



3000788 Intro to Comp Molec Biol

Lecture 18: Biological networks

Fall 2025



Sira Sriswasdi, PhD

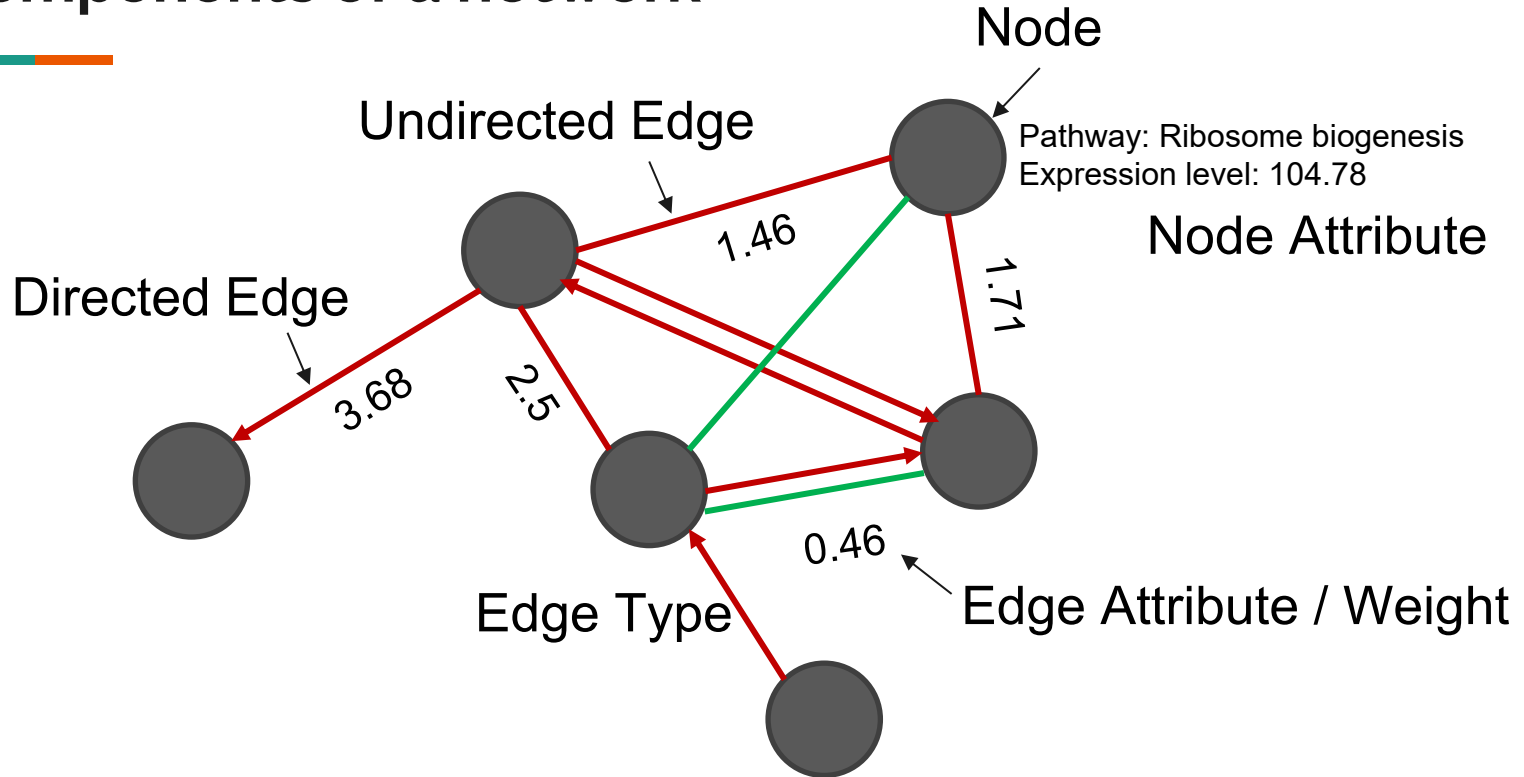
- Research Affairs
- Center of Excellence in Computational Molecular Biology (CMB)
- Center for Artificial Intelligence in Medicine (CU-AIM)

Today's agenda



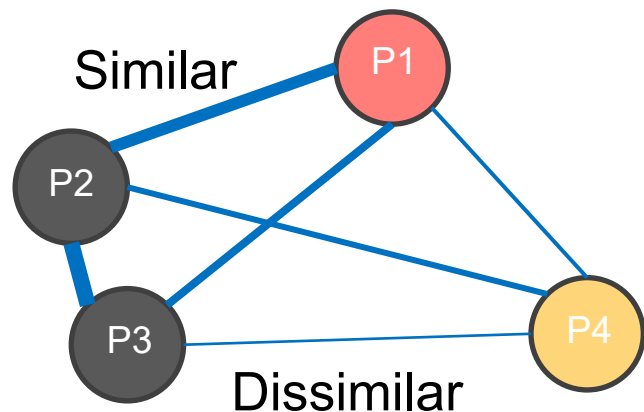
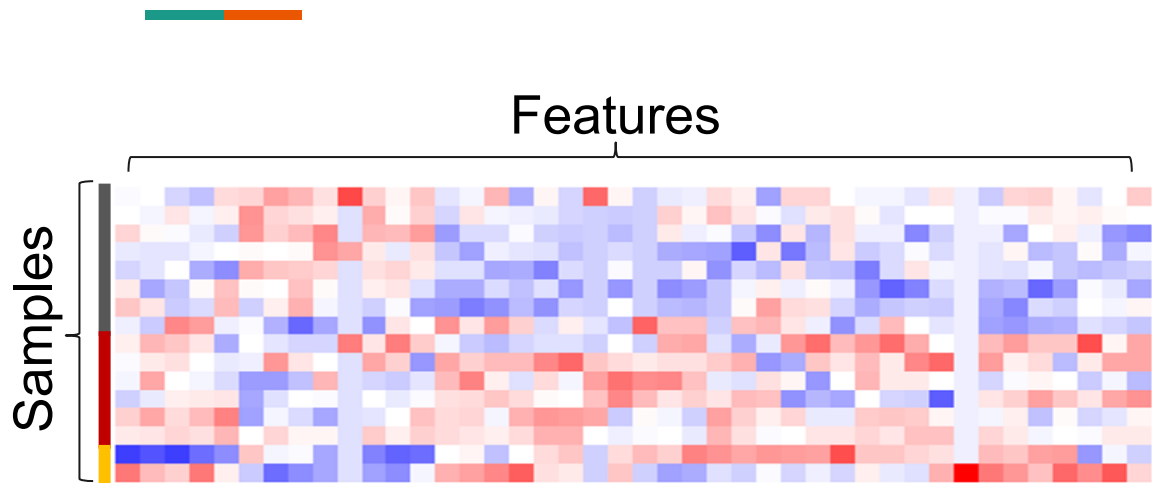
- Biological and other real-world networks
- Properties of a network

Components of a network



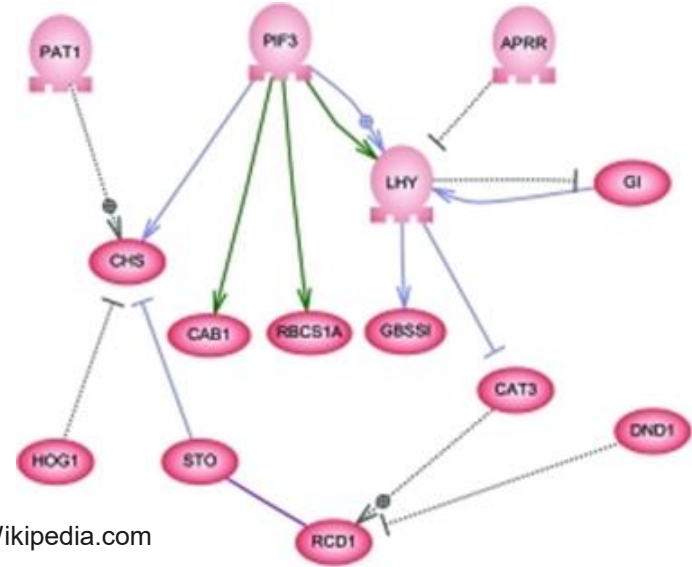
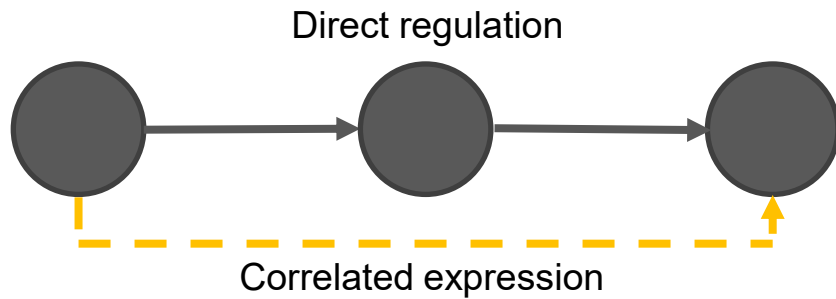
- Describe relationships (edges) between entities (nodes)

Sample-sample similarity as a distance network



- Node = sample
- Edge weight = similarity / distance between samples
- Underlie network clustering and dimensionality reduction approaches

Regulatory and pathway network



Source: Wikipedia.com

- Causation = direct interaction: protein binding, TF-DNA binding, etc.
- Type of interaction: activation, repression

Real-world networks

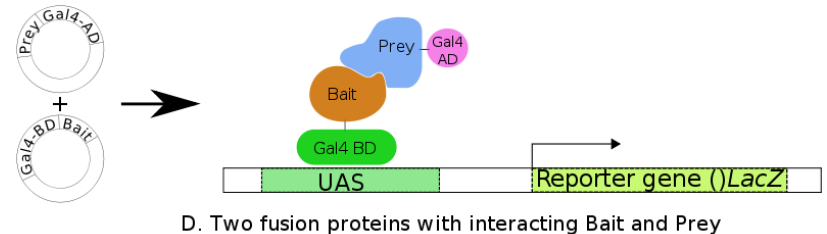
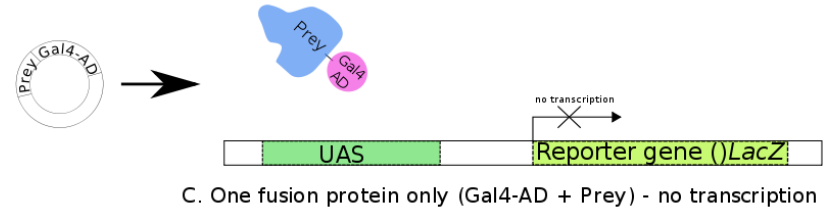
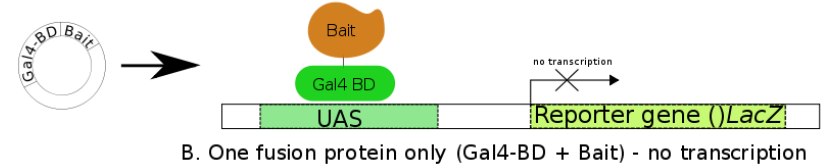
- Computer network
- City-street
- Internet webpages
- Co-authorship
- Friendship
- River & sewage



