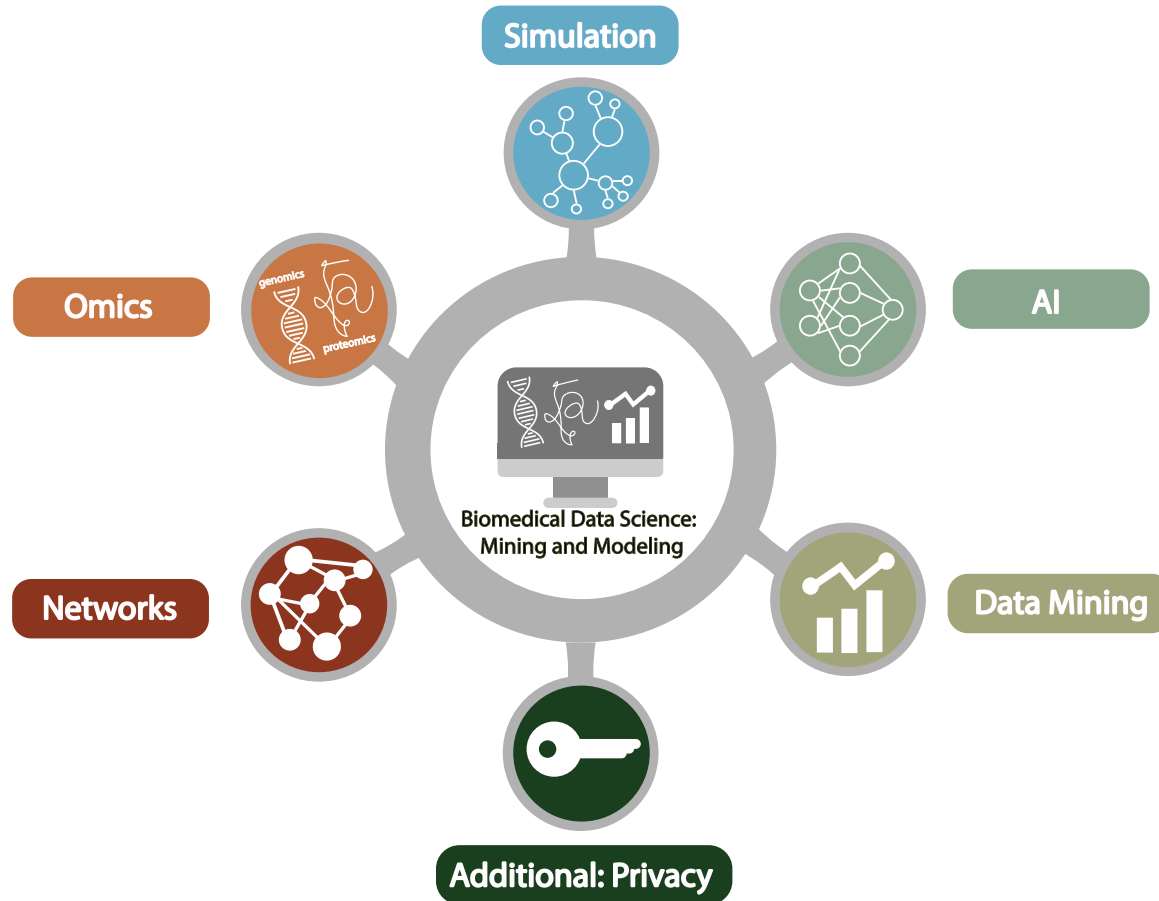


Biomedical Data Science (GersteinLab.org/courses/452)

Network Topology – Central Points (23m10d)

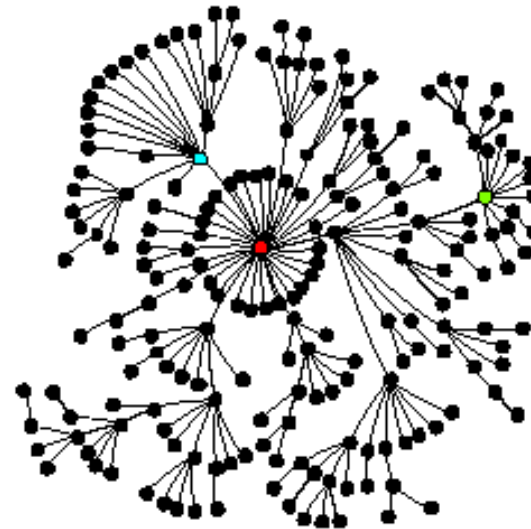
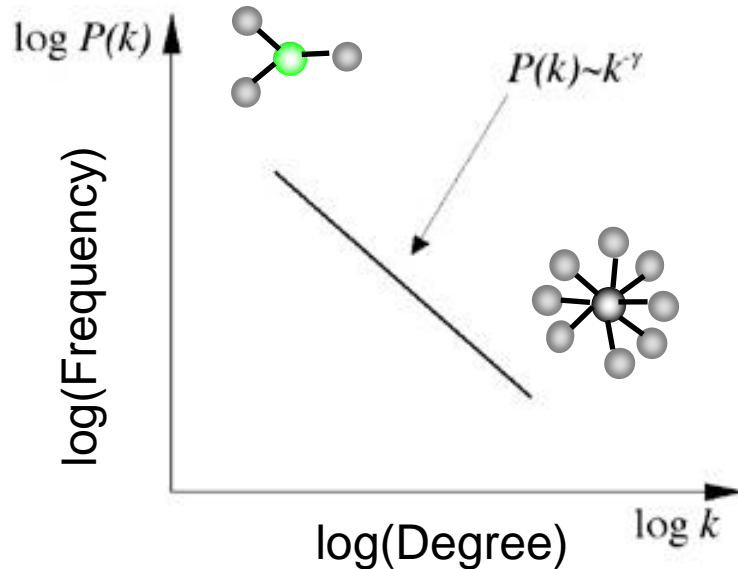


Network Topology

Measures of Centrality: Hubs

Scale-free networks in Biology

Power-law distribution

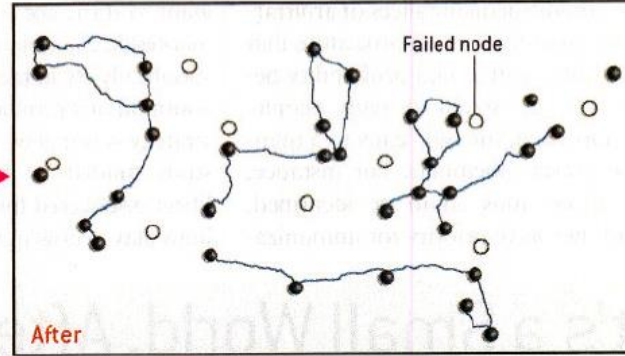
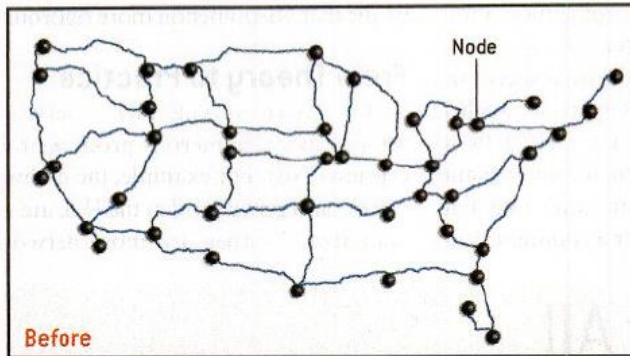


