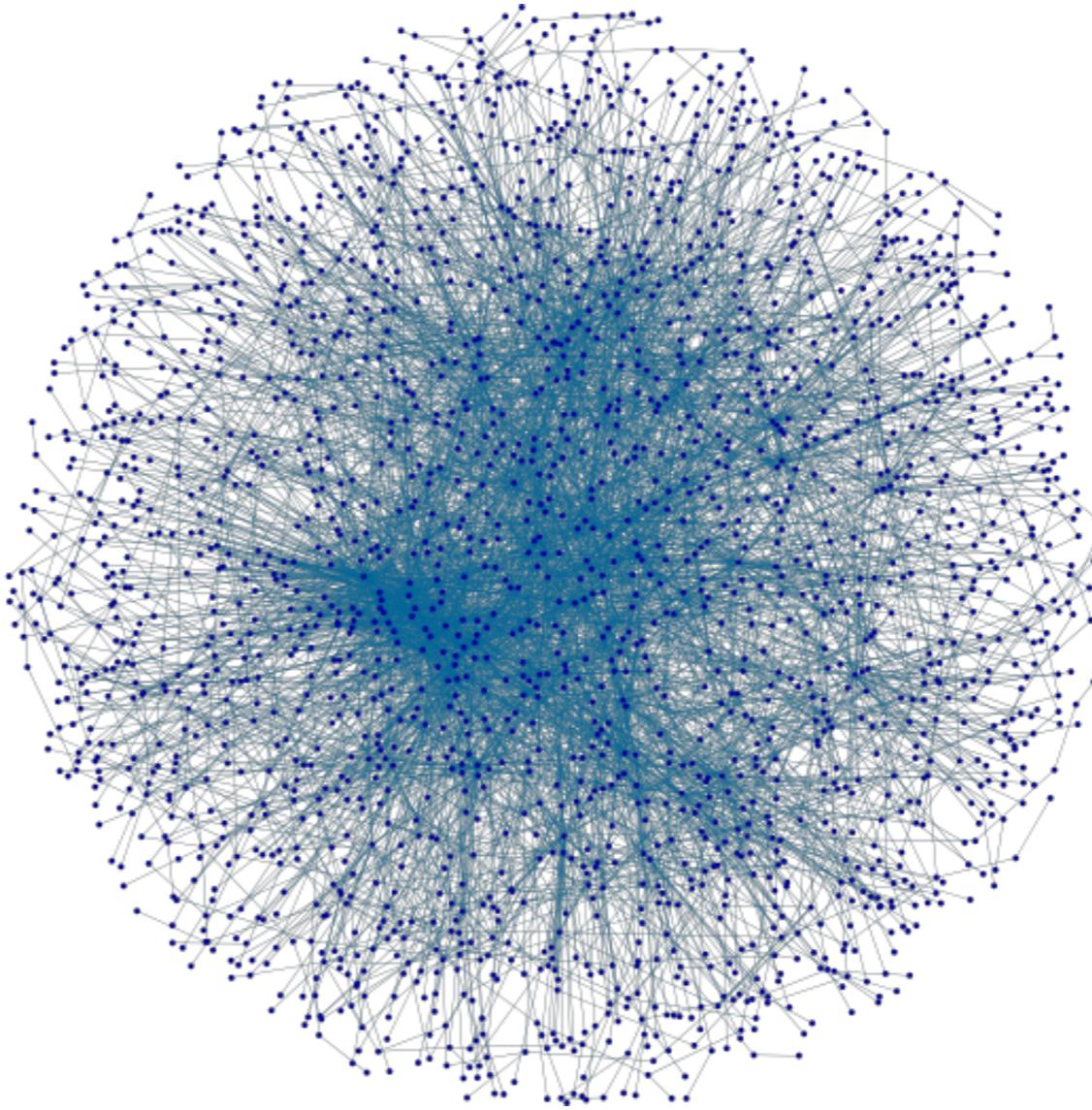
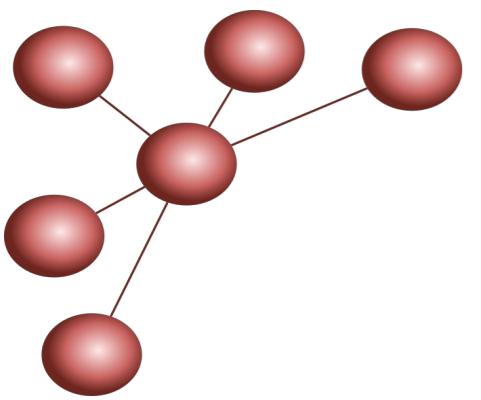
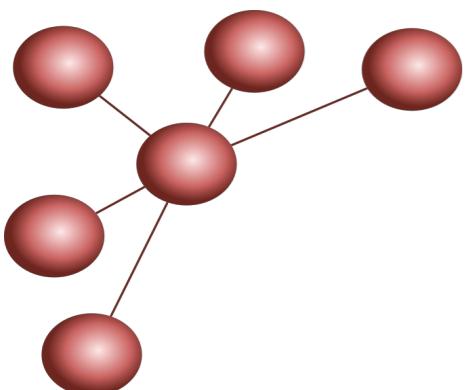


# Integrative Bioinformatics

## Module 6 - Session 6

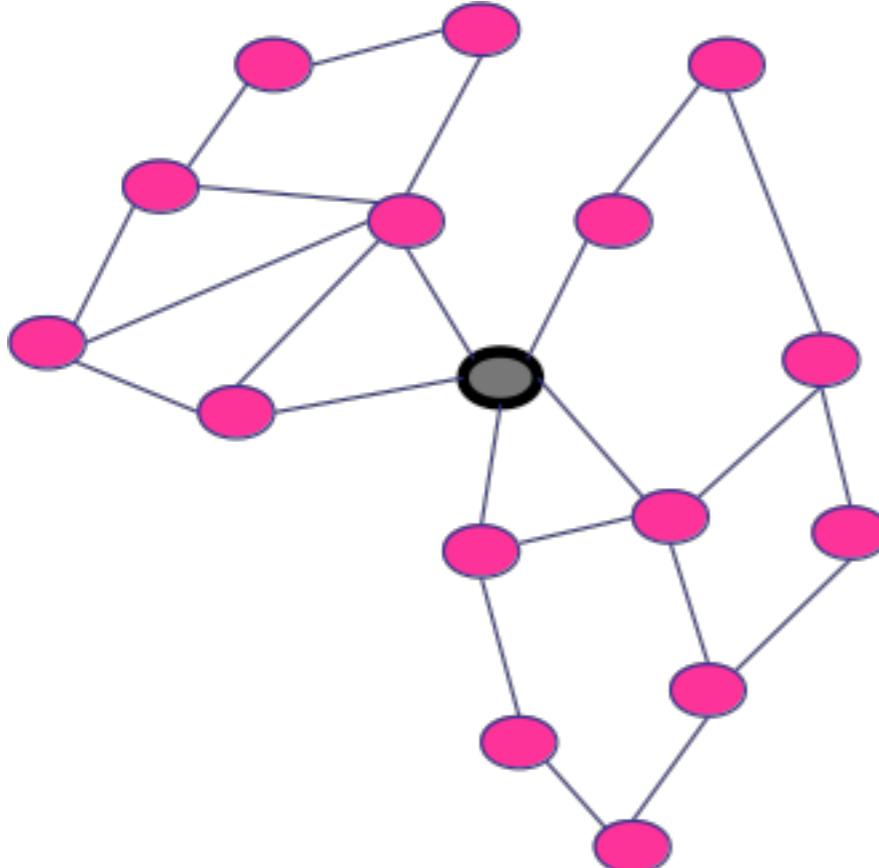


**What next ?**  
**=> Network Analysis**



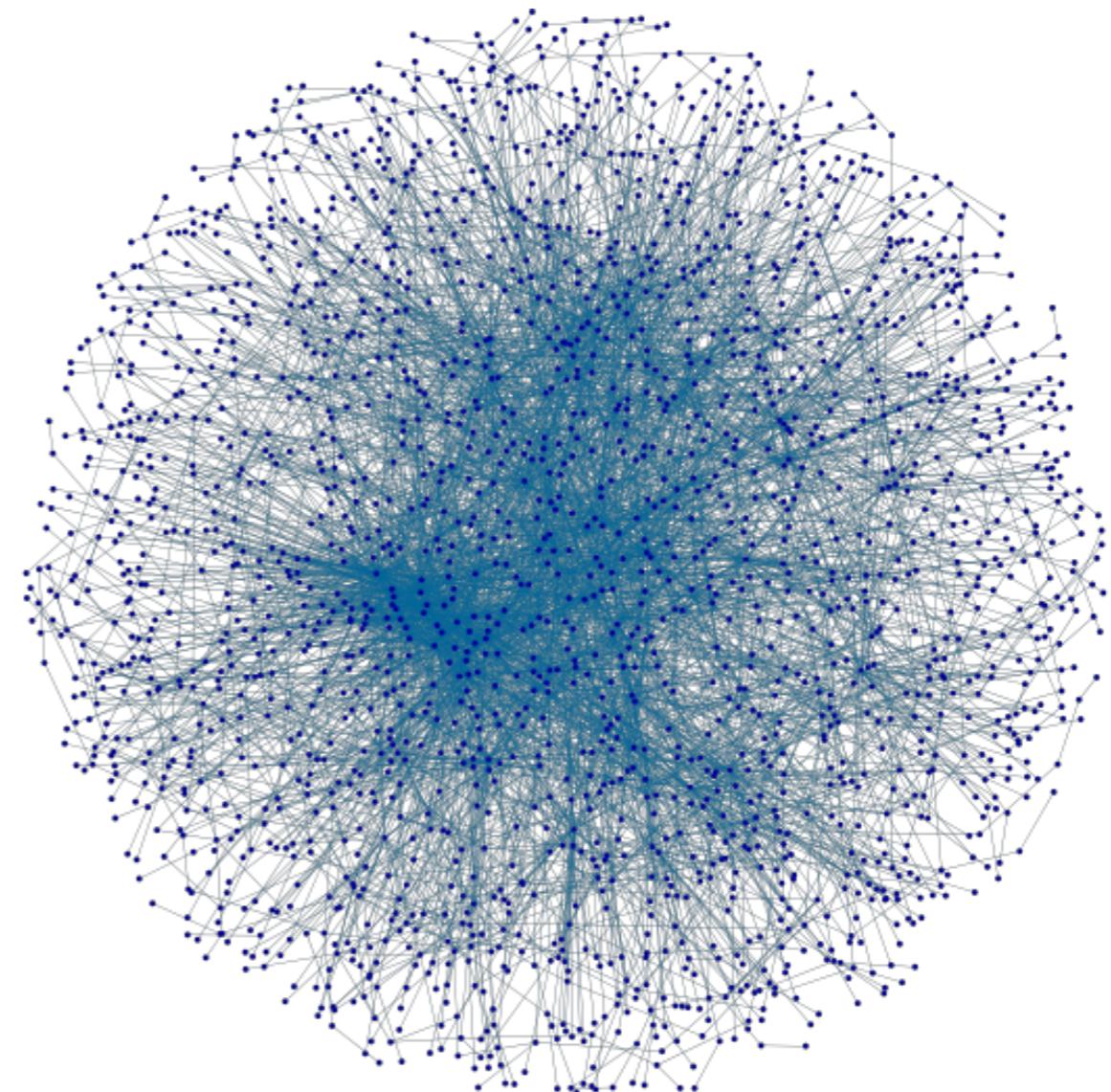
# How to use large-scale biological networks ?

