

Systems Biology Across Scales: A Personal View

XIII. Intra-cellular Systems III: Protein-Protein Interaction

Sitabhra Sinha

IMSc Chennai

Intra-cellular biochemical networks

□ Metabolic networks

Nodes: metabolites (substrates & products of metabolism)

Links: chemical reactions (directed)

□ Genetic regulatory networks

Nodes: Genes & Proteins

Links: regulatory interactions (directed)

□ Protein-Protein interaction network

Nodes: Proteins

Links: physical binding and formation of protein complex (undirected)

□ Signaling network

Nodes: Signaling molecules e.g., kinase, cyclicAMP, Ca

Links: chemical reactions (directed)

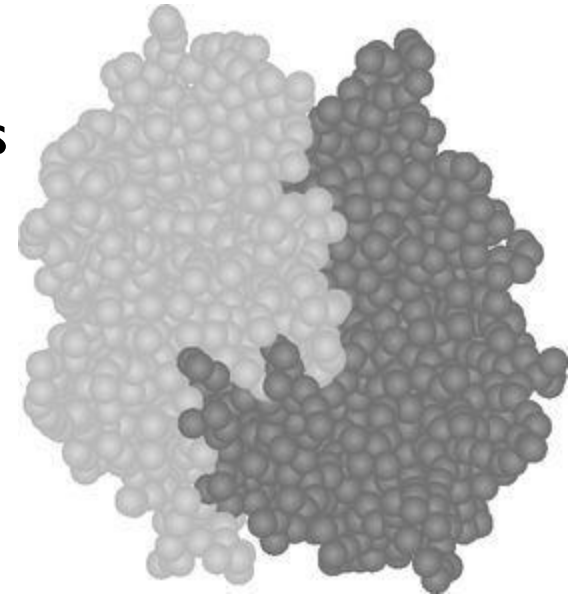
Protein-Protein Interaction

Proteins interact with other proteins primarily through physical binding – locking of 3-D shapes creating *protein complexes*

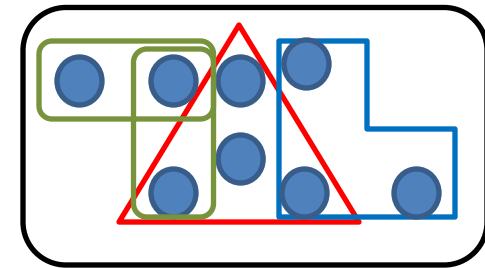
[Other modes of interaction include chemical – e.g., in signaling networks – where small subgroups such as a phosphate group (in phosphorylation) is exchanged.]

Interactions involving more than 2 proteins should ideally be represented using **hypergraphs** – but usually conveyed approximately by a number of pairwise interactions

Set of all such undirected pairwise interactions constitute the Protein-Protein Interaction Network (PPIN)



Newman, *Networks: An Introduction*



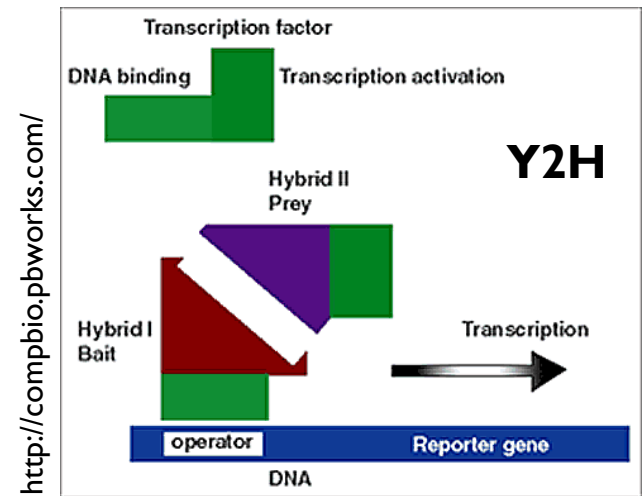
Experimental techniques for PPI

<http://compbio.pbworks.com/>

Methods have different sensitivity and specificity

high sensitivity \Rightarrow many interactions that occur in reality are detected by the screen.

high specificity \Rightarrow most of the interactions detected are also occurring in reality.



Yeast two-hybrid screen investigates interaction between artificial fusion proteins inside the nucleus of yeast - can fish out binding partners of a protein

But has a high false-positive rate

Necessary to verify identified interactions by co-immunoprecipitation

Co-immunoprecipitation (gold standard assay for protein-protein interactions) Protein of interest is fished out of the cells with a specific antibody. Interaction partners sticking to this protein are subsequently identified by western blot

Interactions detected by this approach considered to be real - but can only verify interactions between suspected interaction partners – NOT a screening approach.

Tandem affinity purification (TAP) detects interactions within the correct cellular environment (e.g. in the cytosol of a mammalian cell) (Rigaut et al., 1999) – advantage compared to Y2H approach.

But requires 2 successive steps of protein purification – cannot readily detect transient PPI