Hubs dictate the structure of the network

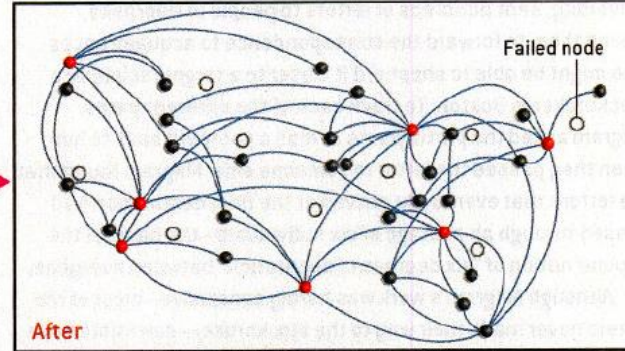
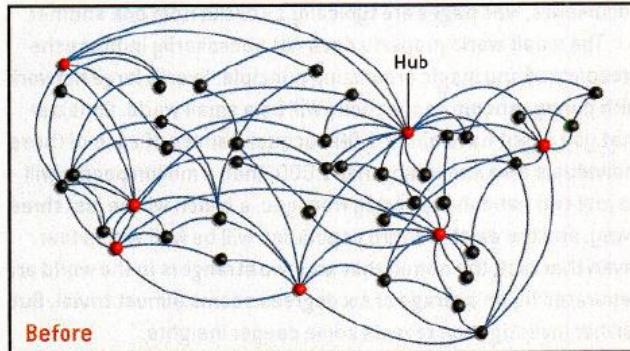
[Barabasi]

Knocking Out Nodes in Scale-free and Random Networks

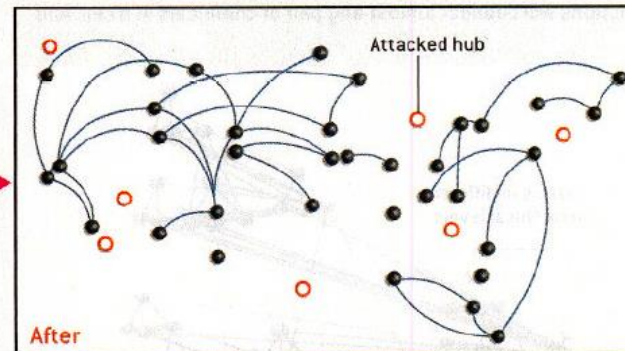
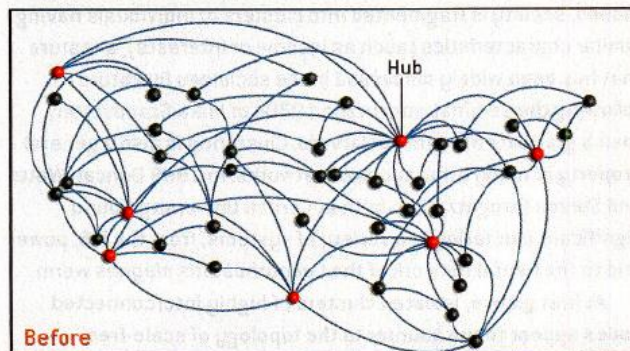
Random Network, Accidental Node Failure



Scale-Free Network, Accidental Node Failure



Scale-Free Network, Attack on Hubs

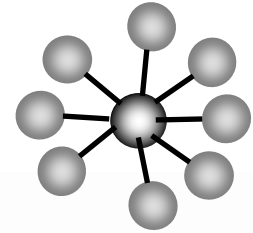


[From Barabasi & Bonabeau, *Sci. Am.*, May '03]

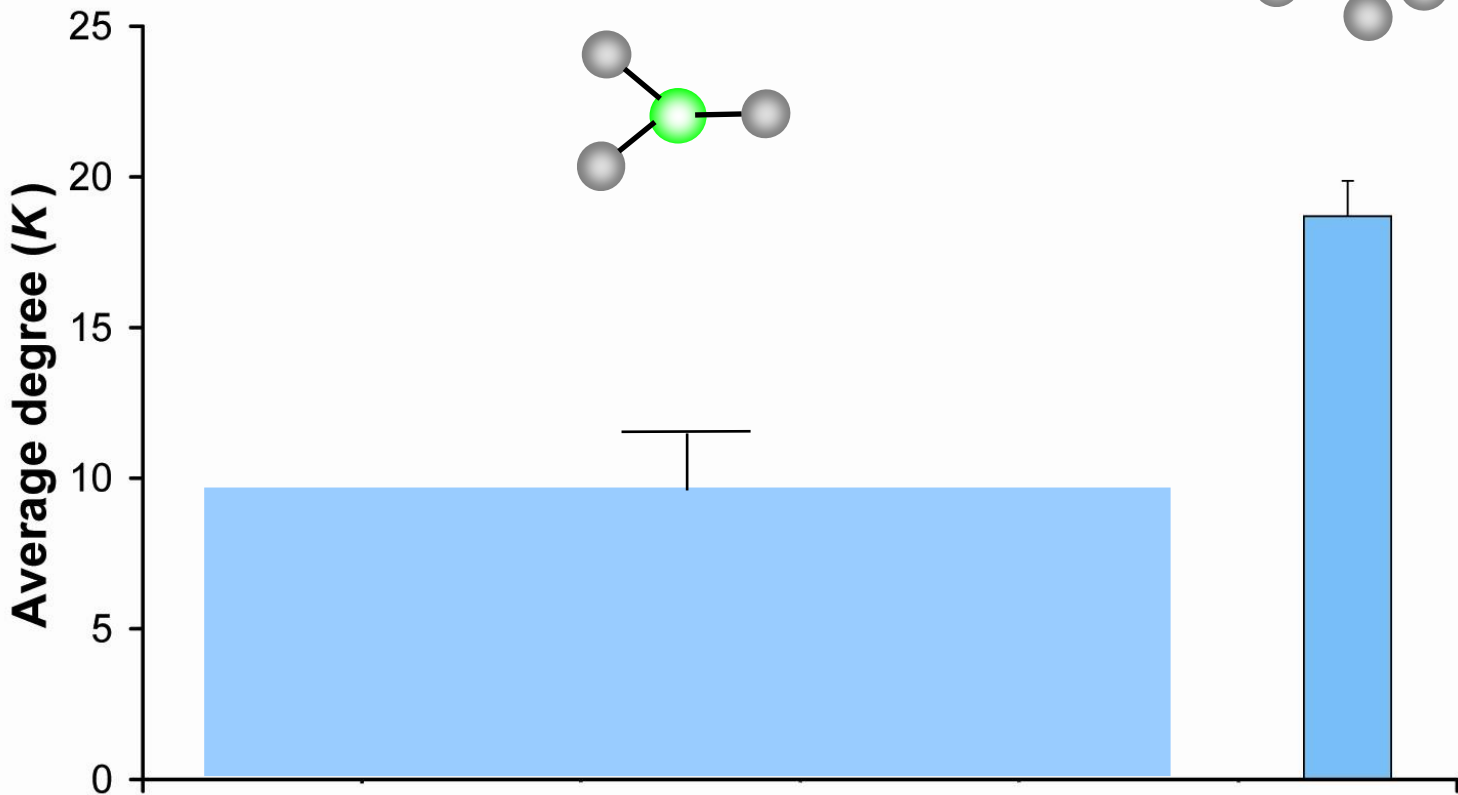
Hubs tend to be Essential

Integrate gene essentiality data with protein interaction network. Perhaps hubs represent vulnerable points?