**Local approaches**

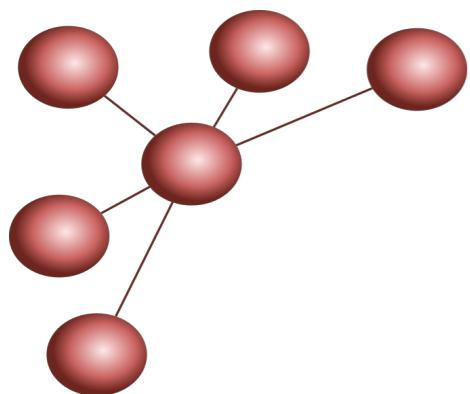


**“guilt by association”**

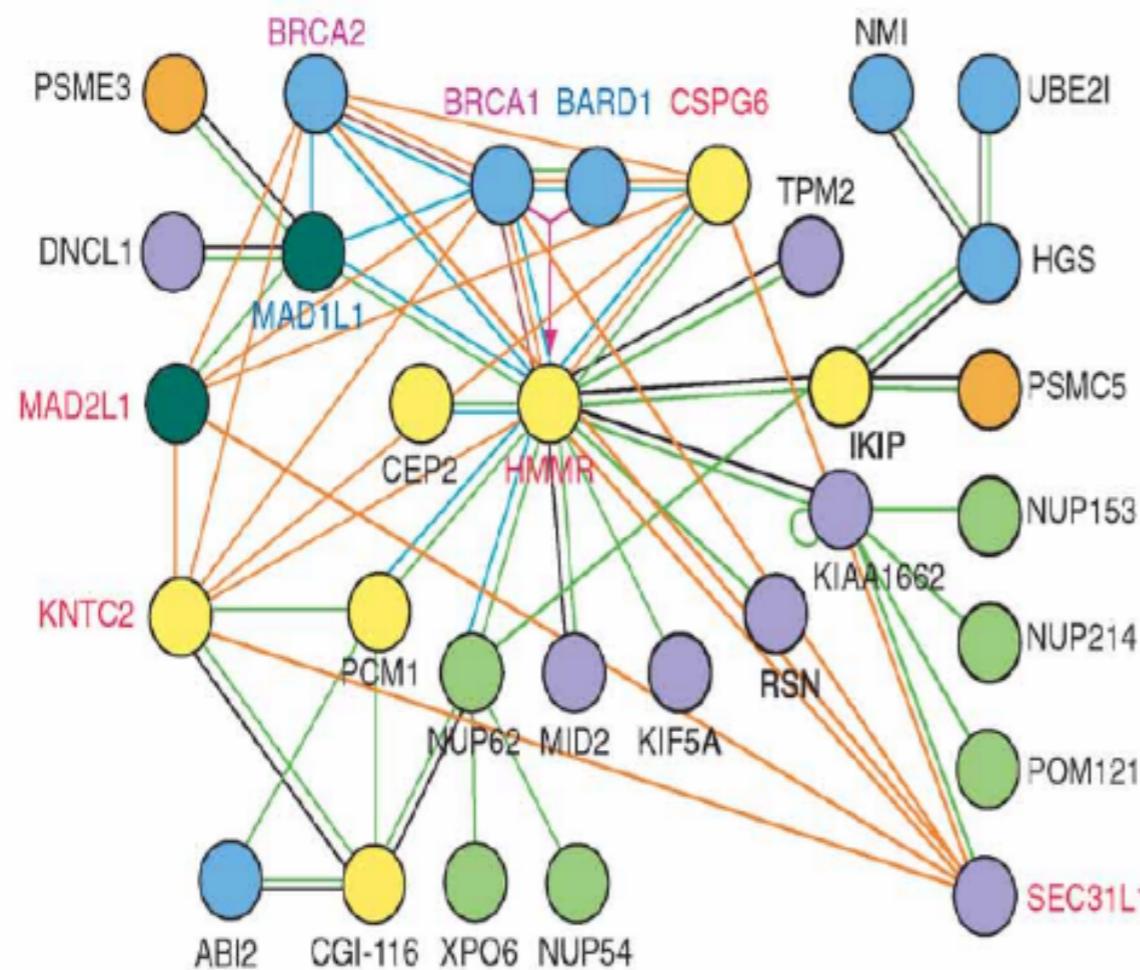
**Global approaches**



**Topological features  
Clustering / communities**



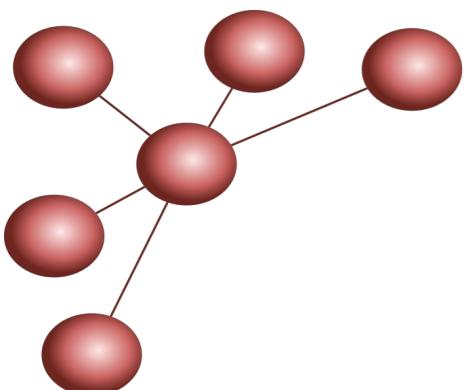
# Identification of a new gene involved in breast cancer



**Nodes correspond to proteins,  
edges to interactions  
identified by different  
experimental techniques**

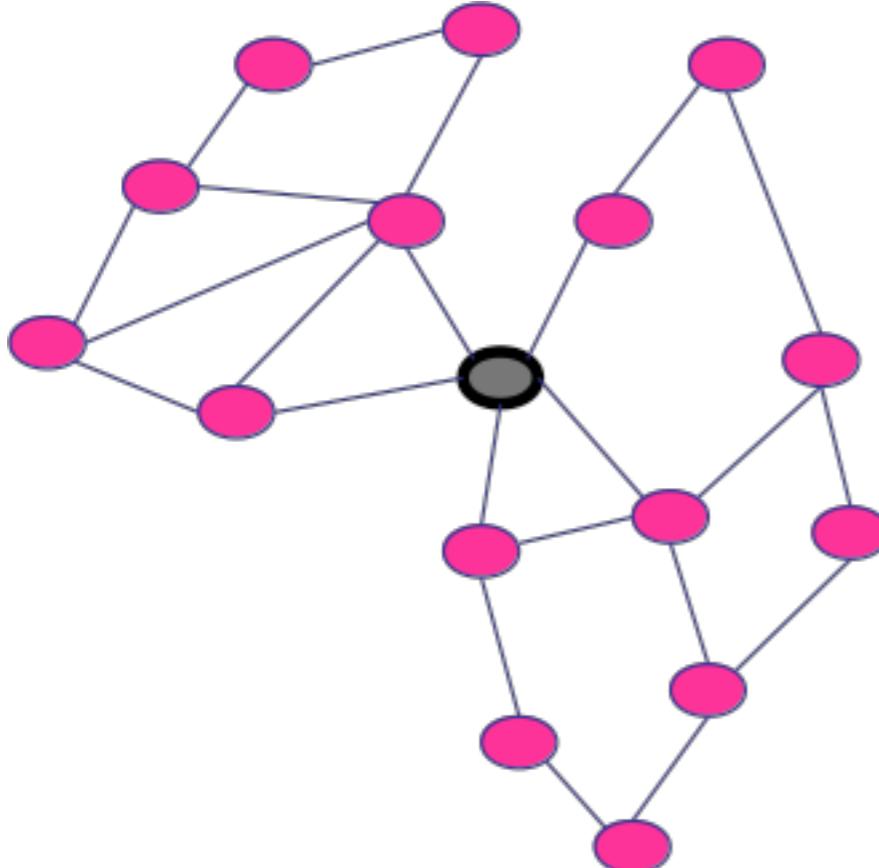
## Functional associations (*n*)

- Expression profiling similarity (20)
- Similar gene deficiency phenotype (2)
- Y2H binary protein interaction (32)
- Protein co-AP (13)
- Protein co-IP (11)
- Biochemical interaction (1)



