Large-scale organization of metabolic networks
Jeong et al.

CS 466

Saurabh Sinha

Complex network in cell

- Cellular processes generating mass & energy, transfering information, specifying cell fate
- Integrated through complex network of several constituents and reactions
- Can we look at such networks and learn something about biology and evolution?

Key findings

- “Metabolic” networks analyzed
- Compared among 43 species from all three domains of life (archaea, bacteria, eukarya)
- Noticed the same topologic scaling properties across all species
- Metabolic organization identical for living organisms, and is “robust”
- May extrapolate to other cellular networks

Introduction

- Fundamental design principles of cellular networks?
- Example: Dynamic interactions of various constituents impart “robustness” to cellular processes
  - If one reaction did not happen optimally, does not necessarily mess up the whole process

Introduction

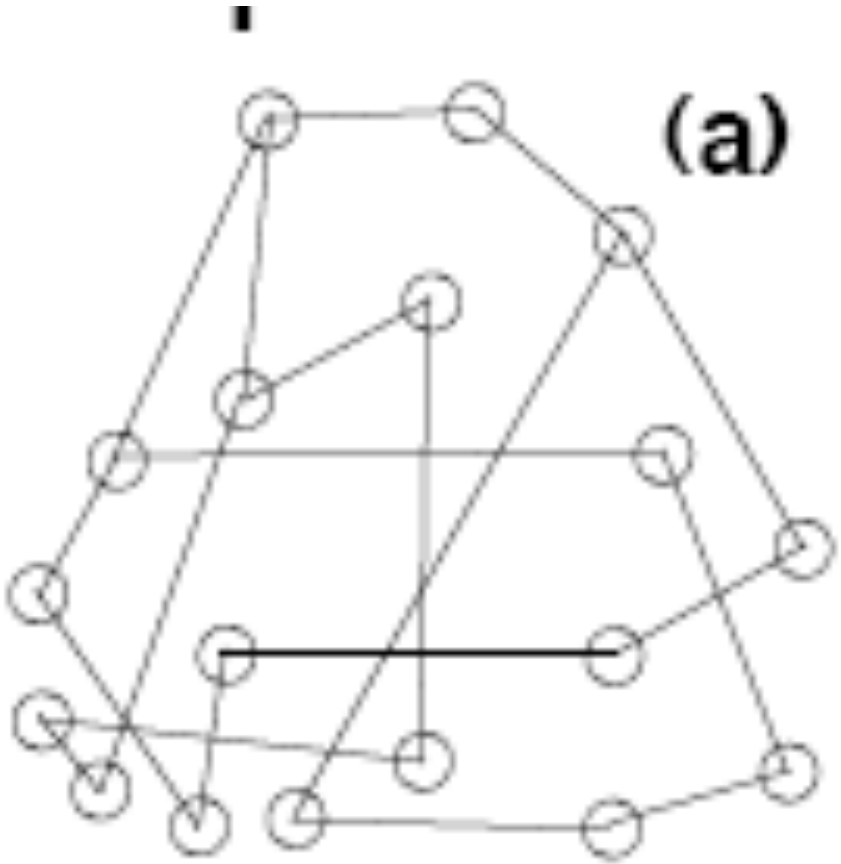
- Constituents of networks: DNA, RNA, proteins, small molecules
- High throughput biology has helped develop databases of networks, e.g., metabolic networks
- Maps are extremely complex
- Fundamental features of network topology?