Image from <https://www.flickr.com/photos/caseorganic/4935751455>

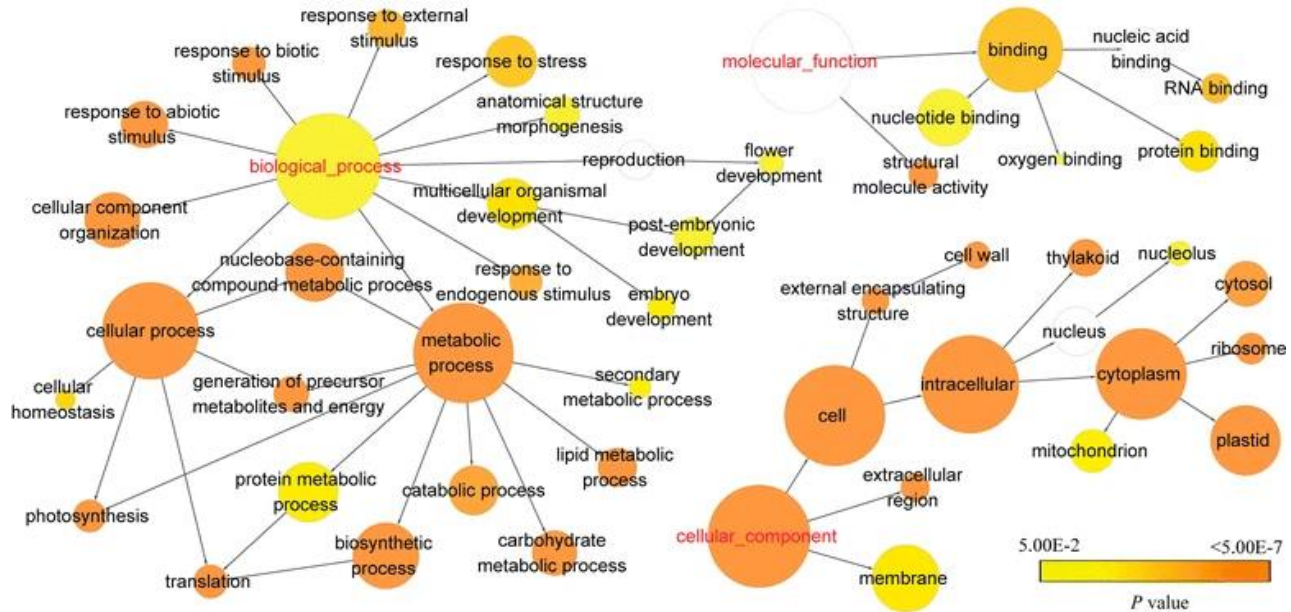
Networks from omics data

- Yeast-2-hybrid → protein-protein interaction networks
- Immunoprecipitation-MS → protein complex
- ChIP-seq → TF-gene regulatory network
- RNA-seq → gene co-expression network



Source: Wikipedia.com

Gene ontology network



Gao, B. et al. BMC Genomics 16:416 (2015)

- Visualize groups of related terms with enrichment scores

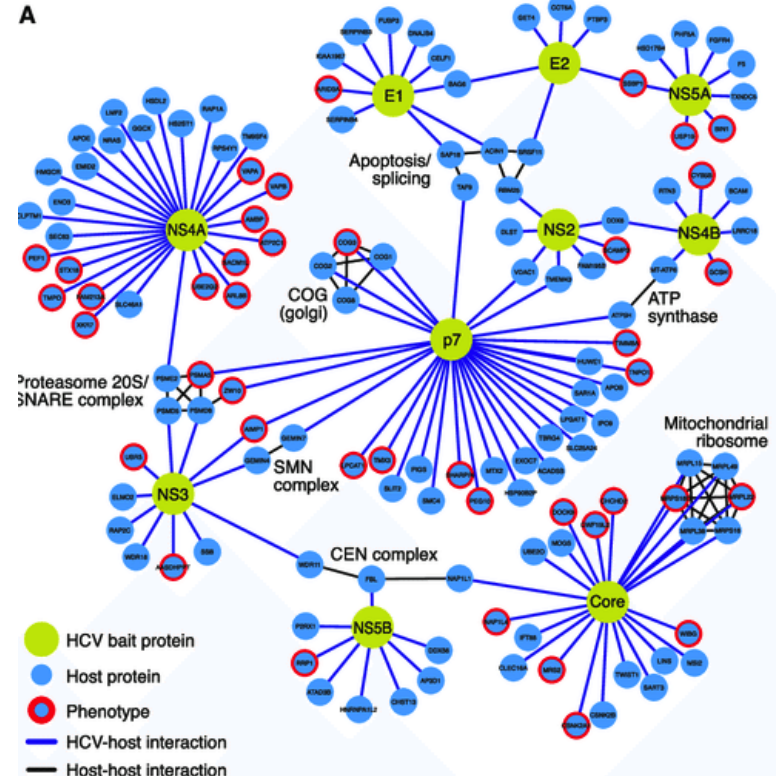
The Connectivity Map

The Connectivity Map (CMap) at the Broad Institute is creating a genome-scale library of cellular signatures that catalogs transcriptional responses to chemical, genetic, and disease perturbation. To date, the library contains more than 1 Million profiles resulting from perturbations of multiple cell types.

- >1M gene expression profiles of cell lines with various disease states and treated with various small molecules
- Use transcriptome similarity to characterize effect of new drugs

Host-viral protein interaction

- Two node types
- Two edge types
- Node attribute: affected by infection
- Propose mechanisms underlying the effect of infection
- Prioritize targets for antibody design

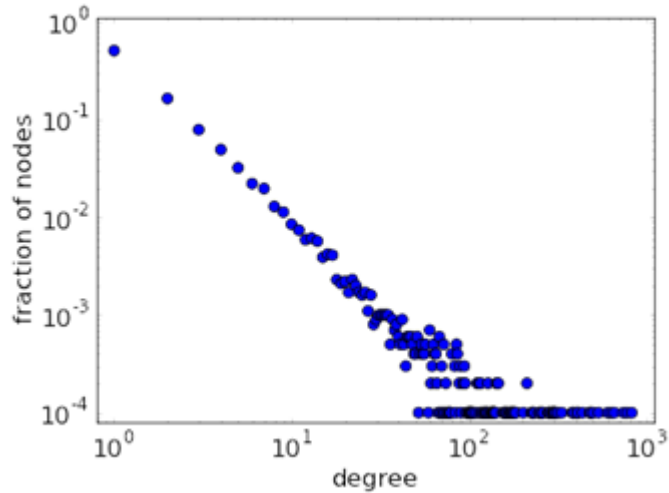


Source: Ramage et al. Mol Cell (2015)

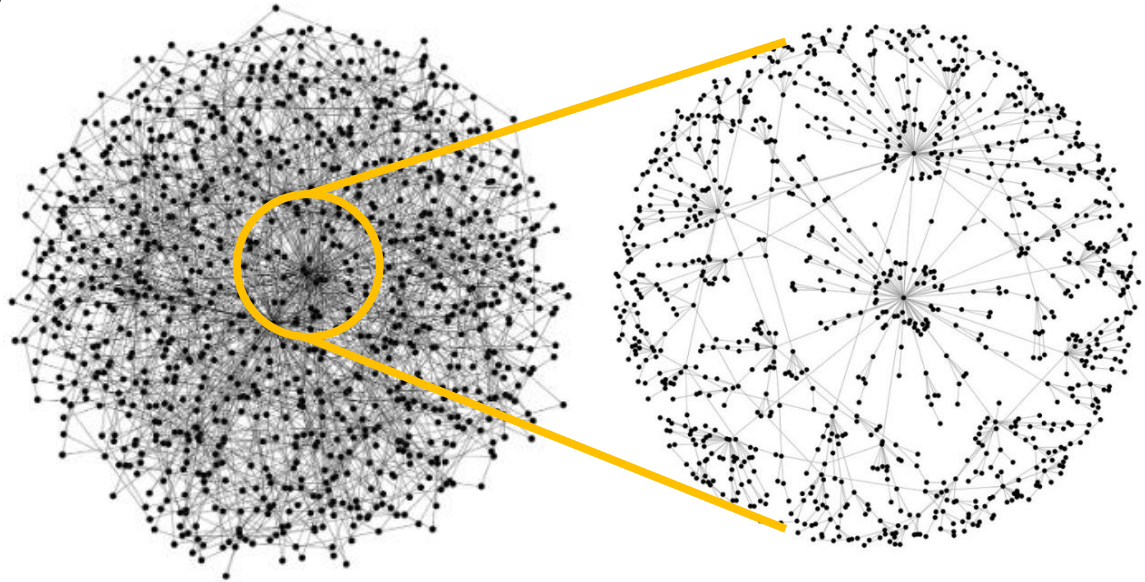


Properties of real-world networks

Scale-free property



Source: mathinsight.org

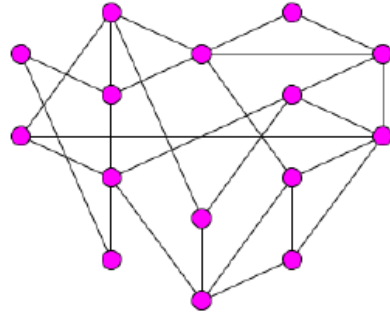
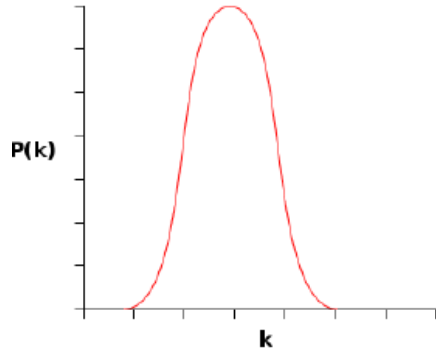


Source: flickr.com & pacojariego.me

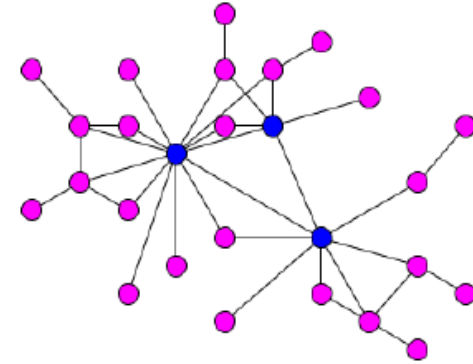
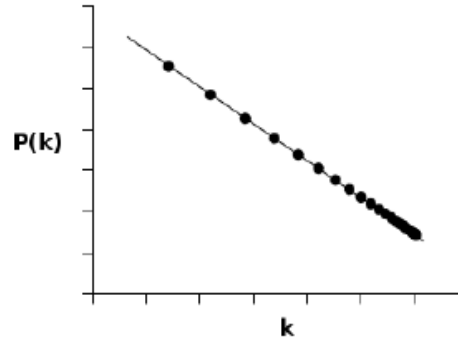
- **Power law:** Proportion of a node connected to k edges $\sim 1 / k^n$
- Same local structure as global structure when zoomed in
 - Proportion of nodes connected to $k\%$ of other nodes

