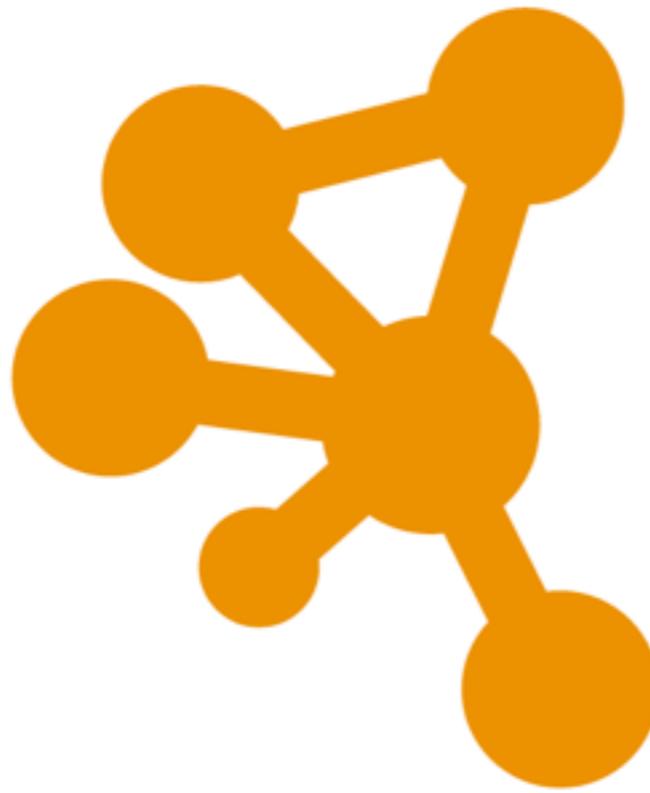
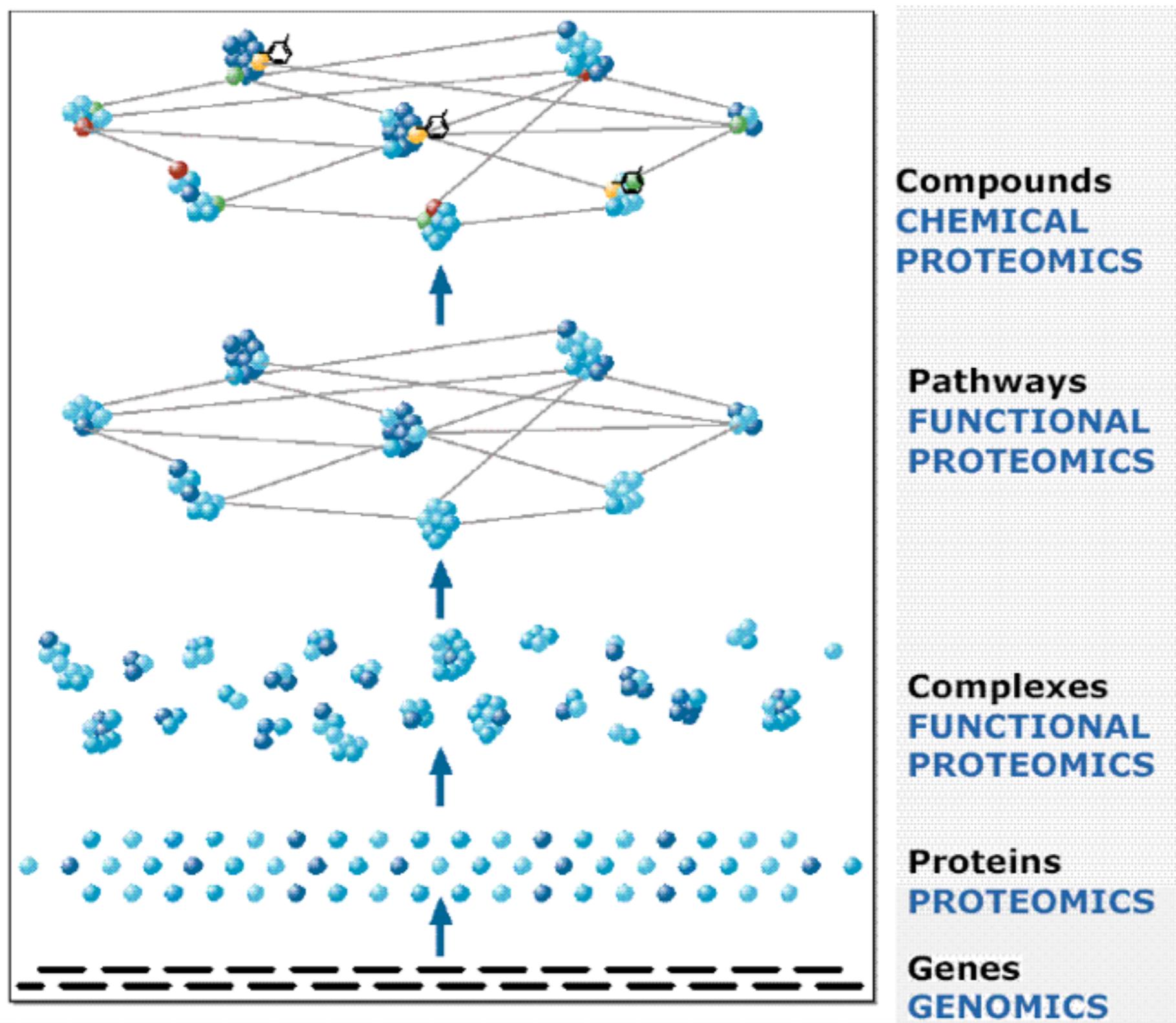


Biological Networks

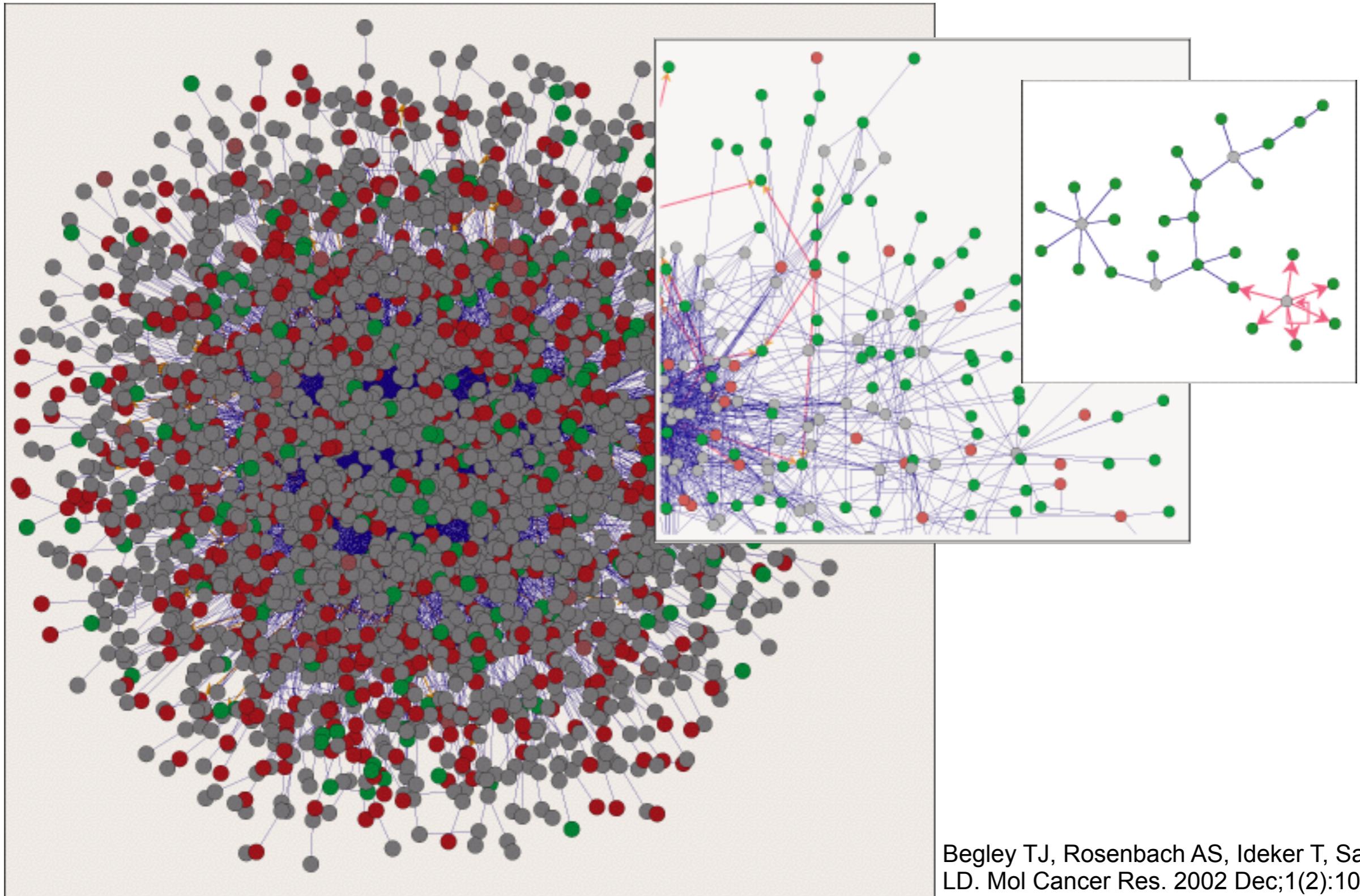
Helen Cook
March 5, 2016



Levels of information



Complexity



Begley TJ, Rosenbach AS, Ideker T, Samson LD. Mol Cancer Res. 2002 Dec;1(2):103-12.

Metabolic networks

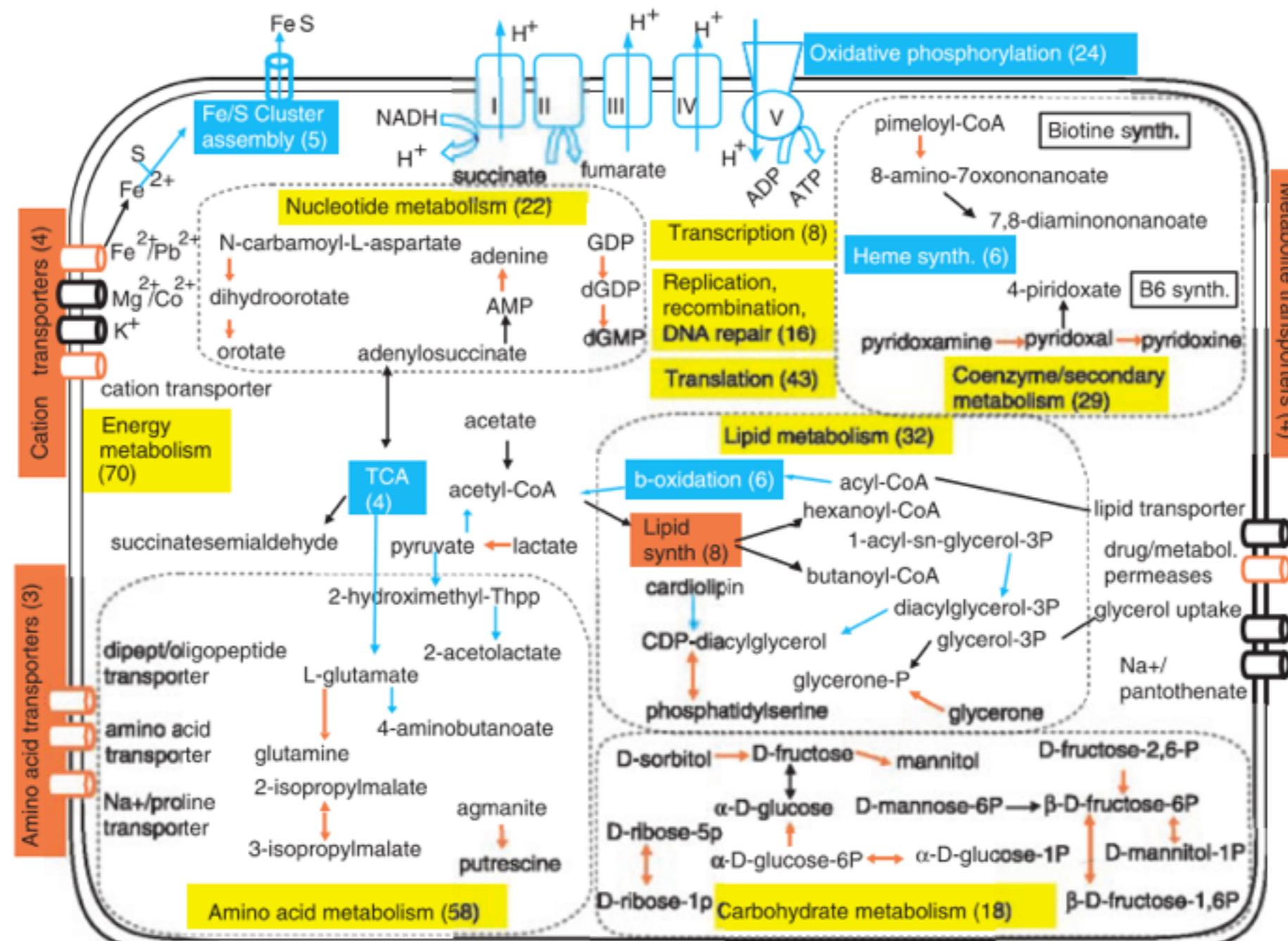


Fig. 1. An overview of metabolism and transport in the proto-mitochondrion, as deduced from the orthologous groups present in its estimated proteome. Yellow boxes indicate the number of groups in that COG functional class. Boxes, arrows, and cylinders indicate pathways, enzymes, and transporters, respectively. Blue: proteins are mitochondrial in yeast or human. Orange: human and yeast orthologs have not been observed in mitochondria. Black: there is no human or yeast representative of this orthologous group.

Regulation of transcription factors in *E. coli*

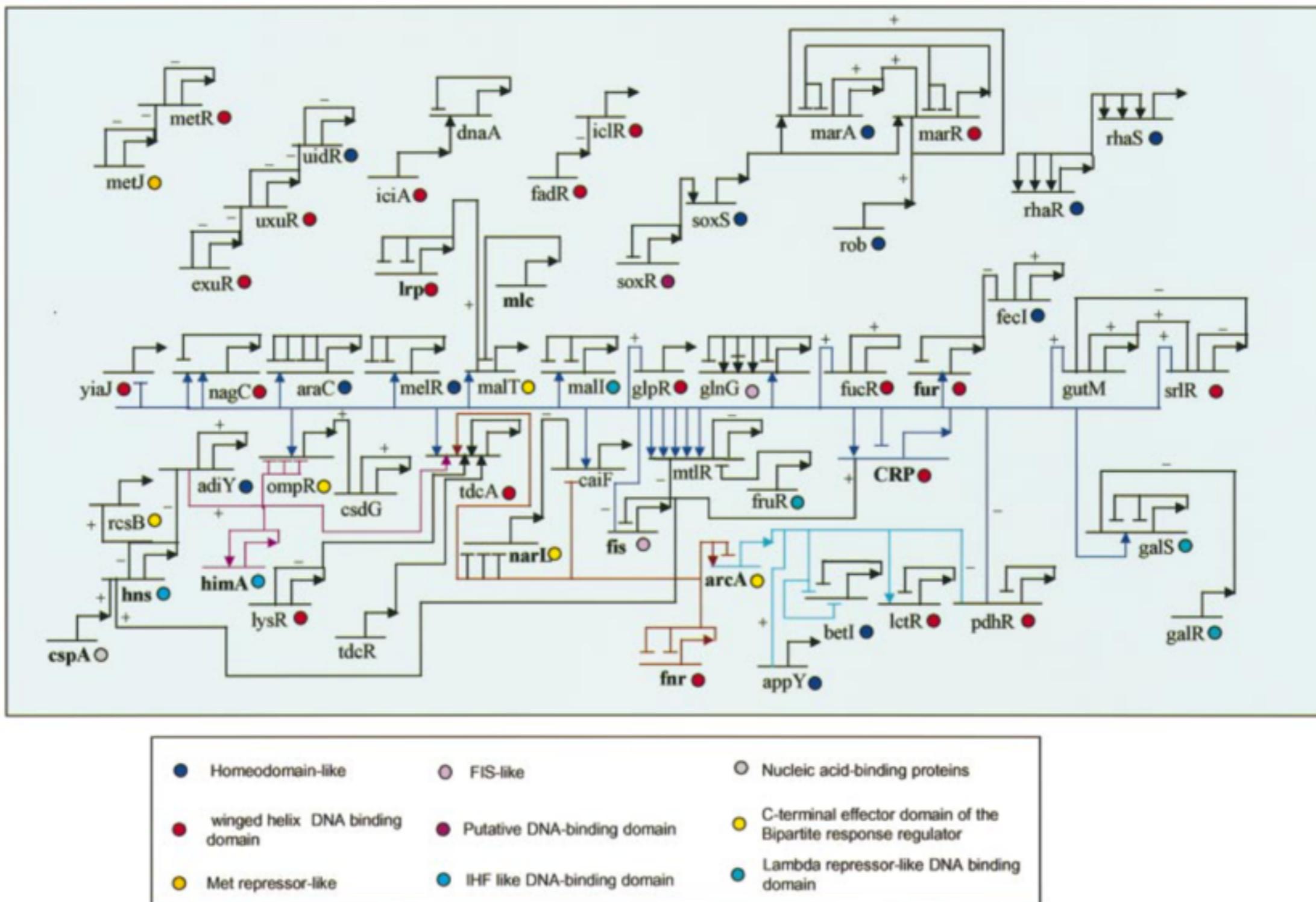
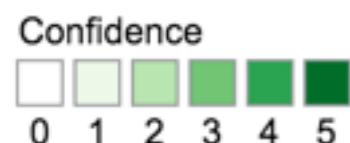
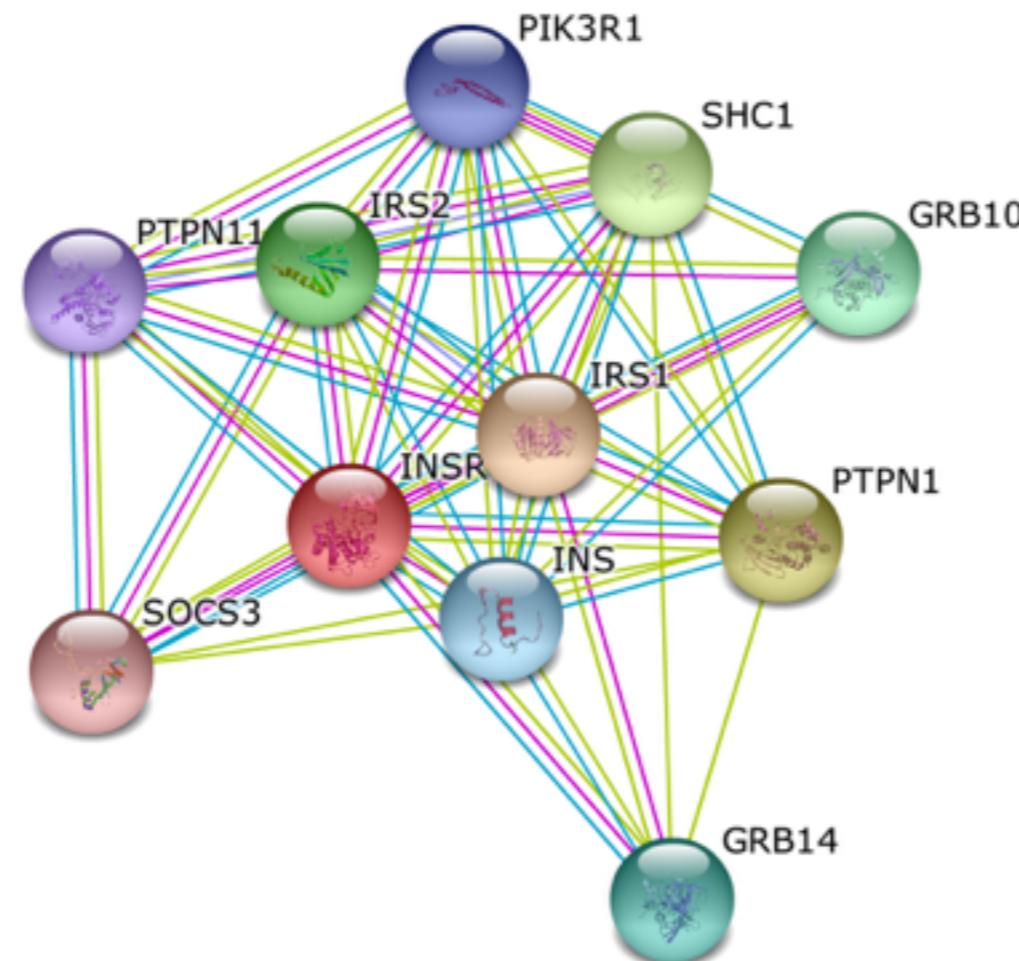