Network models

Erdos-Renyi random graph

- Start with a fixed number of nodes, no edges
- Each pair of nodes connected by an edge with probability p
- This leads to a “statistically homogeneous” graph or network
- Most nodes have same degree

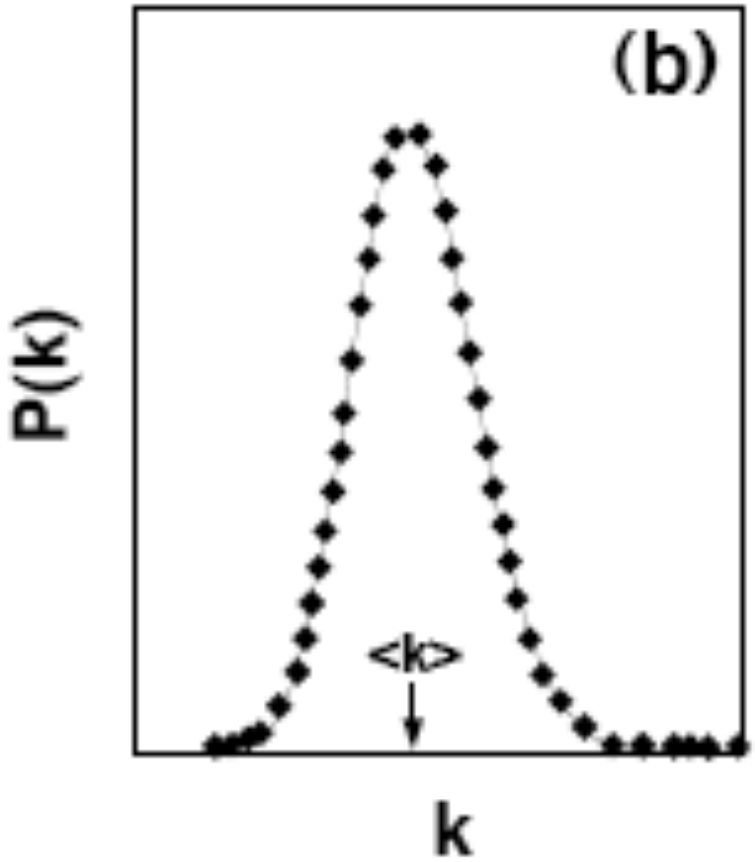
http://arxiv.org/pdf/cond-mat/0010278



Erdos-Renyi random graph

- Degree distribution is Poisson with strong peak at mean <k>
- Therefore, probability of finding a highly connected node decays exponentially

http://arxiv.org/pdf/cond-mat/0010278



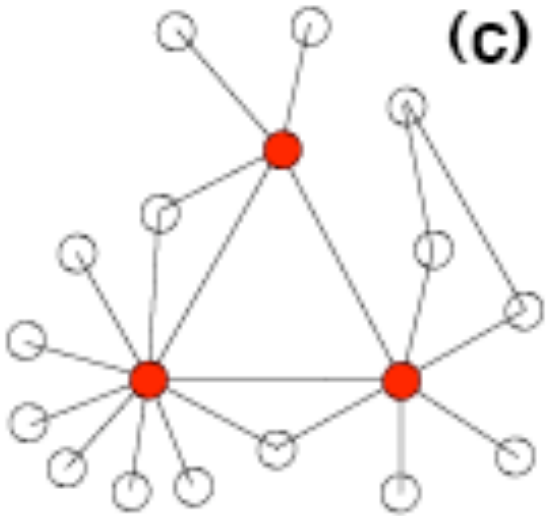
Empirical graphs

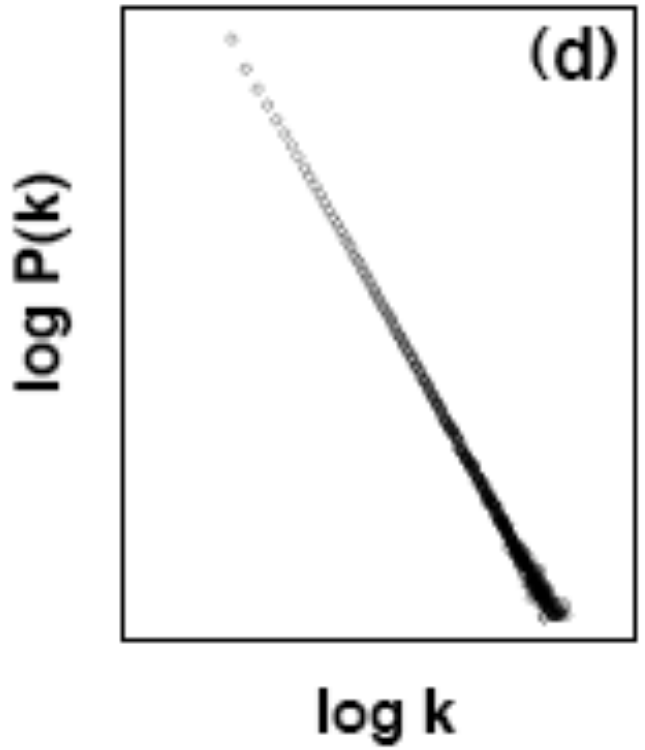
- World-wide web, internet, social networks have been studied
- Serious deviations from random structure of E-R model
- Better described by “scale-free” networks