Small-world property (existence of hubs)

(a) Random Network



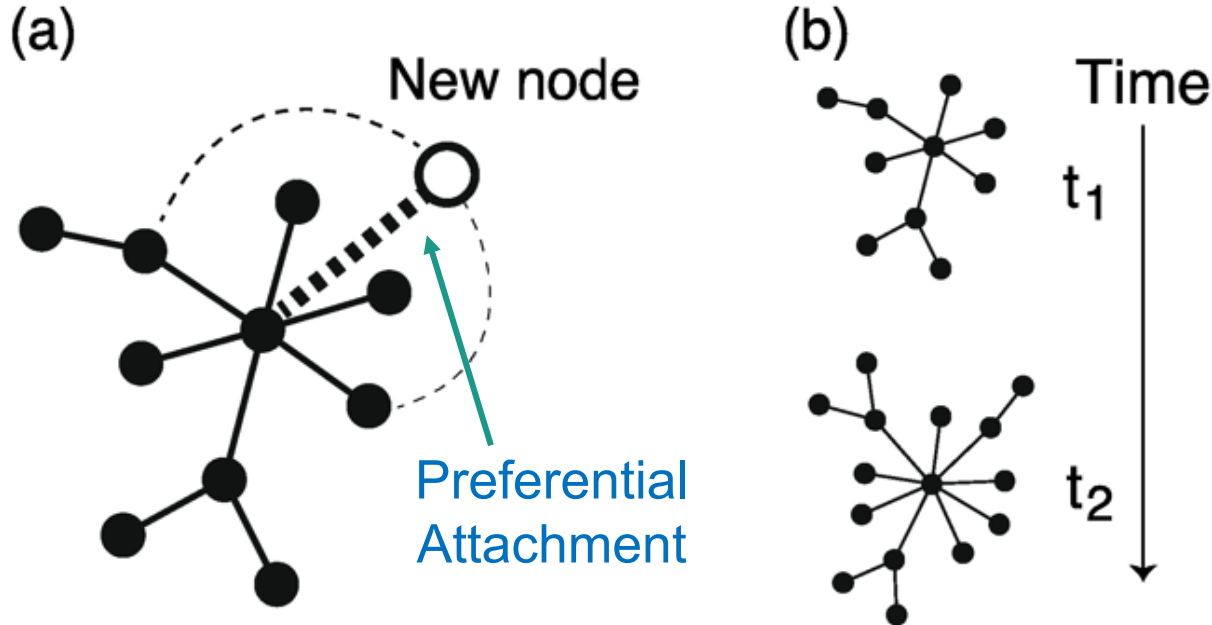
(b) Scale-free Network



Source: Segura-Cabrera *et al.* Analysis of Protein Interaction Networks to Prioritize Drug Targets of Neglected-Diseases Pathogens

- Highly-connected hub nodes act as shortcut for traffic, signaling
 - One transcription factor regulates many genes
 - Major city's airport
 - Social media influencers on the internet

Preferential attachment: A model of real-world networks



Takemoto. Metabolites, 2:429-457 (2012)

- New nodes preferentially attach to existing nodes with many edges

Rationale for preferential attachment

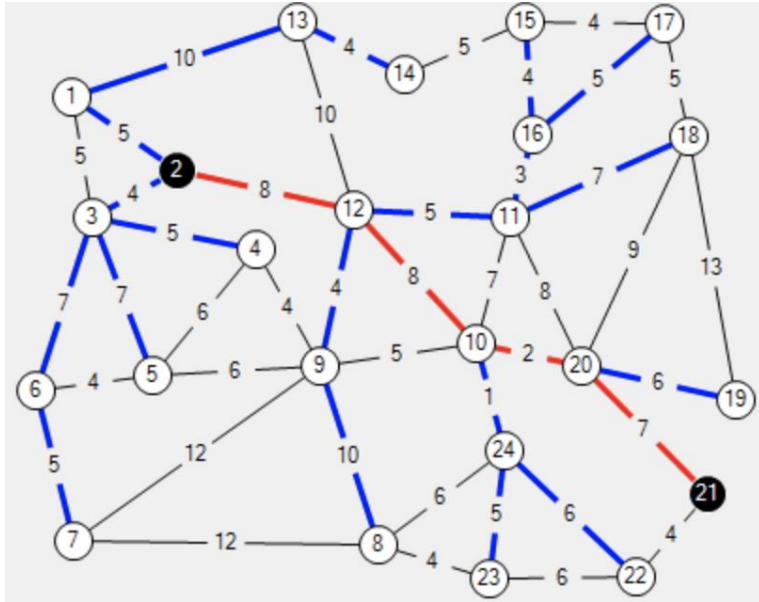


- Transcription factors that can bind to many locations on the genome have more chance to regulate newly emerged genes
- More airlines are likely to fly to existing major cities
- More people are likely to follow an already famous influencers
- Consequences
 - **Small-world:** Can connect any two nodes through a few hubs
 - **Scale-free:** There are global hubs and local hubs



Network connectivity measures

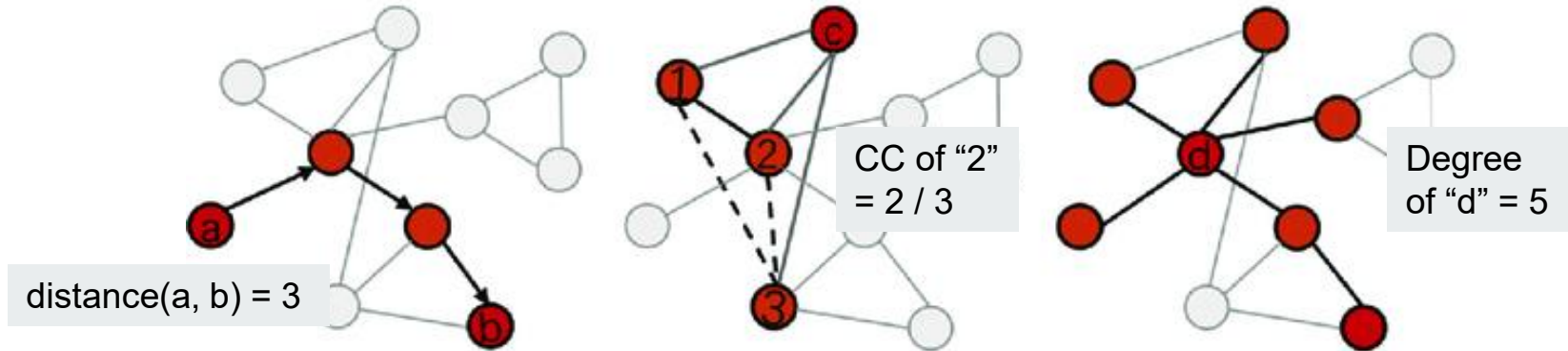
Importance of connectivity



Source: CsharpHelper.com

- Network operates by transmitting signals from one node to another
- Perturbation of a node or edge can propagate through the network
- **Important nodes and edges**
 - Highly connected
 - Involved in paths between other nodes
 - Located in the “center” of the network

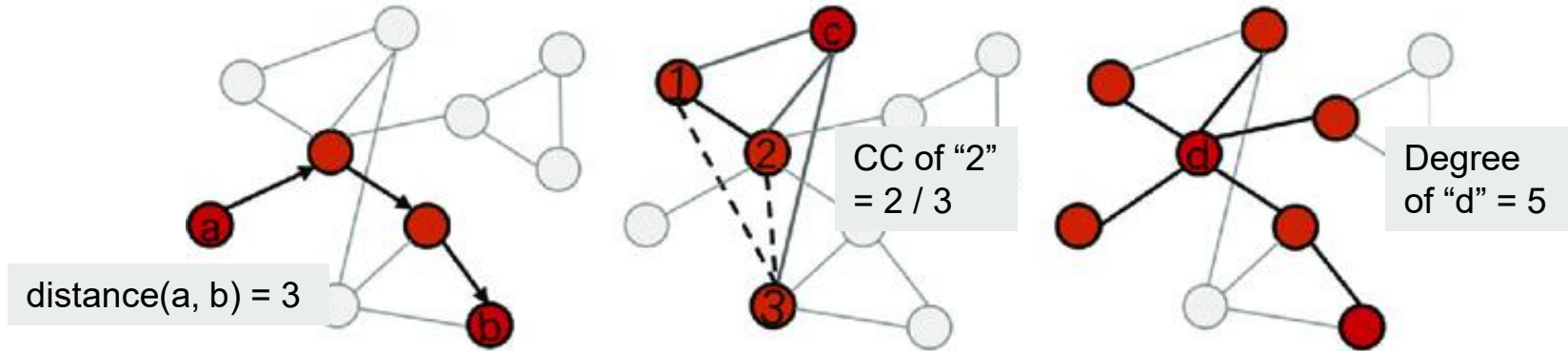
Degree, clustering coefficient, and path



Cai and Niu. Dev Cog Neuroscience (2018)

- **Degree:** number of edges connected to a node
- **Clustering coefficient:** proportion of neighbors of a node that are also connected themselves
 - Indicate the extent of local connectivity / redundancy of the network
- **Path:** (shortest) connection between node through edges

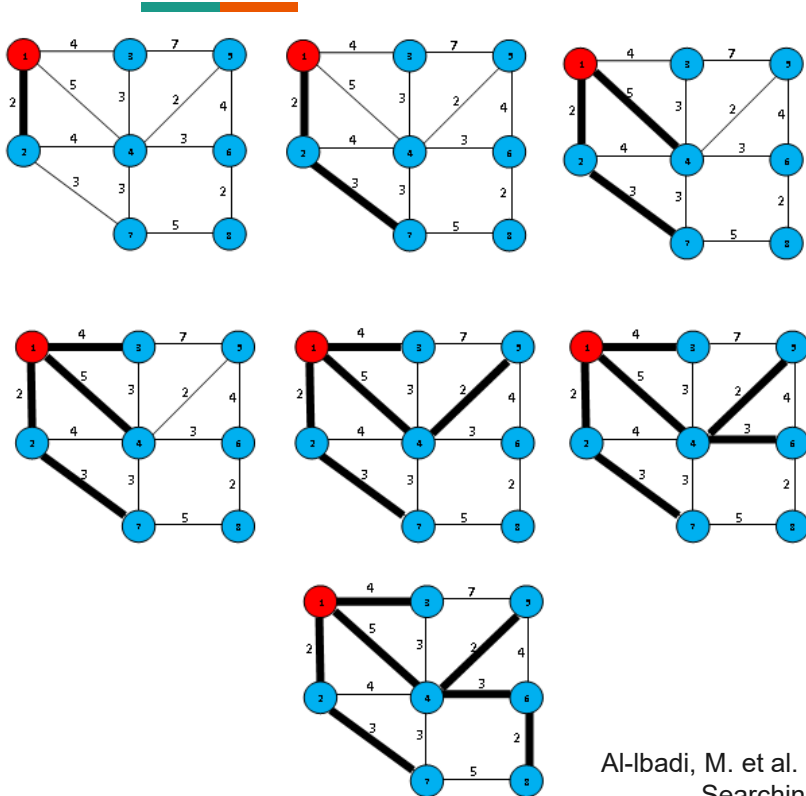
Interpretation of basic properties



Cai and Niu. Dev Cog Neuroscience (2018)