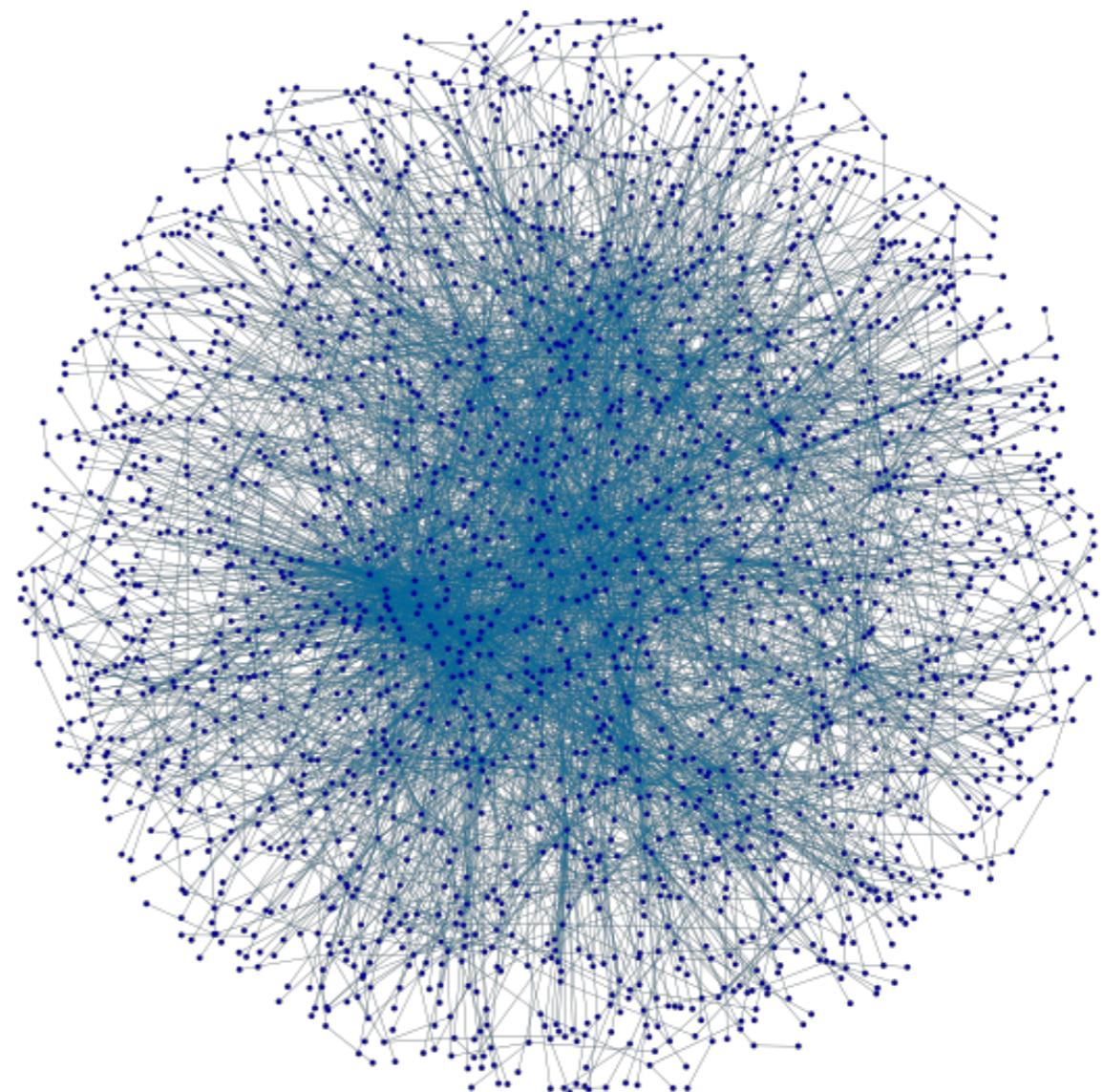
# How to use large-scale biological networks ?

**Local approaches**



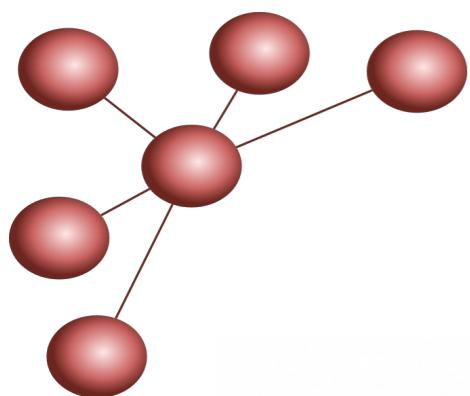
**“guilt by association”**

**Global approaches**

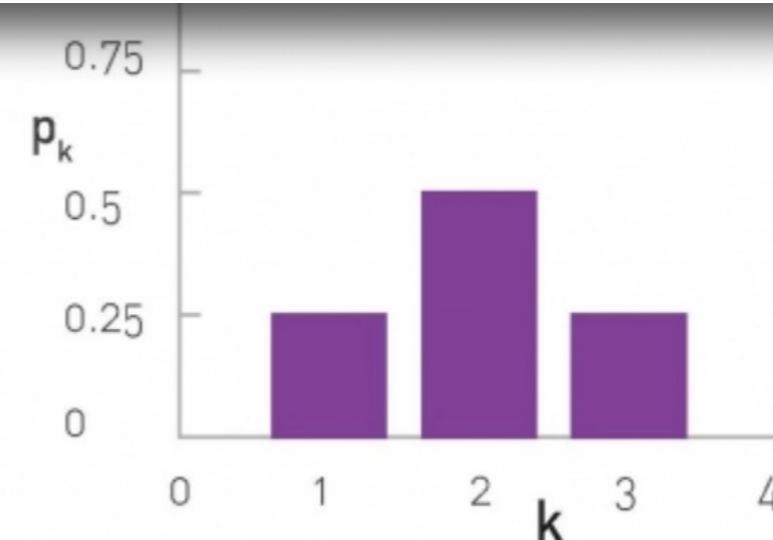
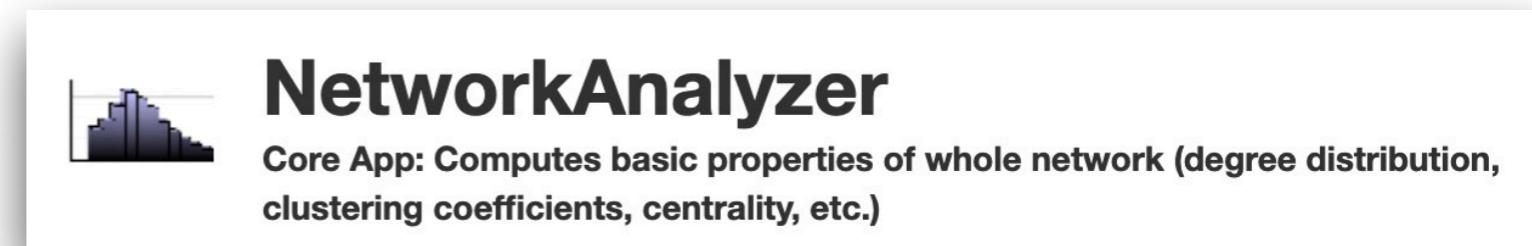
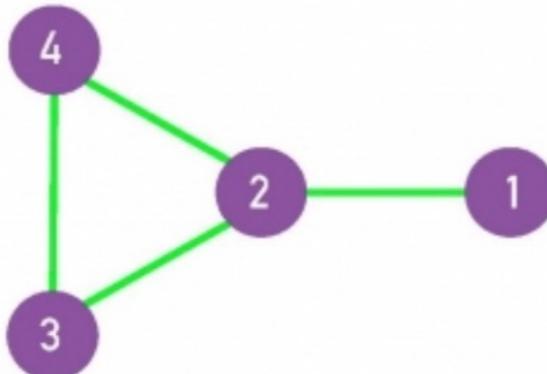


**Topological features  
Clustering / communities**

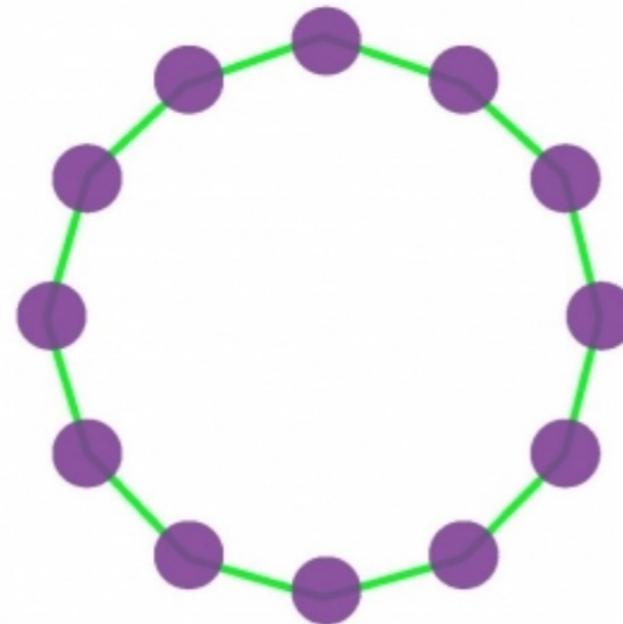
# Degree distribution



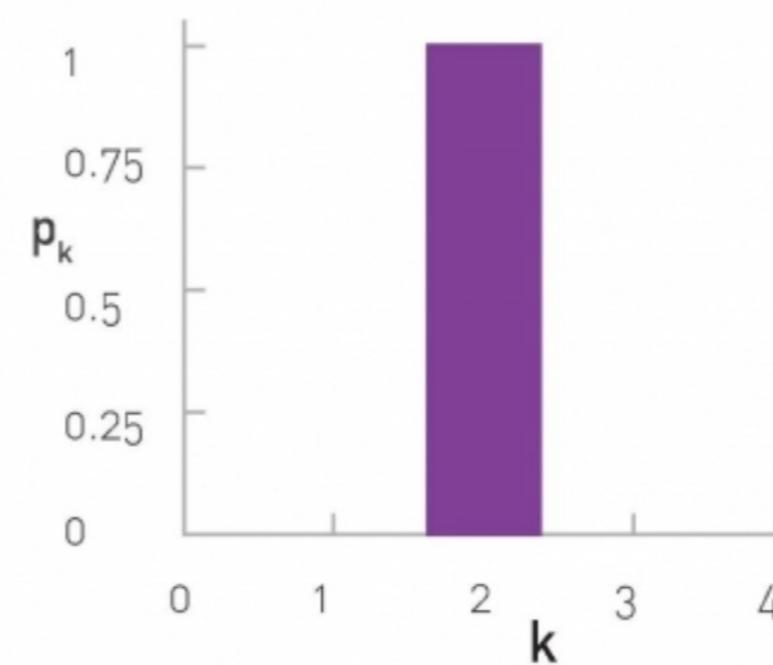
a.

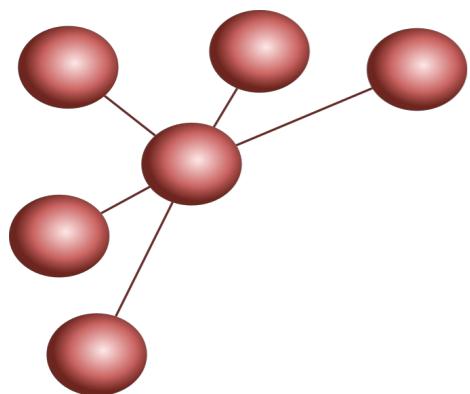


c.

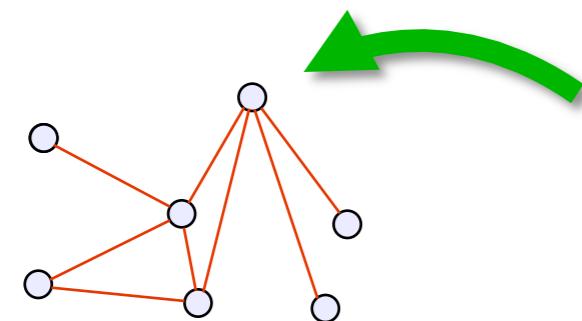


d.