[Lauffenburger, Barabasi]



"hubbiness"



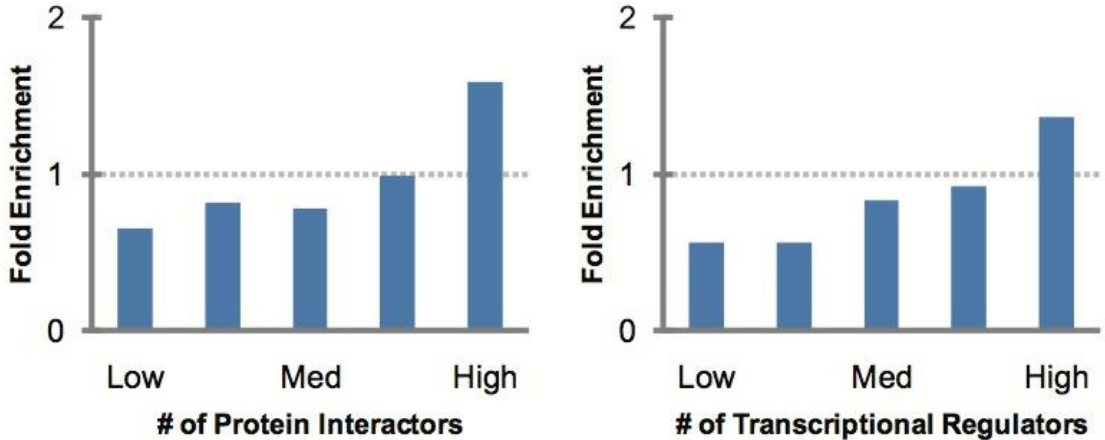
[Yu et al., 2003, TIG]

Non-Essential

Essential

- Notably, the most highly conserved proteins were highly connected in *E.coli* protein interaction network. (*E.coli*)
- Butland et al. 2004, Nature
- Connectivity of well-conserved proteins in network is negatively correlated with their rate of evolution. (*Saccharomyces cerevisiae*)
- Fraser et al. 2002, Science
- Fraser et al. BMC Evol. Biol. (2003)
- Proteins that have a more central position evolve more slowly and are more likely to be essential for survival. (*Saccharomyces cerevisiae*; *C. elegans*; *Drosophila*)
- Hahn et al. 2005, Mol Biol Evol
- More miRNA regulation of a target gene associated with lower dN/dS ($r = -.21$) (*M. musculus*)
- Cheng et al. 2009, BMC Genomics
- Slowly evolving proteins tend to have more interaction partners. (*Saccharomyces cerevisiae*)
- Xia et al. 2009, Plos Comput Biol

**More Connectivity,
More Constraint :
A theme borne out
in many studies**



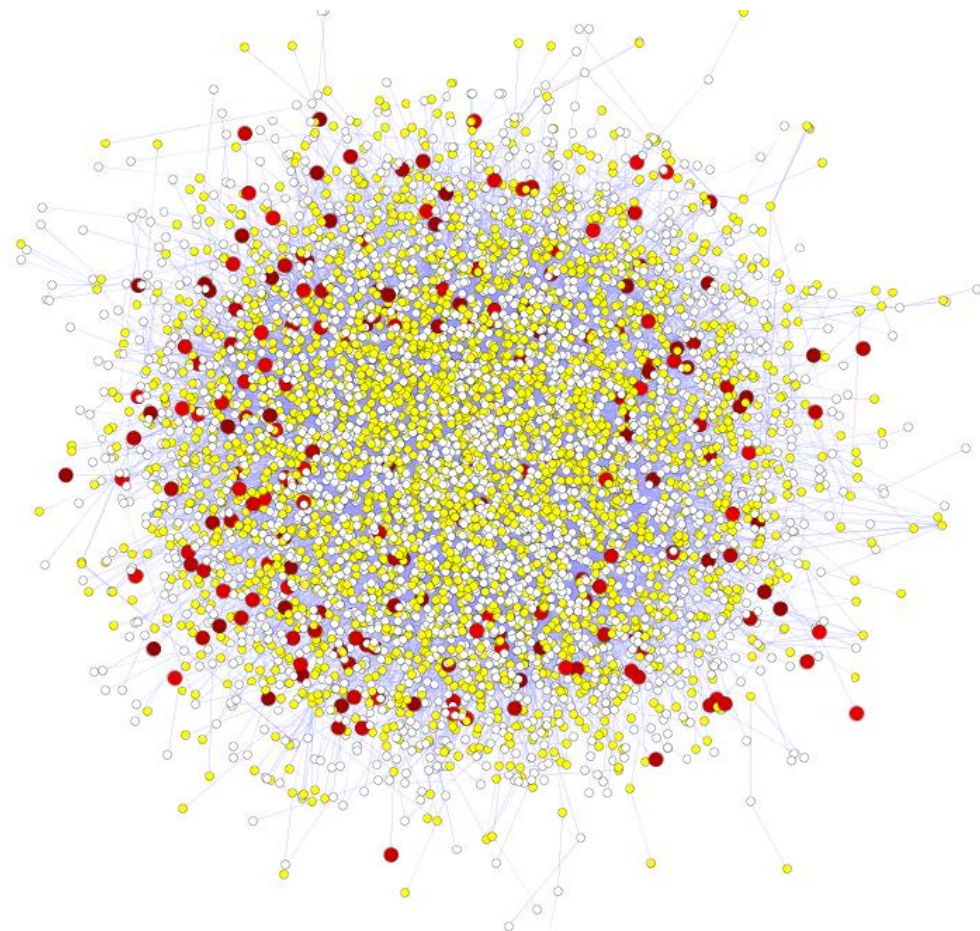
Fold Enrichment of slowly evolved proteins - Xia et al.