Figure 3. The transcription factor regulatory network in *E.coli*. When more than one transcription factor regulates a gene, the order of their binding sites is as given in the figure. An arrowhead is used to indicate positive regulation when the position of the binding site is known. A horizontal bar is used to indicate negative regulation when the position of the binding site is known. In cases where only the nature of regulation is known, without binding site information, + and - are used to indicate positive and negative regulation, respectively. These examples may be indirect rather than direct regulation. The DBD families are indicated by circles of different colours as given in the key. The names of global regulators are in bold.

Differential expression in tissues, compartments



Santos, A., Tsafou, K., Stolte, C., Pletscher-Frankild, S., O'Donoghue, S. I., & Jensen, L. J. (2015). Comprehensive comparison of large-scale tissue expression datasets. PeerJ, 3, e1054.



Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., ... Von Mering, C. (2015). STRING v10: Protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Research, 43(D1), D447–D452.



Binder, J. X., Pletscher-Frankild, S., Tsafou, K., Stolte, C., O'Donoghue, S. I., Schneider, R., & Jensen, L. J. (2014). COMPARTMENTS: unification and visualization of protein subcellular localization evidence. Database : The Journal of Biological Databases and Curation, 2014, 1–9.

Protein-Protein interactions

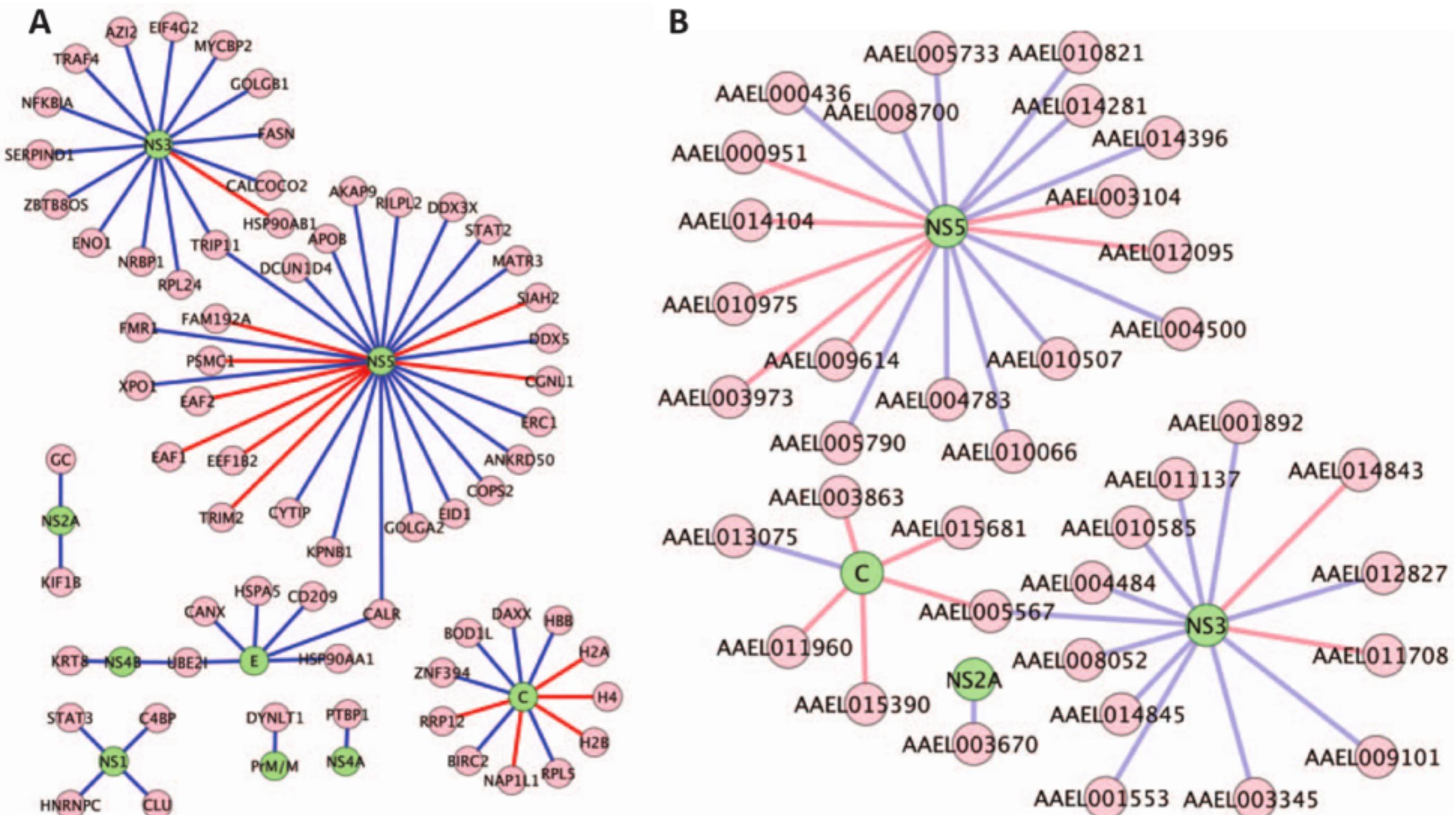
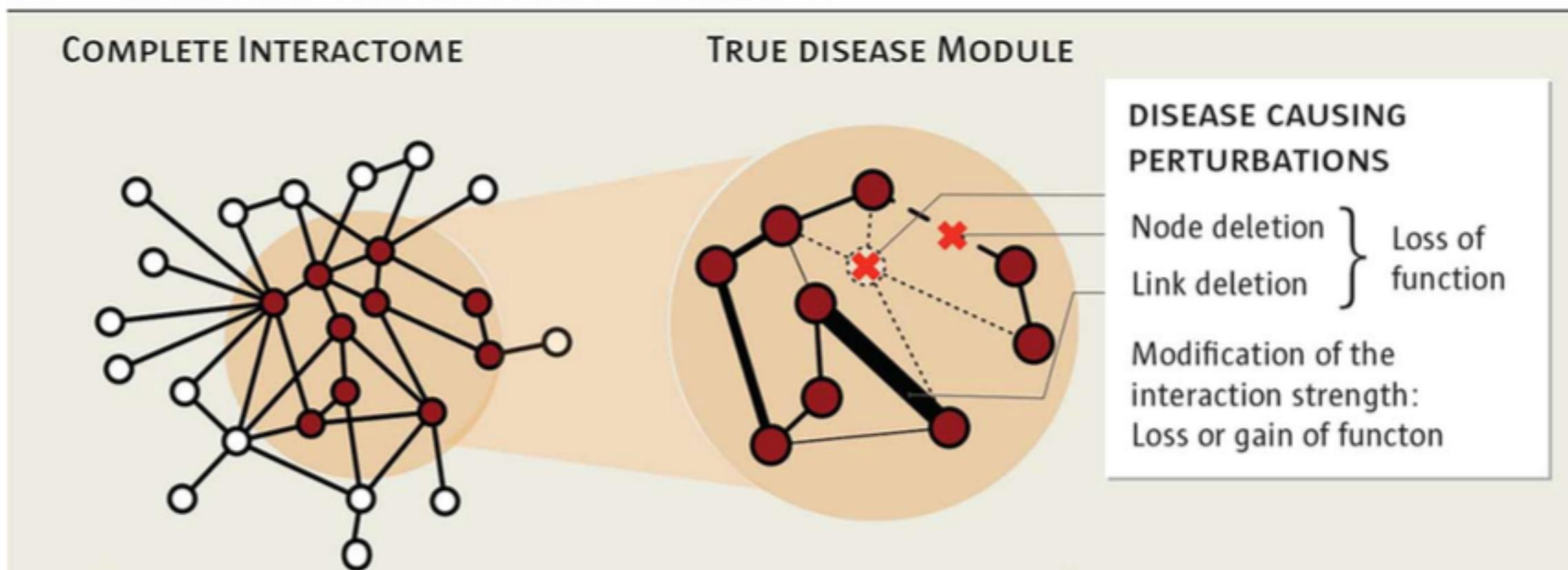


Figure 5. Dengue-host interactions supported by multiple forms of evidence. (A) Dengue-human interactome. (B) Dengue-mosquito interactome. Pink nodes represent host proteins. Green nodes represent dengue proteins. Red edges represent PPI with conserved interologs. Additional details are available in Cytoscape files in supplemental data (Data S1).
doi:10.1371/journal.pone.0053535.g005

Diseases perturb protein-protein interactions

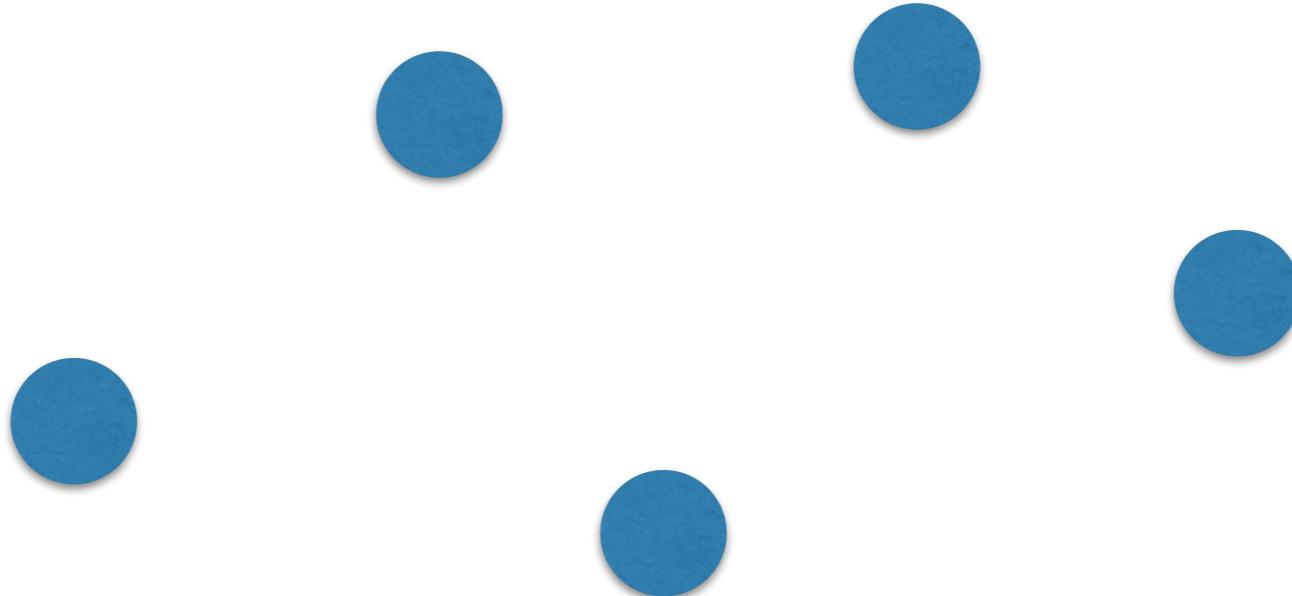
A DISEASE MODULES AS LOCAL PERTUBATIONS



Jörg Menche,^{1,2,3} Amitabh Sharma,^{1,2} Maksim Kitsak,^{1,2} Susan Dina Ghiassian,^{1,2}
Marc Vidal,^{2,4} Joseph Loscalzo,⁵ Albert-László Barabási^{1,2,3,5*}

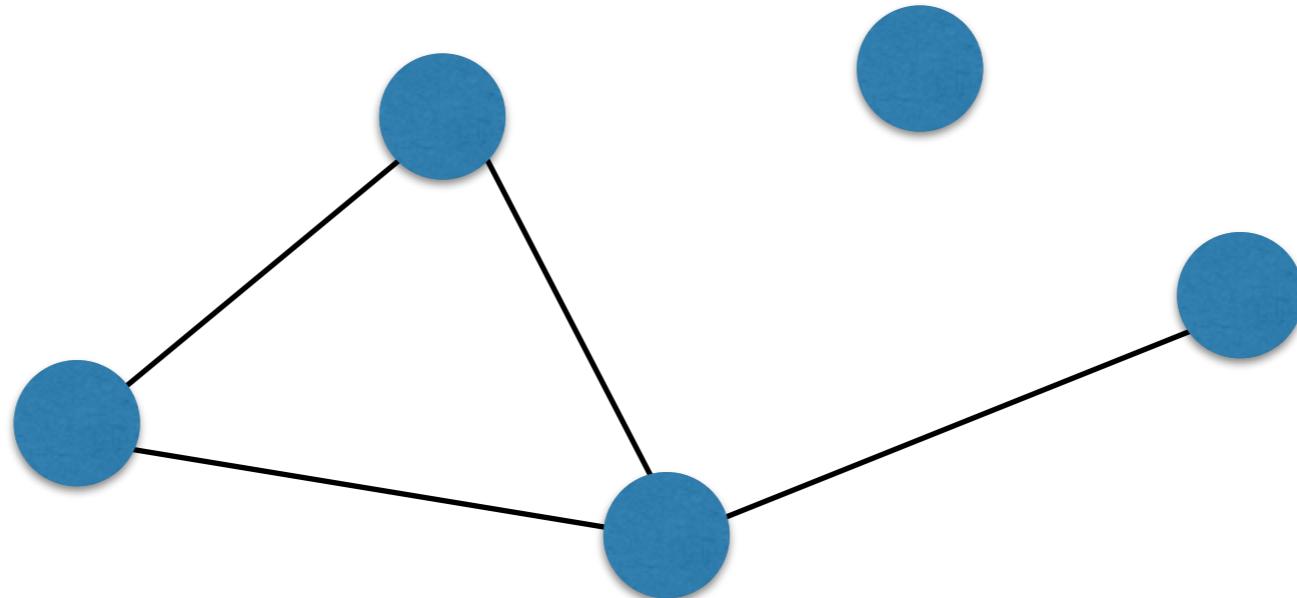
Graph theory

- Graph = Nodes + Edges
- Nodes, sometimes called vertices (node/vertex)