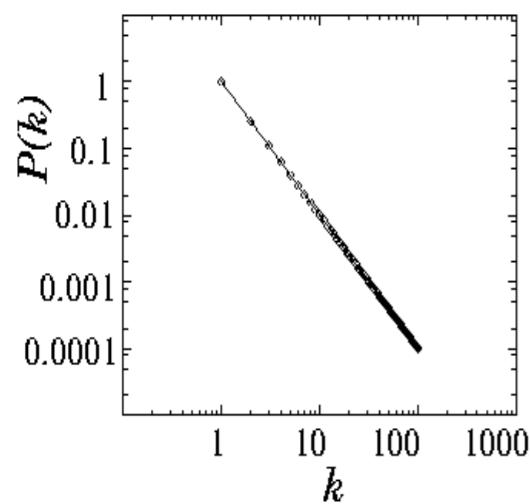


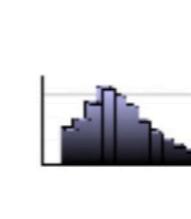
# Protein degree distribution : interactomes are scale-free and small-world

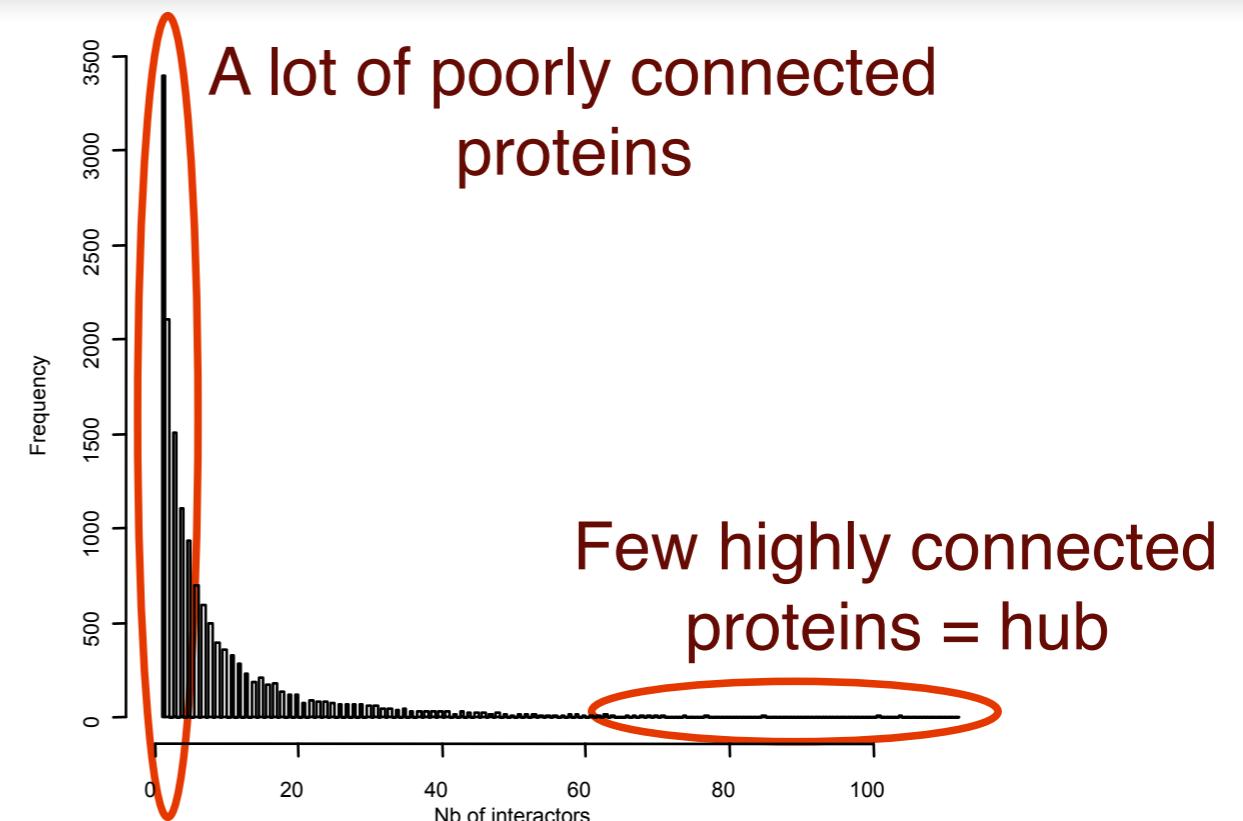


$k = 4$

Power-law distribution



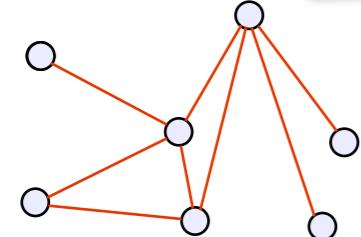
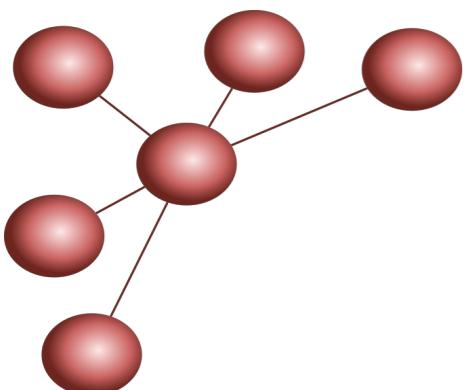
 **NetworkAnalyzer**  
Core App: Computes basic properties of whole network (degree distribution, clustering coefficients, centrality, etc.)



Biological interpretation?

Robust to random attack, sensitive to targeted attacks  
Growth with preferential attachment (“rich get richer”) => create “hubs”

# Protein degree distribution : interactomes are scale-free and small world

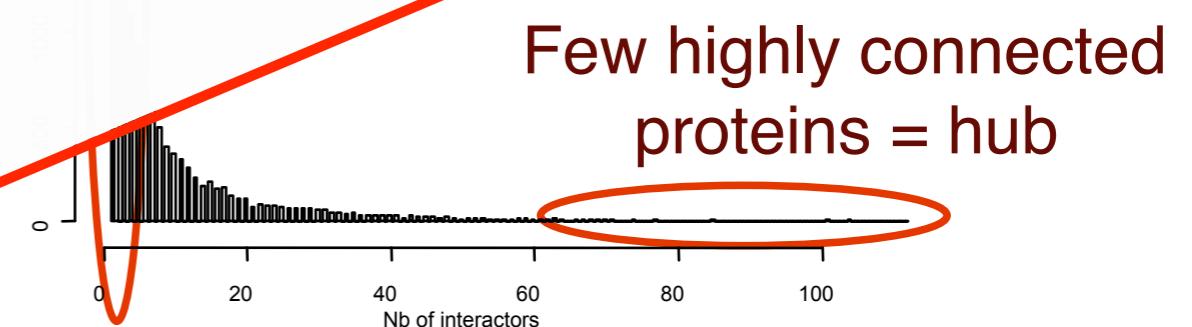
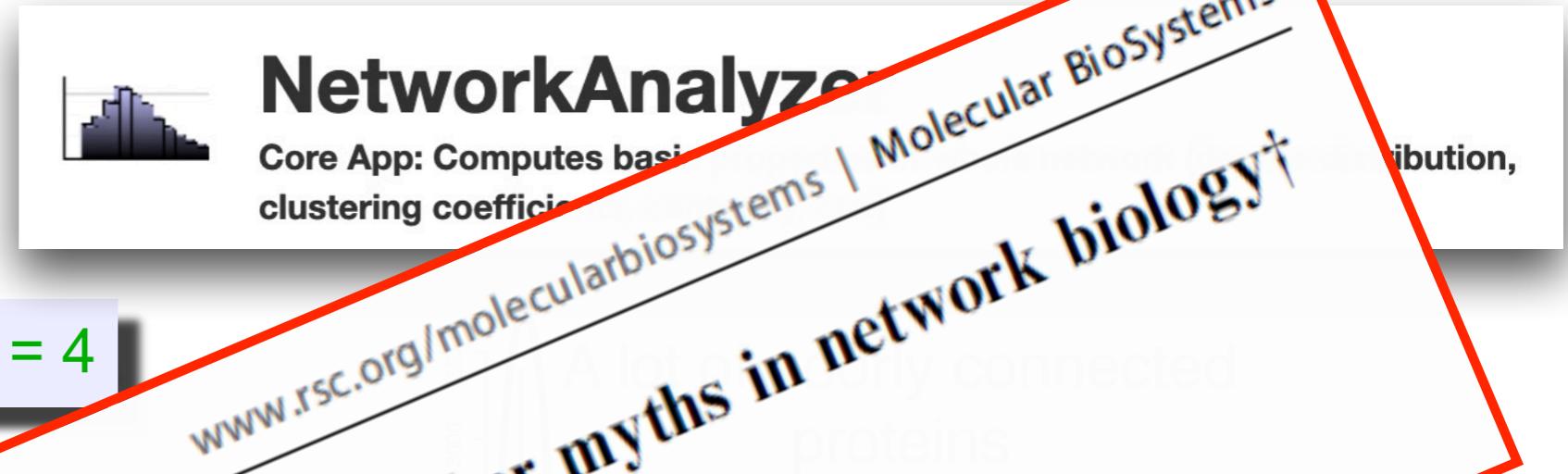


$k = 4$

REVIEW

The powerful law of the power law and other myths in network biology†  
Gipsi Lima-Mendez\* and Jacques van Helden\*

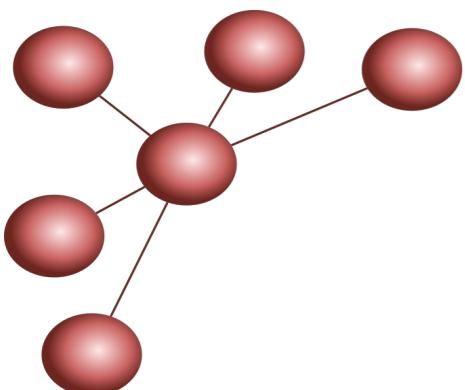
Received 5th May 2009, Accepted 12th August 2009  
First published as an Advance Article on the web  
DOI: 10.1039/b908681a



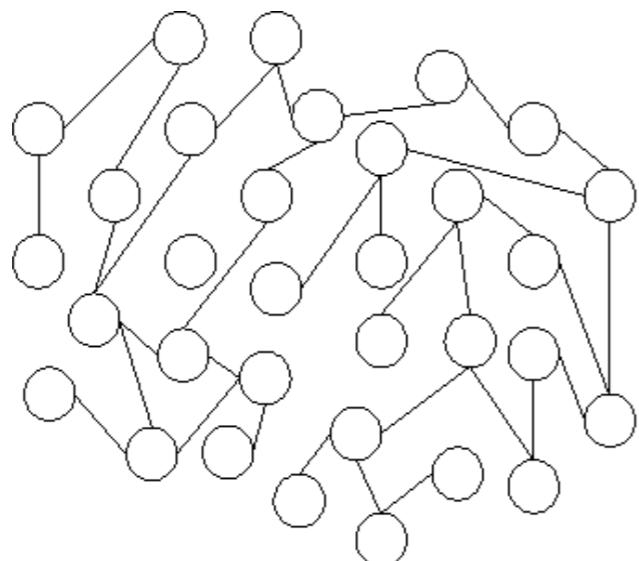
Biological interpretation?