Scale-free networks

- Degree distribution P(k) follows power law distribution
  - $P(k) \sim k^{-x}$
- Scale-free networks are extremely heterogeneous:
  - a few highly connected nodes (hubs)
  - rest of the (less connected) nodes connect to hubs
- Generated by a process where new nodes are preferentially attached to already high-degree nodes

http://arxiv.org/pdf/cond-mat/0010278





What does this tell us?

- Difference between Erdos-Renyi and scale-free graphs arise from simple principles of how the graphs were created
- Therefore, understanding topological properties can tell us how the cellular networks were created

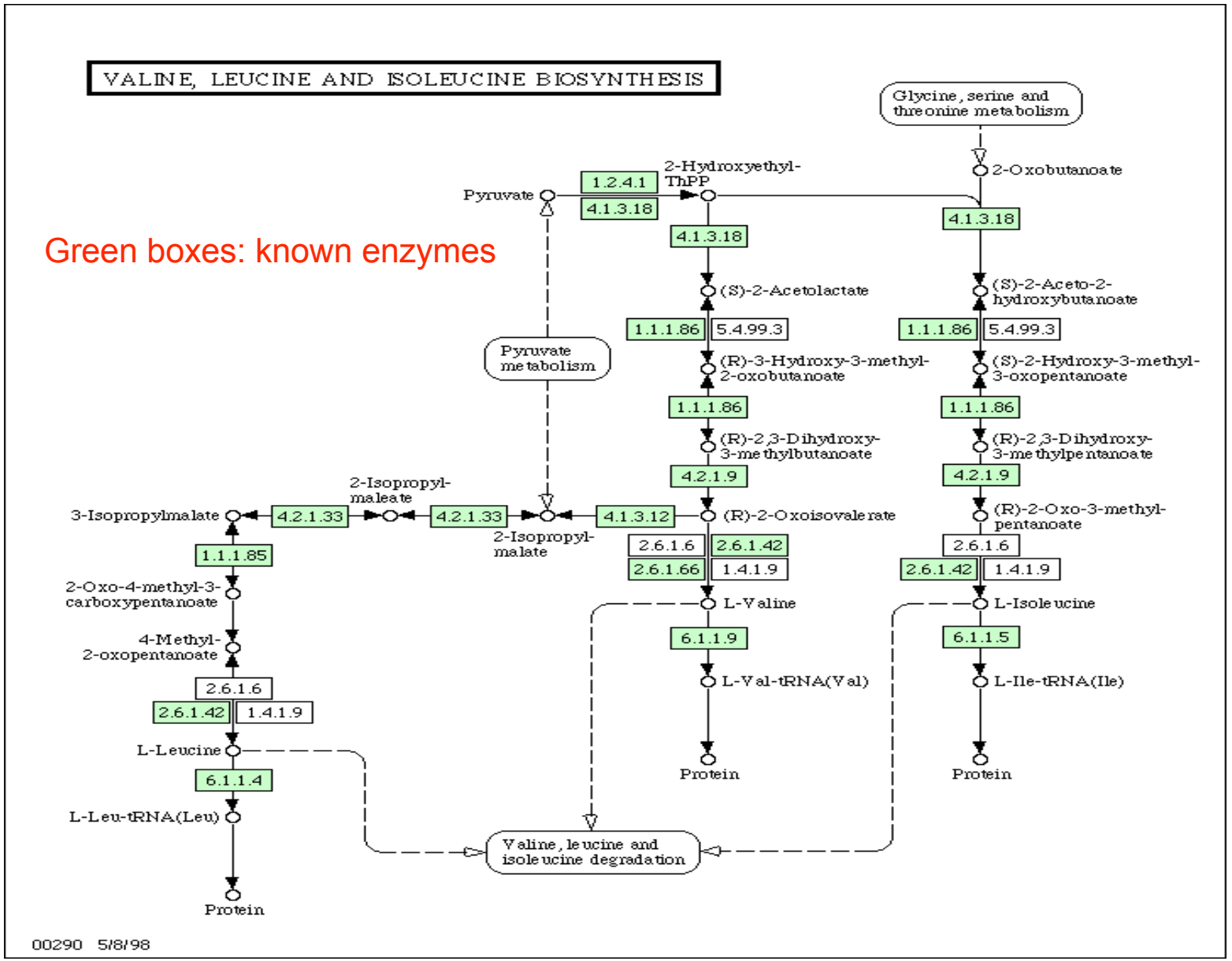
Data

Metabolic networks

- Core metabolic network of 43 different organisms (WIT database)
- 6 archaea, 32 bacteria, 5 eukarya
- Nodes = substrates, edges = metabolic reactions, additional nodes = enzymes
- Based on firmly established data from biochemical literature
- Sufficient data for statistical analysis