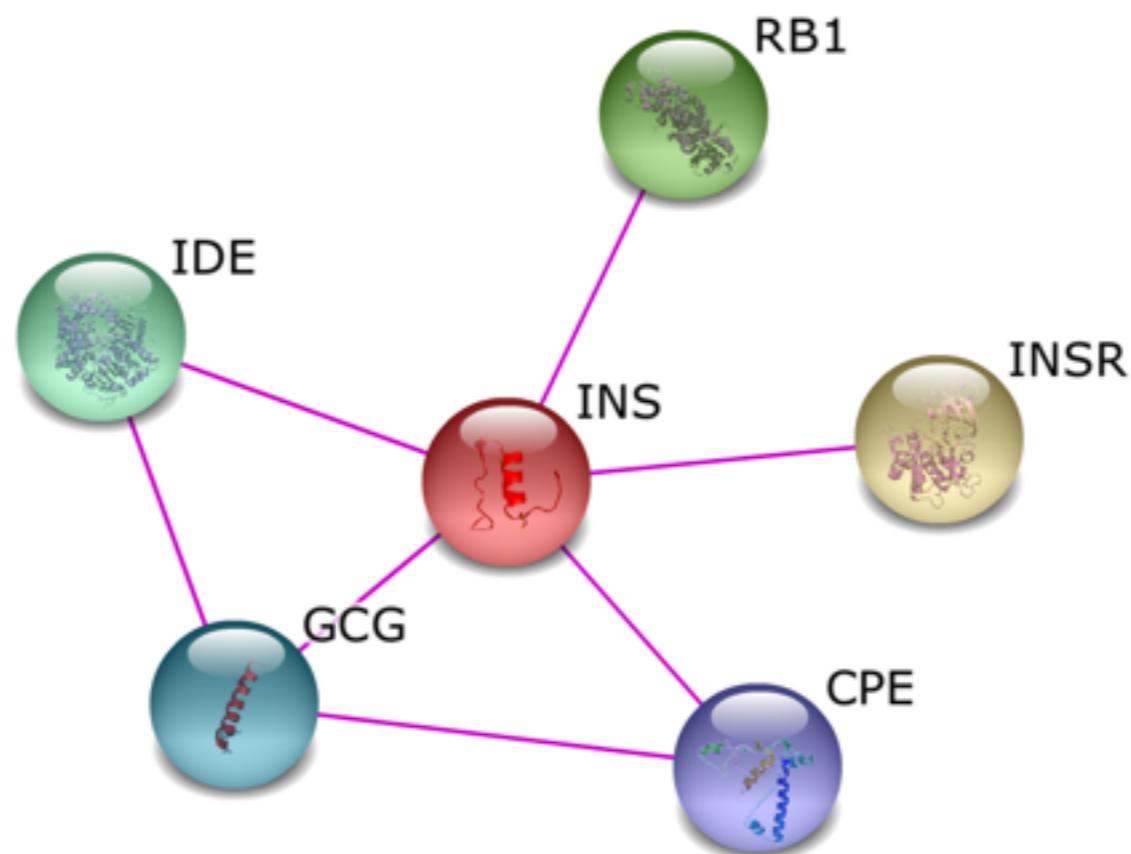


Graph theory

- Graph = Nodes + Edges
- Nodes, also called vertices (node/vertex)

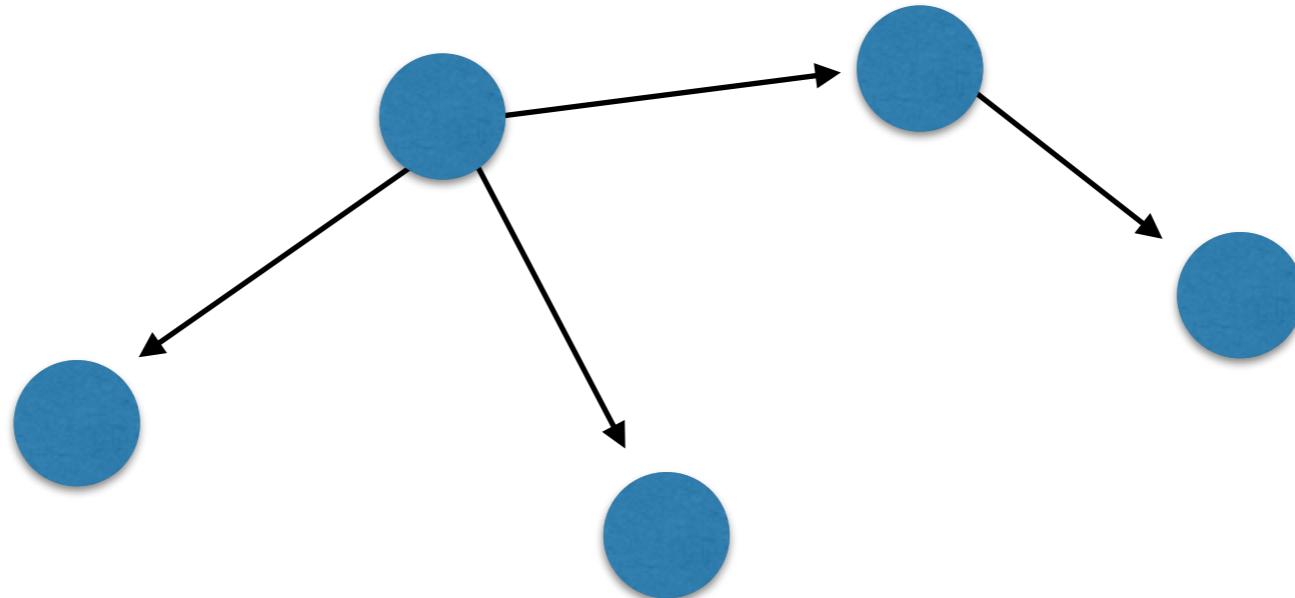


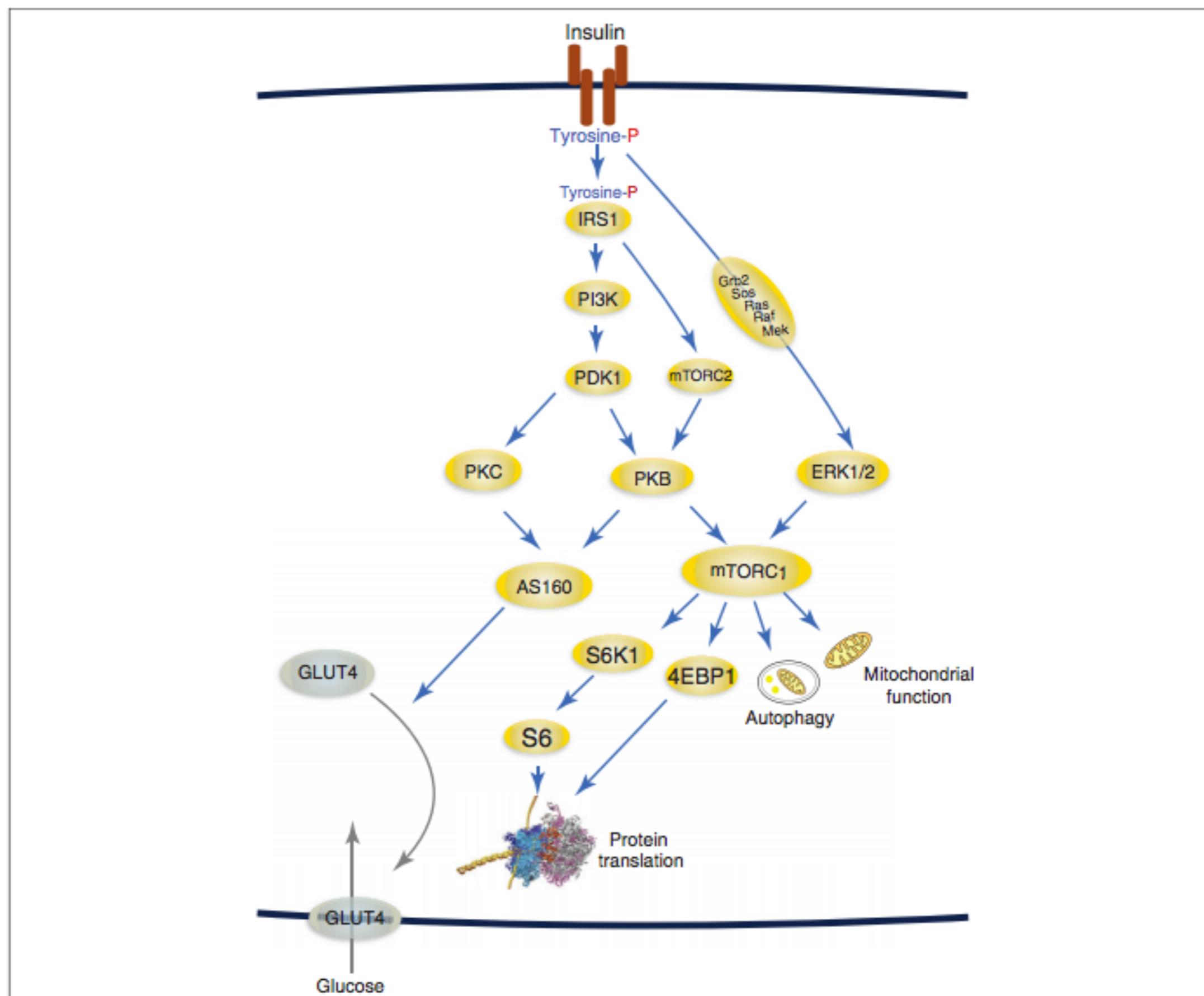
Insulin interactions (experimental evidence)



Graph theory: directedness

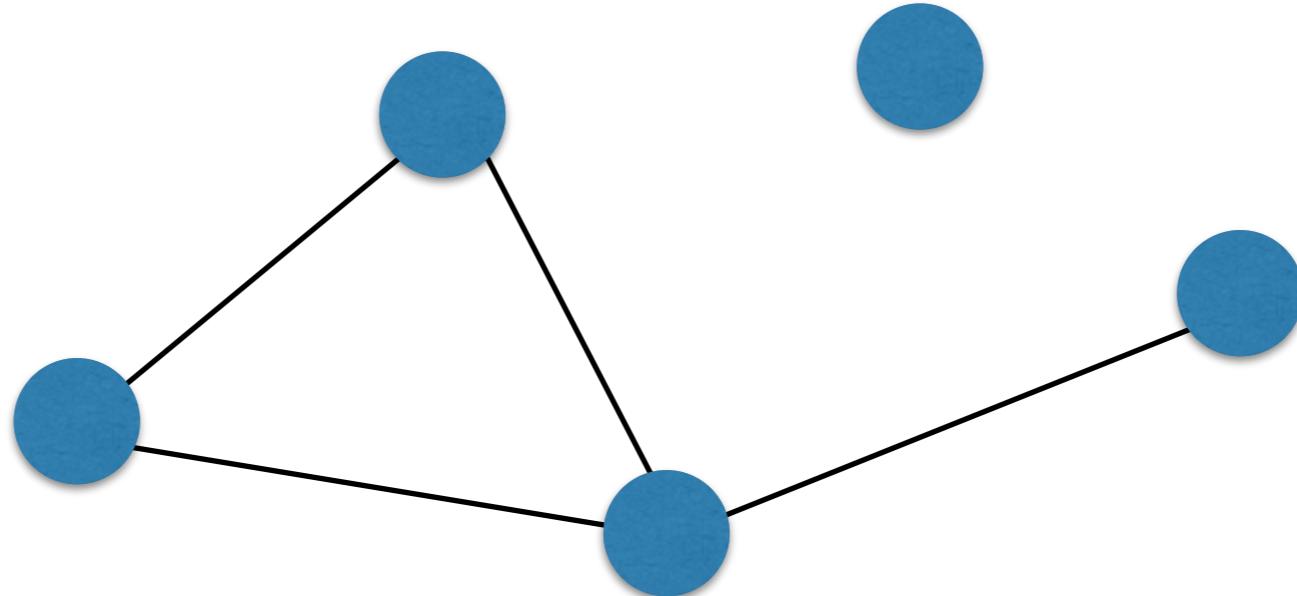
- Edges can be directed





Graph theory: degree

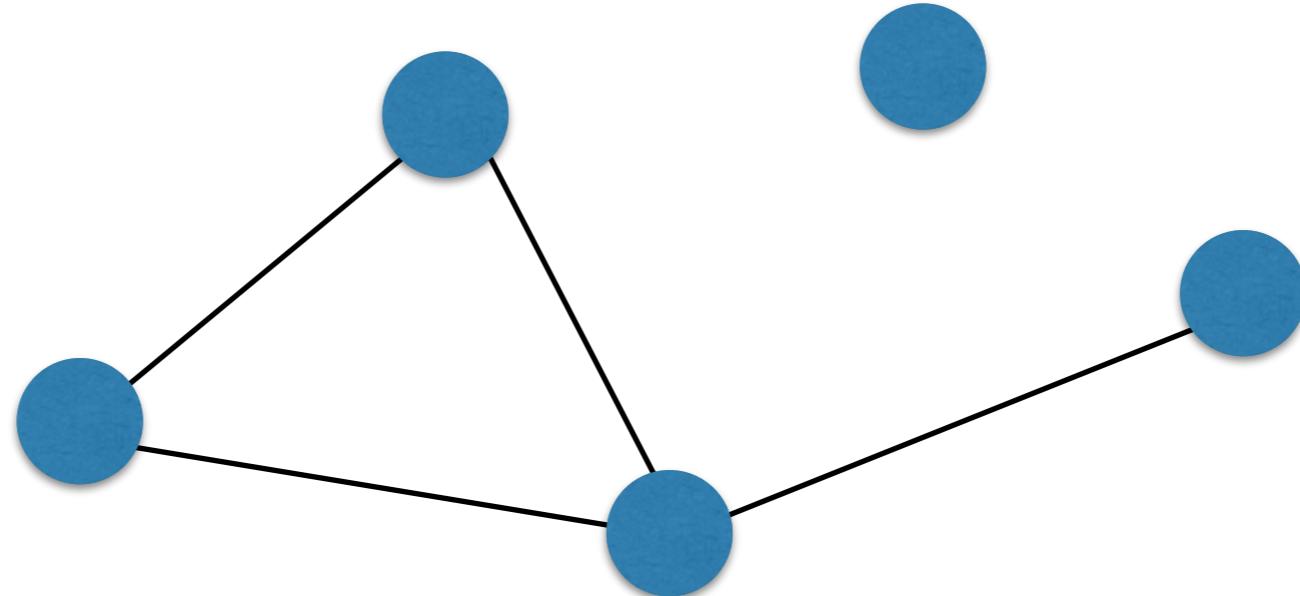
- Degree of a node is the number of edges that connect to it



Graph theory: degree

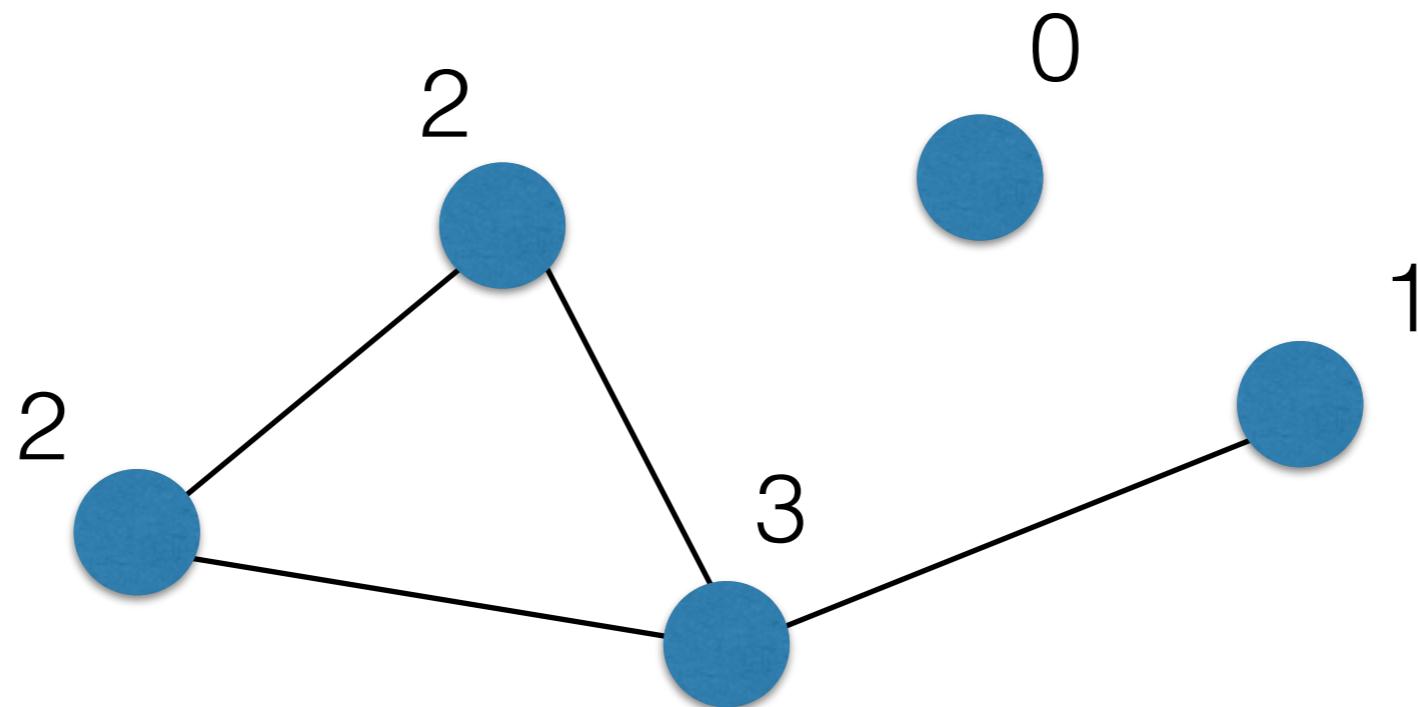
- Degree of a node is the number of edges that connect to it

What is the degree of each node?