Rapid Evolution in the interaction network takes place at the periphery

- **Pos. sel. v. deg. centrality:**

$\rho = -.06$, $P < 1.2e-6$

- Effect is independent of any bias due to gene expression differences
- Update w. 1000G Phase I SNP dens. v. centrality:
 $\rho = -.1$, $p < 2.2e-16$



- High likelihood of positive selection
- Lower likelihood of positive selection
- Not under positive selection
- No data about positive selection

Source: Nielsen et al. *PLoS Biol.* (2005), HPRD, and Kim et al. *PNAS* (2007)

Network Topology

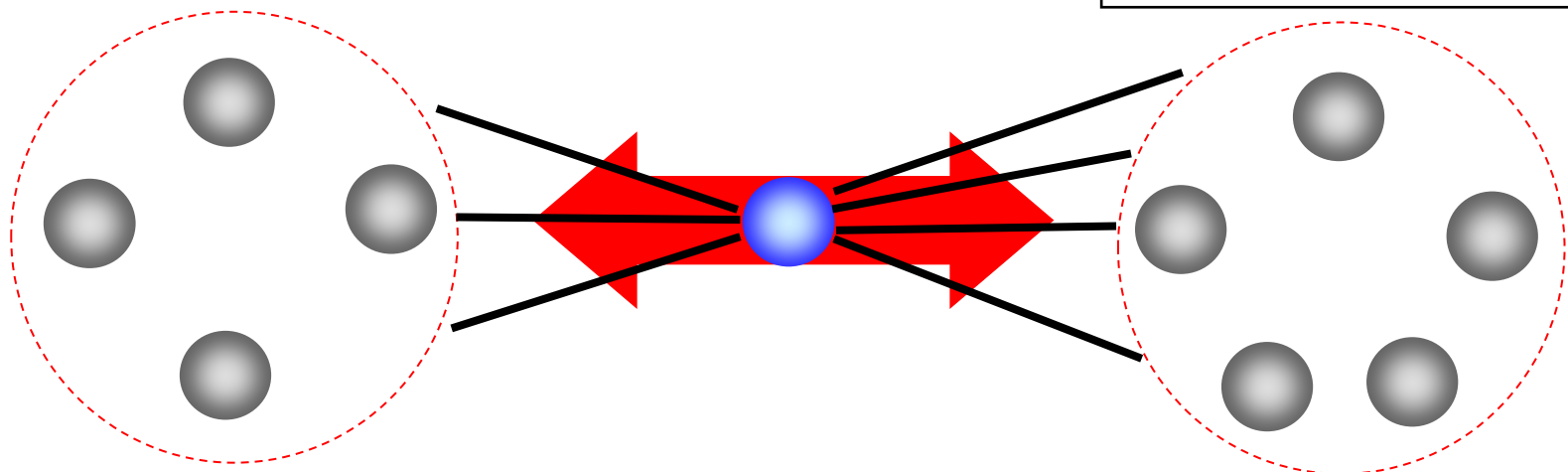
**Other Measures of
Centrality besides Hubs:
Bottlenecks**

Another measure of Centrality: Betweenness centrality

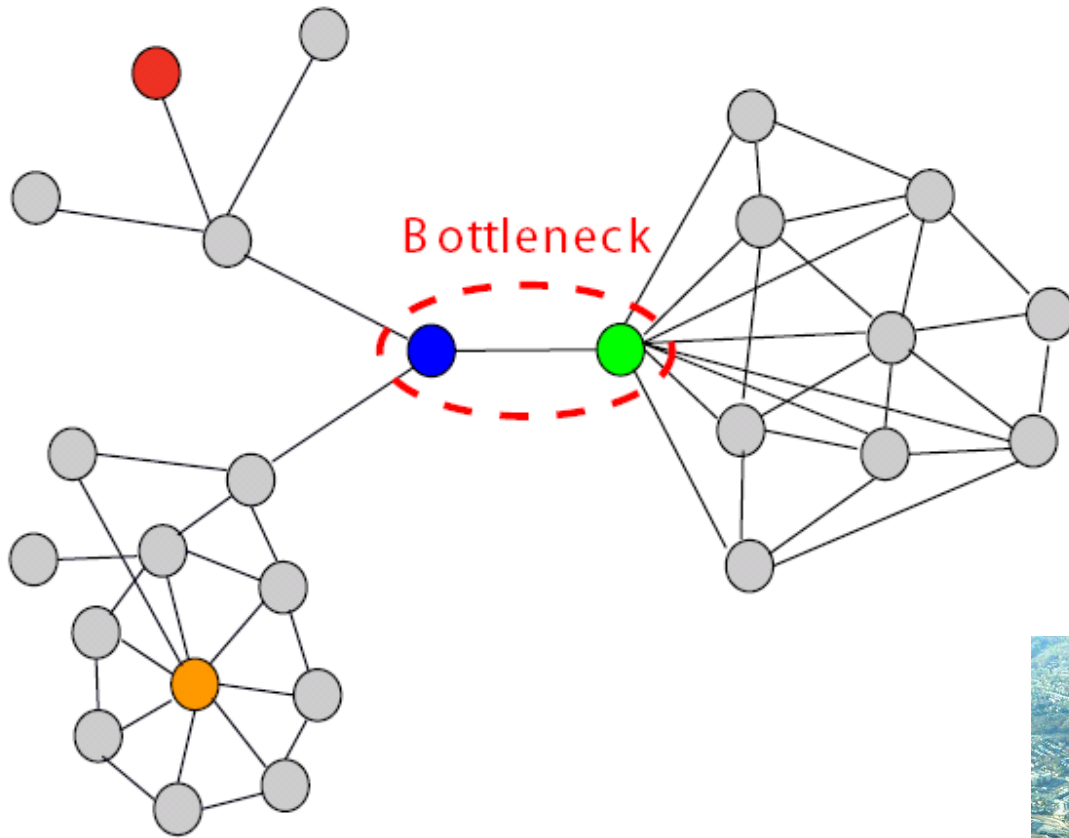
Betweenness of a node is the number of shortest paths of pairs of vertices that run through it -- a measure of information flow.





Freeman LC (1977) Set of measures of centrality based on betweenness. *Sociometry* 40: 35–41.

Girvan & Newman (2002) *PNAS* 99: 7821.



