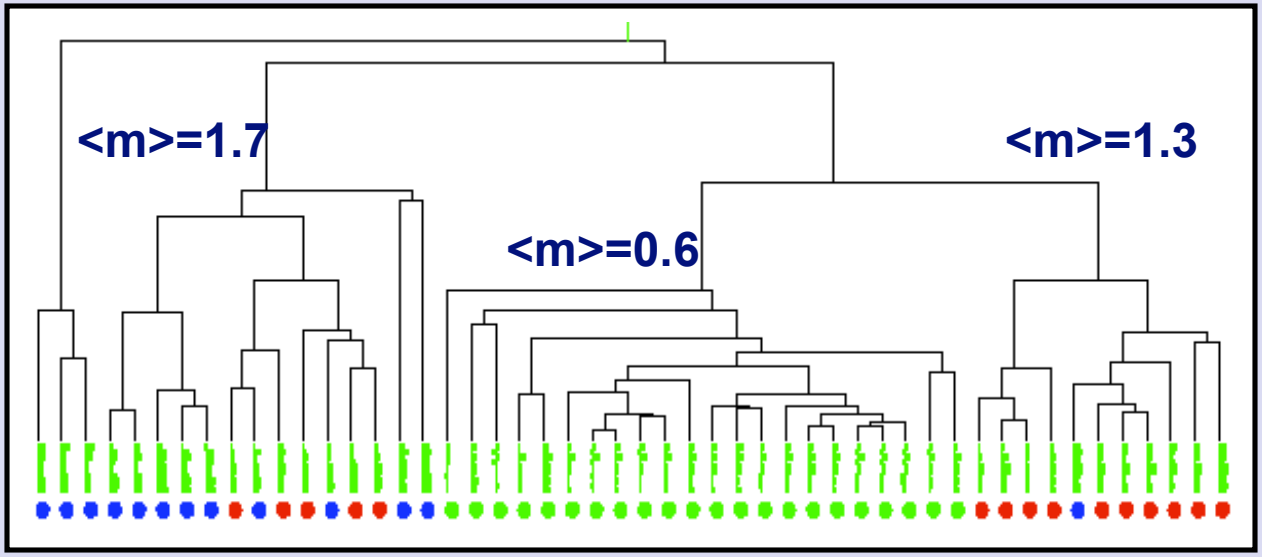


- **Available data on *E. Coli***
  - ❖ Gene content and activation-repression interactions for 400 operons
  - ❖ mRNA concentration of all E. Coli genes in 50+17 experimental conditions

- **Operon activation states:**
  - ❖ **Observation:**
    - ▶ mRNA levels of genes in an operon can be very different
    - ▶ Relative levels are very similar across some experiments

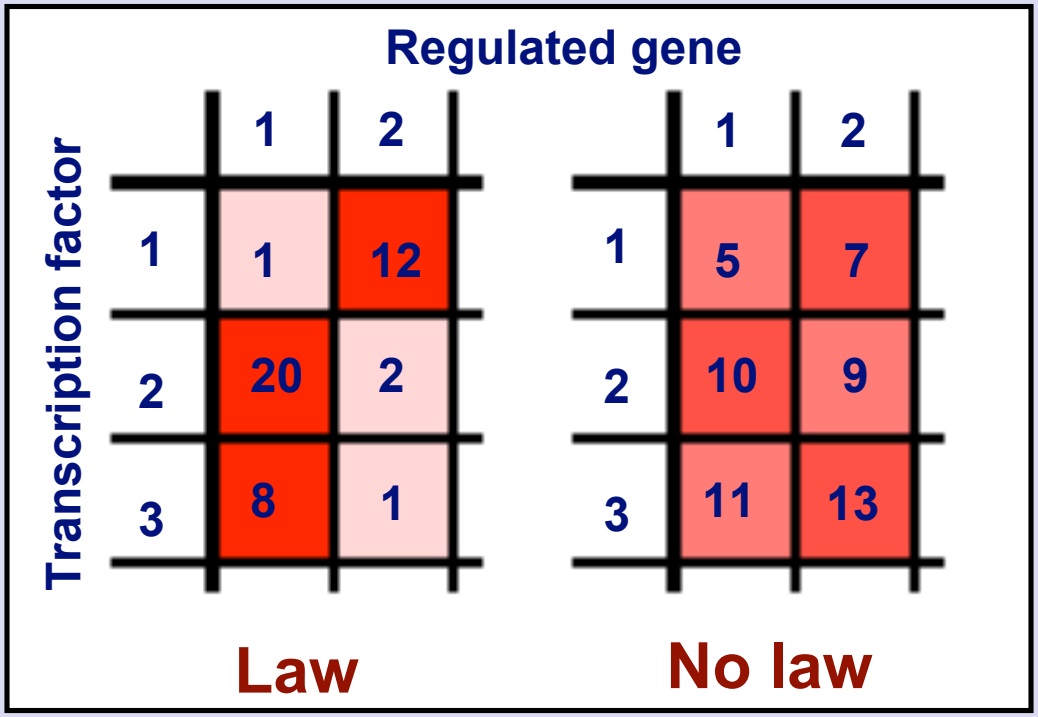


- ▶ clustering based on profile similarity leads to 2-3 clusters



- ▶ clusters have different average expression levels!