Graph theory: degree

- Degree of a node is the number of edges that connect to it

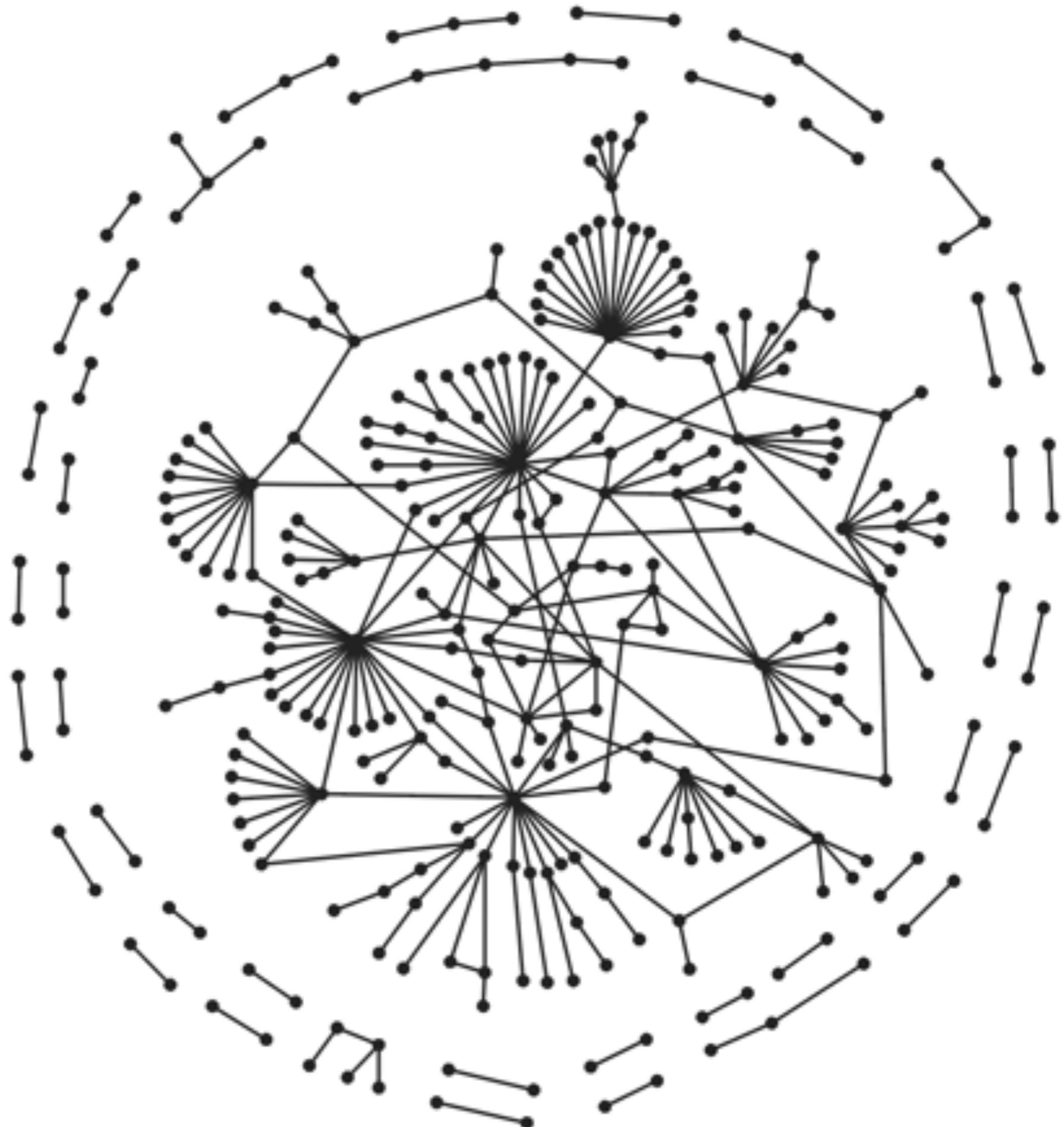


Specificity and Stability in Topology of Protein Networks

Sergei Maslov¹ and Kim Sneppen²

3 MAY 2002 VOL 296 SCIENCE

Fig. 1. Network of physical interactions between nuclear proteins. Here, we show the part of the network reported in (2), consisting of all proteins that are known to be localized in the yeast nucleus (3), and which interact with at least one other protein in the nucleus. This subset consists of 318 interactions between 329 proteins. Note that most neighbors of highly connected nodes have rather low connectivity.

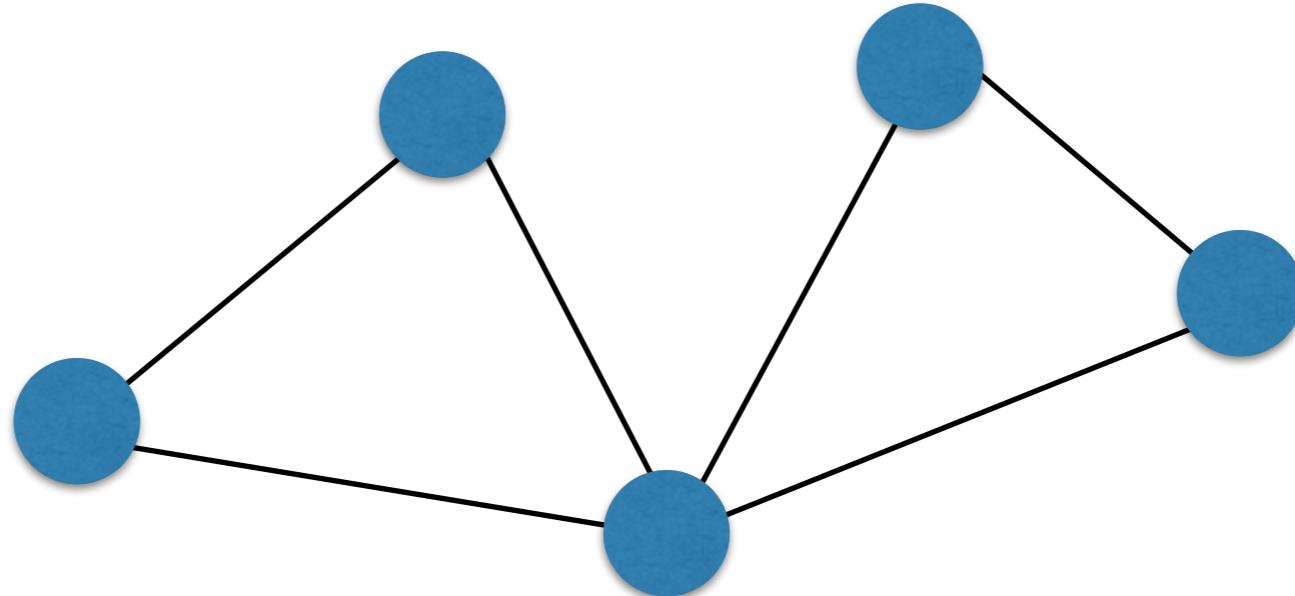


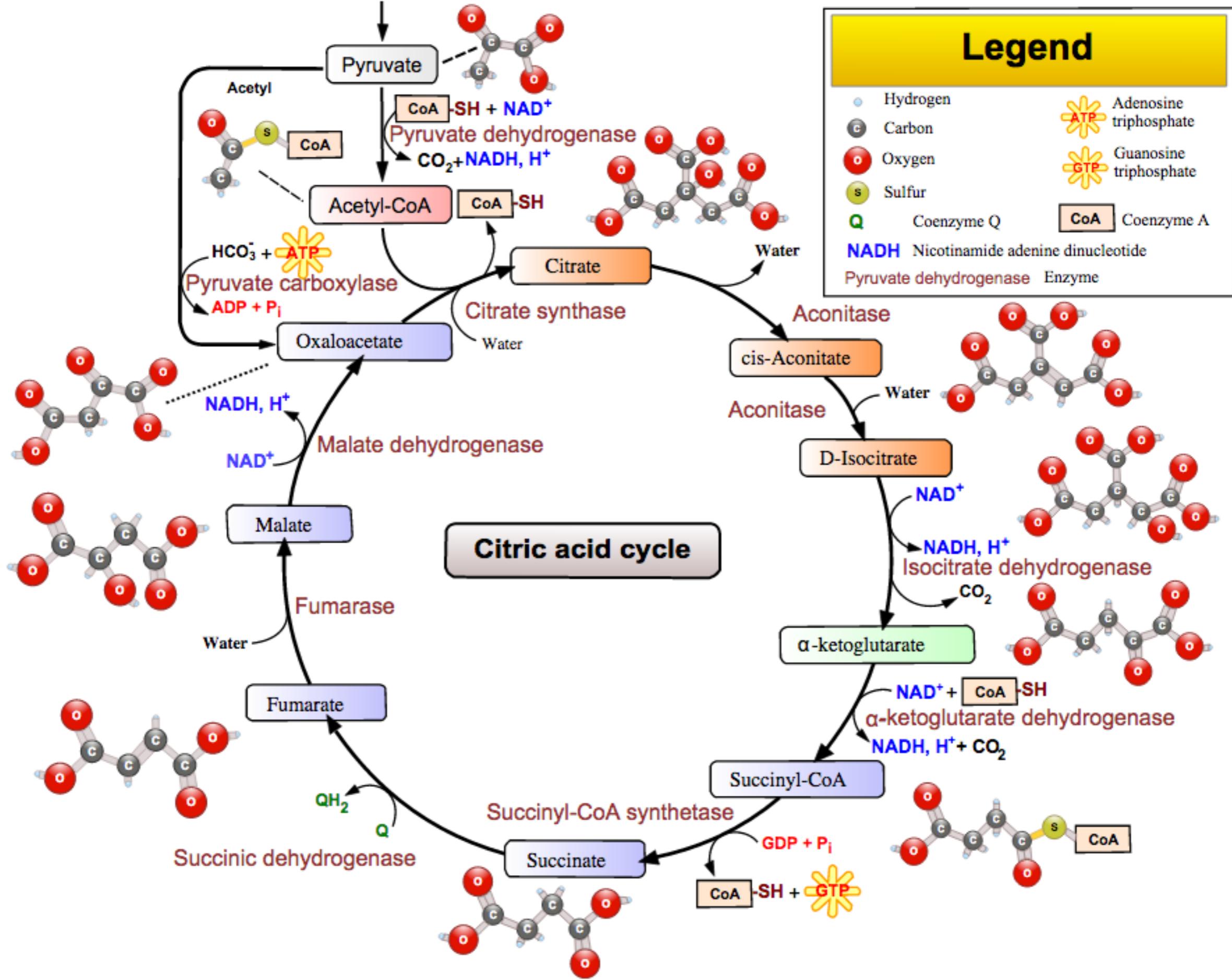
Hubs

- Hubs tend to not interact directly with other hubs.
- Hubs tend to be older proteins that are more highly conserved.

Graph theory: cycles

- Cycles/circuits are loops in the network





"Citric acid cycle with aconitate 2" by Narayane, WikiUserPedia, YassineMrabet, TotoBoggins

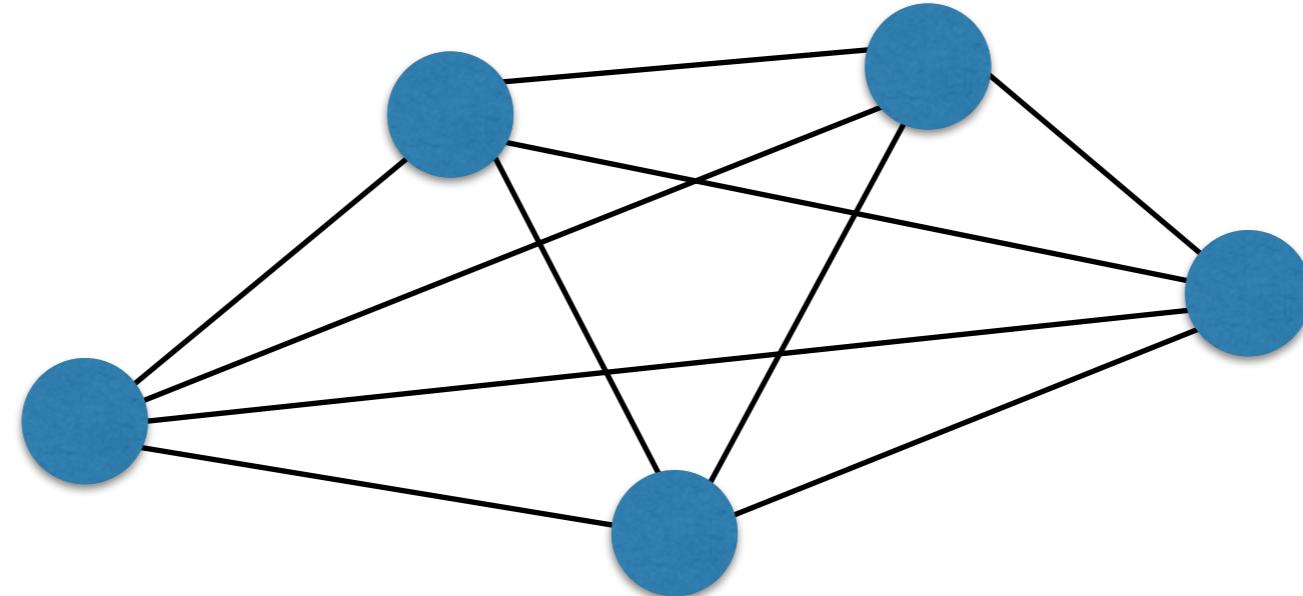
<http://biocyc.org/META/NEW-IMAGE?type=PATHWAY&object=TCA>. Image adapted from :Image:Citric acid cycle noi.svg(uploaded to Commons by wadester16).

Licensed under CC BY-SA 3.0 via Commons

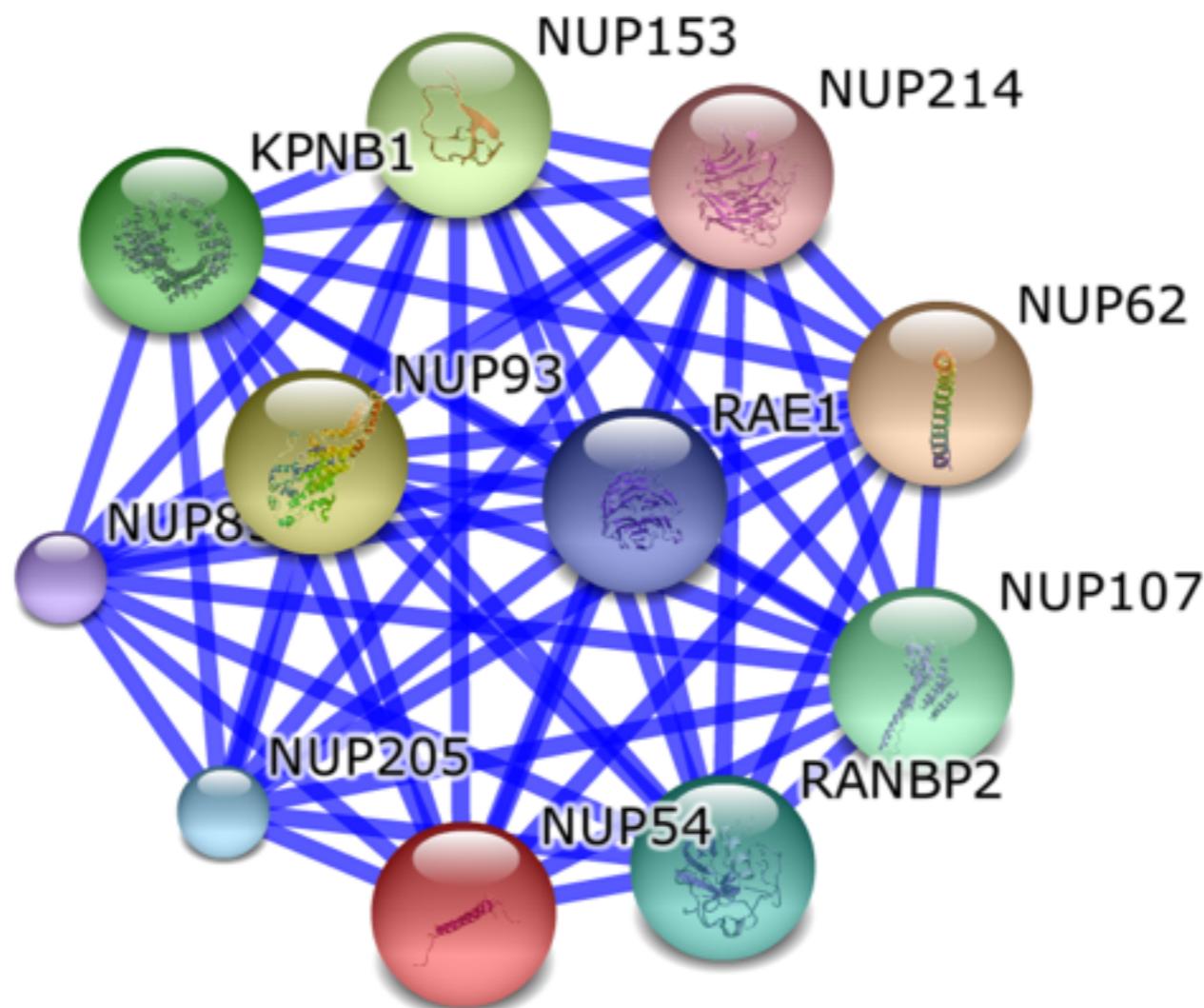
https://commons.wikimedia.org/wiki/File:Citric_acid_cycle_with_aconitate_2.svg#/media/File:Citric_acid_cycle_with_aconitate_2.svg