Yeast protein interaction network

$N = 1870$ nodes, $L = 2240$ links

PPIs determined by Y2H method

Node color represents phenotypic effect of removing corresponding protein

Red: lethal

Green: nonlethal

Orange: slow growth

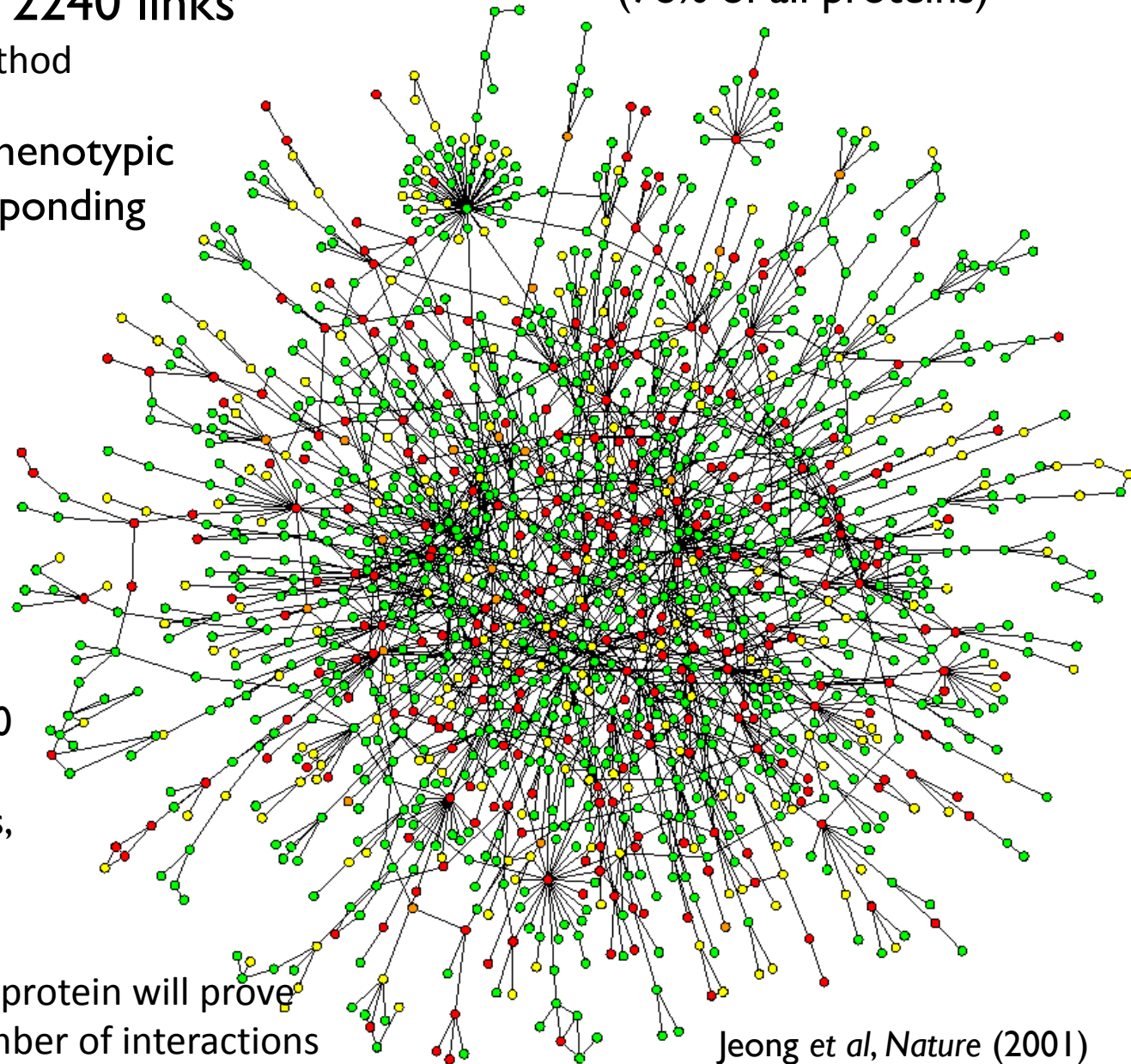
Yellow: unknown

Degree distribution $P(k)$ is power-law having exponent ~ 2.4 over a limited range – exponential cut-off at $k = 20$

Tolerance to random errors, fragility w.r.t. removal of most connected nodes

Likelihood that removal of a protein will prove lethal correlates with its number of interactions

Largest connected cluster (78% of all proteins)

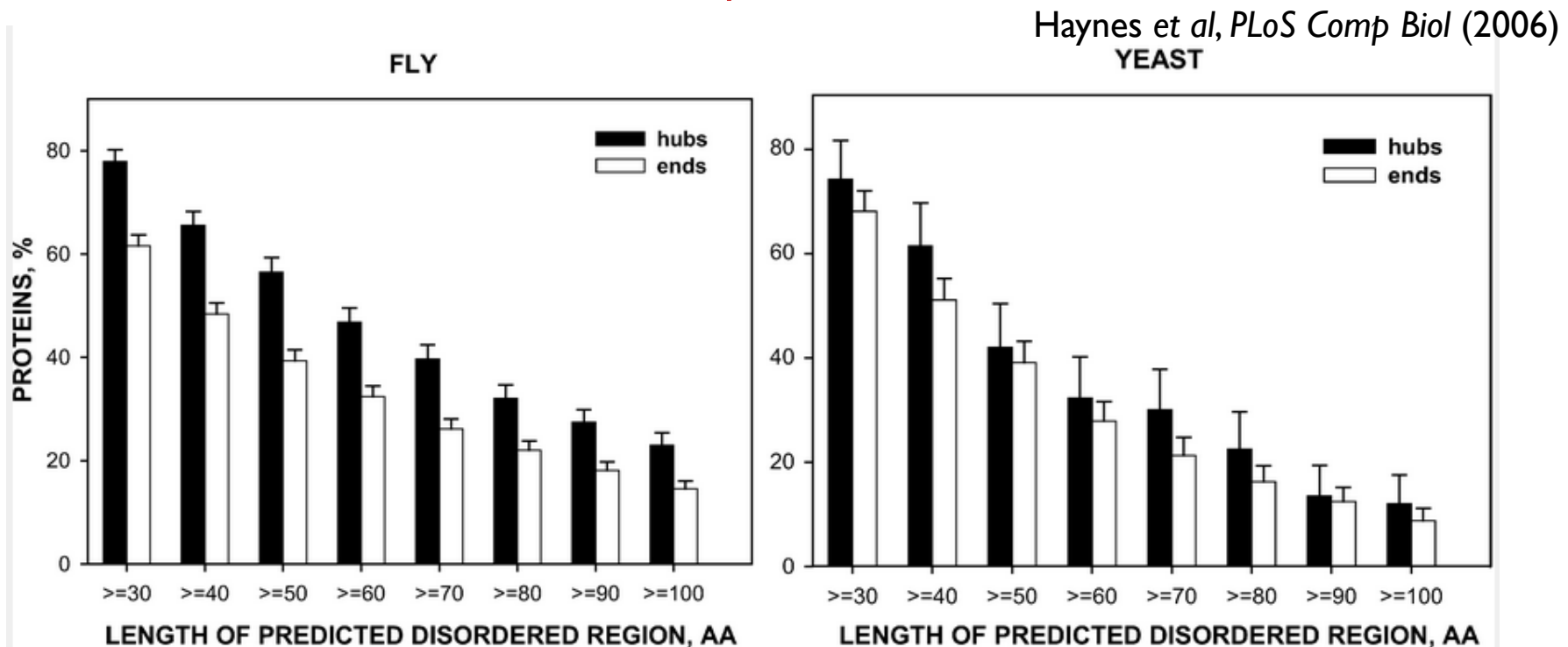


Jeong *et al*, *Nature* (2001)