Robust to random attack, sensitive to targeted attacks

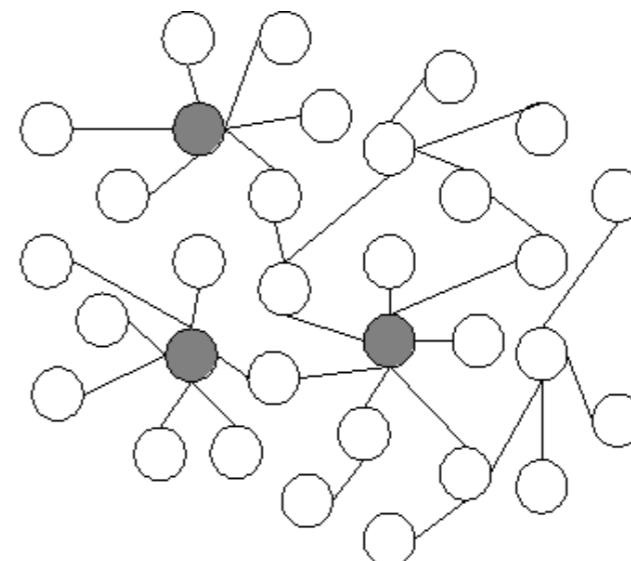
Growth with preferential attachment (“rich get richer”) => create “hubs”



# Network topological structure : Small-world property

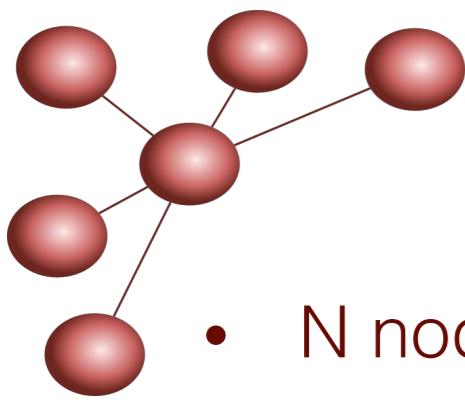


(a) Random network



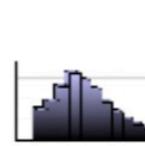
(b) Scale-free network

- Milgram, 6 degrees of separation



- N nodes, V edges
- Network size
- Adjacency matrix
- Degree, degree distribution
- Path, shortest path, distances
- Connectivity, clustering coefficient
- Betweenness
- Motifs

# Metrics on graphs

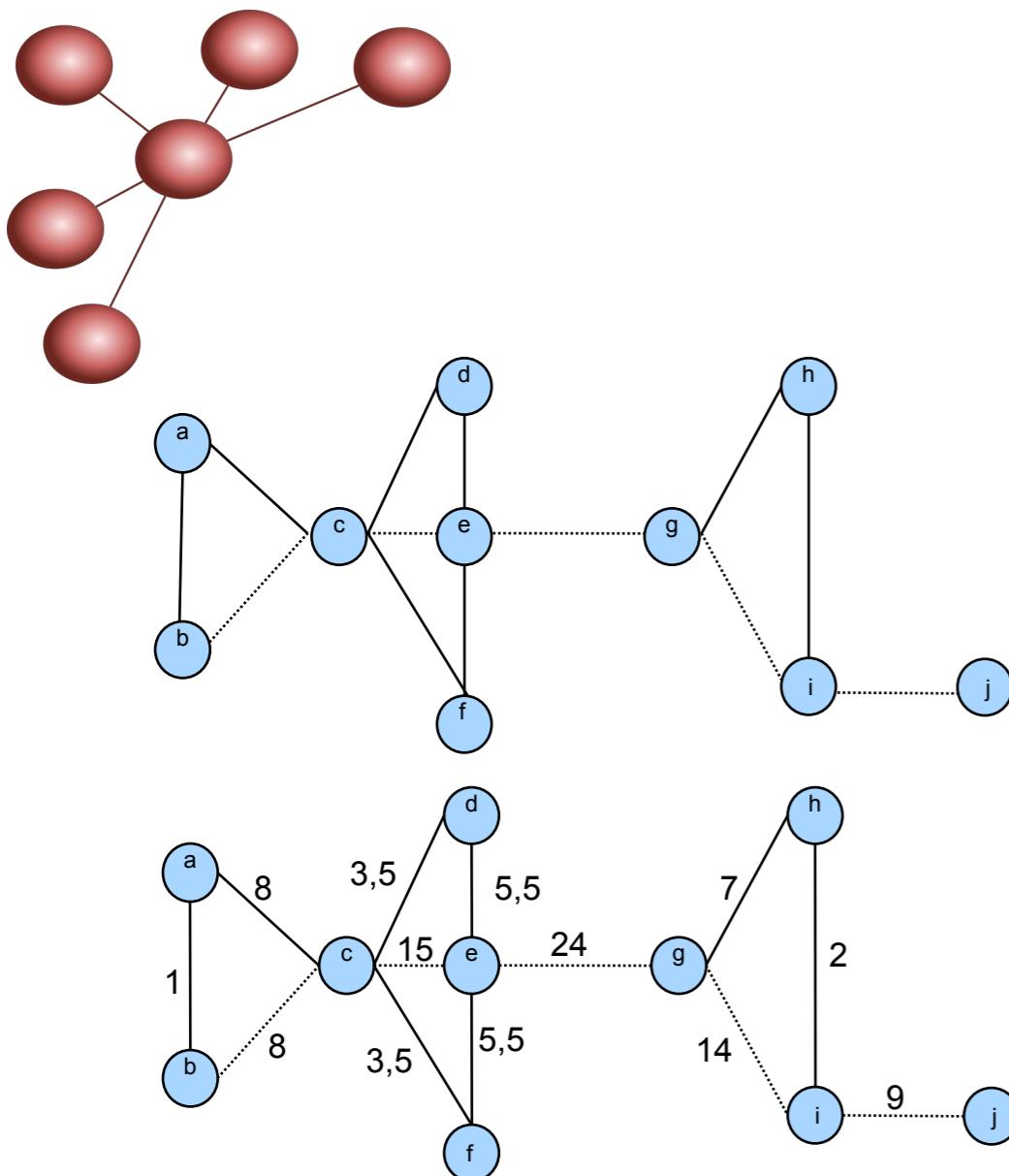


## NetworkAnalyzer

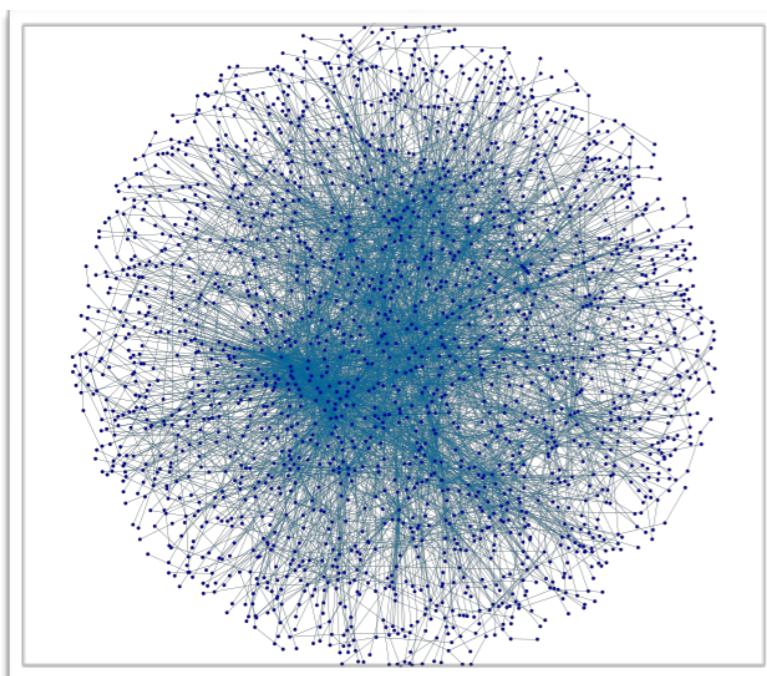
Core App: Computes basic properties of whole network (degree distribution, clustering coefficients, centrality, etc.)

| NETWORK MEASURES  |  |  |   |   |
|---|--|--|---|---|
| Degree/<br>connectivity (k)                                 | Clustering coefficient/<br>interconnectivity (C)   | Assortativity/average nearest<br>neighbor's connectivity (NC)                                | Shortest path (SP)<br>between two nodes           | Betweenness/<br>centrality (B)  |
| <p><math>k_A = \text{Nb of edges through } A = 5</math></p> | <p><math>C_A = \frac{\text{Actual links between } A\text{'s neighbors (black)}}{\text{Possible links between } A\text{'s neighbors (orange)}}</math></p> <p><math>C_A = n_A / [k_A(k_A - 1)/2] = 2 / [4 \times (4 - 1)/2] = 0.333</math></p> | <p><math>NC_A = (k_B + k_C + k_D + k_E + k_J) / 5 = (5 + 2 + 2 + 3 + 1) / 5 = 2.6</math></p> | <p><math>SP_{FH} = (F, D, A, B, H) = 4</math></p> | <p><math>B_4 = \text{Fraction of SPs passing through } A = 0.090</math></p> |

# “Betweenness”



Number of shortest paths  
running through an edge  
= “bootleneck”

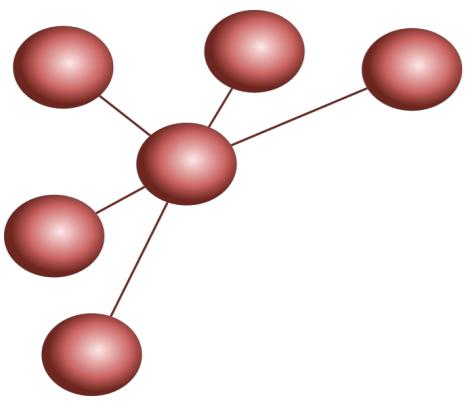


Biological interpretation ?  
Correlation with gene essentiality, gene  
involvement in diseases, importance in  
flux transmission ...