Graph theory: connectedness

- Fully connected
- Terminology: “dense” if nearly fully connected

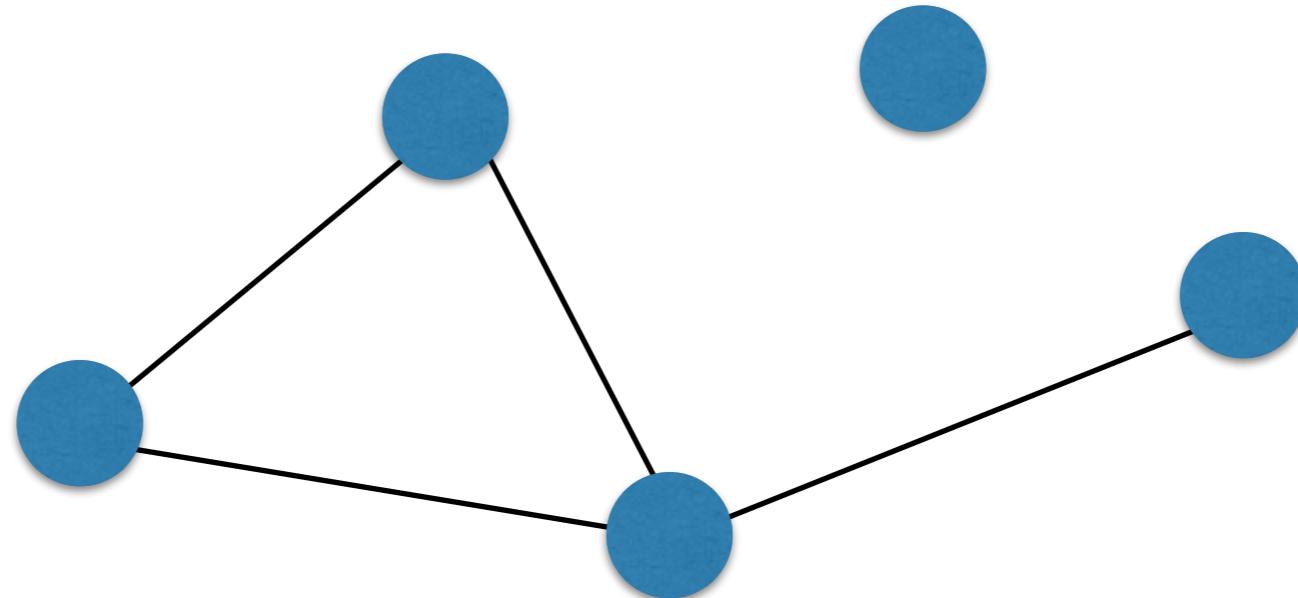


Components of the nuclear pore complex



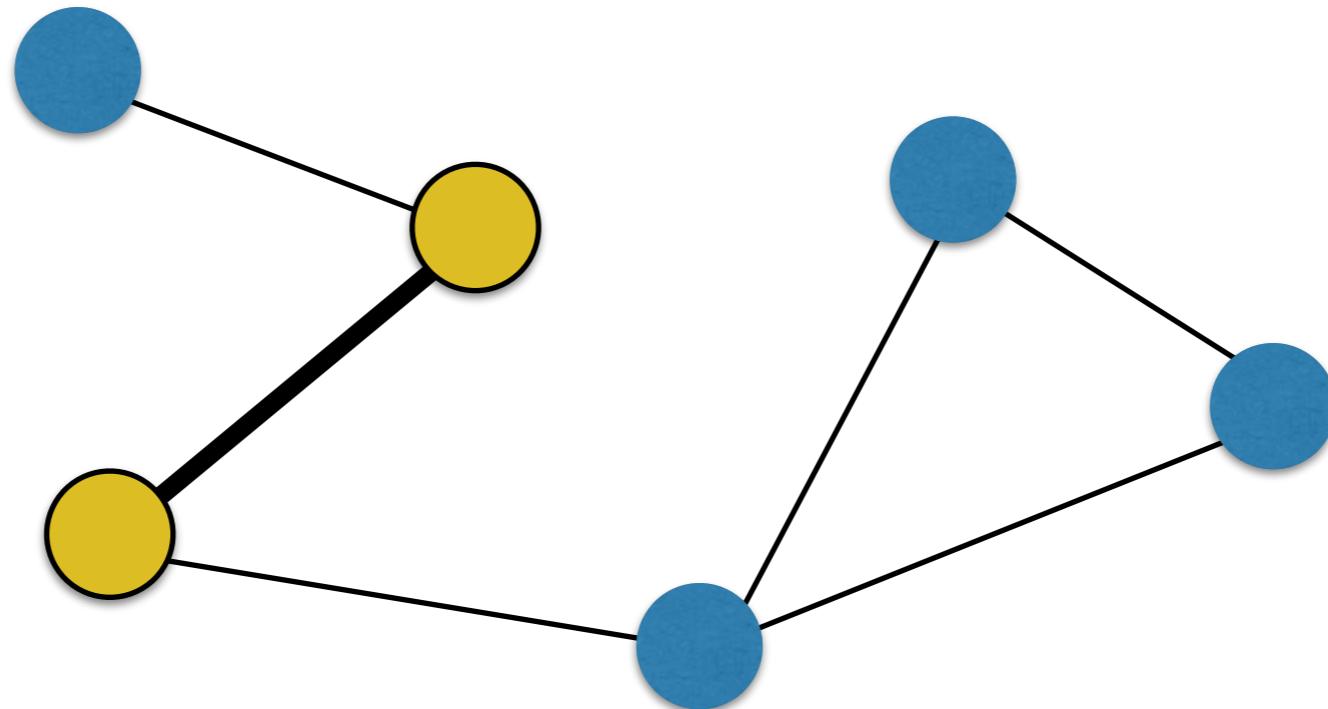
Graph theory: connectedness

- Connected components can be traversed by following the edges
- 2 connected components:



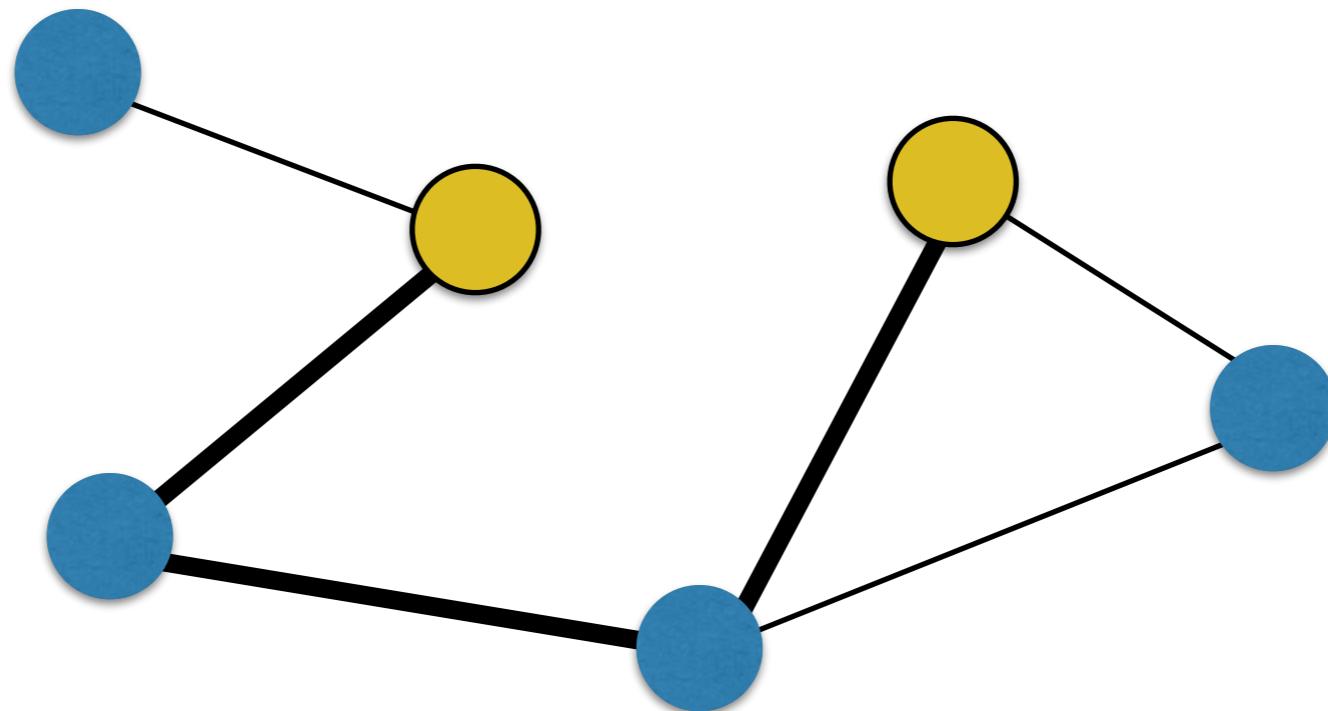
Graph theory: shortest path

- Shortest path between nodes



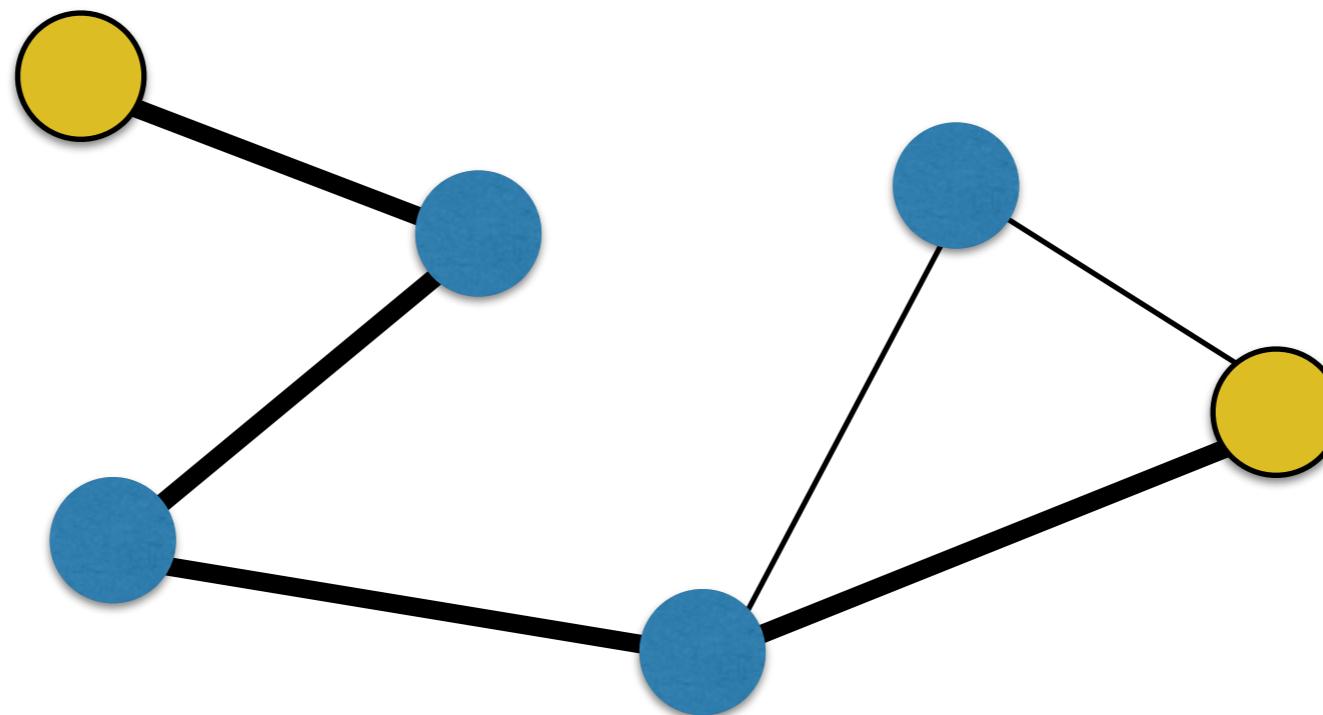
Graph theory: shortest path

- Shortest path between nodes

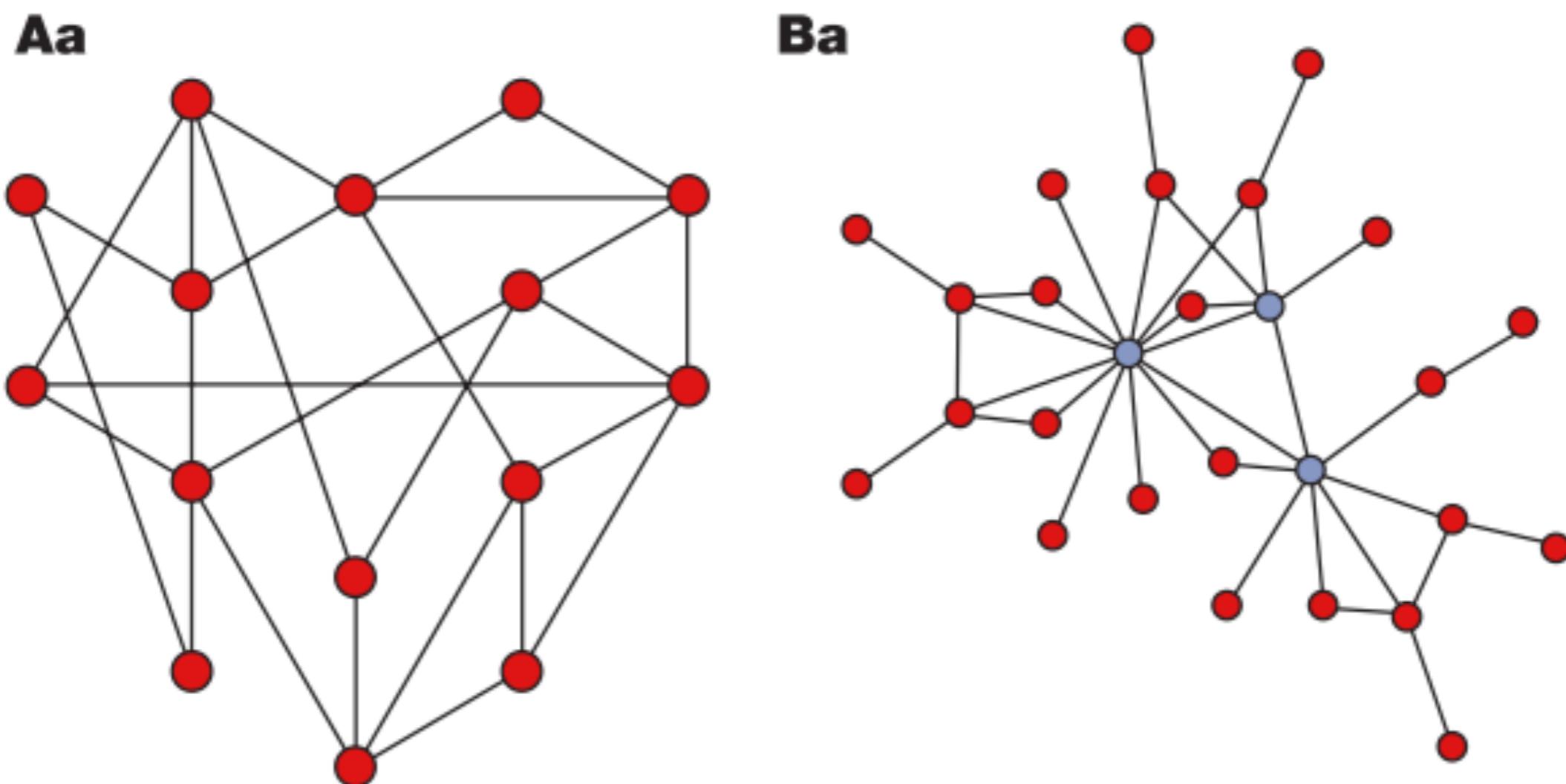


Graph theory: diameter

- Longest shortest path = 4
- Terminology: network diameter

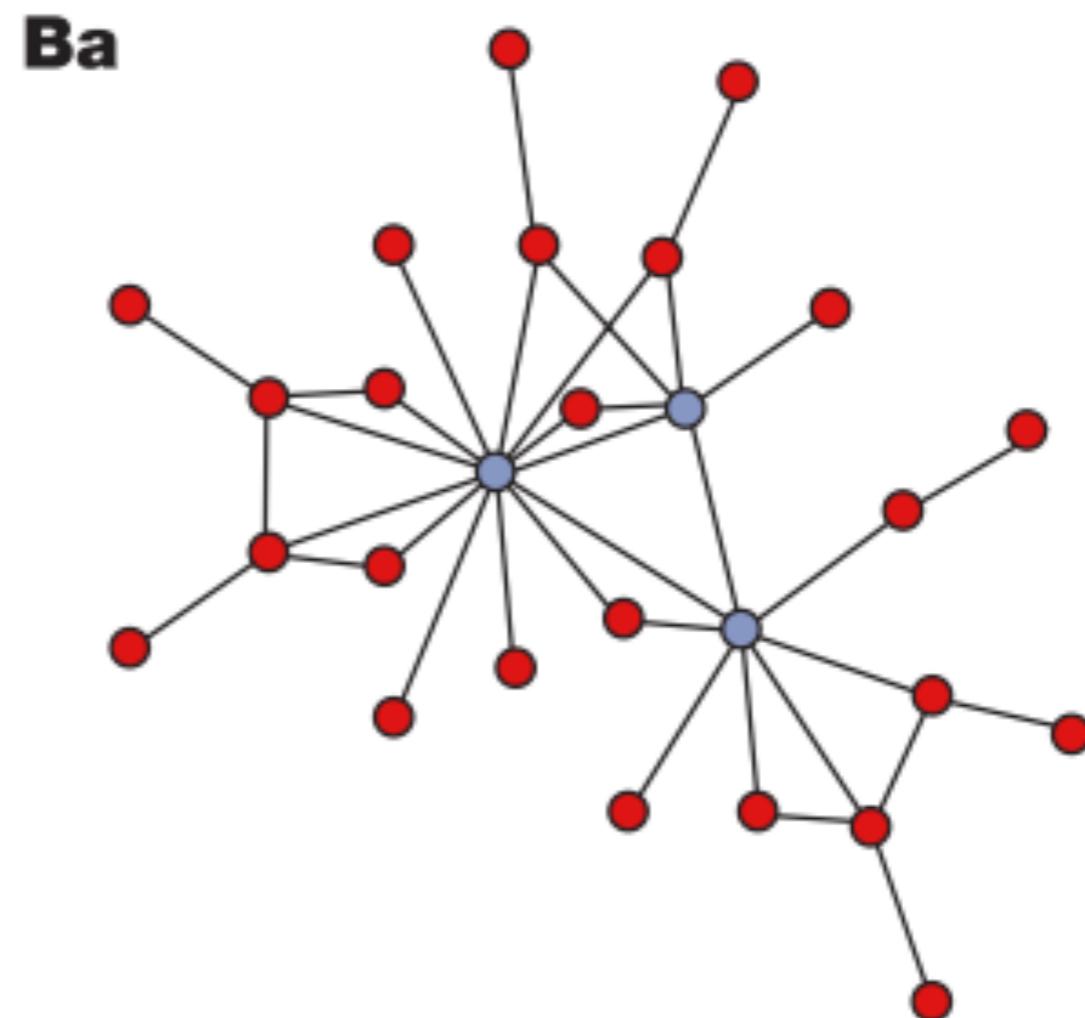
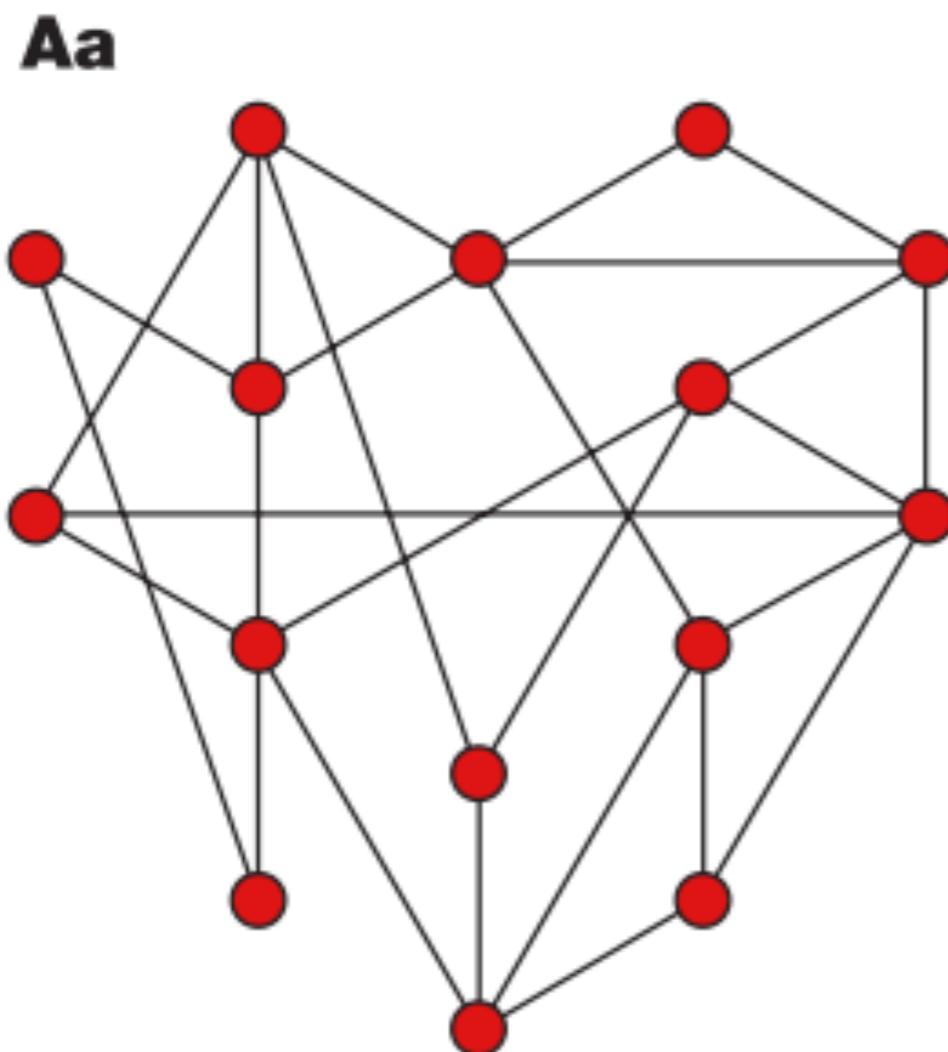


Network Organisation



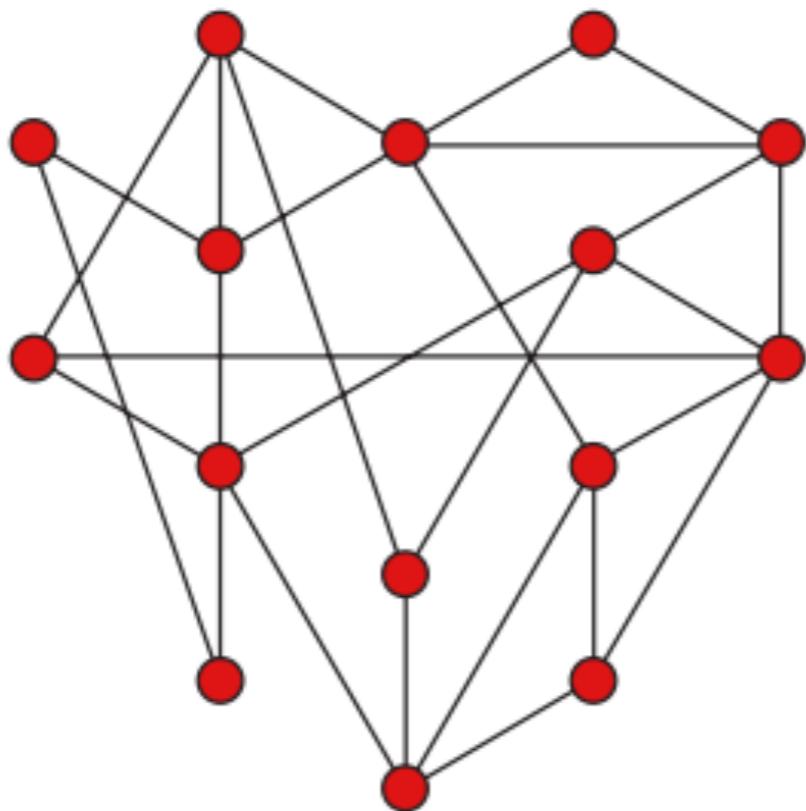
Network Organisation

What is different between these networks?



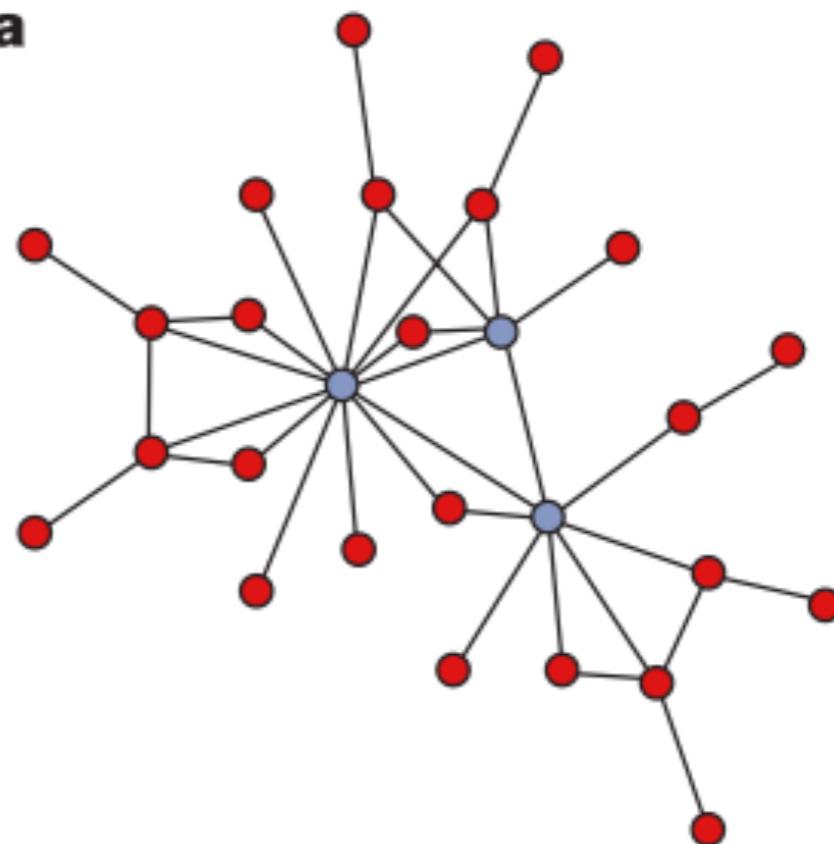
A Random network

Aa

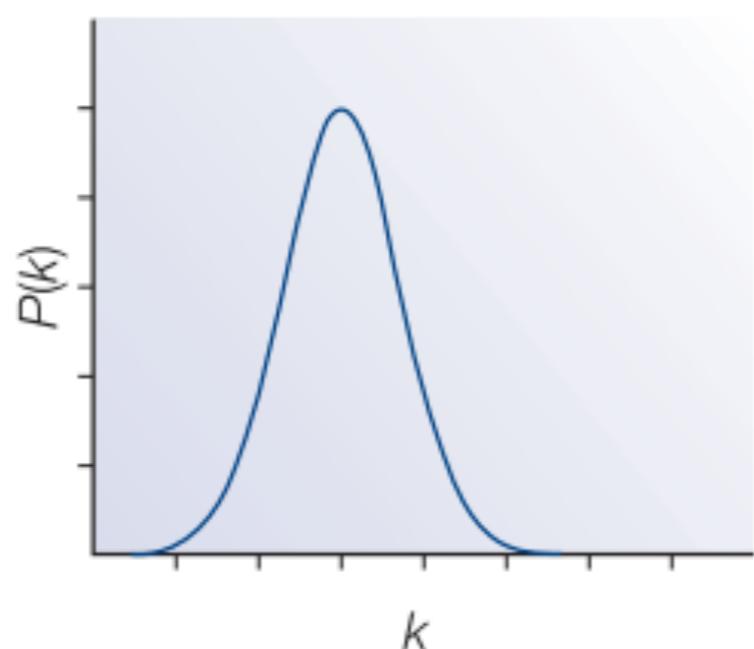


B Scale-free network

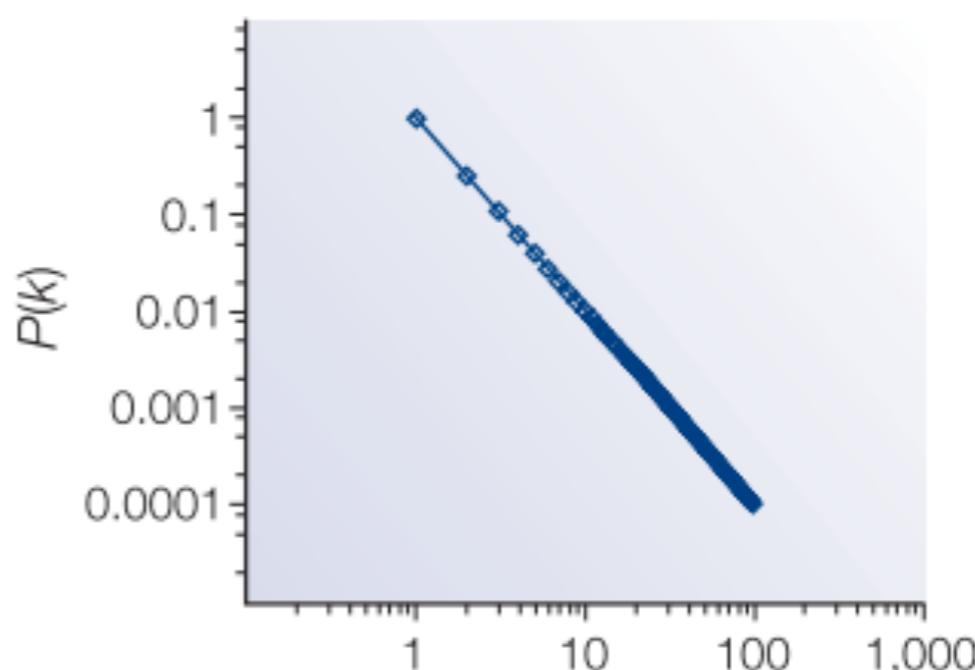
Ba



Roll
dice

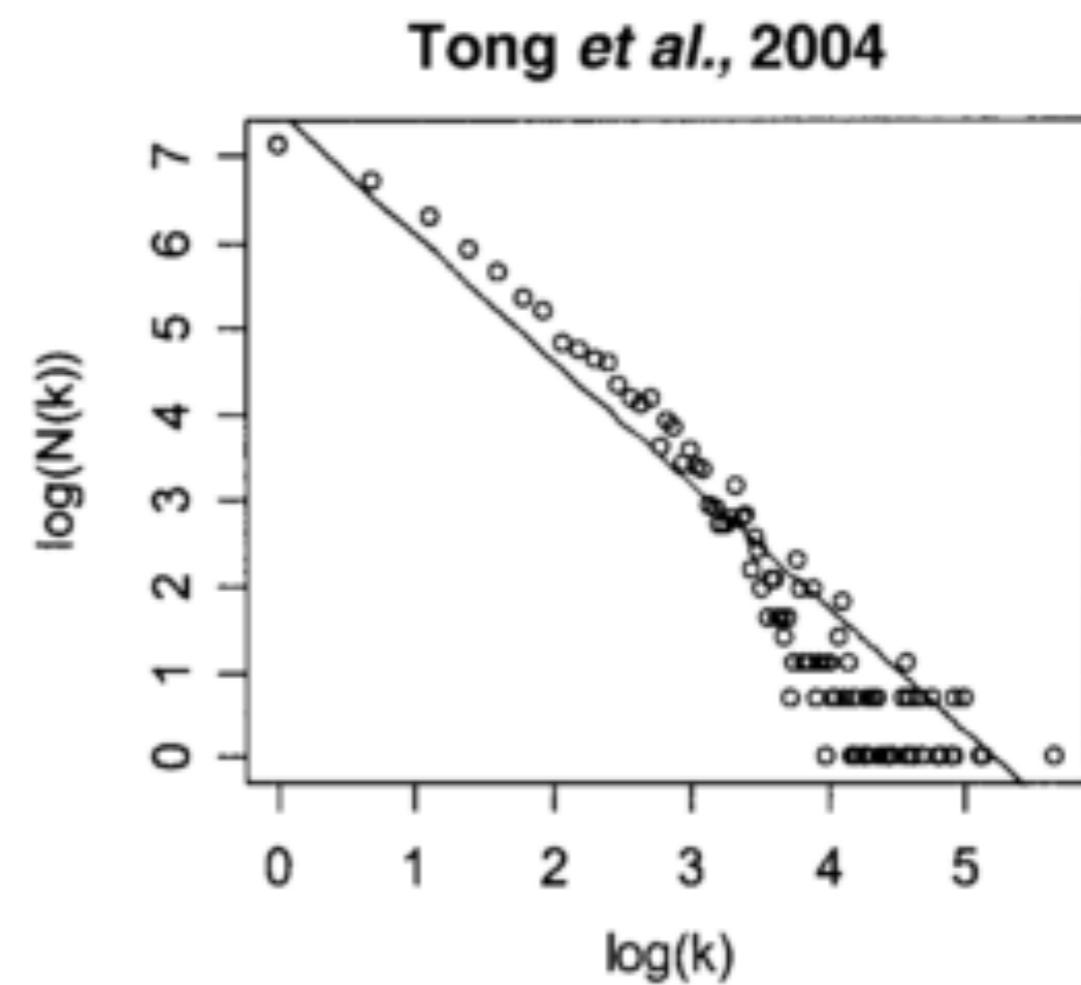
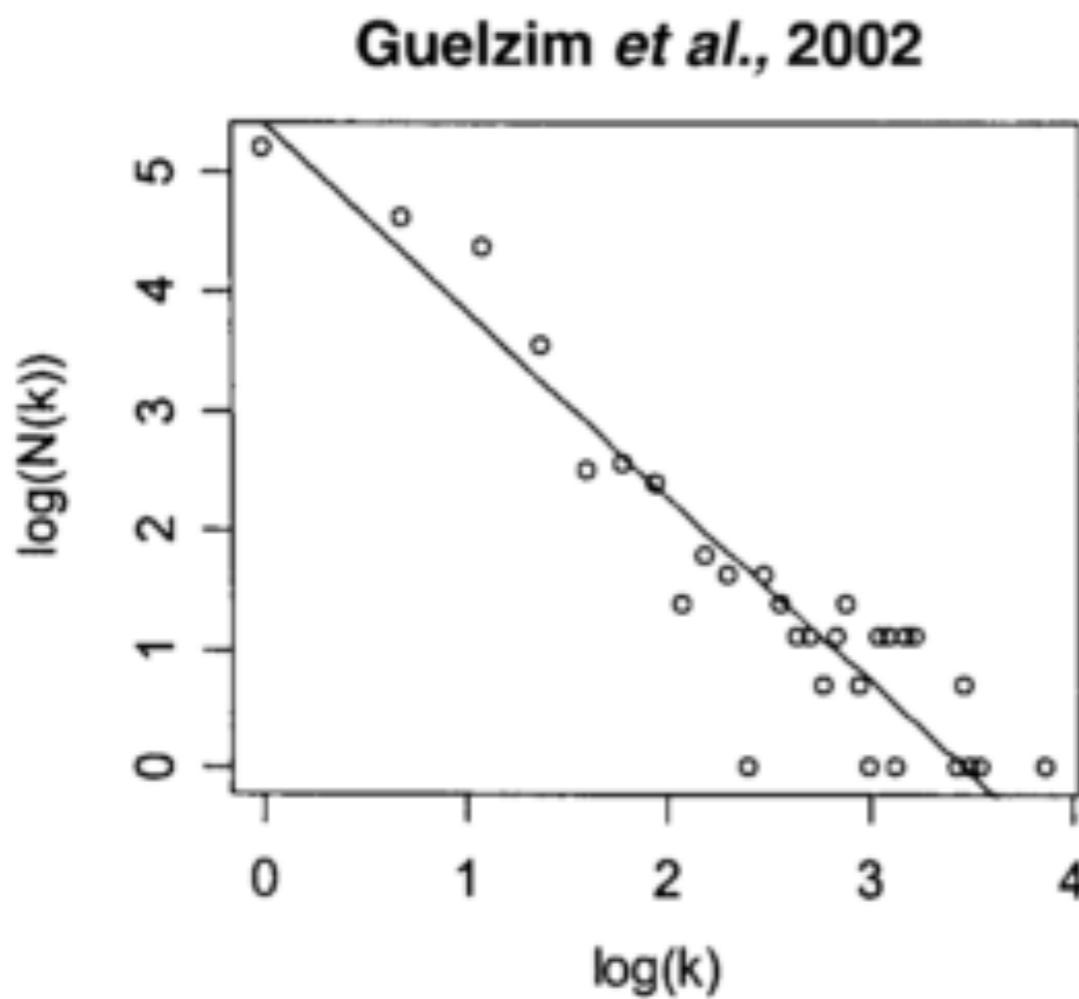