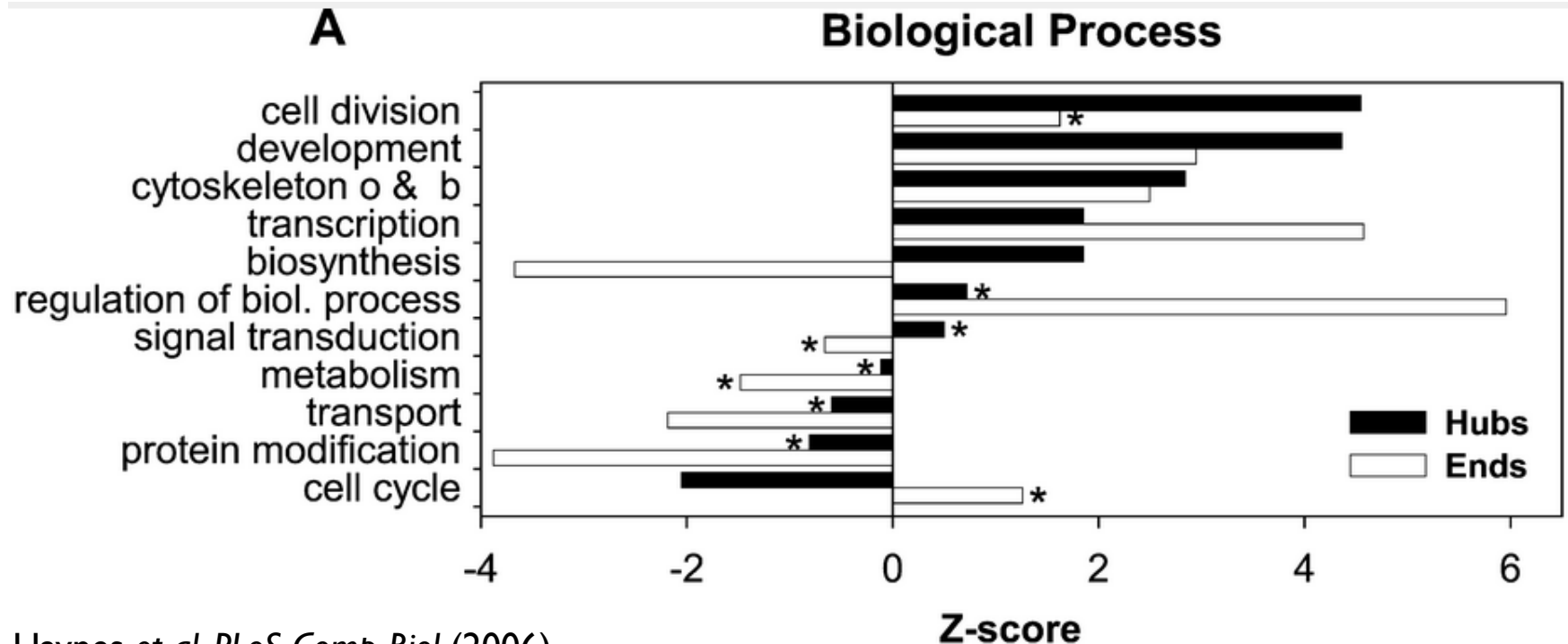
Are hub proteins special ?

Hub proteins are more likely to be intrinsically disordered (ID)

ID proteins and protein regions lack a unique 3-D structure and exist in a dynamic ensemble of conformations. Many ID proteins shown to mediate interactions via disorder-to-order transition on binding to their biological targets
Ability to recognize multiple binding partners with distinct interaction surfaces \Rightarrow more efficient hubs relative to ordered proteins



What processes are they involved in?



Haynes *et al*, *PLoS Comp Biol* (2006)

A positive (negative) Z-score indicates that more (less) disorder is associated with the indicated biological process than would be expected by chance

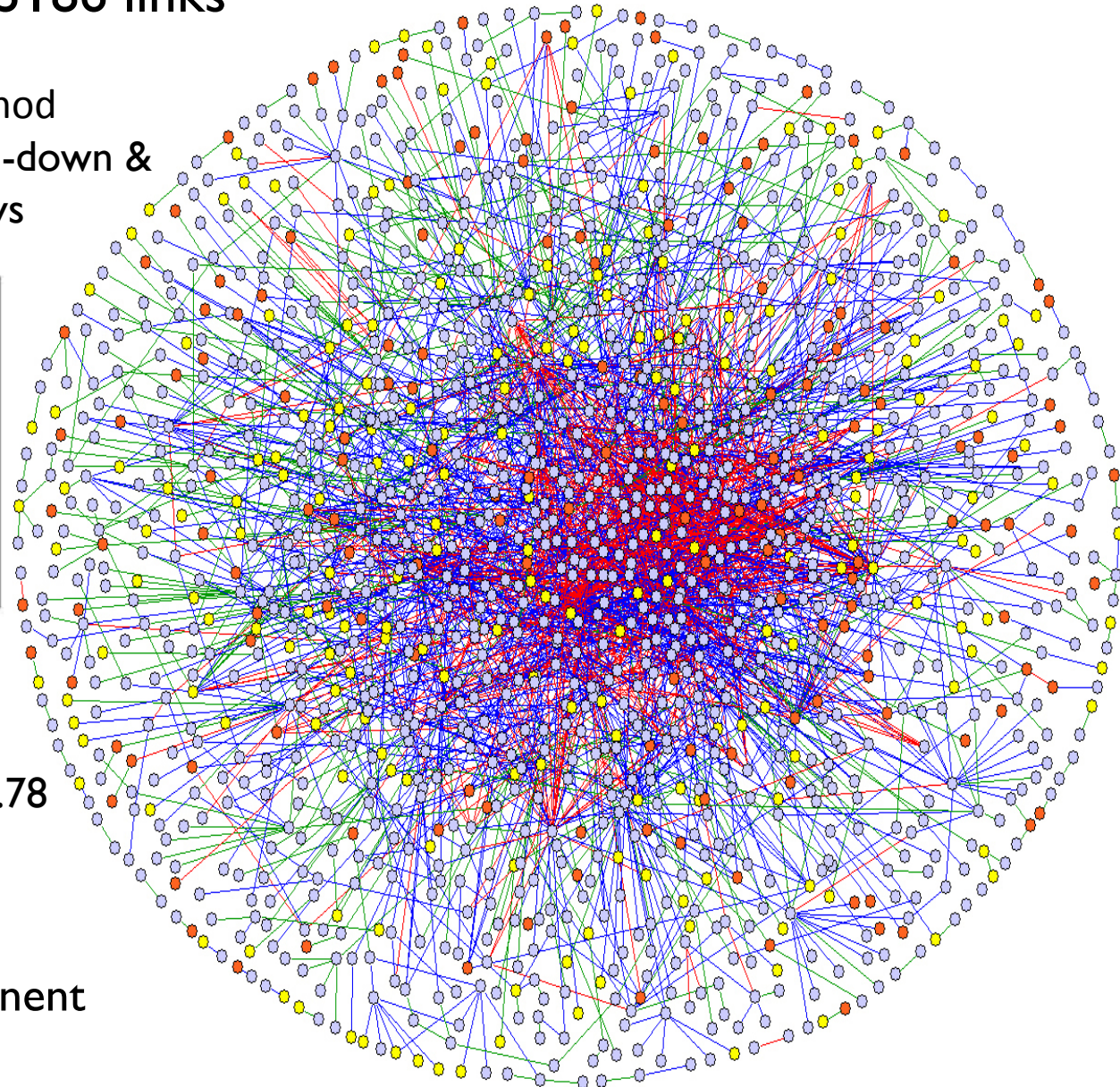
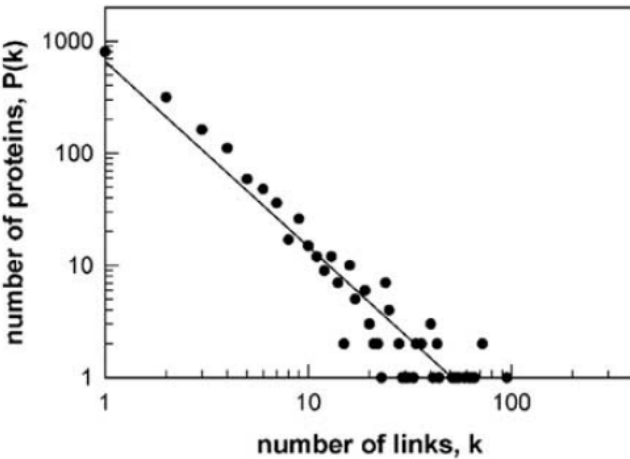
Disorder is enriched in both hubs and ends for several processes including development, cytoskeleton organization and biogenesis, and transcription: consistent with the hypothesis that disorder is highly involved in functions specific to eukaryotes. Hubs are only significantly depleted for cell cycle, whereas the ends are significantly depleted for biosynthesis