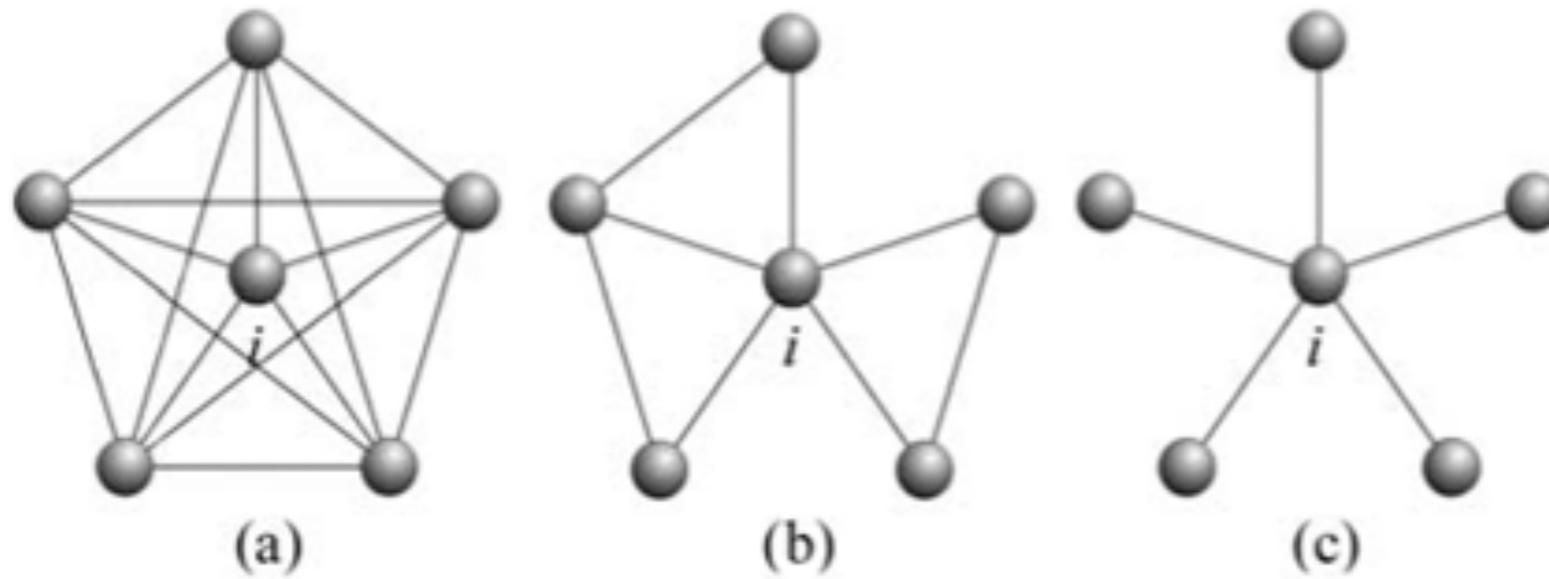


**NetworkAnalyzer**

Core App: Computes basic properties of whole network (degree distribution, clustering coefficients, centrality, etc.)



# Clustering coefficient / modularity

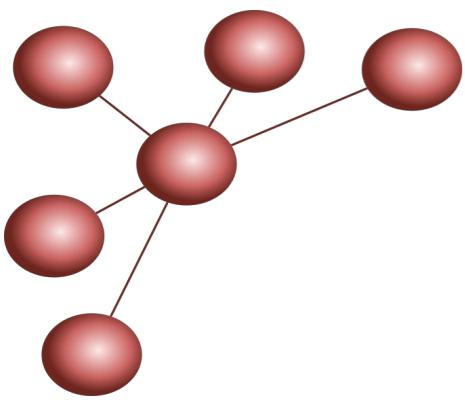


**Actual links between neighbours / Possible links between neighbours**

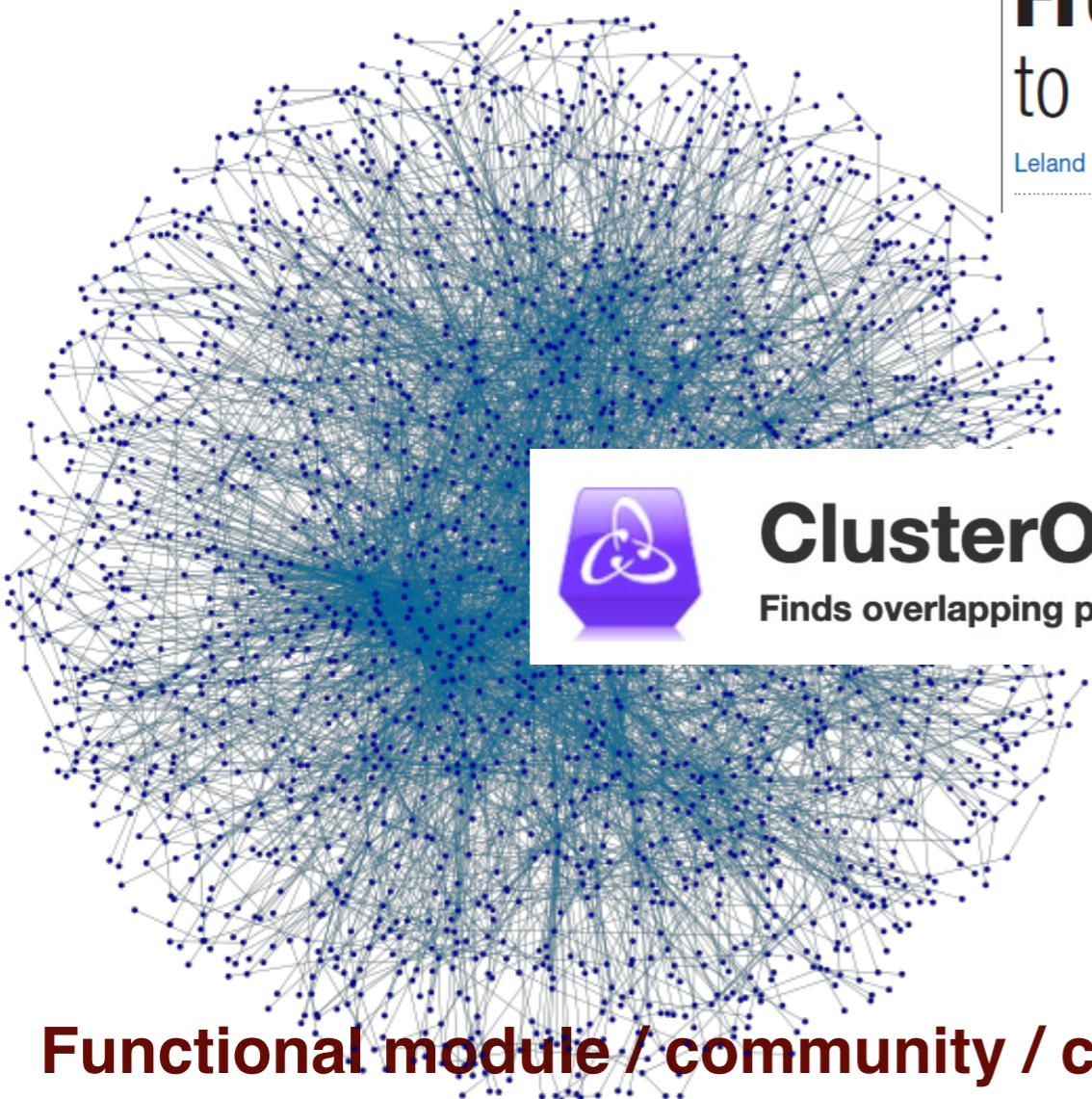


**NetworkAnalyzer**

Core App: Computes basic properties of whole network (degree distribution, clustering coefficients, centrality, etc.)