Bottlenecks & Hubs

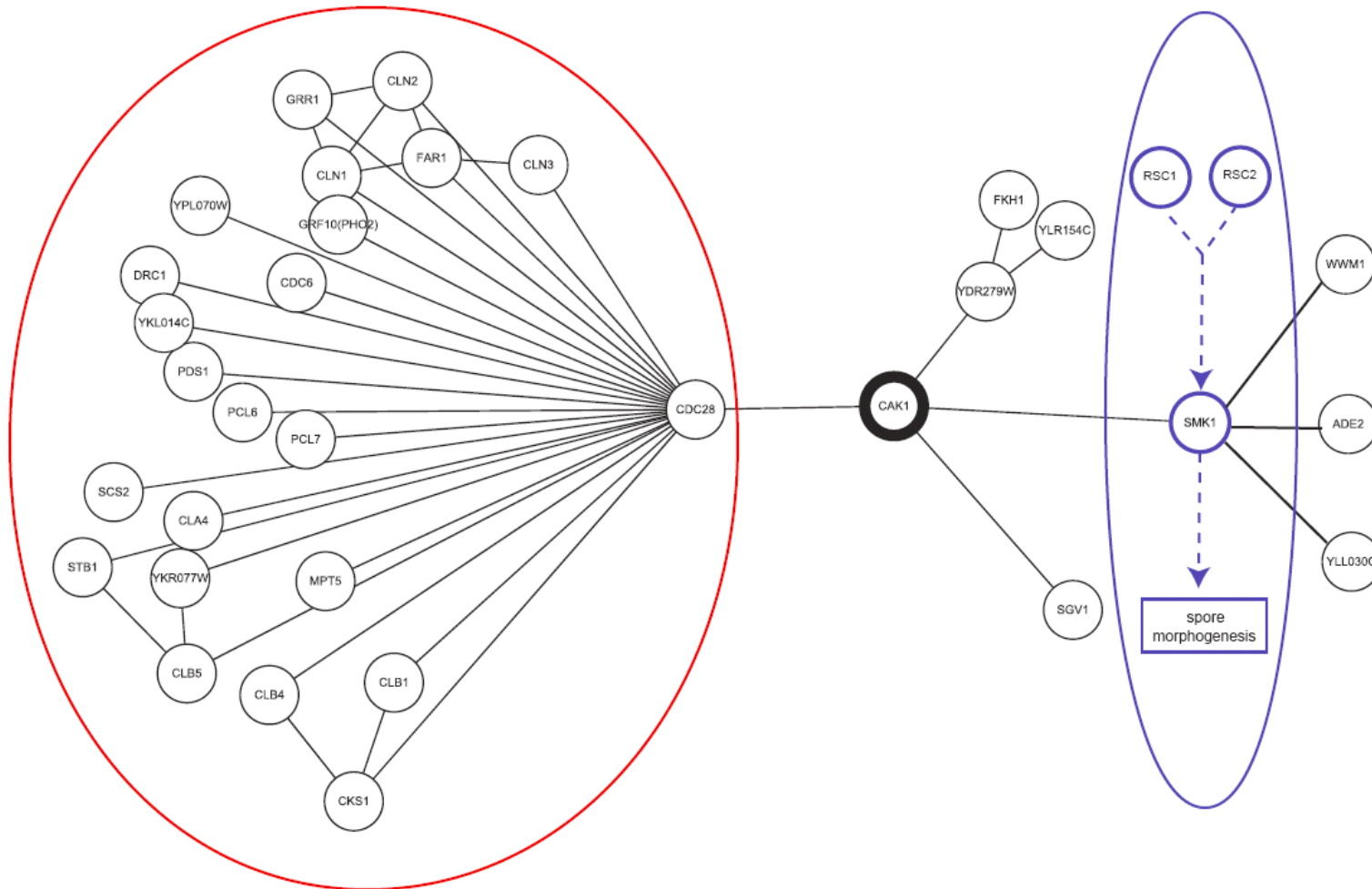


-  Hub-bottleneck **node**
-  Non-hub-bottleneck **node**
-  Hub-non-bottleneck **node**
-  Non-hub-non-bottleneck **node**



[Yu et al., PLOS CB (2007)]