- **High degree** nodes are hubs that interact with many partners
- **High clustering coefficient** implies redundant connectivity, could be resilient to perturbation or network disruption
- **Short path** means signal & effect from one node can propagate to another

Dijkstra algorithm for shortest path calculation



- Dynamic programming approach
- From a start node, traverse the edges and record current distances for visited nodes
- Update distance $d(i, j)$ between nodes i and j if a shorter path is found

Network connectivity via signal flows

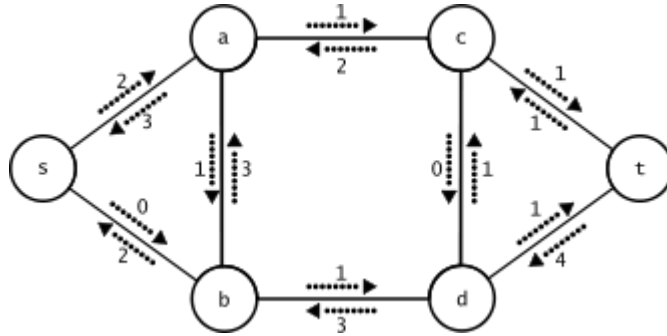


Image from https://en.wikipedia.org/wiki/Flow_network

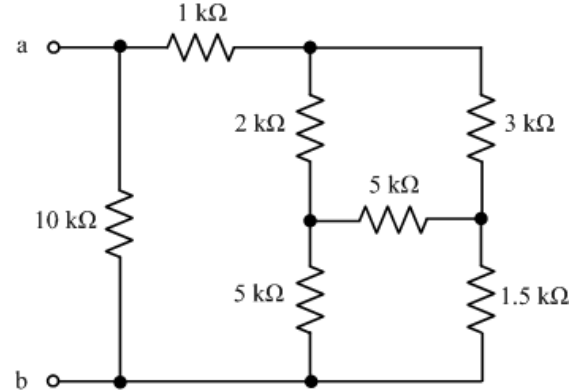
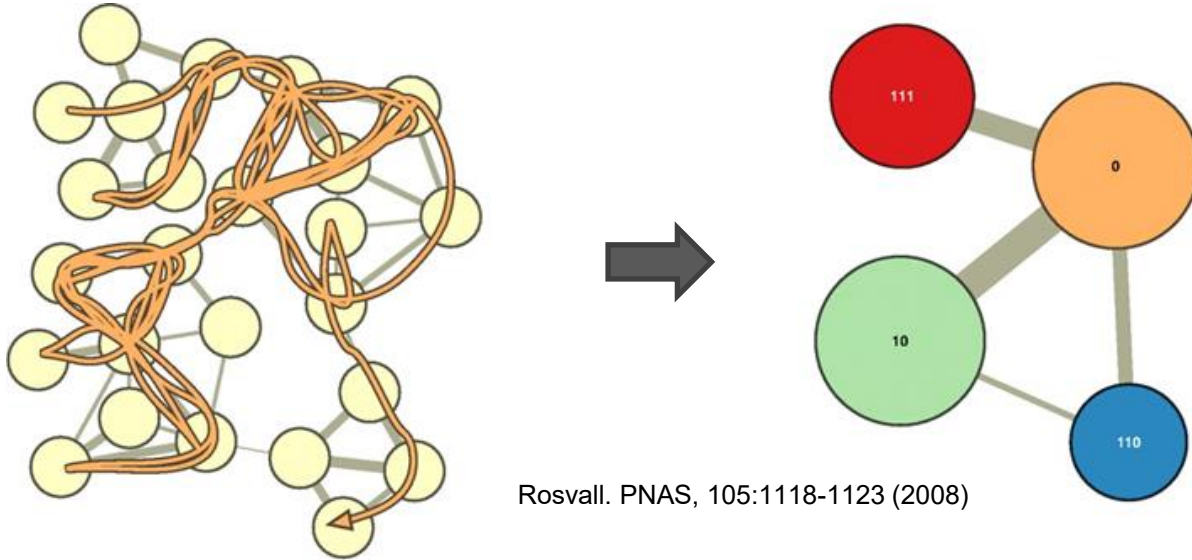


Image from <http://www.rose-hulman.edu/CLEO/browse/?path=1/2/79/91/92/19>

- If 1M liters of water is poured into a node, how much water will be present at each node at equilibrium?
- If 100 voltage is applied between two specific nodes, how much electric current will flow through each edge?

Network connectivity via random walks



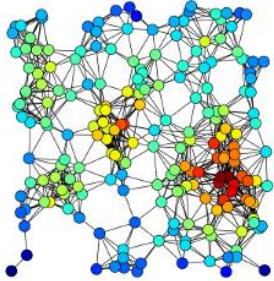
Rosvall. PNAS, 105:1118-1123 (2008)

- Particles can travel from node to node with probability (edge weight)
- How frequent is a node or an edge visited?
- Are particles stuck in some local areas? → Network clustering

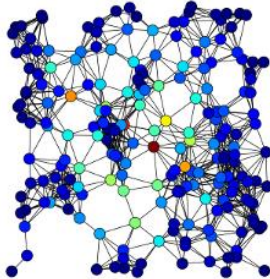
Centrality scores



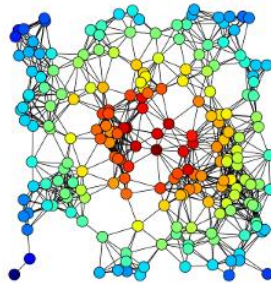
Degree



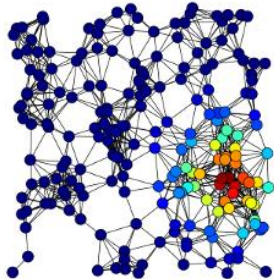
Betweenness



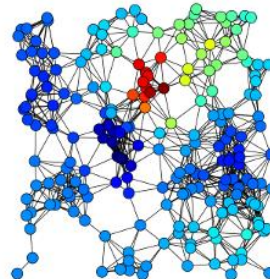
Closeness



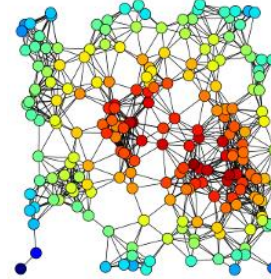
Eigenvector



Katz



Harmonic



- Indicate the importance of a node in the context of the connectivity of the network

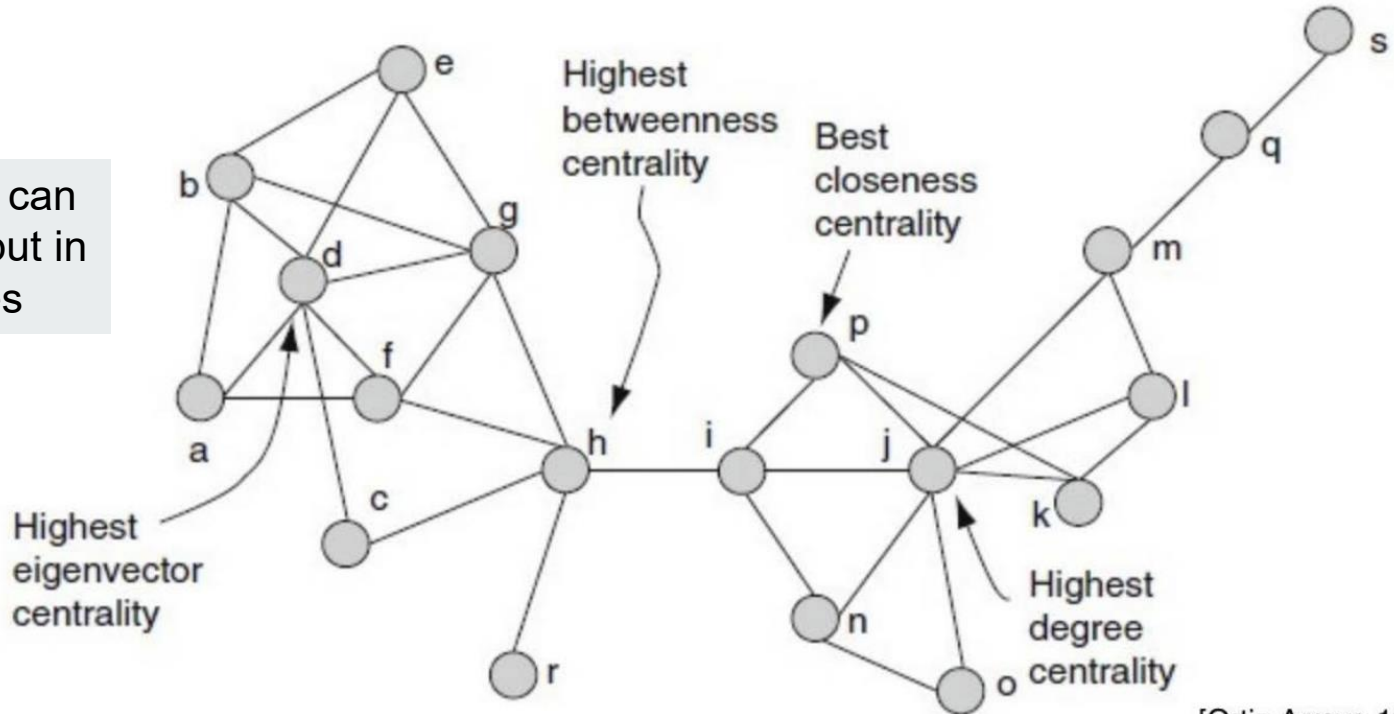
- **Degree:** local connectivity

- **Betweenness:** fraction of shortest paths involved

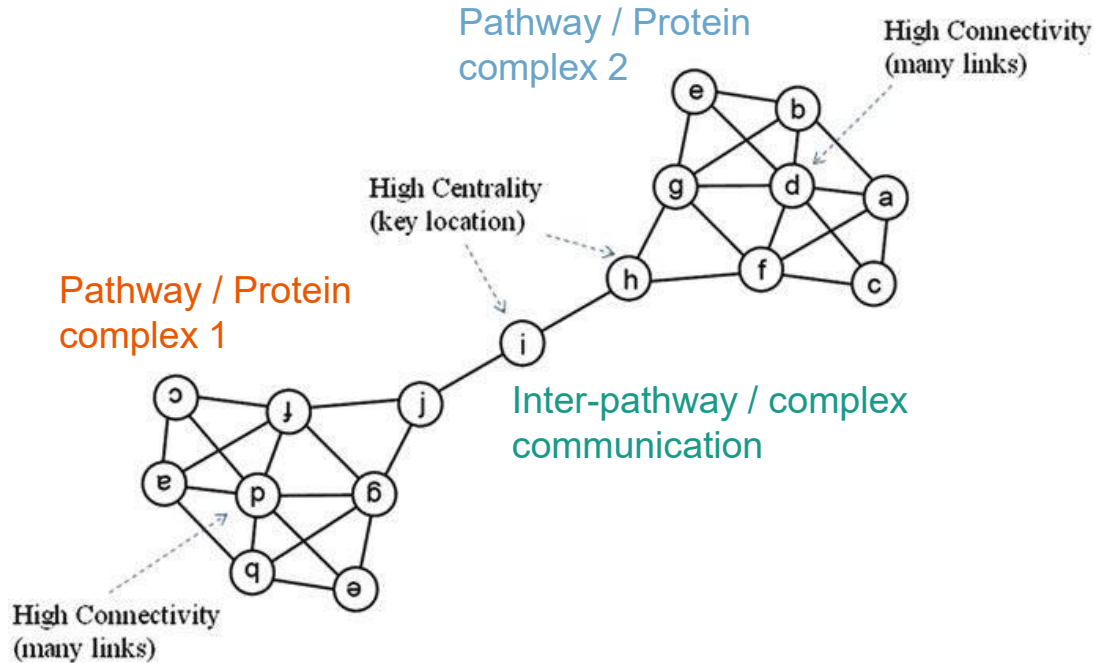
- **Closeness:** average distance to other nodes