Bb



Power law
 $P(k) = k^{-\alpha}$

Biological networks (approx) follow
a power law and are scale free

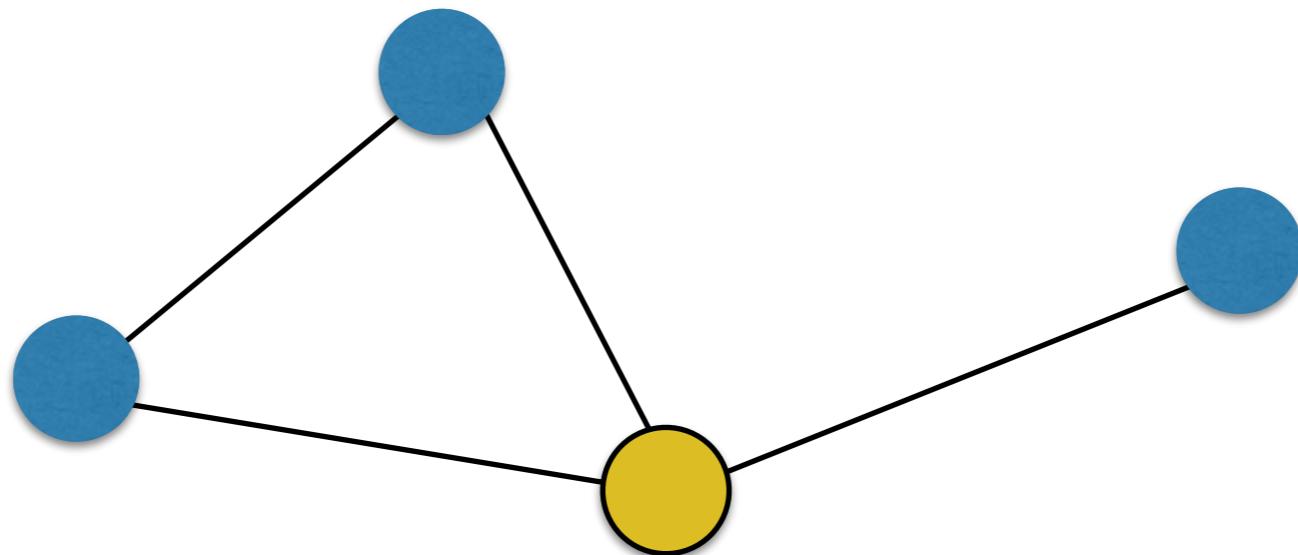


Scale free

- Each node can be reached from any other node in a small number of hops (small network diameter)
- Robust: deleting many random nodes will maintain most network connectivity.
- But vulnerable to deletion of hubs.

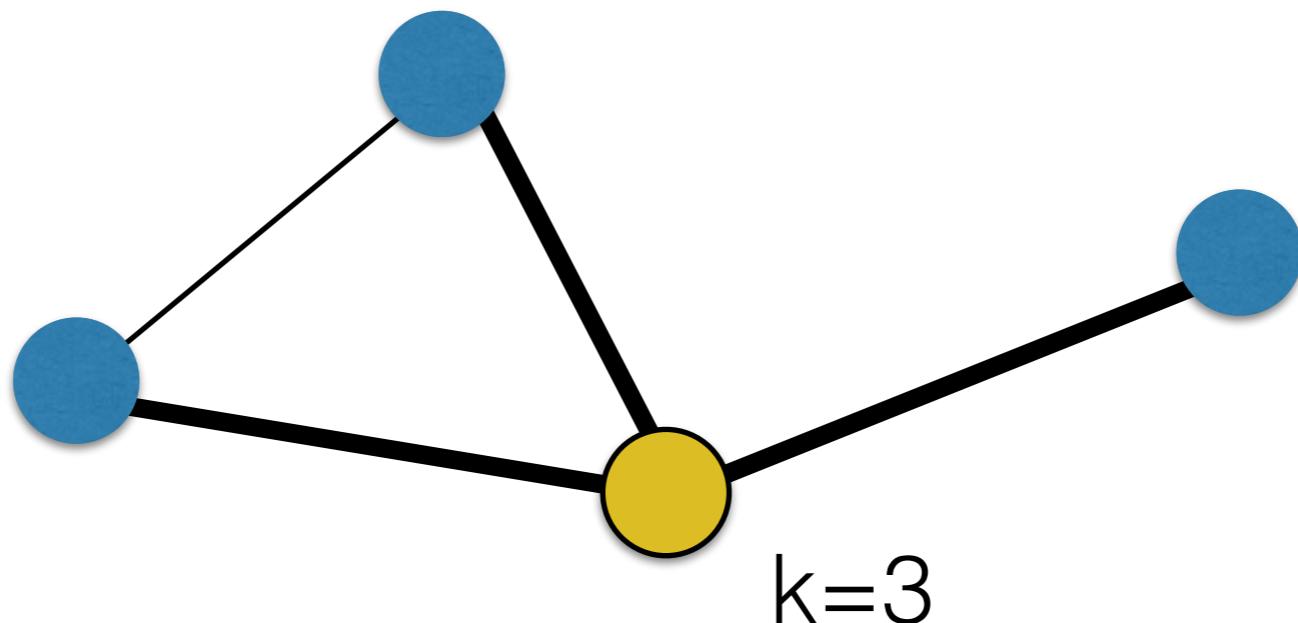
Clustering coefficient

- Are my friends friends with each other?



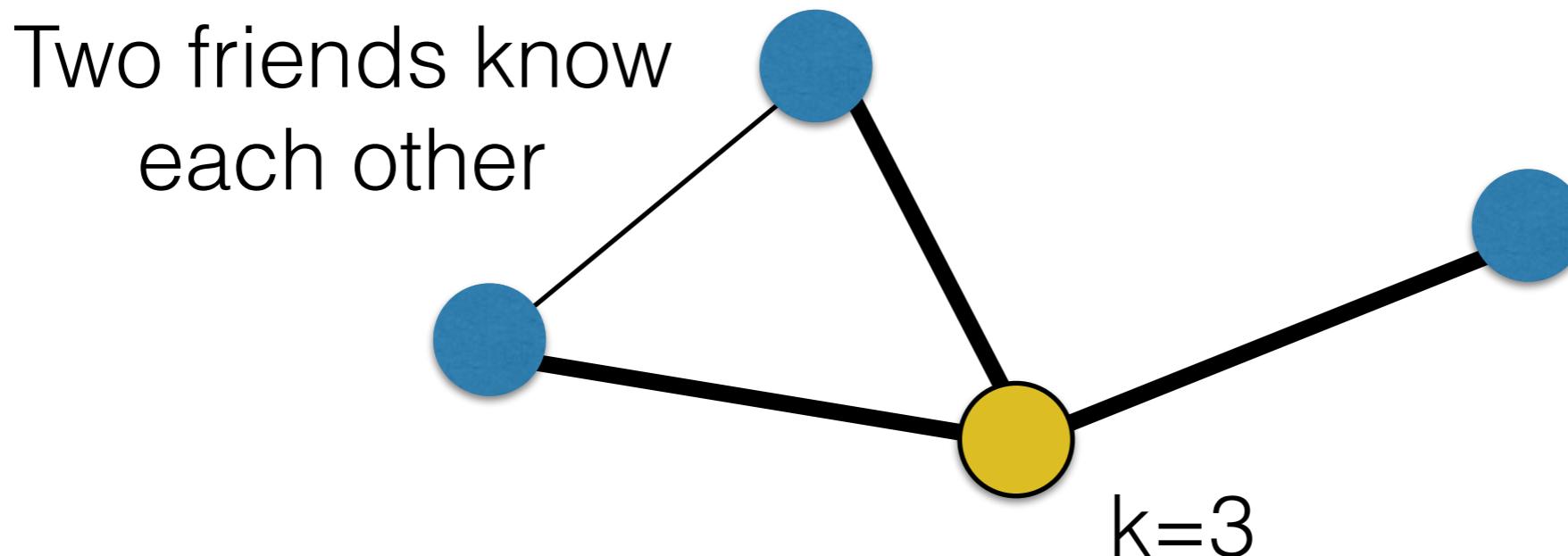
Clustering coefficient

- Are my friends friends with each other?



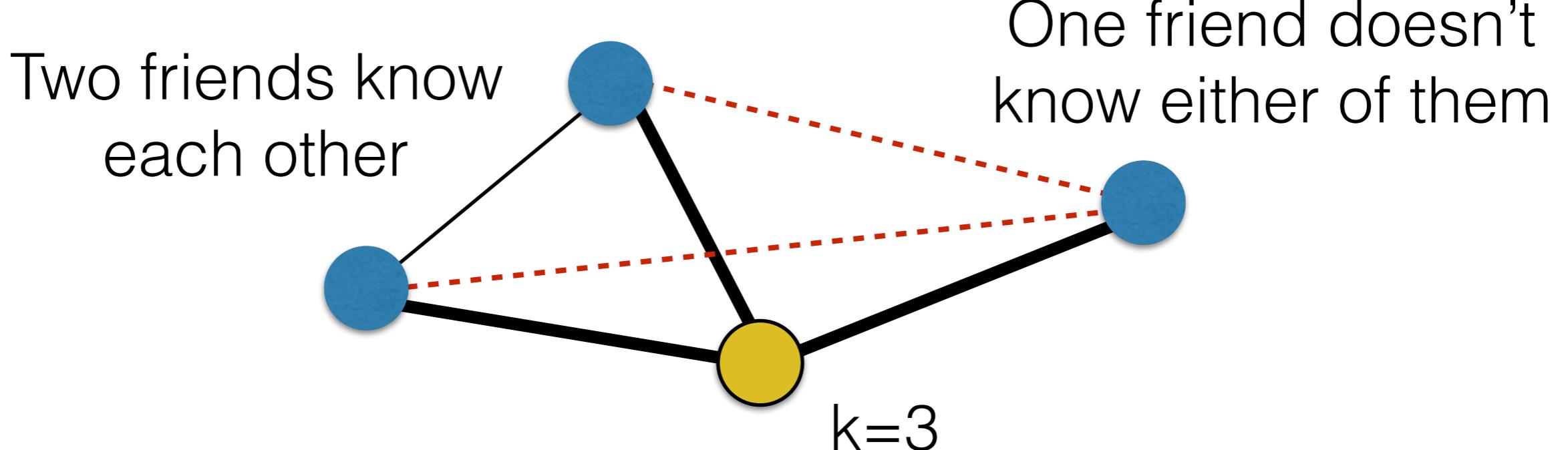
Clustering coefficient

- Are my friends friends with each other?



Clustering coefficient

- Are my friends friends with each other?

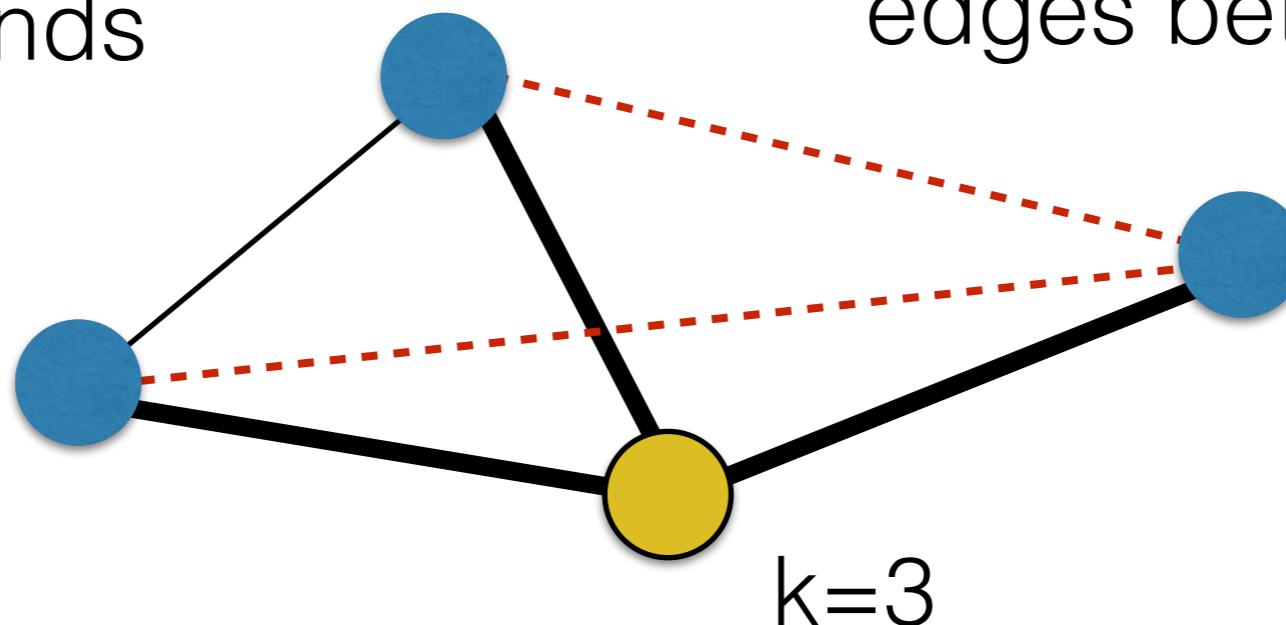


Clustering coefficient

- Are my friends friends with each other?

Number of edges
between friends

Number of possible
edges between friends



Clustering coefficient

- Are my friends friends with each other?

Number of edges
between friends

$$C_i = \frac{\text{Number of edges between friends}}{\text{Number of possible edges between friends}} =$$

Number of possible
edges between friends

k is the number of edges of node i

n_i is the number of edges between neighbours of node i

Clustering coefficient

- Are my friends friends with each other?

Number of edges
between friends

n_i

$$C_i = \frac{\text{Number of edges between friends}}{\text{Number of possible edges between friends}} = \frac{n_i}{\binom{k}{2}}$$

k is the number of edges of node i

n_i is the number of edges between neighbours of node i

Clustering coefficient

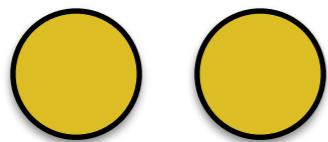
- Are my friends friends with each other?