Example of a metabolic network

http://www.avatar.se/strbio2001/metabolic/what.html



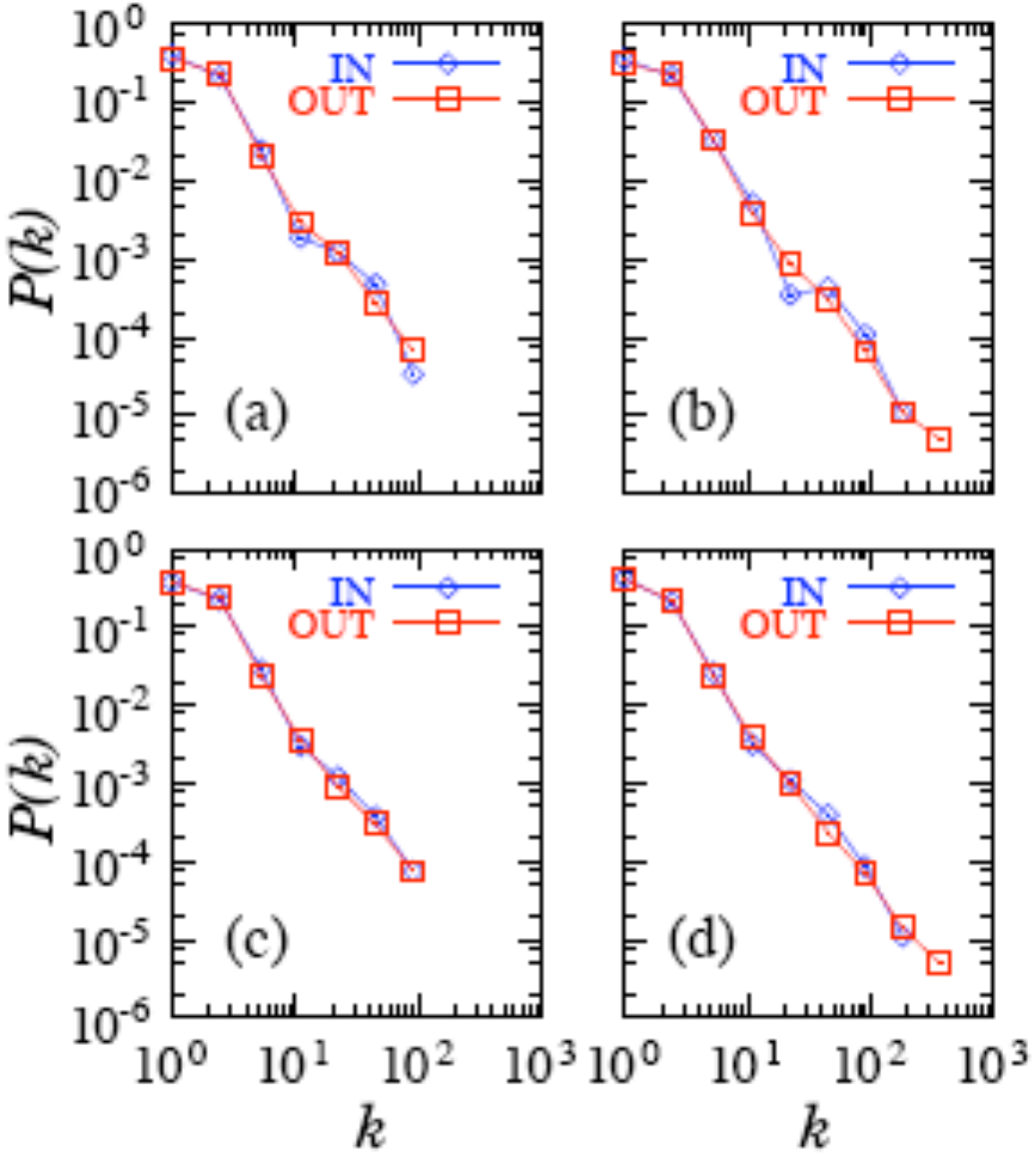
Results

Topology

- Is the topology described by E-R model or by scale-free model?
- Observed that degree distribution follows a power law. Therefore, scale-free networks

http://arxiv.org/pdf/cond-mat/0010278

A. fulgidus



E. coli

C. elegans

Average

Small-world property

- General feature of many complex networks: any two nodes can be connected by relatively short paths