Different meaning of centrality scores

Multiple nodes can be important, but in different senses



A biological interpretation of centrality



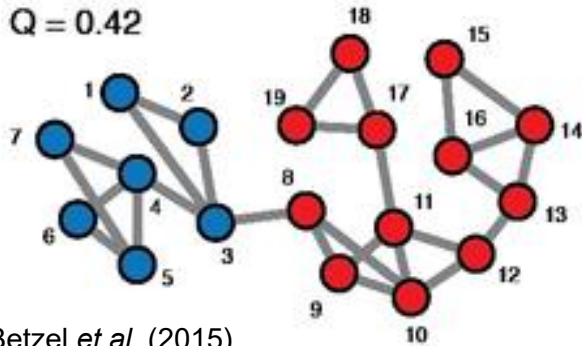
- **Low degree, High betweenness nodes** may transmit signal across functional pathways
- **High degree, Low betweenness nodes** may be part of a protein complex or transcription factors with multiple downstream targets



Network clustering (recap)

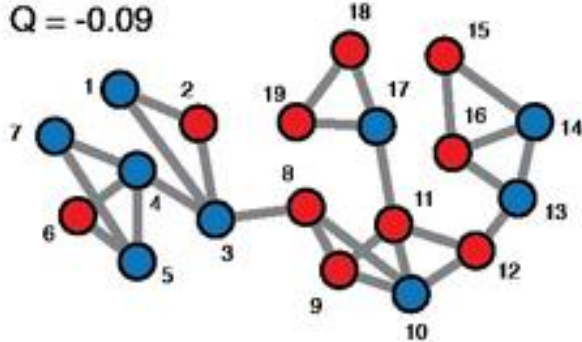
Modularity score

$Q = 0.42$

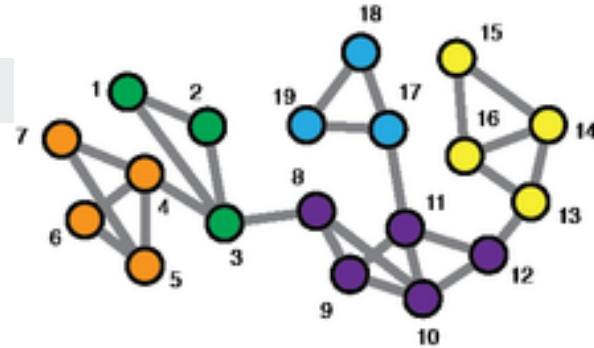


Betzel *et al.* (2015)

$Q = -0.09$



Which is better?



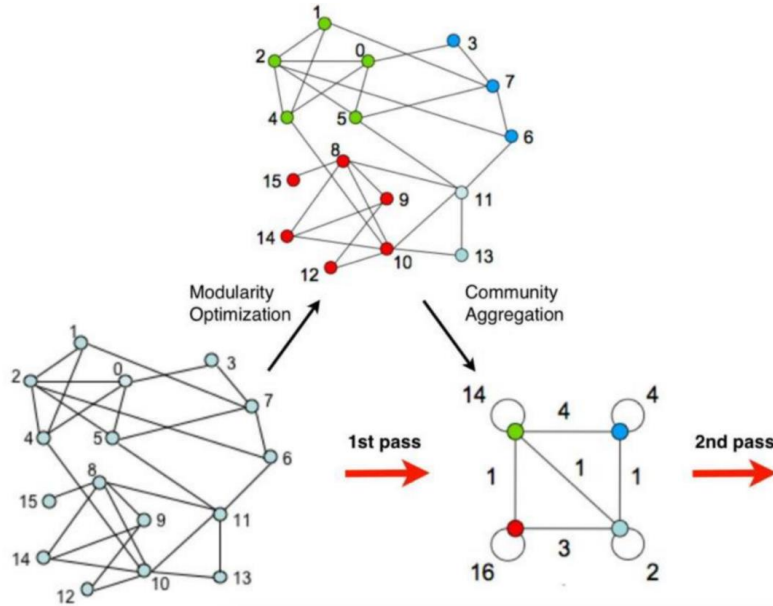
- Number of within-cluster edges compared to expectation (based on number of nodes and global number of edges)
- Depending on the “resolutions”

A variant of modularity score



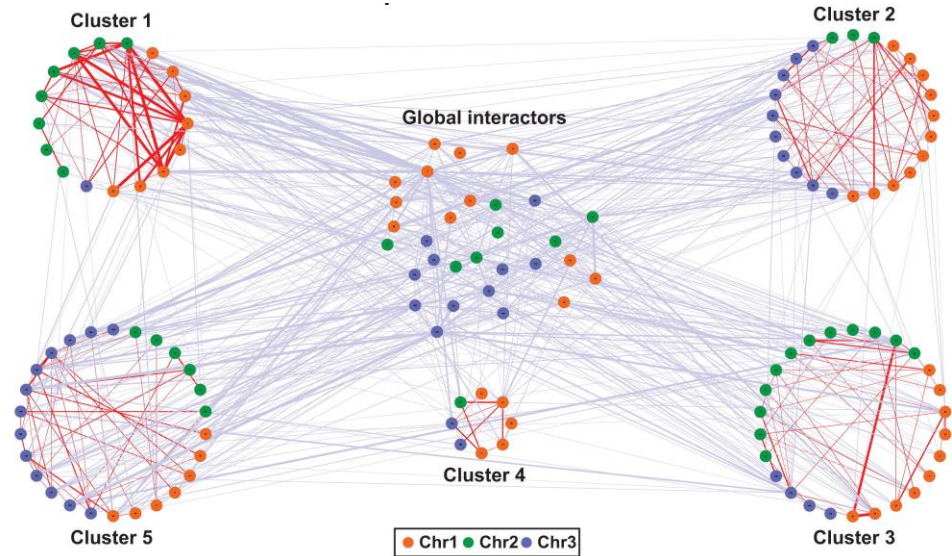
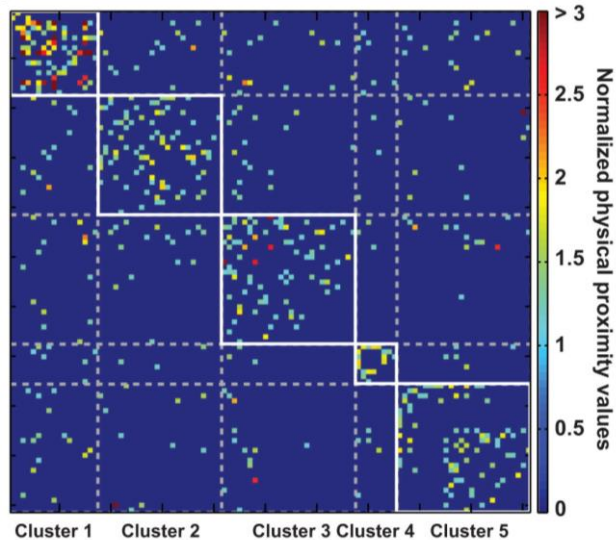
- Node N_i with degree d_i
- Node N_j with degree d_j
- $P(\text{edge between } N_i \text{ and } N_j \text{ by chance}) \approx \frac{d_i d_j}{2|E|}$
 - Proportional to the product of degrees
 - Normalized by the total degree = $2|E|$
- Modularity score of a cluster of nodes (N_1, N_2, \dots, N_n)
 - $Q = \# \text{ within-cluster edges} - \sum_{i,j} \frac{d_i d_j}{2|E|}$

Louvain / Leiden algorithm



- Iteratively partition nodes into clusters that **maximize the modularity score**
- Collapse nodes in each cluster into a representative node to **simplify the network**
- Aggregate representative nodes to further **simplify the clusters**

Network clustering on chromatin interactions



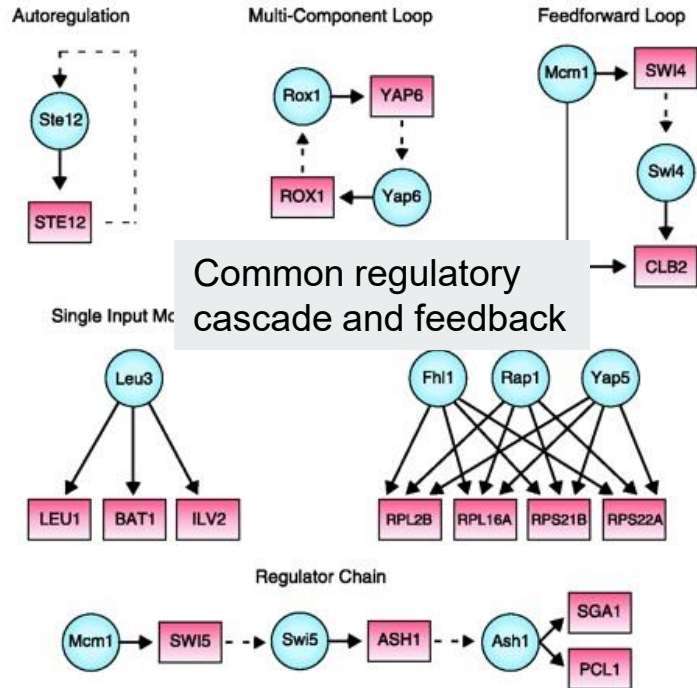
Tanaka *et al.* Molecular Cell, 48:532-46 (2012)