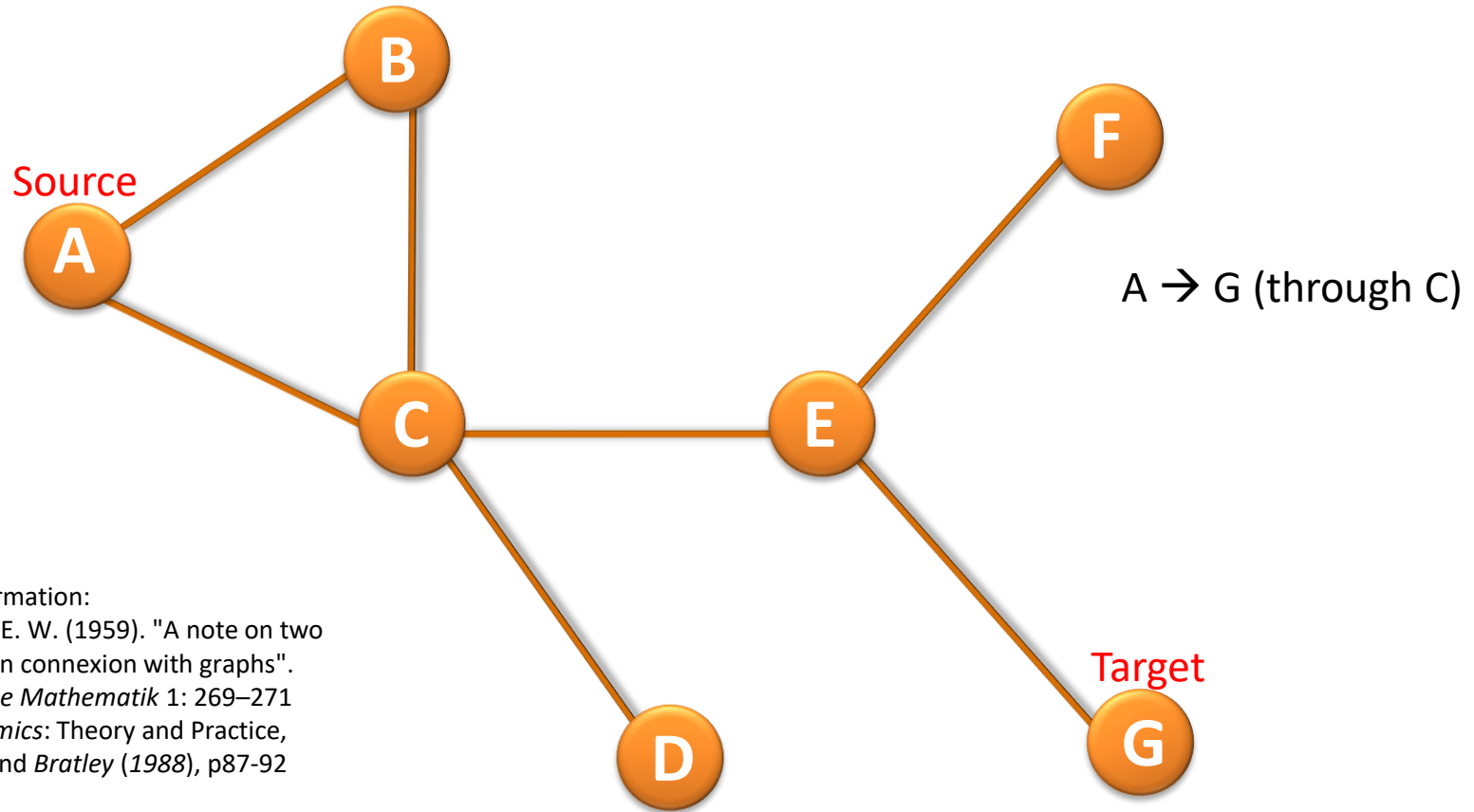
Bottleneck bridging between processes



Regulation of mitotic cell cycle

MAP Kinase pathway
regulating spore morphogenesis

Shortest Paths: The Dijkstra Algorithm



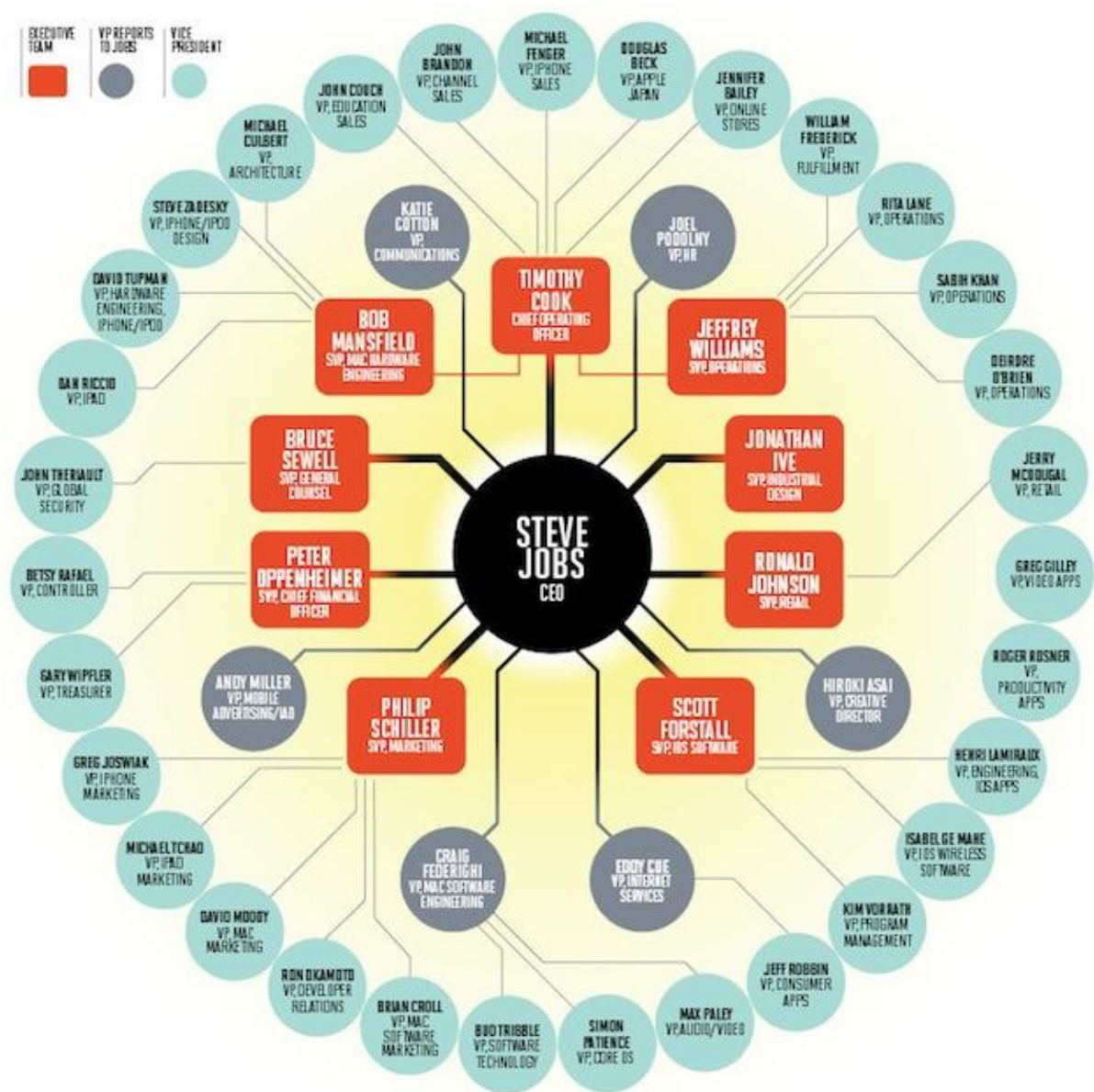
More information:

1. Dijkstra, E. W. (1959). "A note on two problems in connexion with graphs". *Numerische Mathematik* 1: 269–271
2. *Algorithmics: Theory and Practice*, Brassard and Bratley (1988), p87-92

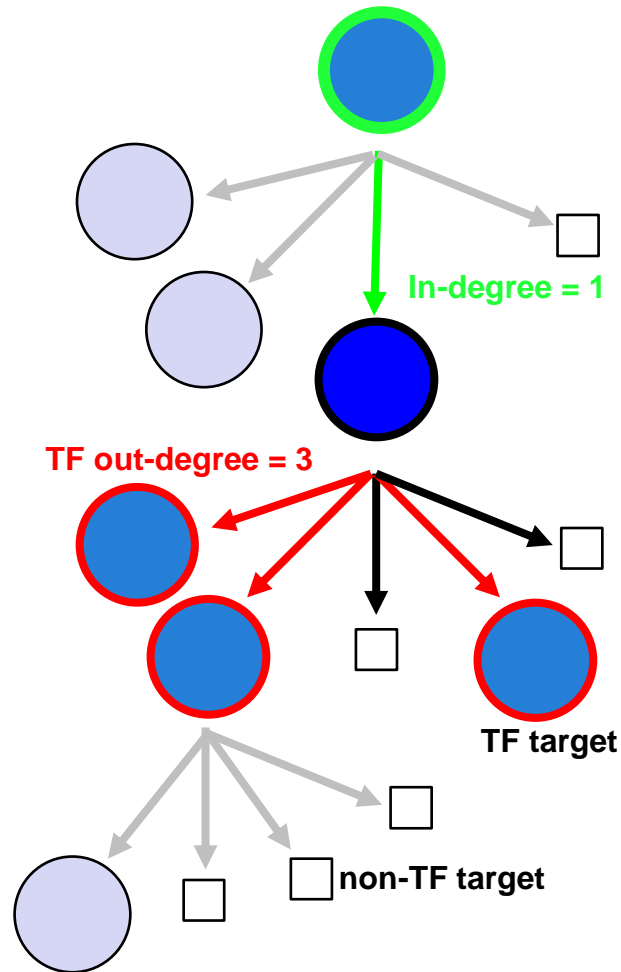
Network Topology

**Other Measures of
Centrality besides Hubs:
Hierarchy**

Network Hierarchy



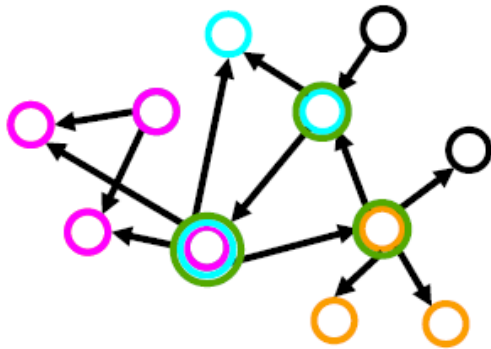
Network Stats to Identify Hierarchy



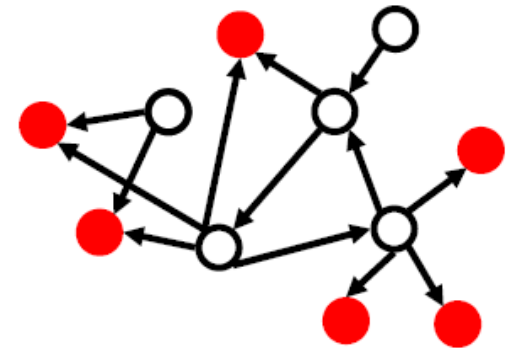
**Hierarchy Height Statistic =
(normalized TF Out deg. – In deg.)**