- In metabolic network, a path is the biological “pathway” connecting two substrates
- Characterized by “network diameter”
  - Shortest path, over all pairs of nodes

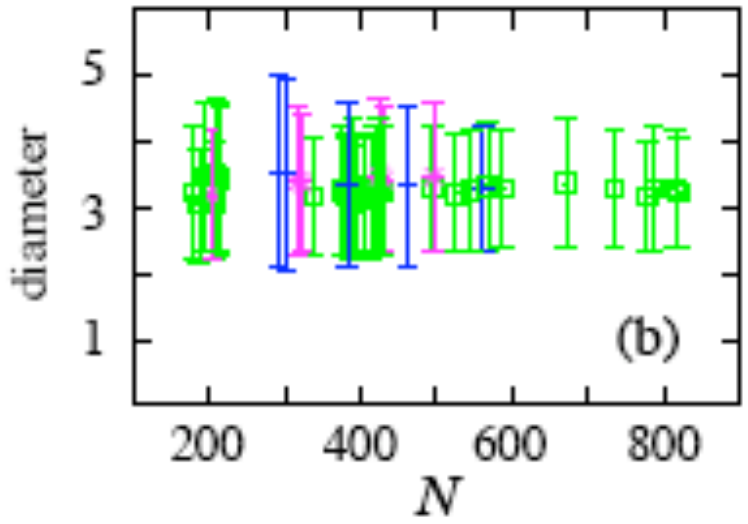
Small-world property

- For non-biological networks, the average degree is usually fixed
- This implies that network diameter increases logarithmically with new nodes being added
- Is this true of metabolic networks ?
- That is, more complex bacterium (more substrates and enzymes) will have larger diameter ?