- Identify clusters of genomic loci that are nearby in 3D + global interactors



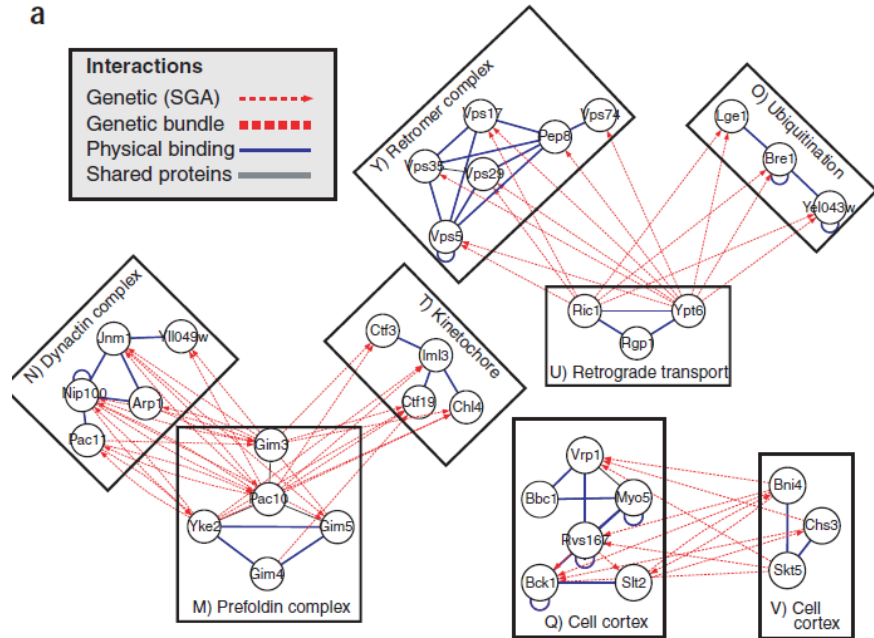
Network motifs

Recurring patterns in network due to biology



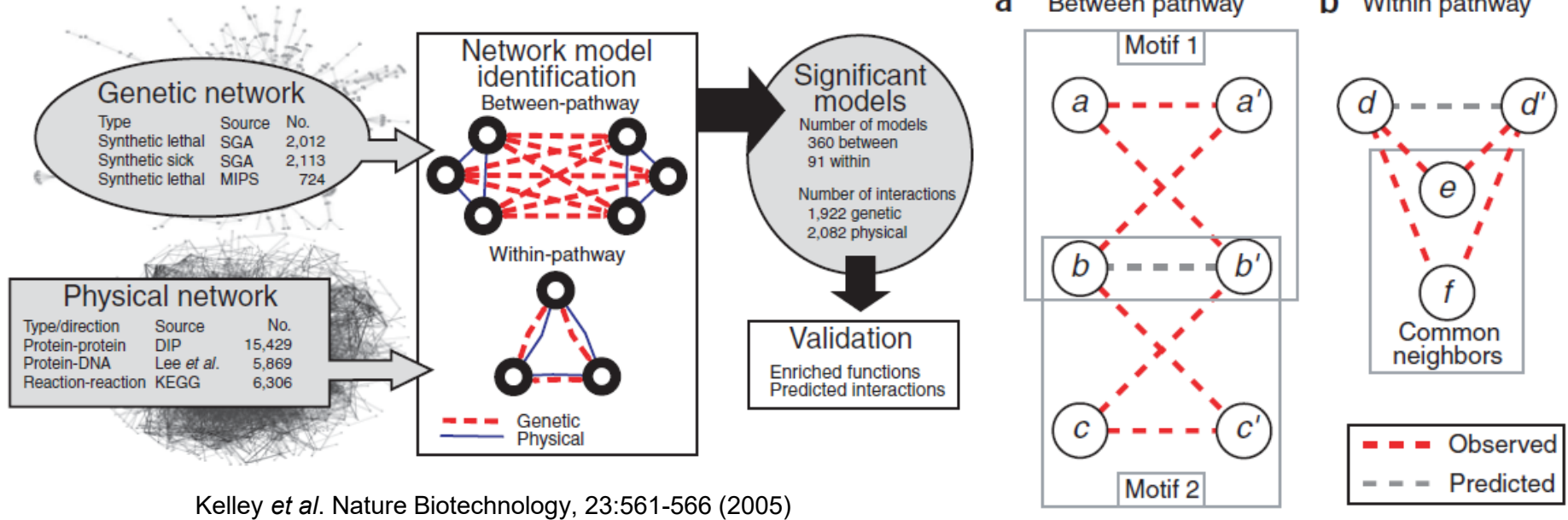
Lee *et al.* Science. 298:799-804 (2002)

Genetic interaction: lethality when disrupted together



Kelley *et al.* Nature Biotechnology, 23:561-566 (2005)

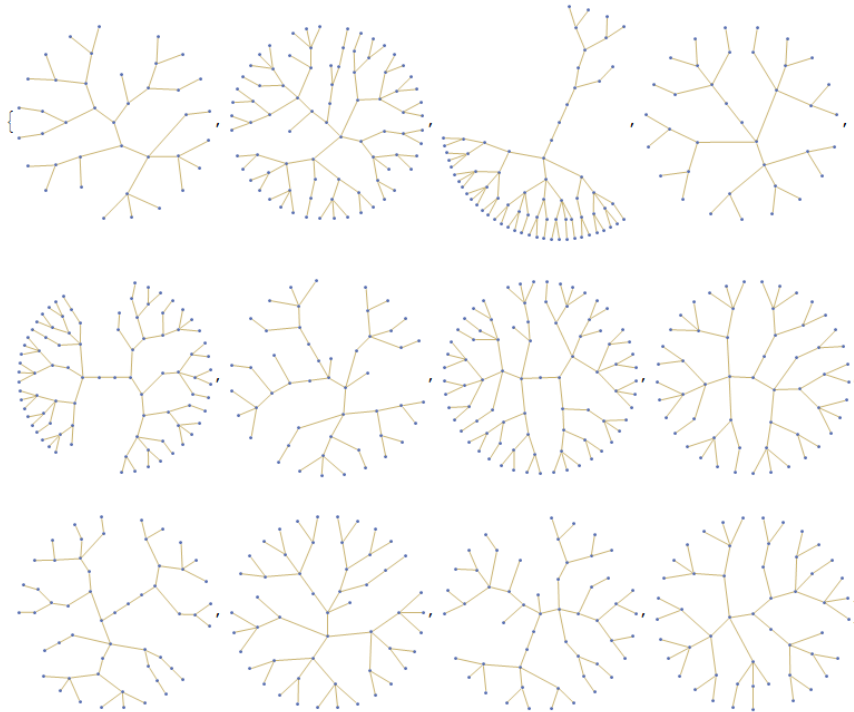
Incomplete motif suggests missing data



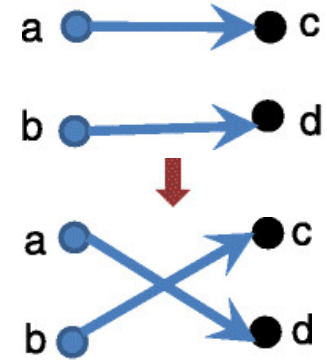
Kelley *et al.* Nature Biotechnology, 23:561-566 (2005)

- **Assumption:** The chance that random associations resemble motifs is low

Can the motif occur by chance?



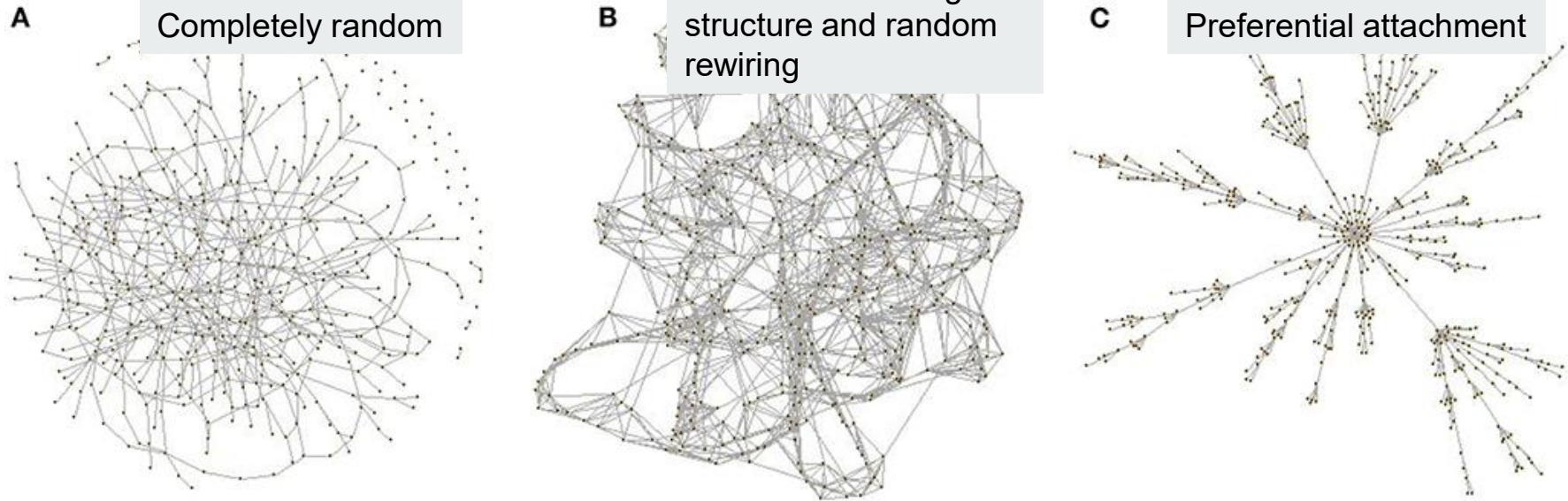
Edge switching



Temate-Tiageru *et al.* BMC Genomics, 17:542 (2016)

- If the motif can occur by chance, they should be frequently observed in random networks

Models for generating random network



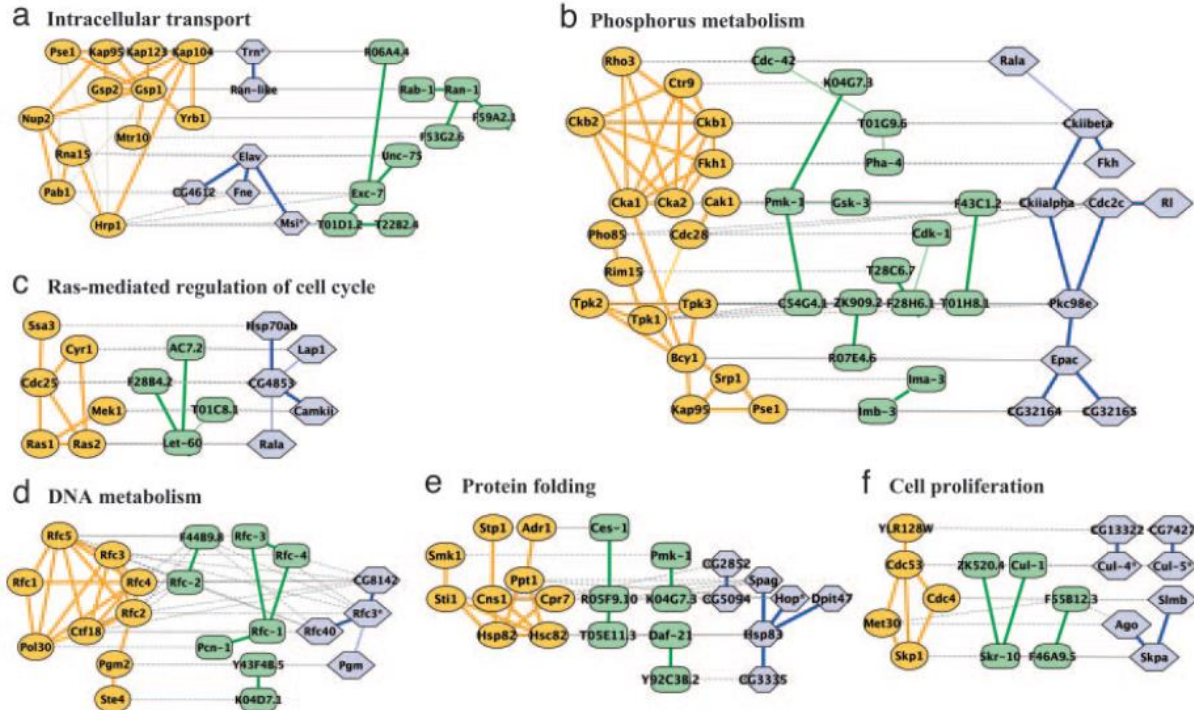
Koutrouli, M. et al. Front. Bioeng. Biotechnol. 31 (2020)