Human protein interaction network

Stelzl *et al*, *Cell* (2005)

$N = 1705$ nodes, $L = 3186$ links

PPIs determined by Y2H method
Validated by independent pull-down &
co-immunoprecipitation assays



Degree distribution $P(k)$ is
power-law with exponent ~ 1.78

Avg degree ~ 1.87

24 hubs with > 30 links

Largest connected component
comprises 1613 proteins

Conserved patterns of protein interaction

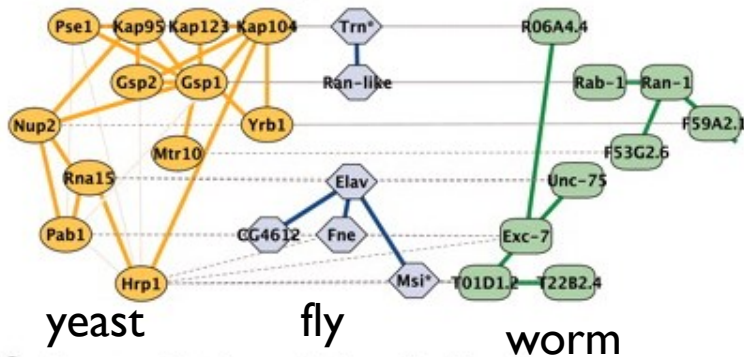
across species

Sharan et al, *PNAS* (2005)

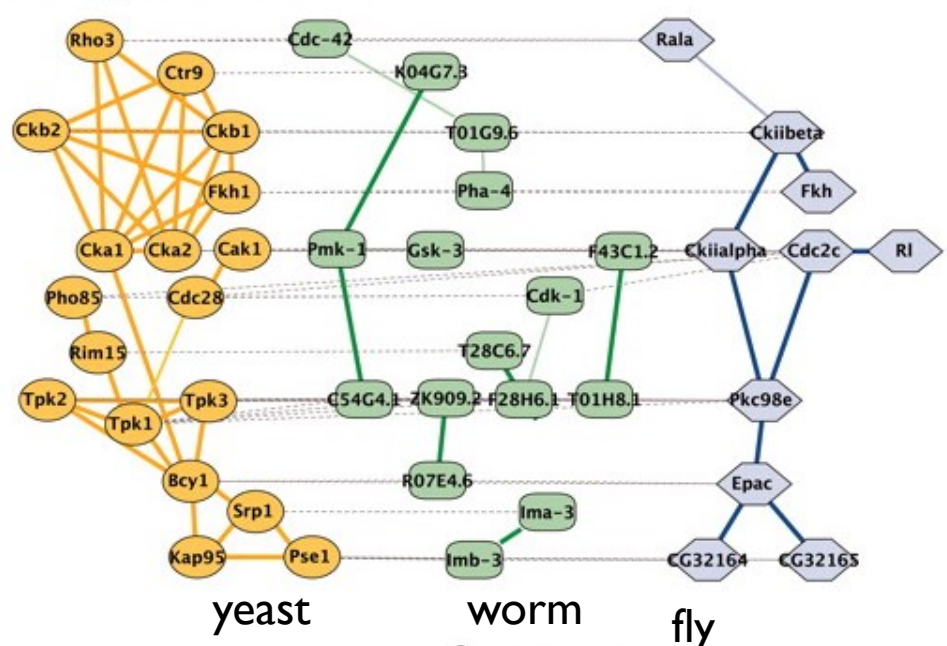
Representative conserved network clusters

Horizontal dotted gray links → cross-species sequence similarity between proteins