Small-world property

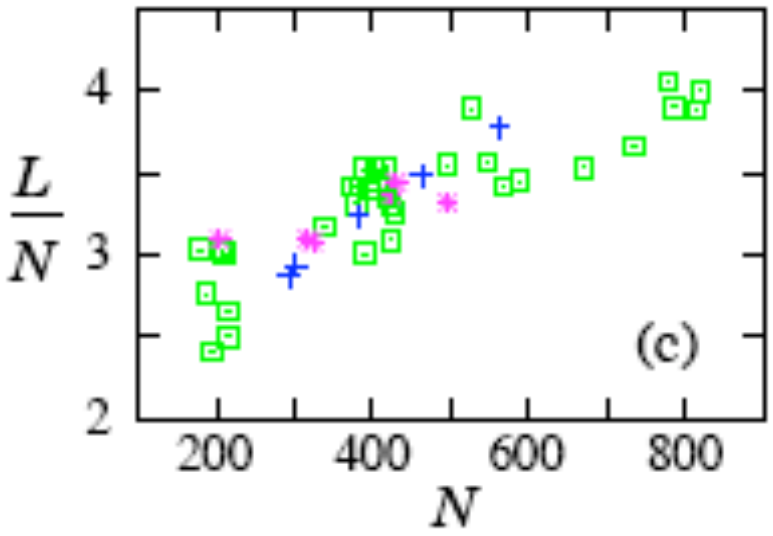
- More complex bacterium (more substrates and enzymes) will have larger diameter ?
- Observed: diameter same across all 43 species !
- A possible explanation: average degree must be higher for more complex organisms
- This is also verified.

http://arxiv.org/pdf/cond-mat/0010278



network diameter for different organisms

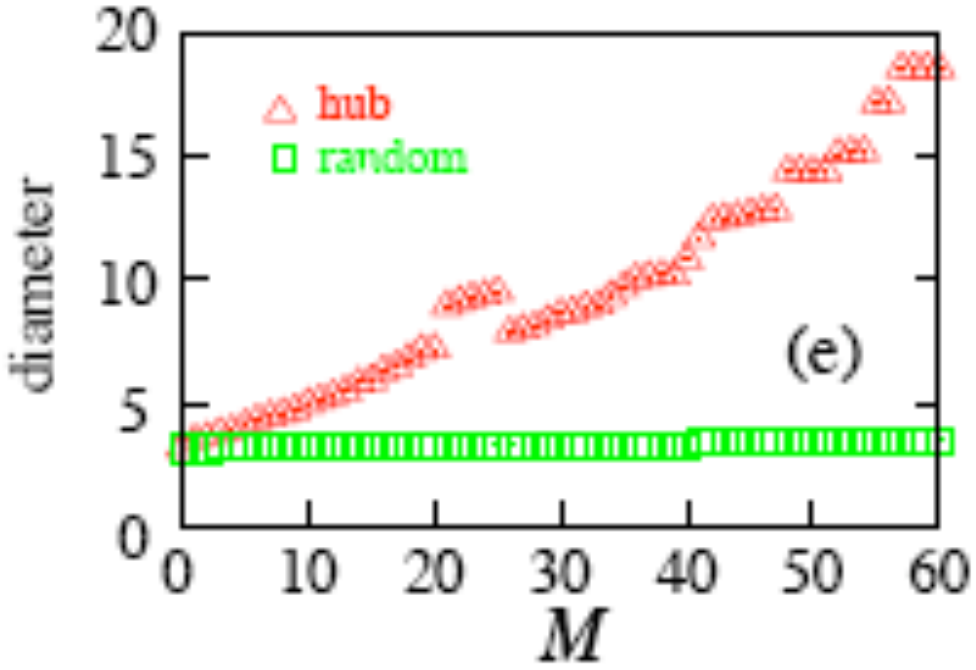
average degree over different organisms



Hubs in network

- Power-law connectivity implies that a few “hub” nodes dominate the overall connectivity
- Sequential removal of hubs => diameter rises sharply
- Observed: metabolic networks show this phenomenon too

http://arxiv.org/pdf/cond-mat/0010278



Diameter after removing M substrates

Hubs in network

- At the same time, scale-free networks are robust to random errors
- In metabolic network, removal of randomly chosen substrates did not affect average distance between remaining nodes
- Fault tolerance to removal of metabolic enzymes also demonstrated through biological experiments