- Erdos-Renyi, Watts-Strogatz, Barabasi-Albert

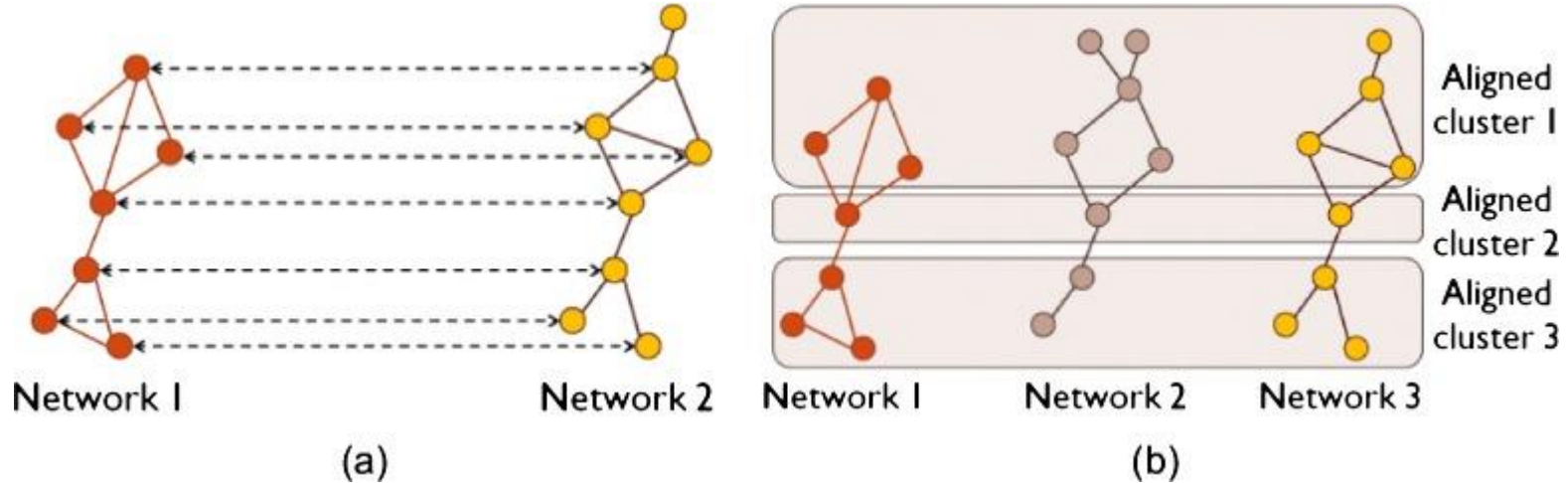


Network alignment

Homology of network structures across species



Evolution as network alterations



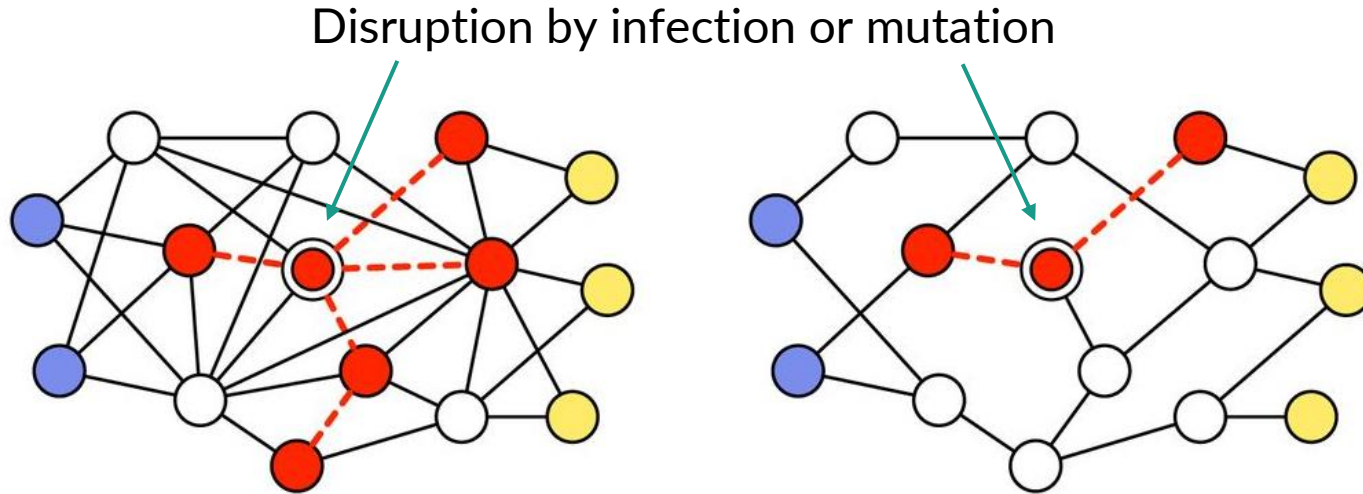
Faisal *et al.* EURASIP J Bioinform Syst Biol (2015)

- Use pathway structures to aid gene alignment
- Identify gene expansion and emergence of new interactions
- Flag missing interactions for validation



Using network to understand the system

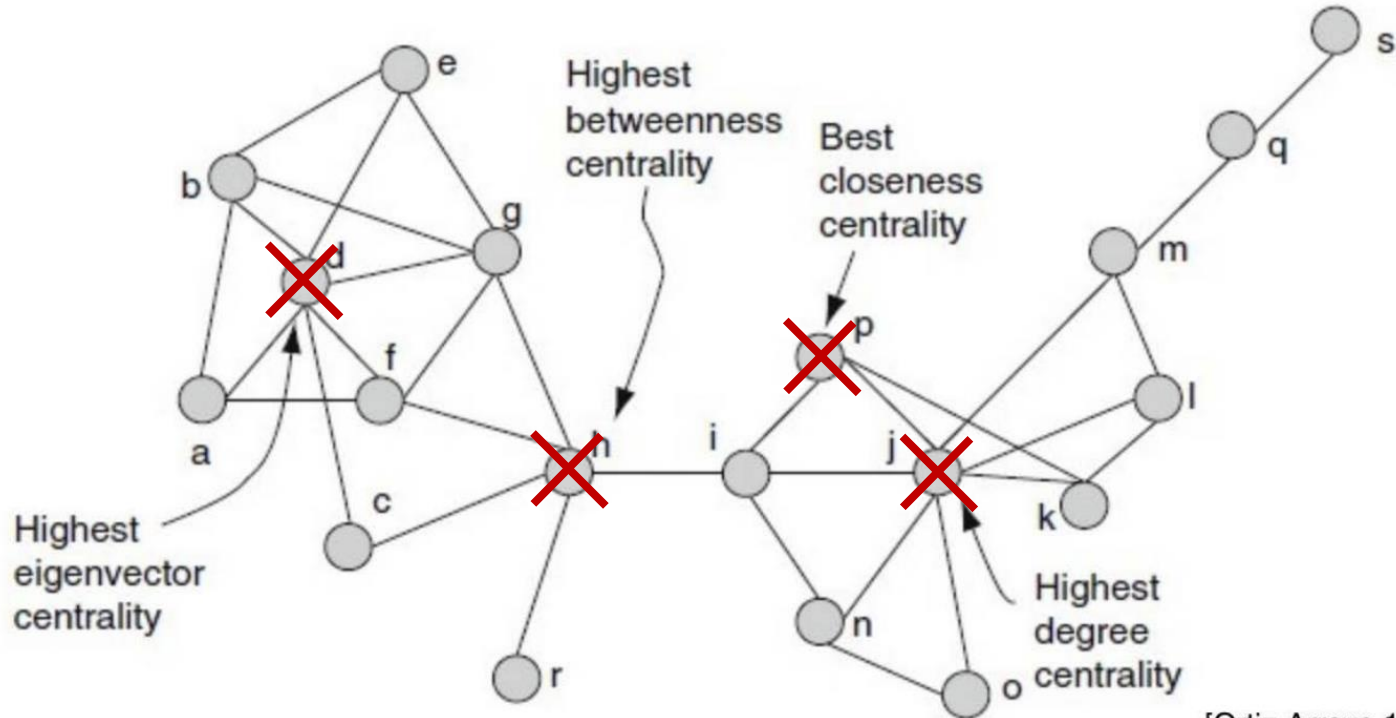
Network perturbation



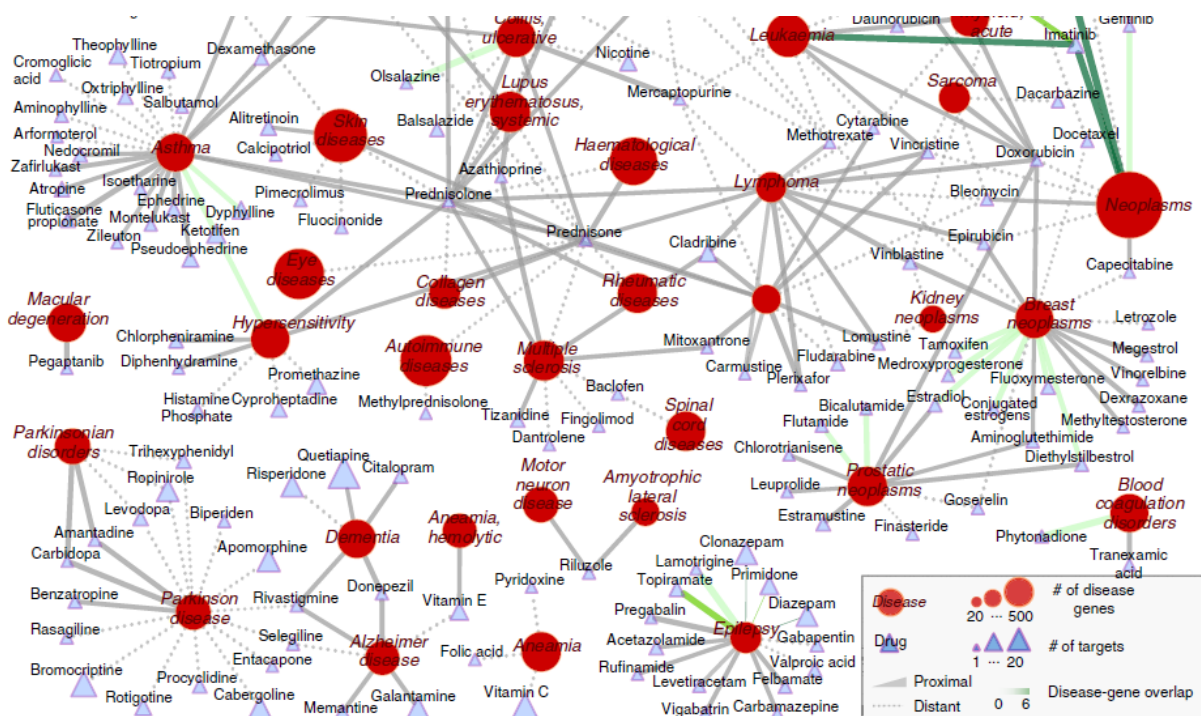
Navlakha *et al.* J of the Royal Society Interface, 11 (2014)