Determination of "Level" in Regulatory Network Hierarchy with Breadth-first Search

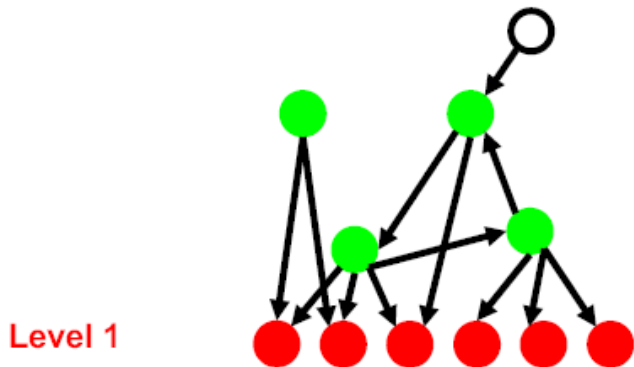
I. Example network with all 4 motifs



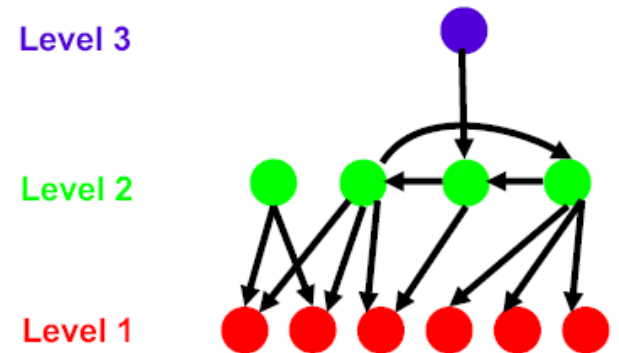
II. Finding terminal nodes (Red)



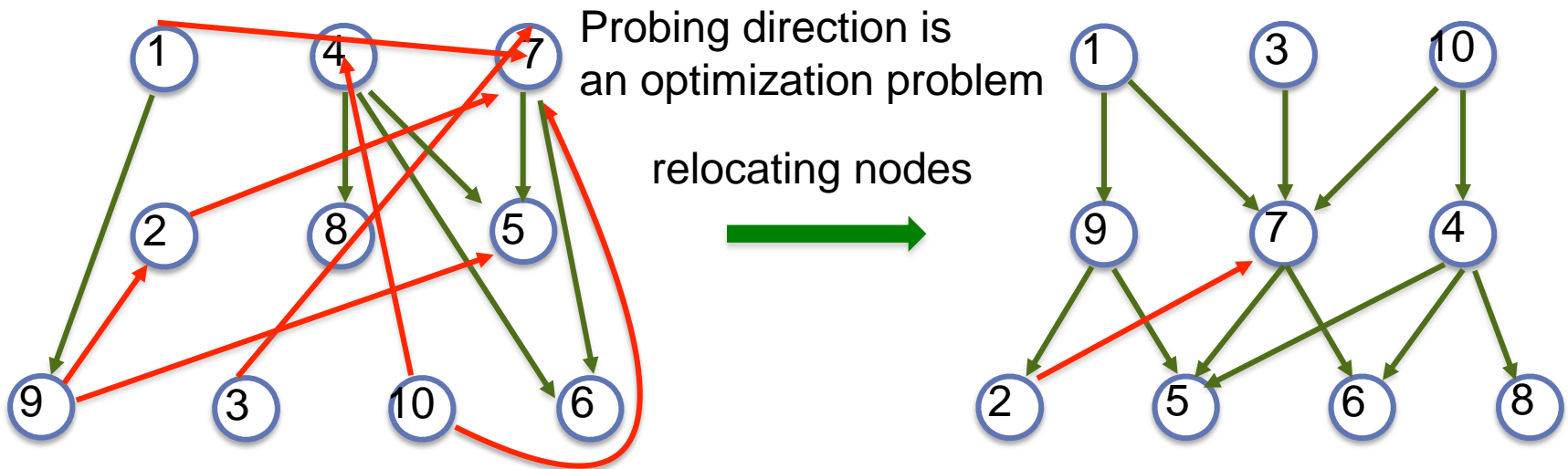
III. Finding mid-level nodes (Green)



IV. Finding top-most nodes (Blue)

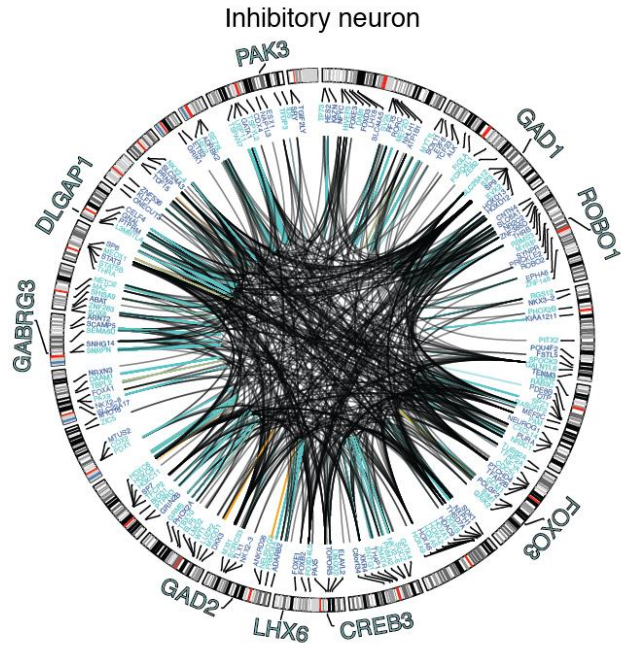


Using Simulated Annealing to Globally Minimize the Number of Upward Pointing Edges