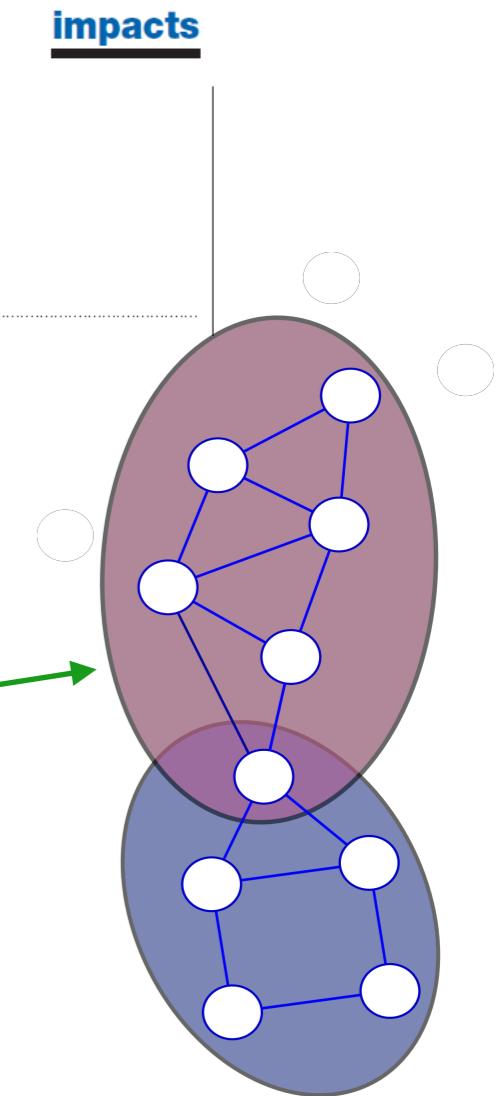


# Global approach - Clustering



**From molecular  
to modular cell biology**

Leland H. Hartwell, John J. Hopfield, Stanislas Leibler and Andrew W. Murray



**Functional module / community / cluster / class : discrete function**

**Modules can be isolated or connected**

**Groups of proteins involved in a common cellular function**



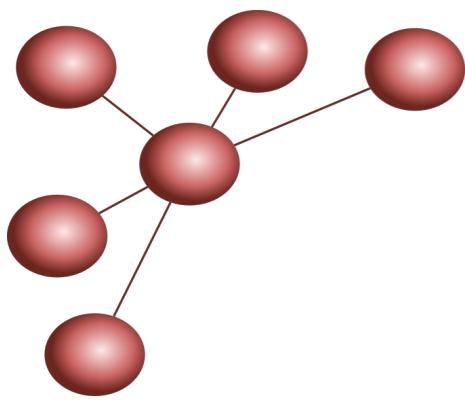
**clusterMaker2**

Multi-algorithm clustering app for Cytoscape

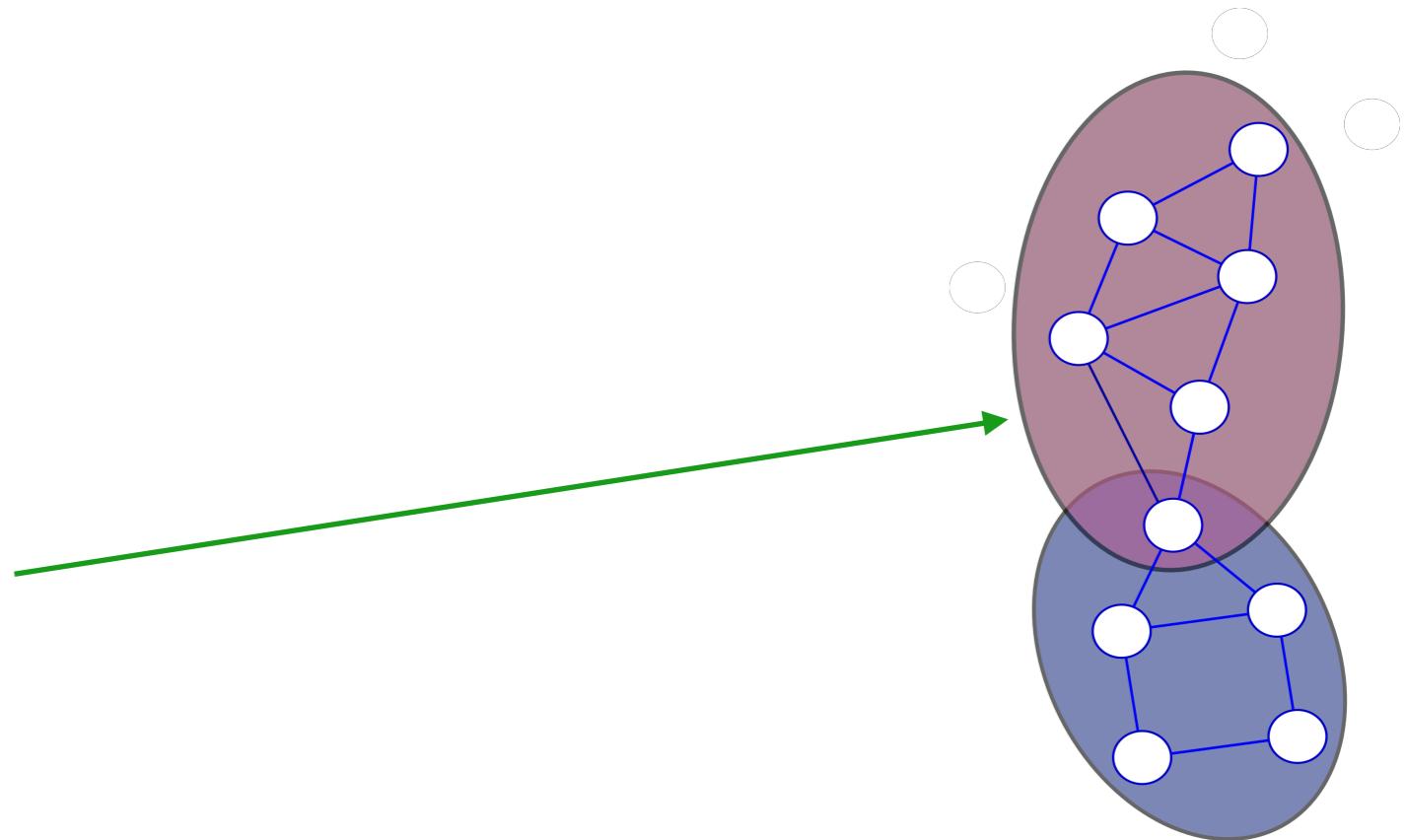
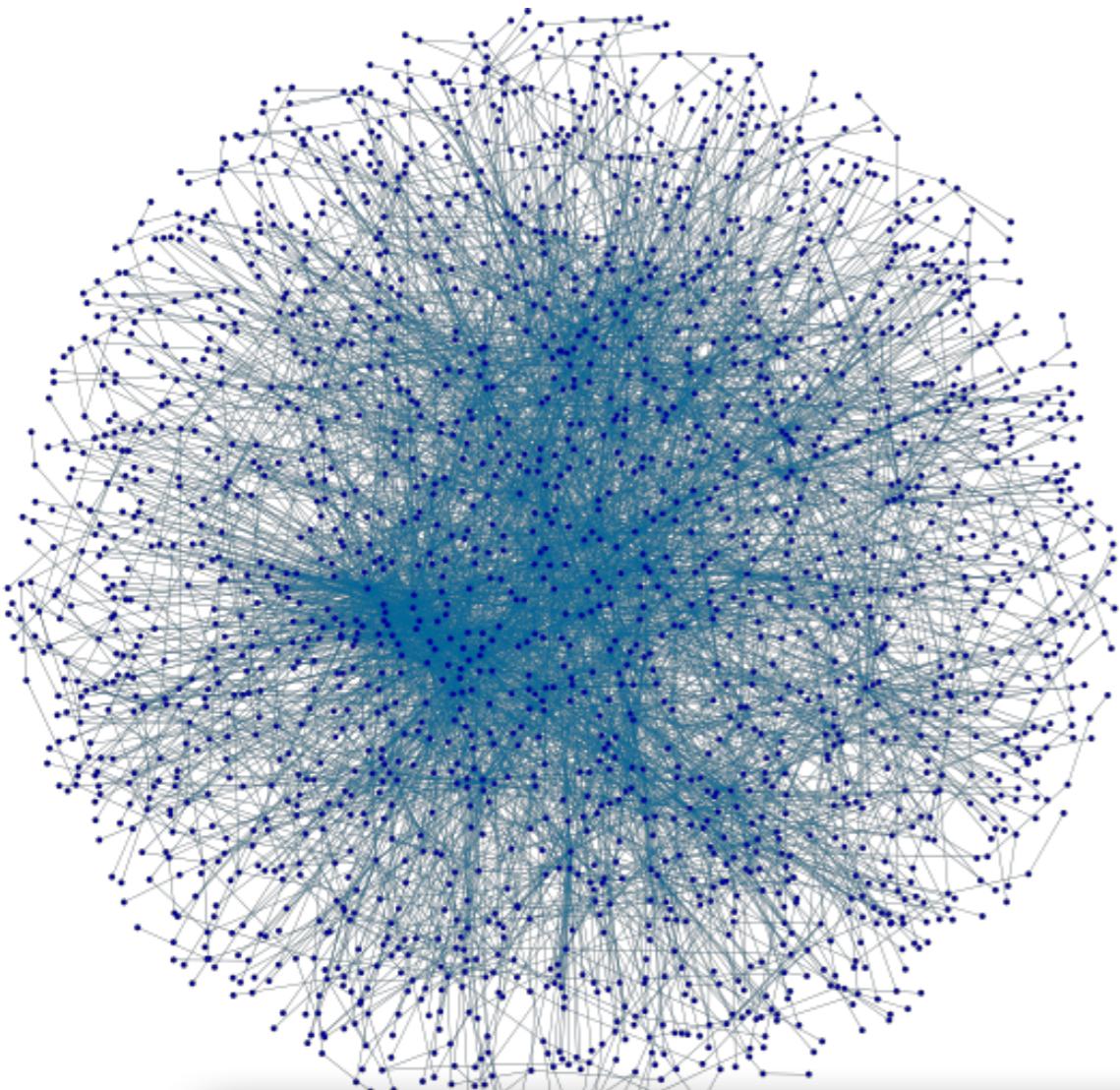


**MCODE**

Clusters a given network based on topology to find densely connected regions.



# Global approach - Clustering



**ClusterONE**

Finds overlapping protein complexes in a protein interaction network.



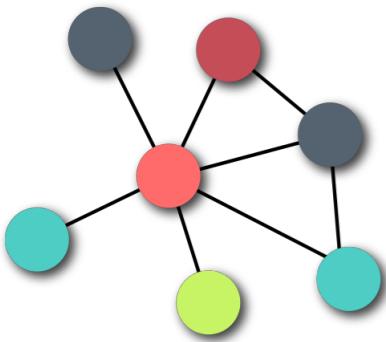
**clusterMaker2**

Multi-algorithm clustering app for Cytoscape



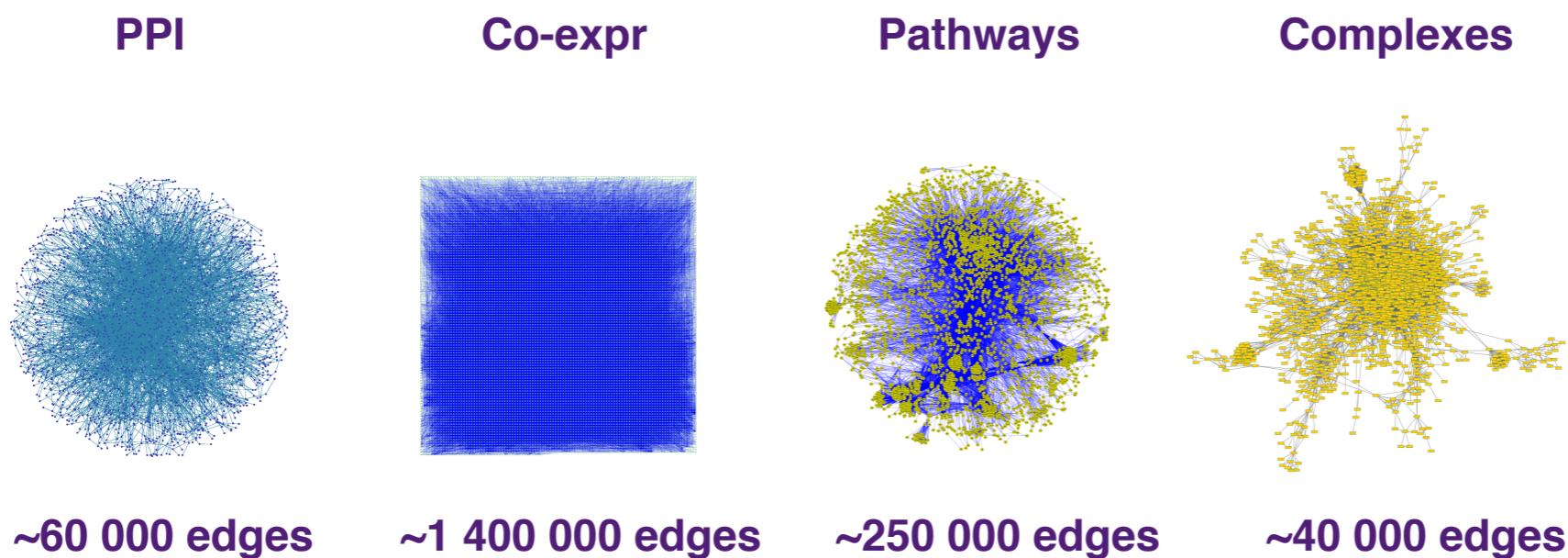
**MCODE**

Clusters a given network based on topology to find densely connected regions.