- Analysis of network-level changes induced by node/edge changes
- Complement static centrality scores

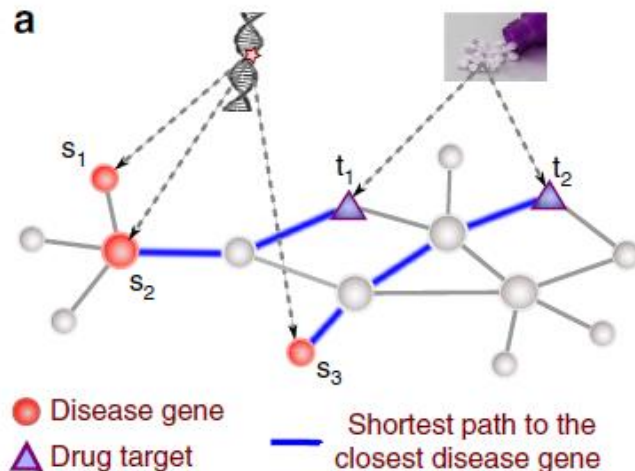
Static properties versus responses to perturbation



Connecting entities through multiple networks



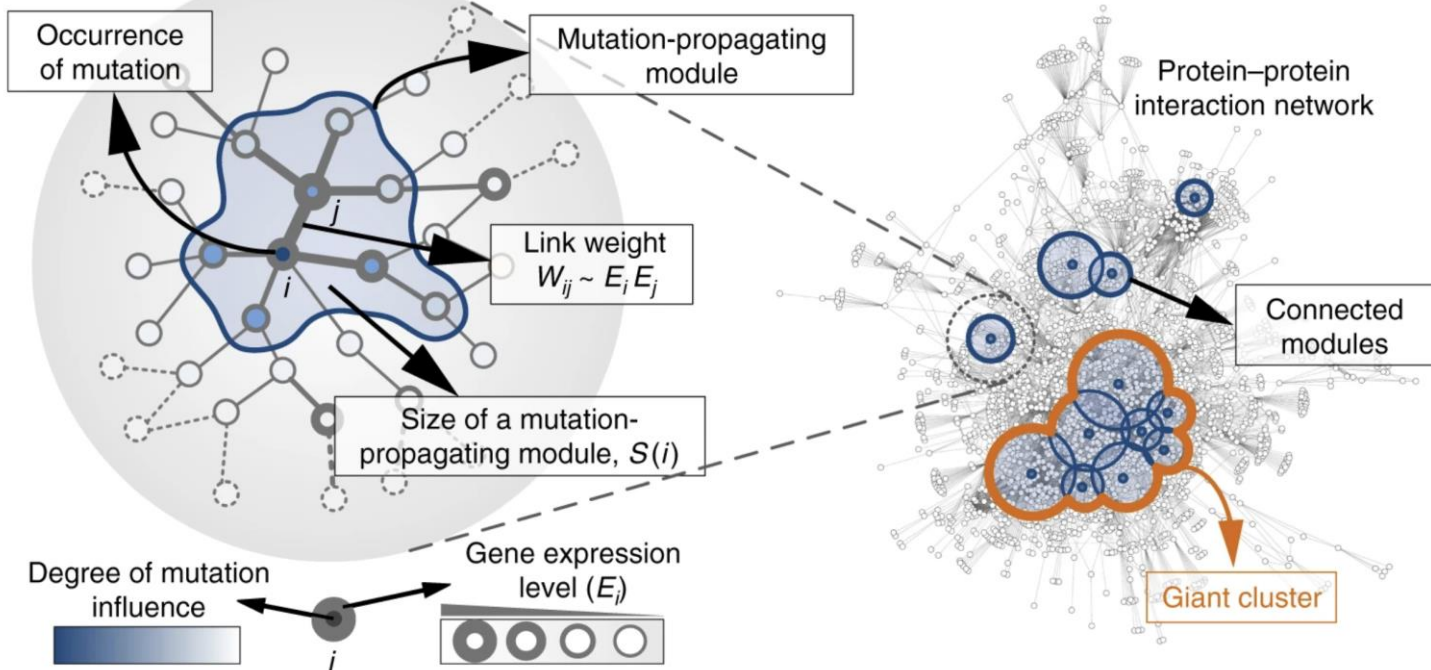
Drug-Gene-Disease



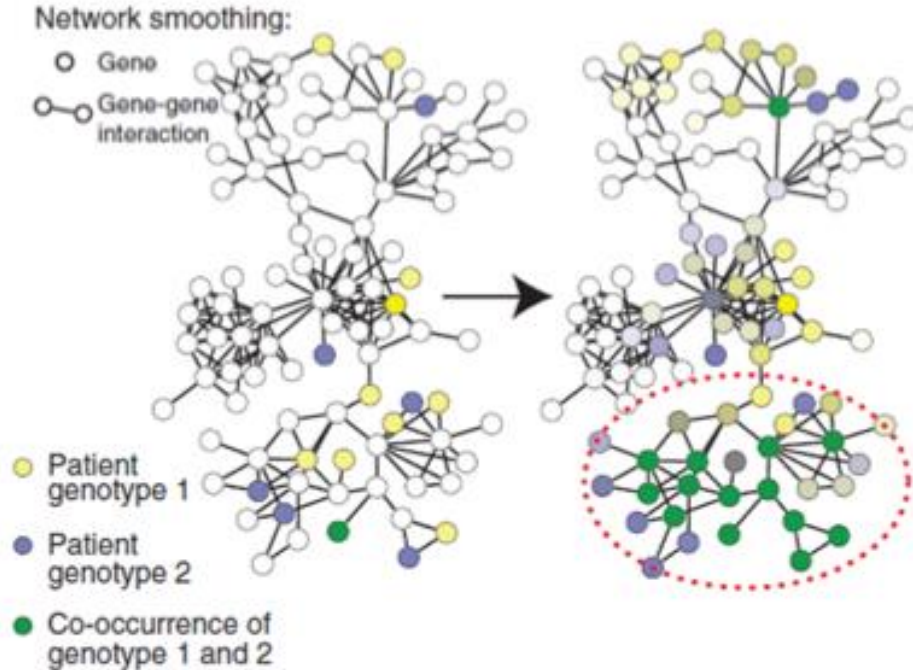
Guney *et al.* Nature Comm, 7:10331 (2016)

Propagate effect of mutations through network

a

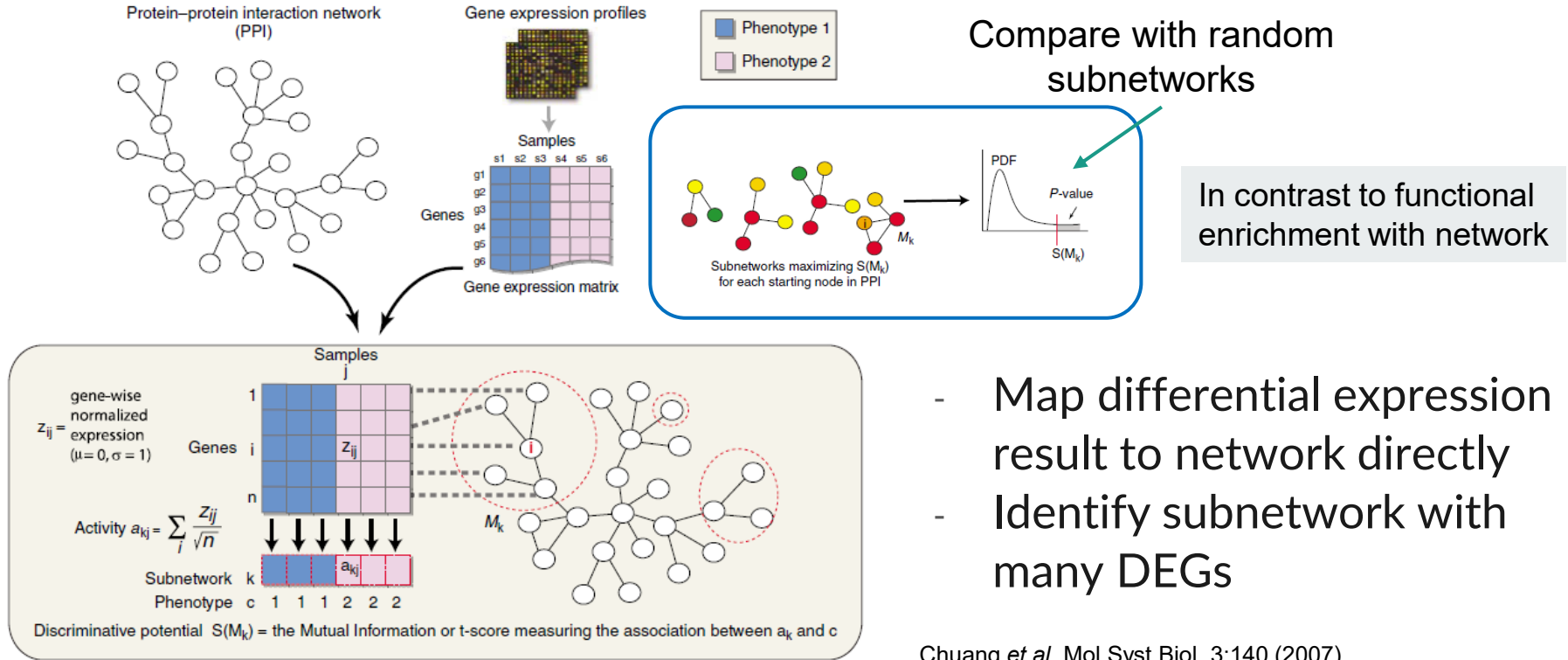


Network-based patient stratification



- No two patients have the same mutation profiles
- Different mutation profiles may have the same impacts on gene-gene network
- Through propagation, identify commonly affected gene-gene subnetworks

Network-based differential expression



Summary



- To understand biological system, we need to study
 - How its components interact, through multi-omics and network diagram
 - How the system behaves over time
- Temporal dynamics can be captured with differential equation models
- Multi-omics integrate mechanistic explanation with data
- Network diagram provides an abstract view of the system
 - Capture consequences of propagation and perturbation

Any question?



- See you next time