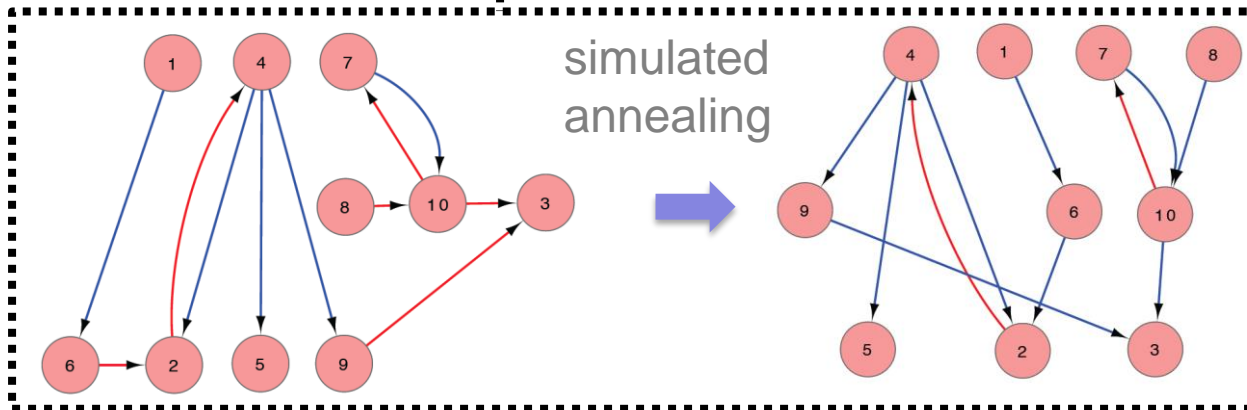
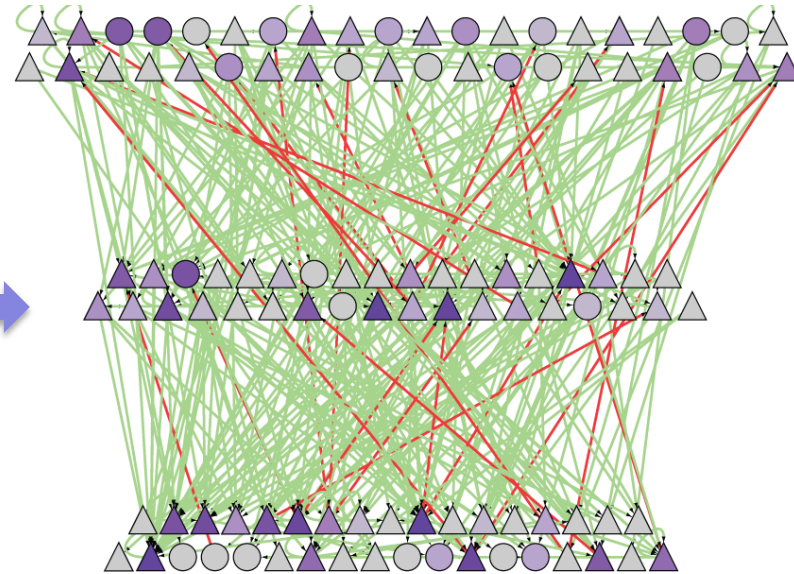


Transforming a Regulatory Network into a Hierarchy

Regulatory Network

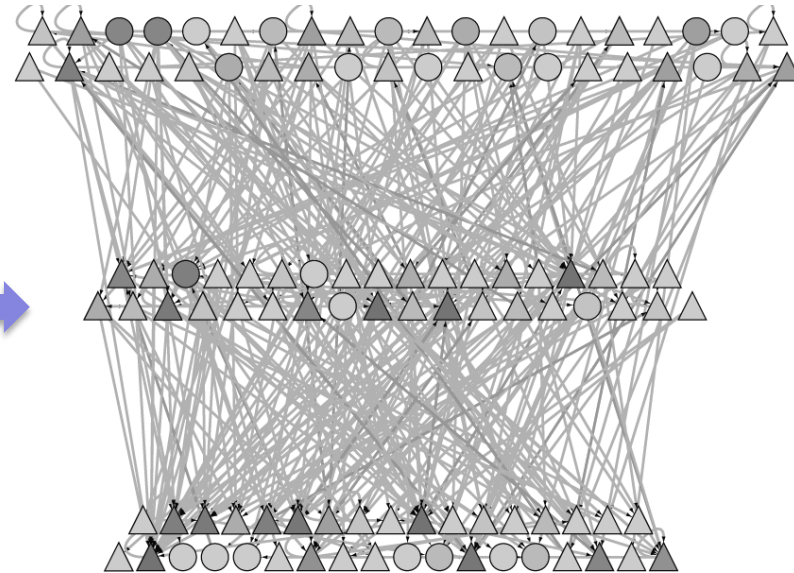
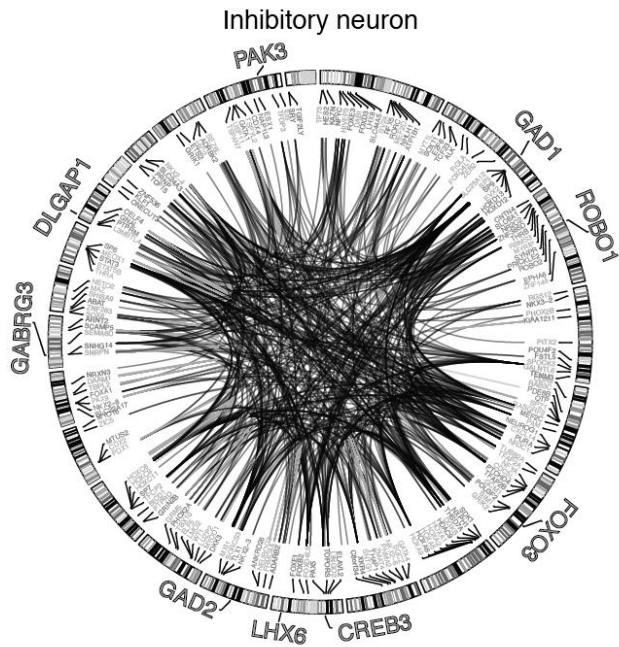


Network Hierarchy



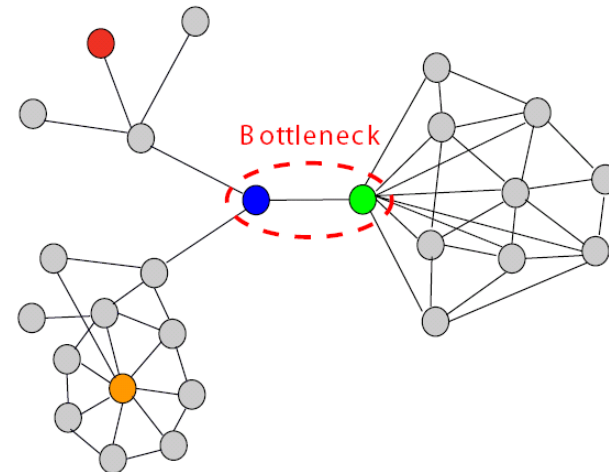
[Wang et al. ('18) Science, Gerstein et al. Nature ('12)]

Mid-level of the hierarchy has many high-connectivity bottlenecks



betweenness

- Hub-bottleneck node
- Non-hub-bottleneck node
- Hub-non-bottleneck node
- Non-hub-non-bottleneck node



[Wang et al. ('18) Science, Gerstein et al. Nature ('12)]