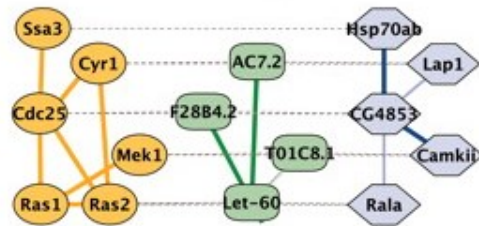
a Intracellular transport



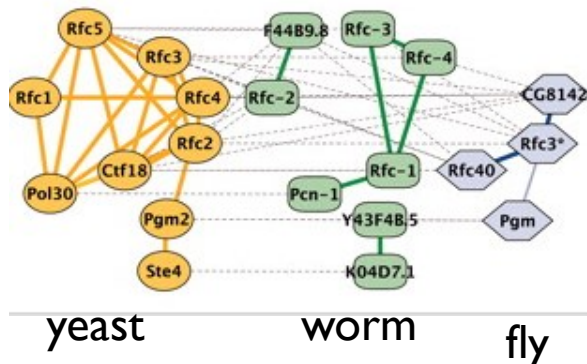
b Phosphorus metabolism



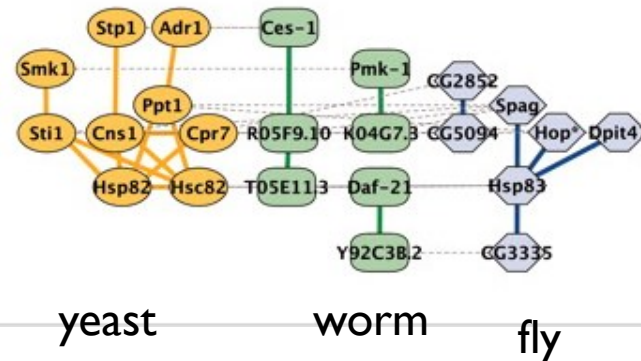
c Ras-mediated regulation of cell cycle



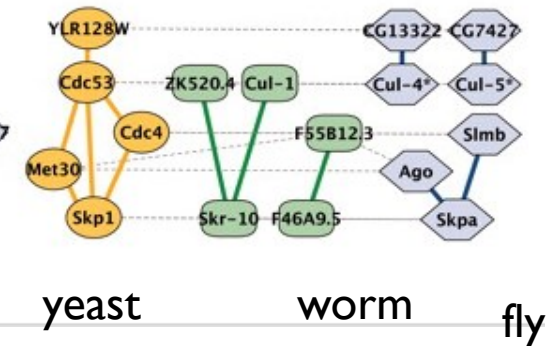
d DNA metabolism



e Protein folding



f Cell proliferation

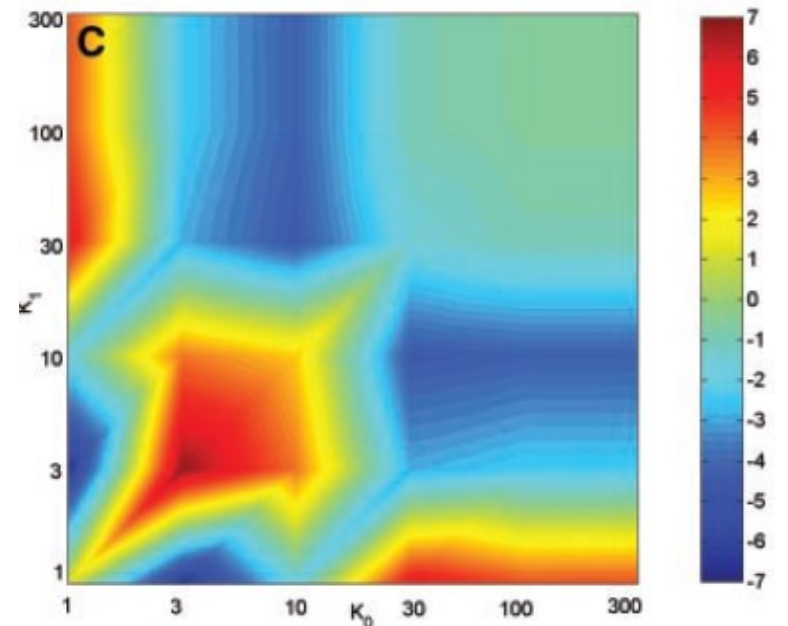
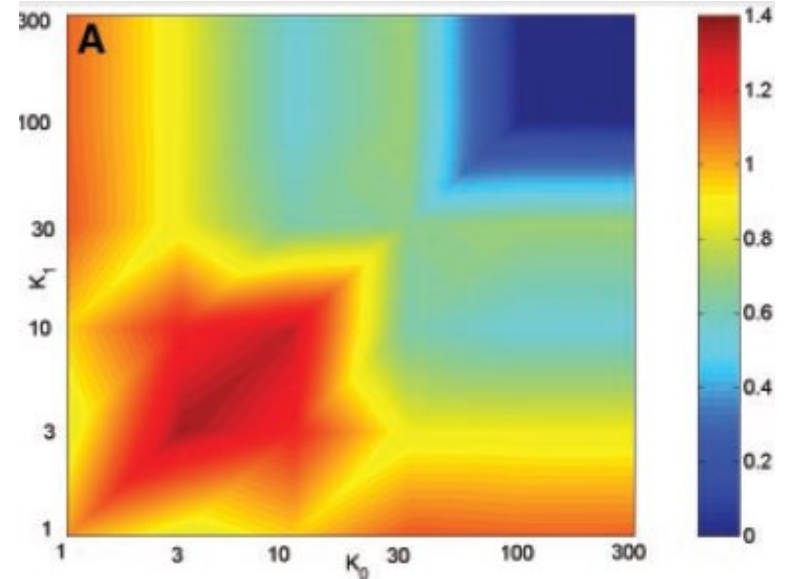


Yeast PPIN is disassortative

Maslov & Sneppen, *Science* (2002)

“links between highly connected proteins are systematically suppressed, whereas those between a highly connected and low-connected pairs of proteins are favored ... decreases the likelihood of cross talk between different functional modules of the cell and increases the overall robustness of a network by localizing effects of deleterious perturbations.”

Ratio of probability of connection between pairs of proteins with degree k_0, k_1 compared to that of a randomized network (top) and the corresponding z-score (bottom)



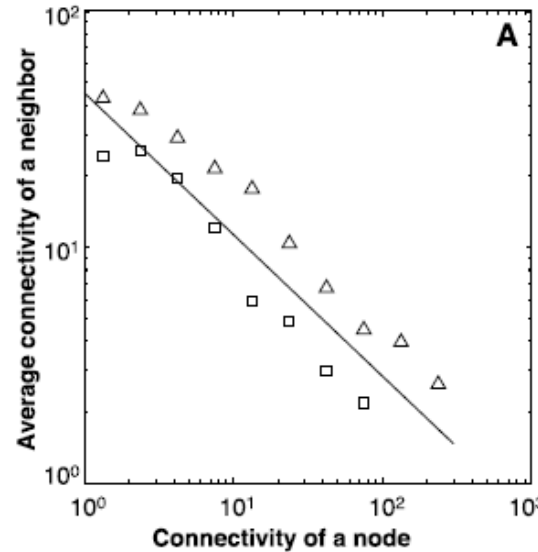
Hub-and-spoke formations

Functional modules clustered around individual hubs

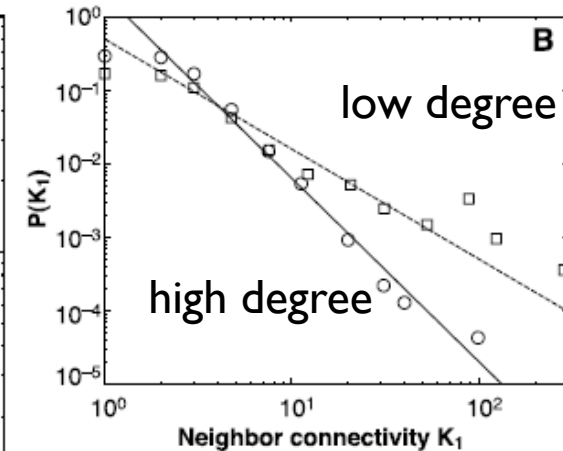


Interaction network between 329 proteins localized in the yeast nucleus that interact with at least one other protein in the nucleus.