- **Next step:**
  - ❖ Only a few “states” of activation?
  - ❖ Proposed test:



  - ❖ Modeling

Publications

— ***"Deterministic scale-free networks",*** **A.-L. Barabási, E. Ravasz and T. Vicsek,** ***Physica A*** **299, 559 (2001).**

— ***"Evolution of the social network of scientific collaborations",*** **A.L. Barabási, H. Jeong, Z. Néda, E. Ravasz, A. Schubert, T. Vicsek,** ***Physica A*** **311, 590 (2002).**

— ***"Hierarchical Organization of Modularity in Metabolic Networks",*** **E. Ravasz, A. L. Somera, D. A. Mongru, Z. N. Oltvai and A.L. Barabási,** ***Science*** **297,1551 (2002).**

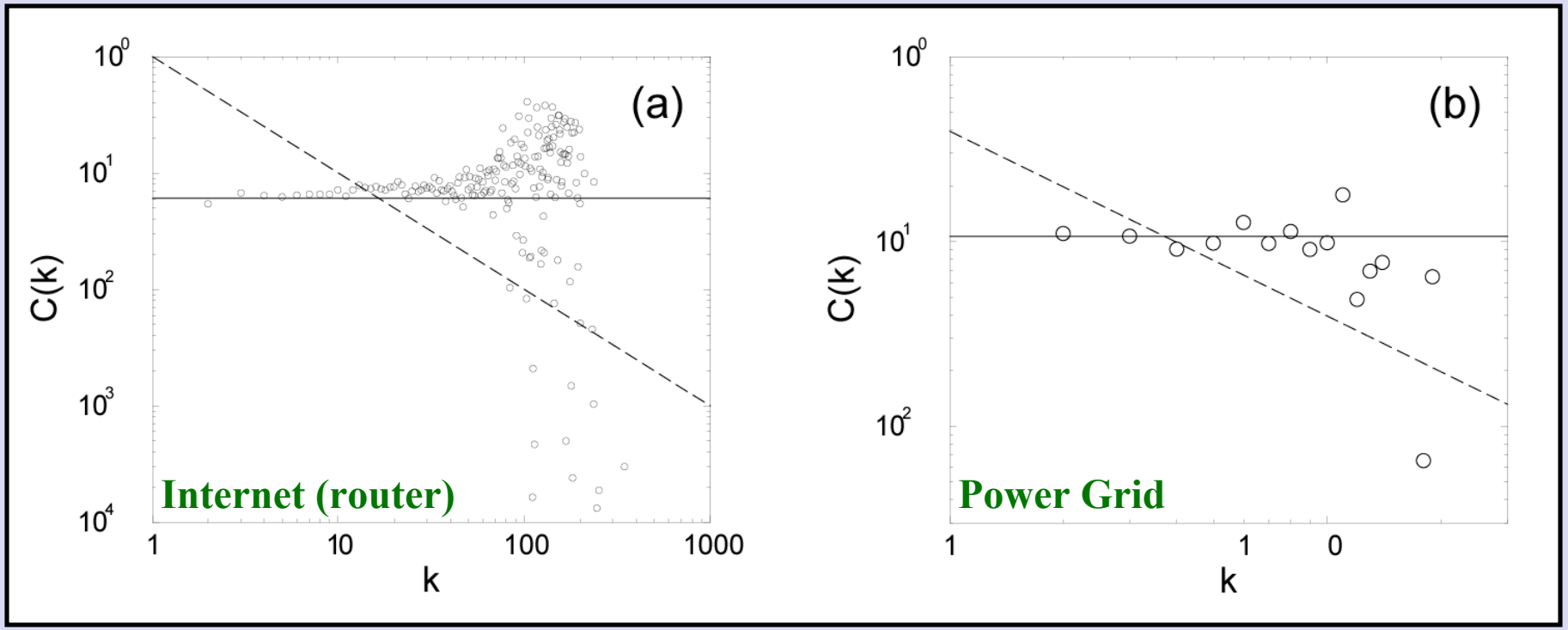
— **"*Hierarchical organization in complex networks",*** **E. Ravasz and A.-L. Barabási,** ***Phys. Rev. E*** **67, 026112 (2003).**

— ***"Networks in life: Scaling properties and eigenvalue spectra",*** **I. Farkas, I. Derenyi, H. Jeong, Z. Neda, Z. N. Oltvai, E. Ravasz, A. Schubert, A.-L. Barabasi and T. Vicsek,** ***Physica A*** **314, 25 (2002).**

— ***"Scale-free and hierarchical structures in complex networks",*** **A.-L. Barabási, Z. Dezső, E. Ravasz, S. H. Yook, and Z. Oltvai,** ***Sitges Proceedings on Complex Networks*****, (2004).**

—***"Experimental Determination and System-Level Analysis of Essential Genes in E. coli MG1655",*** **S.Y. Gerdes et al, submitted.**

- **Some networks are not hierarchical**



➢ **Common features**

- ❖ geographical localization
- ❖ economic pressure to minimize link lengths