$$C_i = \frac{\text{Number of edges between friends}}{\text{Number of possible edges between friends}} = \frac{n_i}{k \text{ choose } 2}$$

k is the number of edges of node i == degree == friends
 n_i is the number of edges between neighbours of node i

k Choose 2

2C2



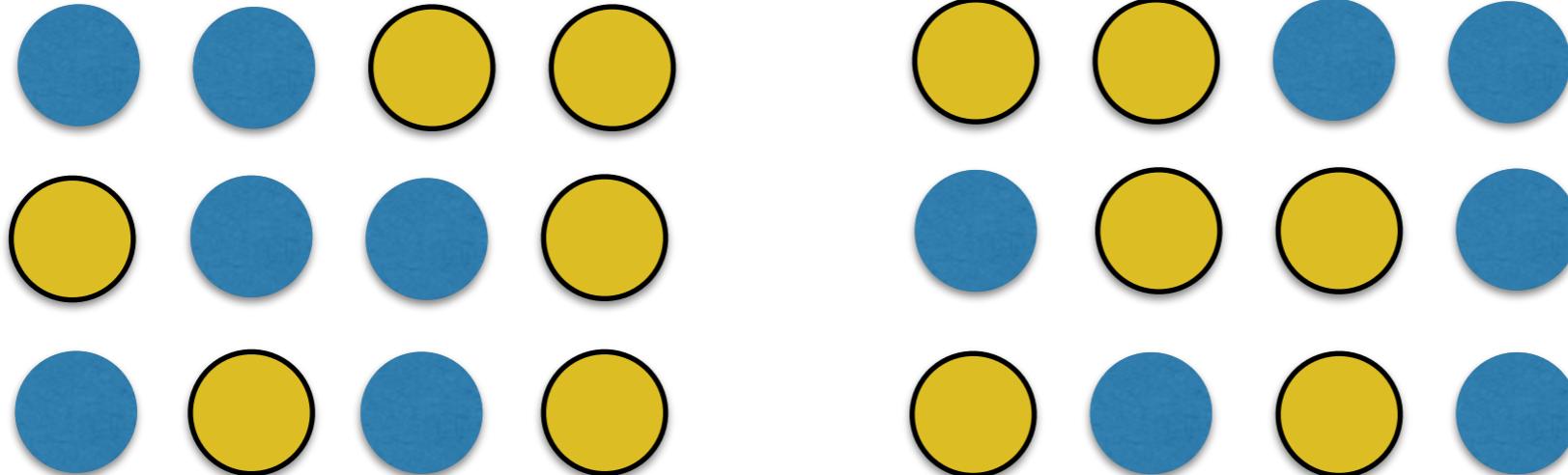
1

3C2



3

4C2



6

k Choose a

(Binomial coefficient)

$$kCa = \frac{k!}{a! (k-a)!}$$

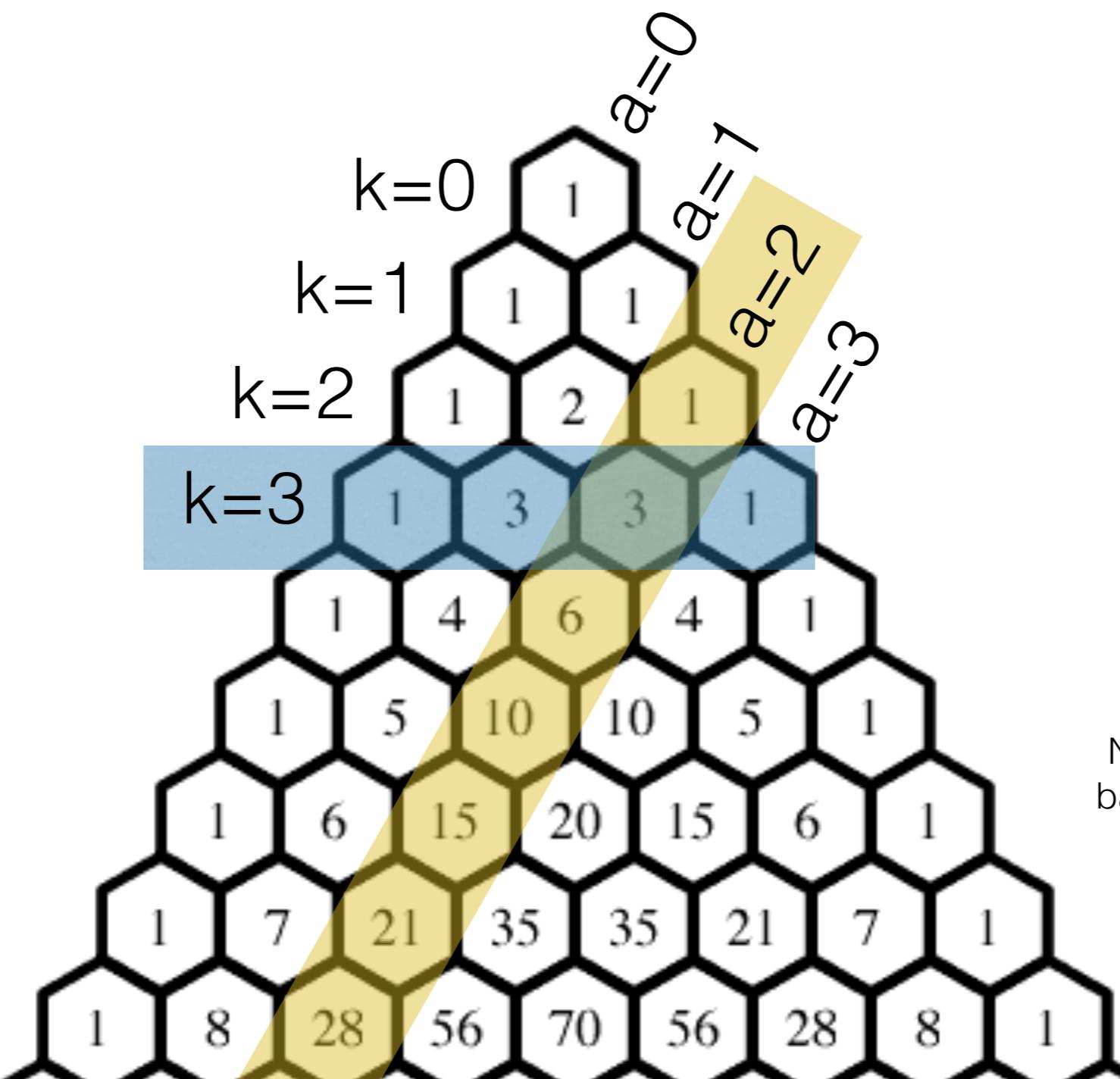
Permuting k of k things, there are:

- k ways to pick the first thing
- $k-1$ ways to pick the second thing
- $k-2$ ways ...
- 1 way to pick the last thing

But we are only picking a things, not all k of them

And the order of the a things doesn't matter

k Choose a (Pascal's Triangle)



Note, this notation is exactly backwards from the standard mathematical notation, because we had already defined n and k

Clustering coefficient

- Are my friends friends with each other?

Number of edges
between friends

$$C_i = \frac{\text{Number of edges between friends}}{\text{Number of possible edges between friends}} = \frac{n_i}{k \text{ choose } 2} = \frac{2n_i}{k(k-1)}$$

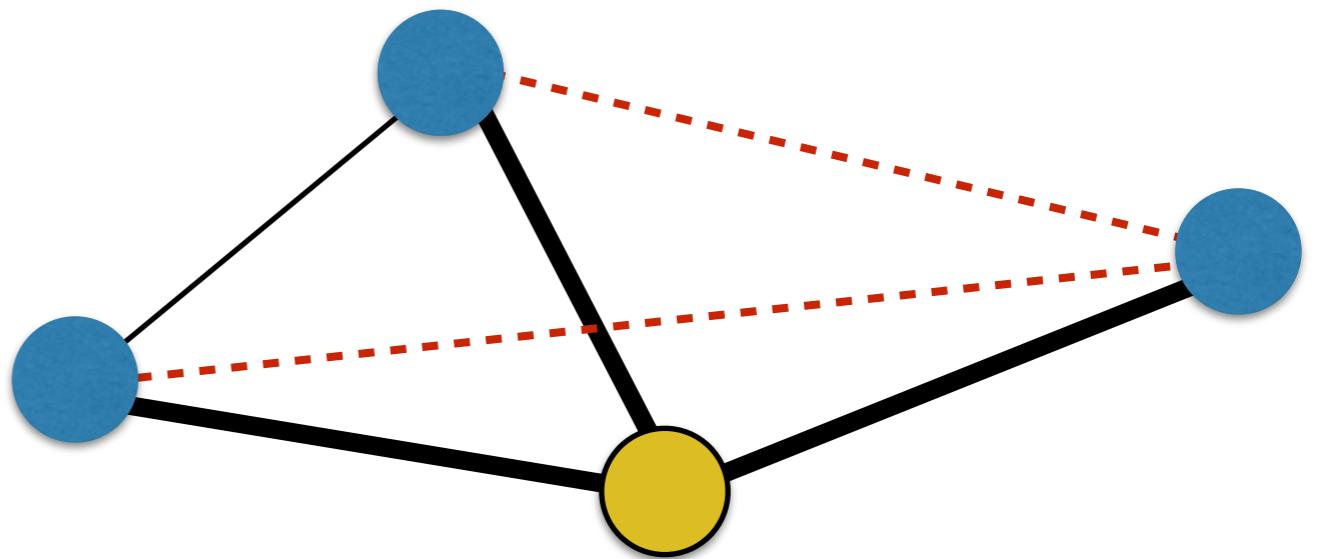
k is the number of edges of node i

n_i is the number of edges between neighbours of node i

Clustering coefficient

- Are my friends friends with each other?

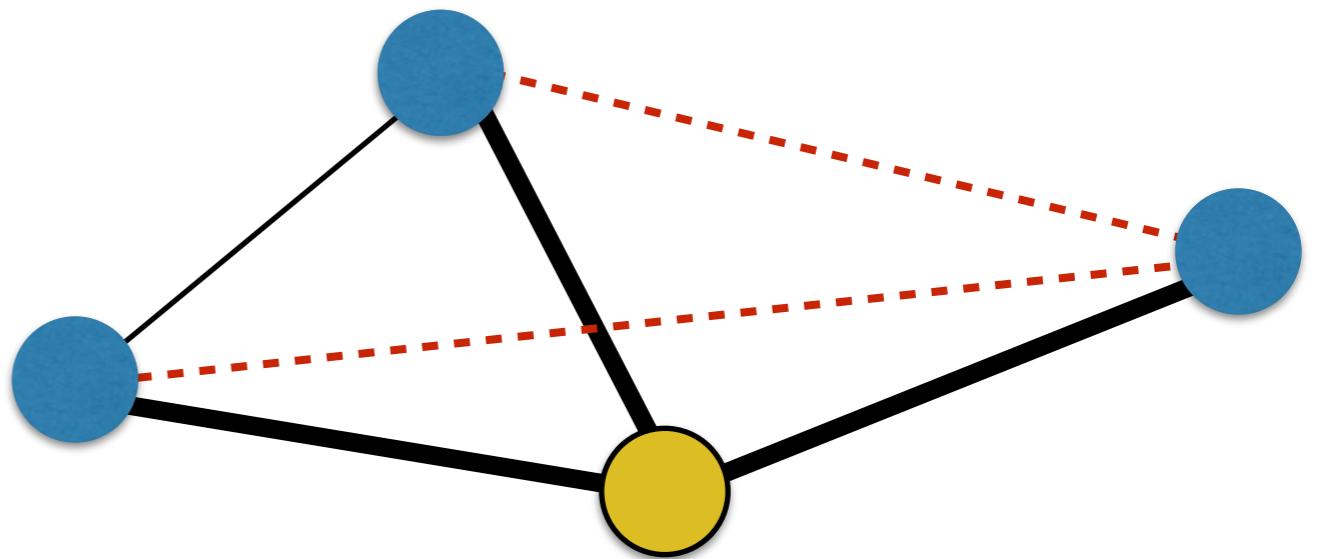
$$C_i = \frac{2n_i}{k(k-1)} =$$

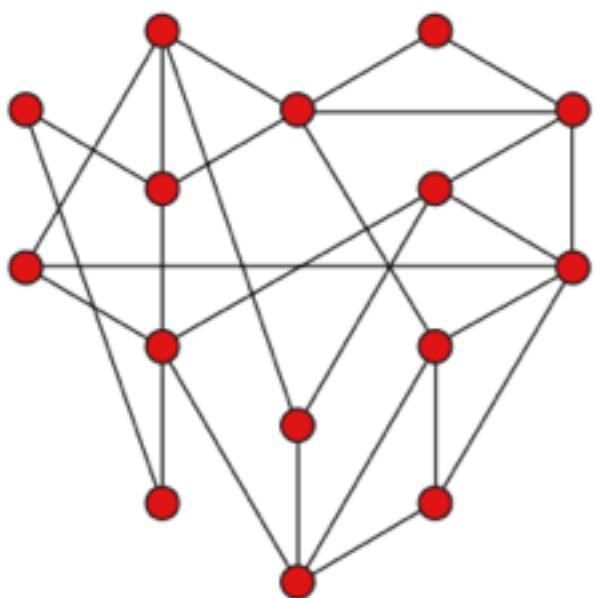
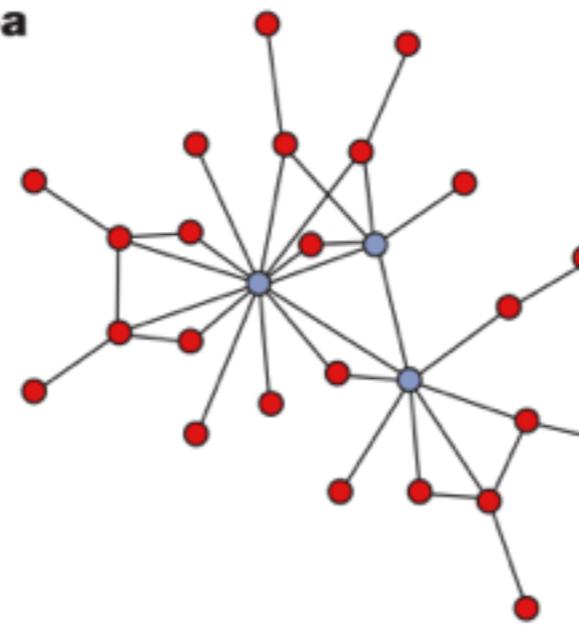
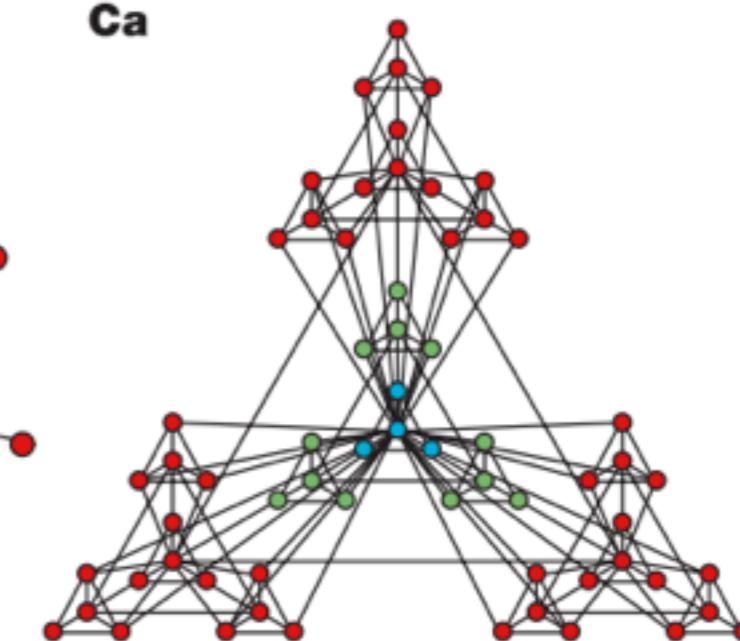
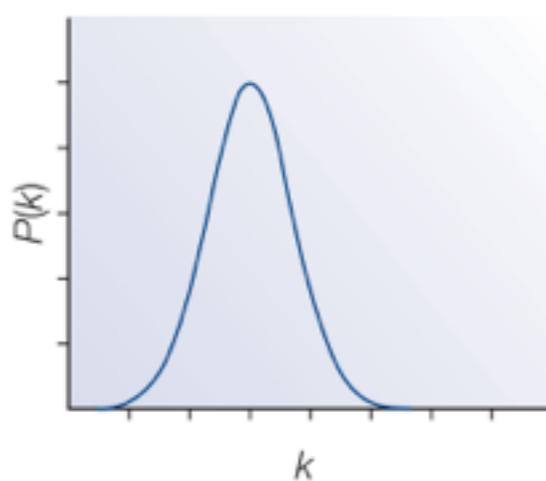


Clustering coefficient

- Are my friends friends with each other?

$$C_i = \frac{2n_i}{k(k-1)} = \frac{2 * 1}{3 * 2} = 1/3$$

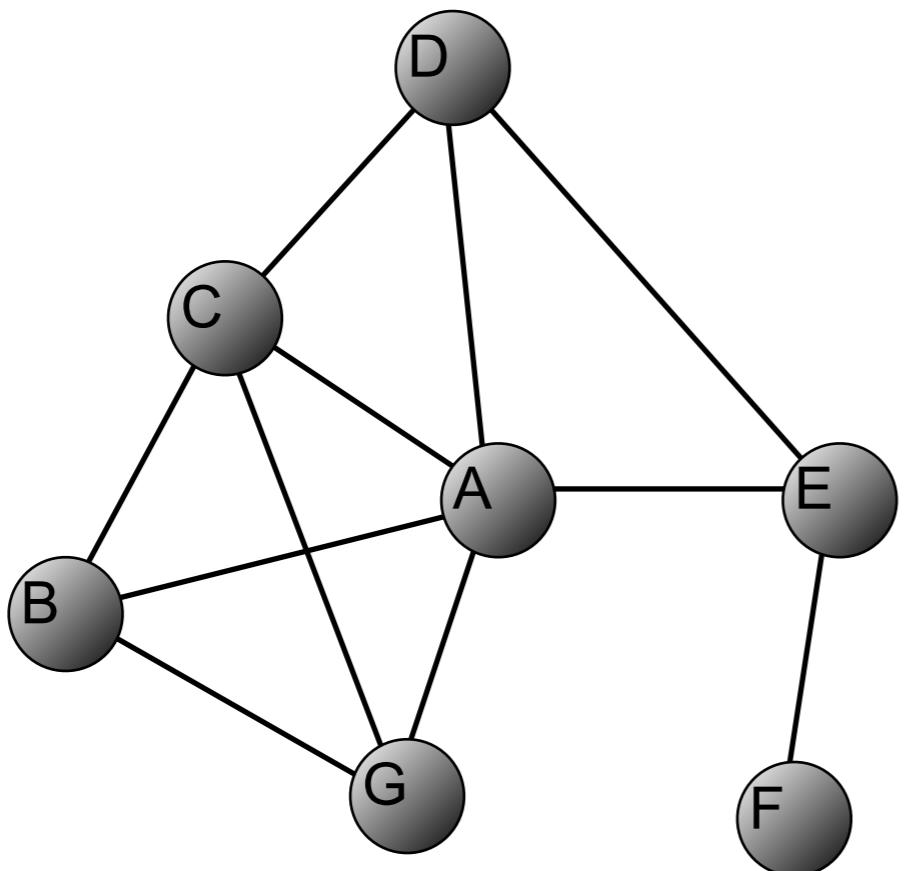


A Random network**Aa****B Scale-free network****Ba****C Hierarchical network****Ca****Ab**

Recap

- Network = nodes + edges
- Degree of a node = # edges that connect to that node
- Scale free networks have a large number of nodes with low degree and few nodes with high degree

Exercise: Network topology measures



Node degree: The number of edges connecting a node to other nodes. Also known as the connectivity.

Clustering Coefficient:

k : neighbors of i

n_i : edges between node i 's neighbors, or number of triangles through node i .