Most neighbors of highly connected nodes have low degree → hub-and-spoke topology



The average degree $\langle k_l \rangle$ of nearest neighbors of a node, as a function of its own degree k_0 shows a power-law decay with exponent 0.6



probability distribution of nearest-neighbor degree k_l shown separately for nodes with low degree $k_0 < 4$ and for high degree $k_0 > 100$ (For uncorrelated networks, this should decay as $k_l / k_l^{\text{degree}}$ distrn exponent i.e., $1/k_l^{1.5}$) For the latter it decays as $1/k_l^{2.5}$.

“Date” hubs and “Party” hubs

Han et al, *Nature* (2004)

Filtered Yeast Interactome
Network of 1379 proteins
LCC: 778 proteins

Considering the temporal structure of PPIs

Interactome hubs characterized by expression profiling across different exptl conditions

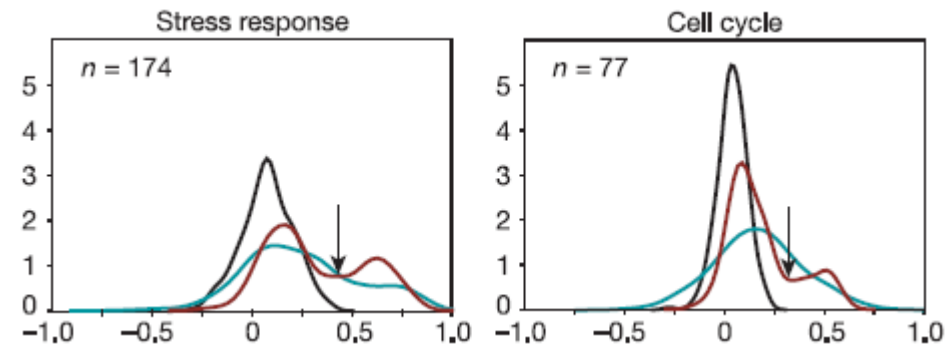
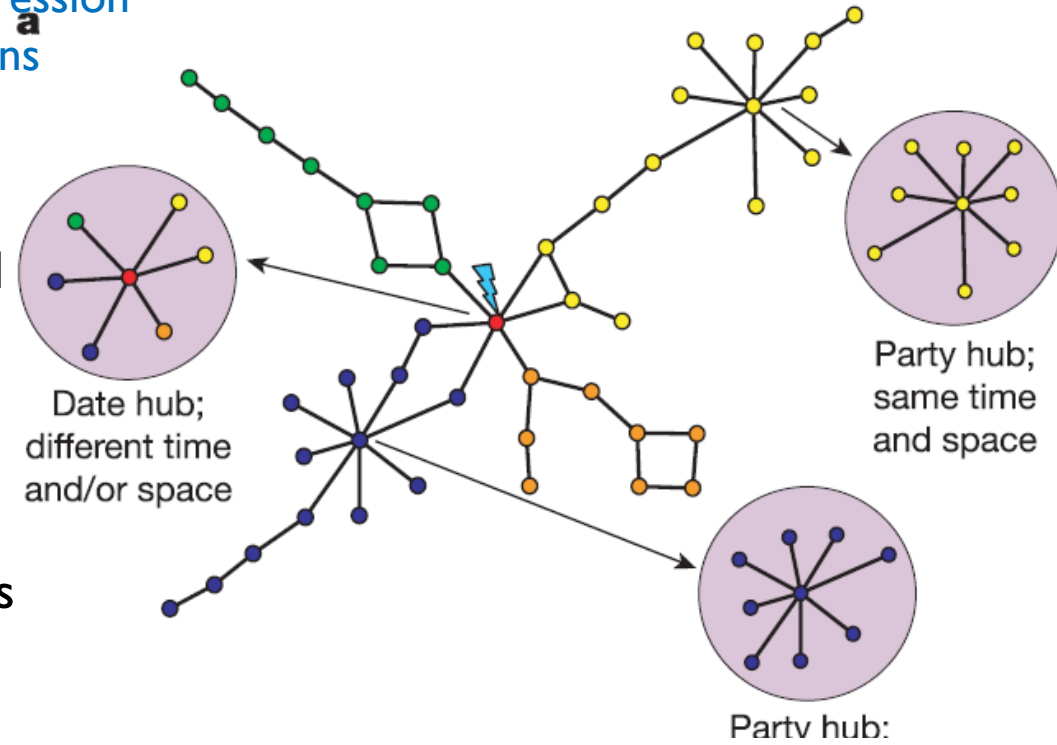
Pearson correlation coefficient (PCC) for mRNA expression of each hub and its partners calculated

PCCs of hubs, defined as nodes (proteins) with degree $k > 5$ follow a **bimodal distribution** (red curve) in contrast to non-hub nodes with $k < 5$ (cyan curve) or in randomized networks (black curve)

Bimodal distribution suggests that hubs can be split into two distinct populations:

Party hubs : relatively high avg PCCs

Date hubs: relatively low avg PCCs

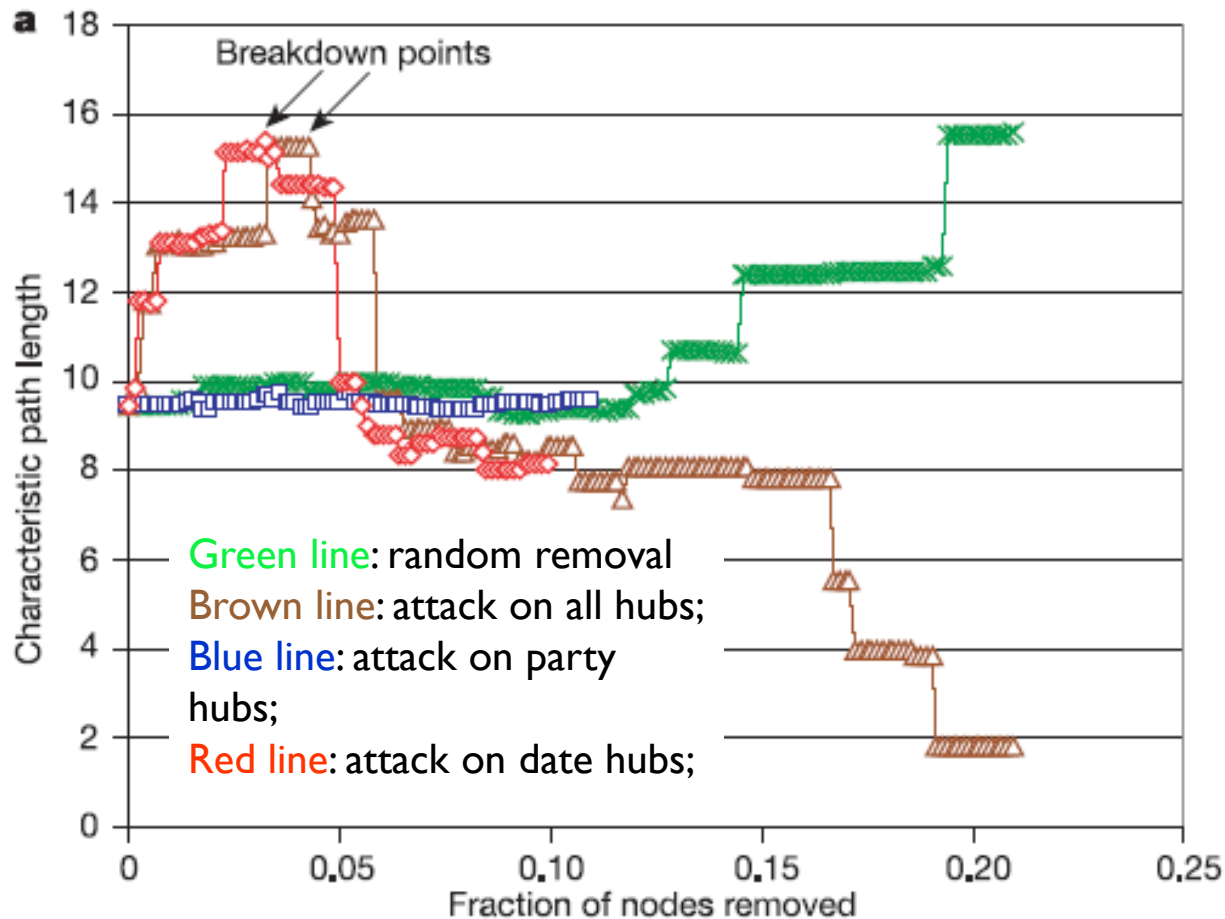


“Date” hubs and “Party” hubs

Han et al, *Nature* (2004)

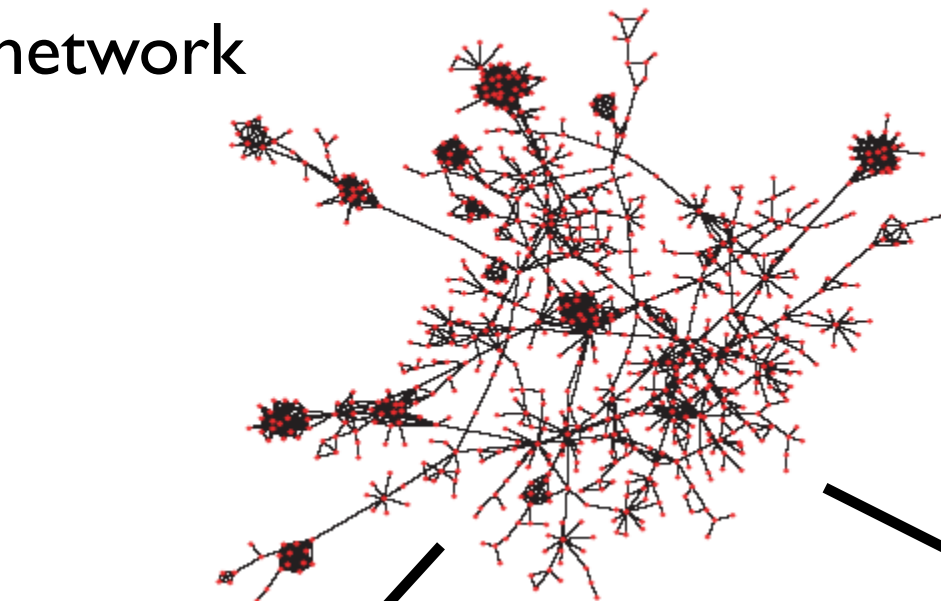
Considering the temporal structure of PPIs

When removed from interactome network, party and date hubs have distinct effects on the overall topology.

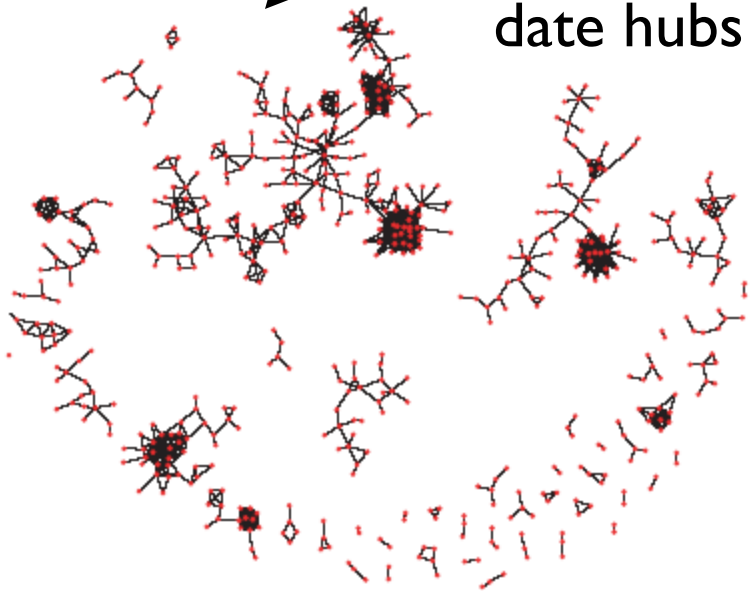


Can be observed from an *in silico* strategy that simulates the effect of specifically removing hubs in the network on the characteristic path length of the main component of the network.