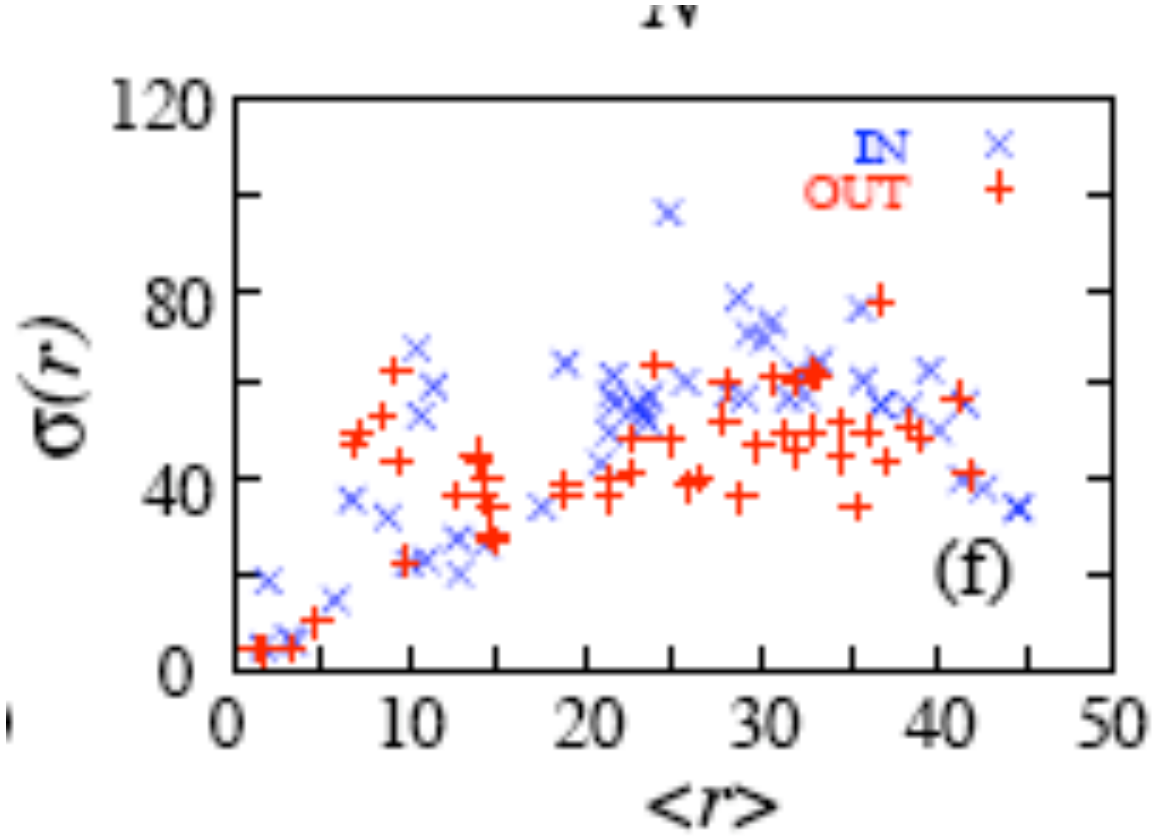
Hubs across networks

- Do the same substrates act as hubs in all organisms?
- Rank all substrates by their degrees
- Ranking of the top substrates is practically same across all species
- For every substrate present in all species, compute rank “r” (in terms of degree) in each species

Hubs across networks

- Compute mean <r> and standard deviation σ_r of rank of each substrate
- Observed: σ_r increases with <r>
- The top-ranking nodes (hubs) have relatively little variance across species

http://arxiv.org/pdf/cond-mat/0010278



Summary

- Other biological networks also hypothesized to be scale-free.
- Evolutionary selection of a robust and error tolerant architecture