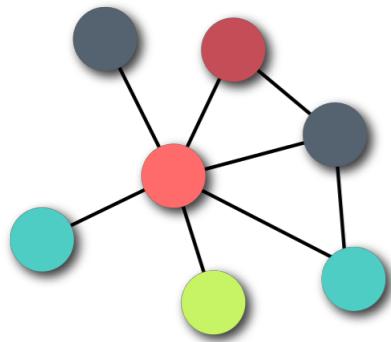


# Integration of networks

- Diverse interaction sources
- Each own features, topology, bias

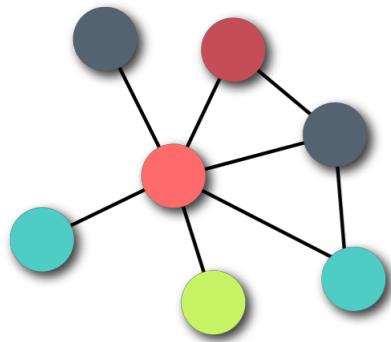


How do we combine many networks /  
interaction sources?



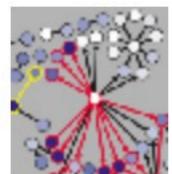
# Networks as scaffolds to integrate

- Knowledge
- Expression data
- Networks of networks



# Integration of networks and expression data

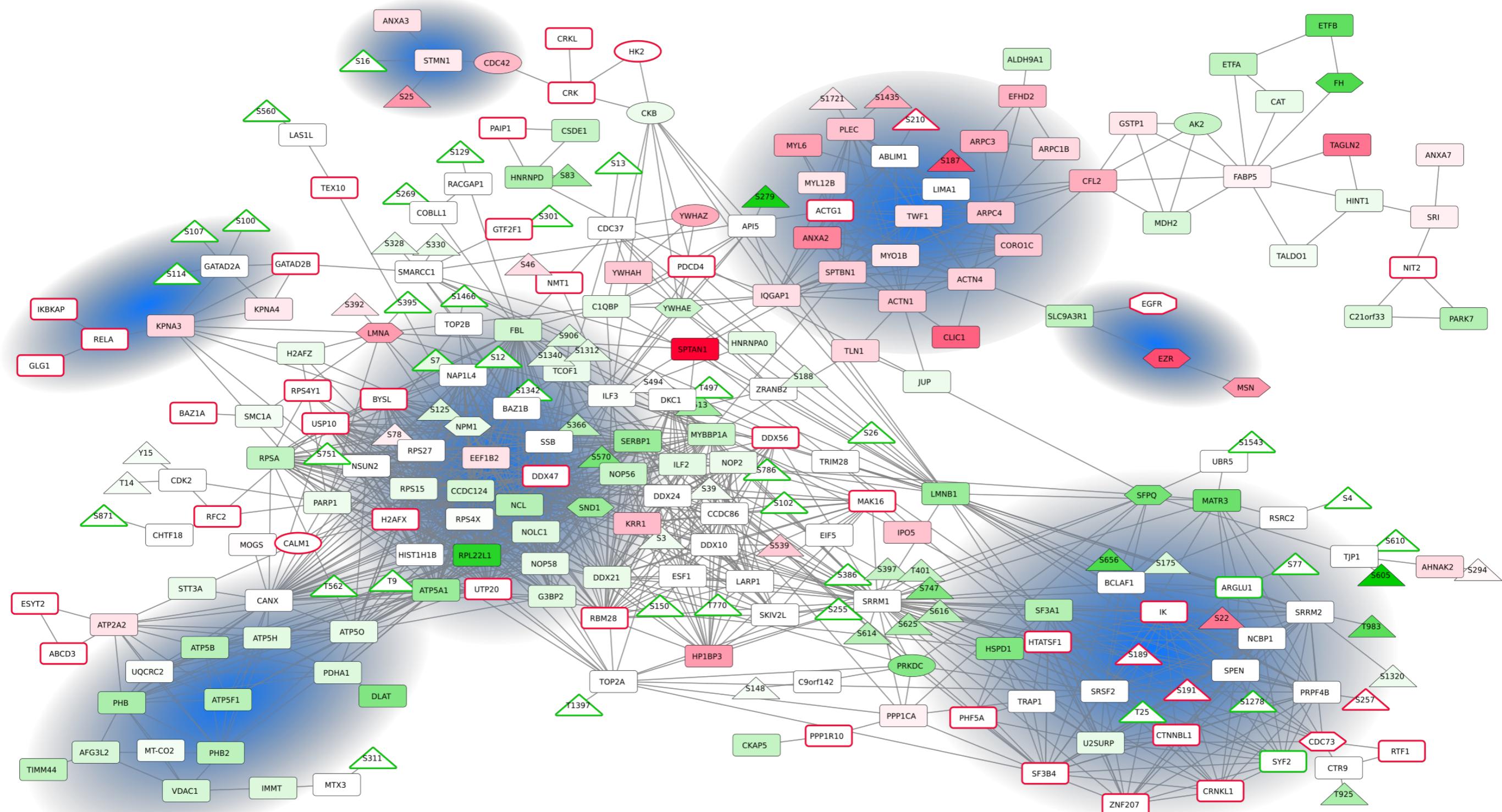
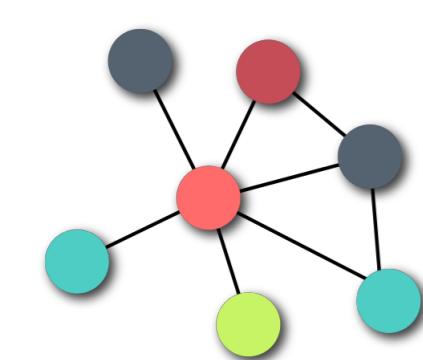
- => identify “active” subnetworks
- App jActive module

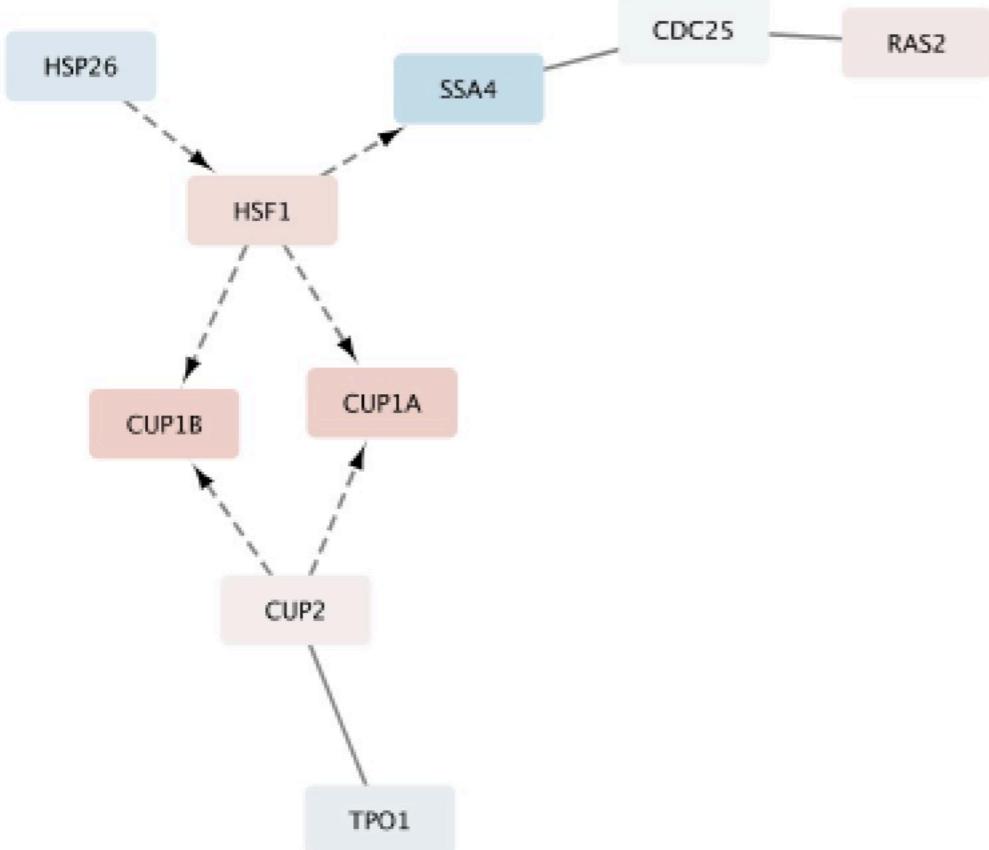
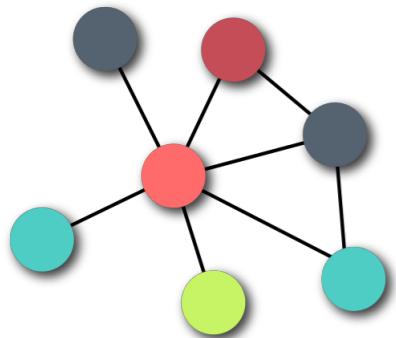


## jActiveModules

Finds clusters where member nodes show significant changes in expression levels.

# Integration of networks and expression data





- **z-score of a subnetwork based on expression p-values**
- **find the bests subnetworks => need heuristic algorithms**