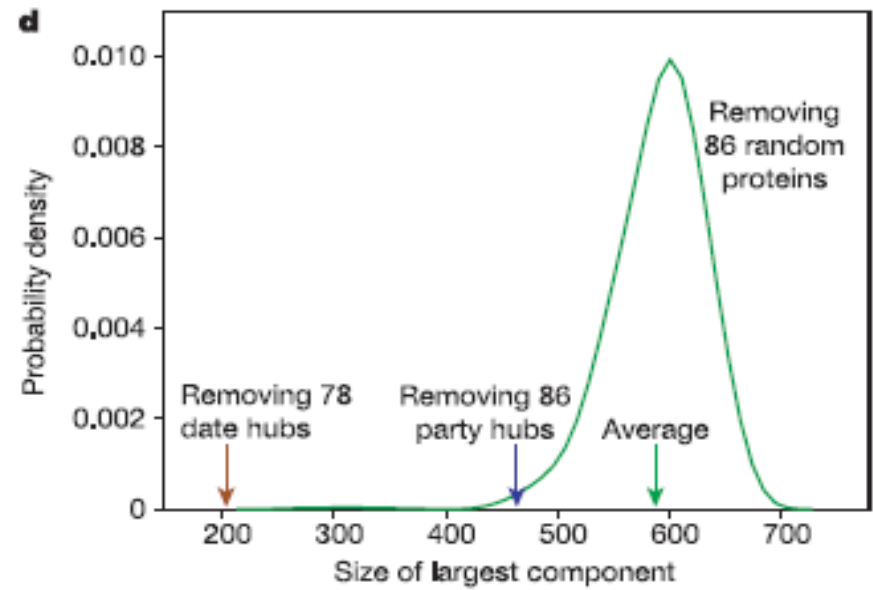
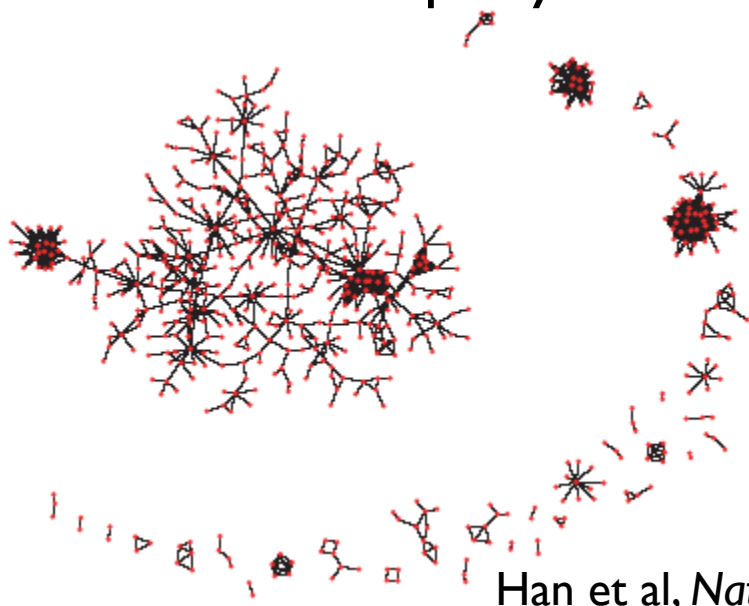
LCC of FYI network



Removal of date hubs

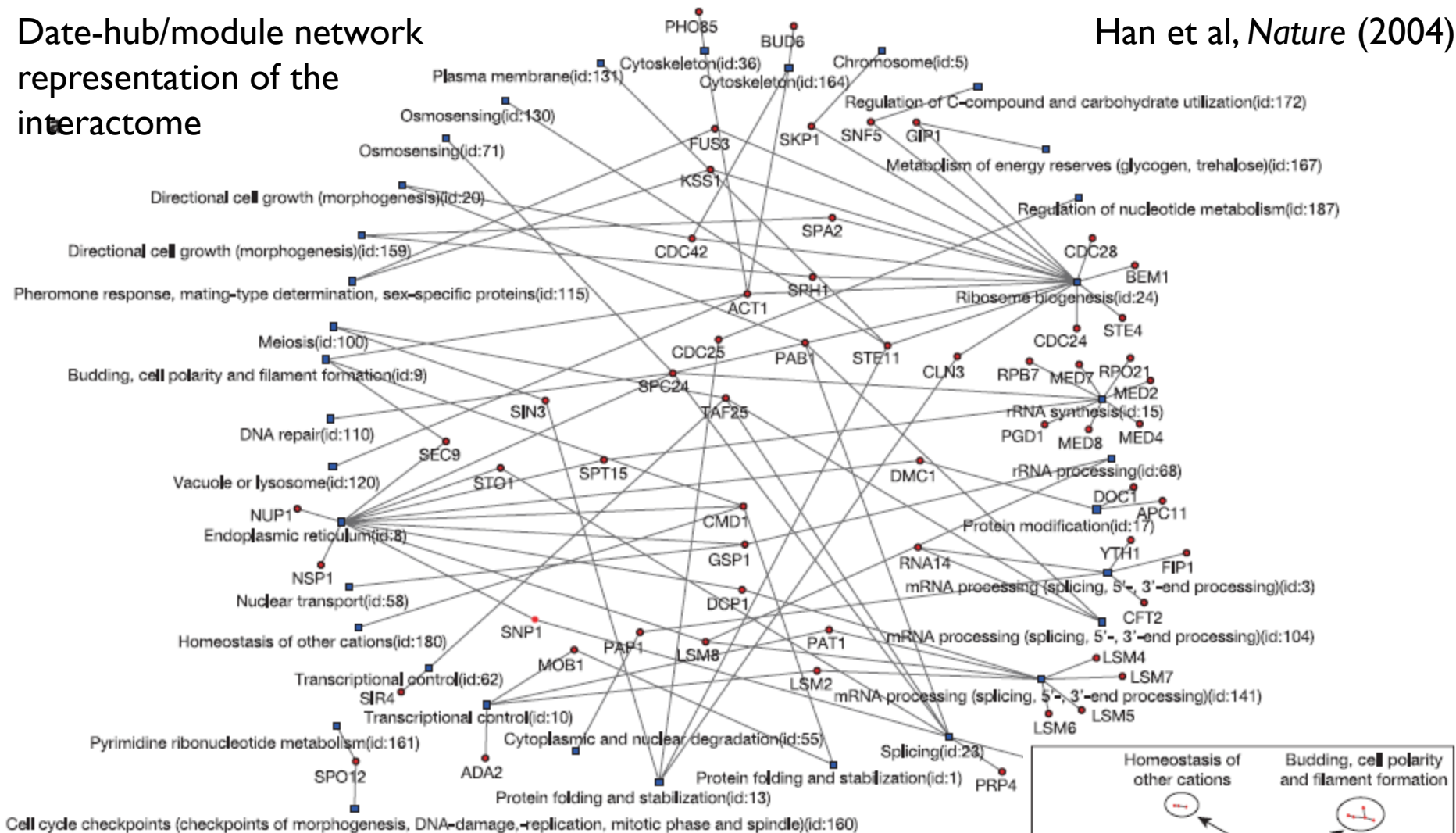


Removal of party hubs



Date-hub/module network representation of the interactome

Han et al, *Nature* (2004)



Date hubs represent global, or 'higher level, connectors between modules (red circles)

Party hubs function inside modules, at a 'lower level' of the organization of the proteome (blue squares)

