Protein-Protein Interaction Network

Lecture 3

Bioinformatic methods

- Homologous method to find Orthology
- Prediction
  - Sequence method
  - Structural based method
- Text mining
- Infer from other networks, such as expression profile, GO annotations.

Text mining

- **Text mining**, sometimes alternately referred to as *text data mining*, refers to the process of deriving high-quality or useful information from text.
- The most famous application of text mining ?
- We want to get protein interaction information from published literatures with text mining methods.

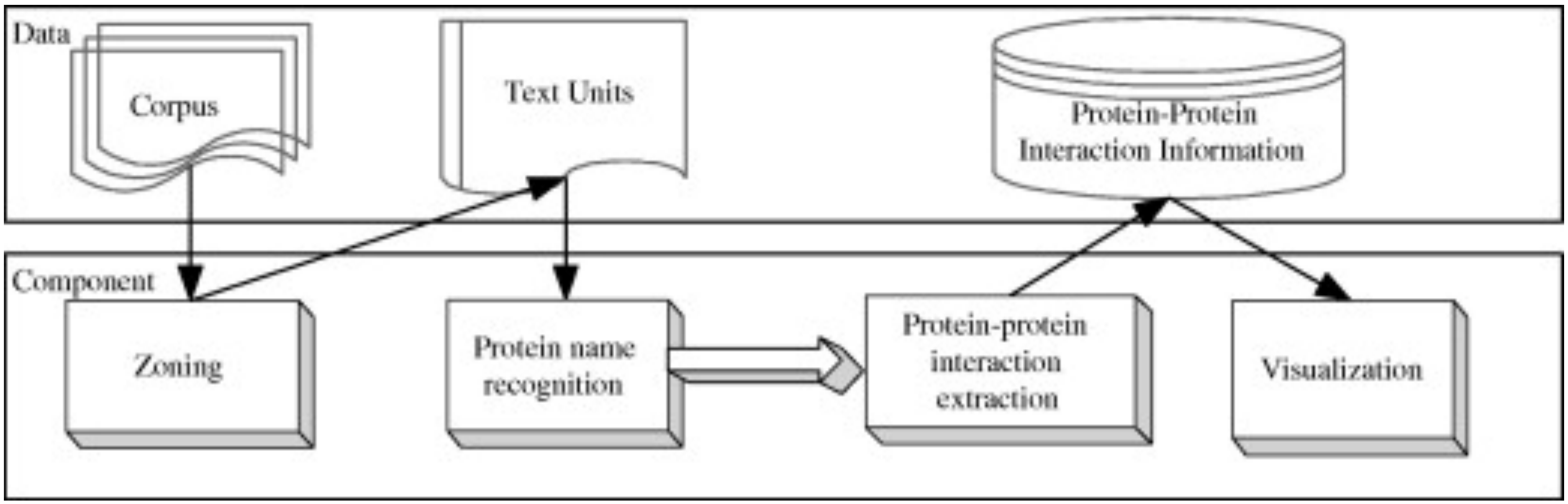
Text mining papers

- Zhou and He (2008), Journal of Biomedical Informatics, 41(2) 393.

- Mining Protein–Protein Interactions from Published Literature Using Linguamatics I2E By: Judith Bandy , David Milward, Sarah McQuay,

- Book Title: Protein Networks and Pathway Analysis Series: Methods in Molecular Biology | Volume: 563 | Page Range: 3-13

General models



Zoning module. It splits documents into basic building blocks for later analysis. Typical building blocks are phrases, sentences, and paragraphs.

Text mining methods

- Computational linguishtics-based method
  - Shallow parsing approaches
  - Deep parsing approaches
- Rule-based methods
- Machine-learning and statistical approaches

Computational linguistics-based methods

- To discover knowledge from unstructured text, it is natural to employ computational linguistics and philosophy, such as syntactic parsing or semantic parsing to analyze sentence structures.
- Methods of this category define grammars to describe sentence structures and use parsers to extract syntactic information and internal dependencies within individual sentences.

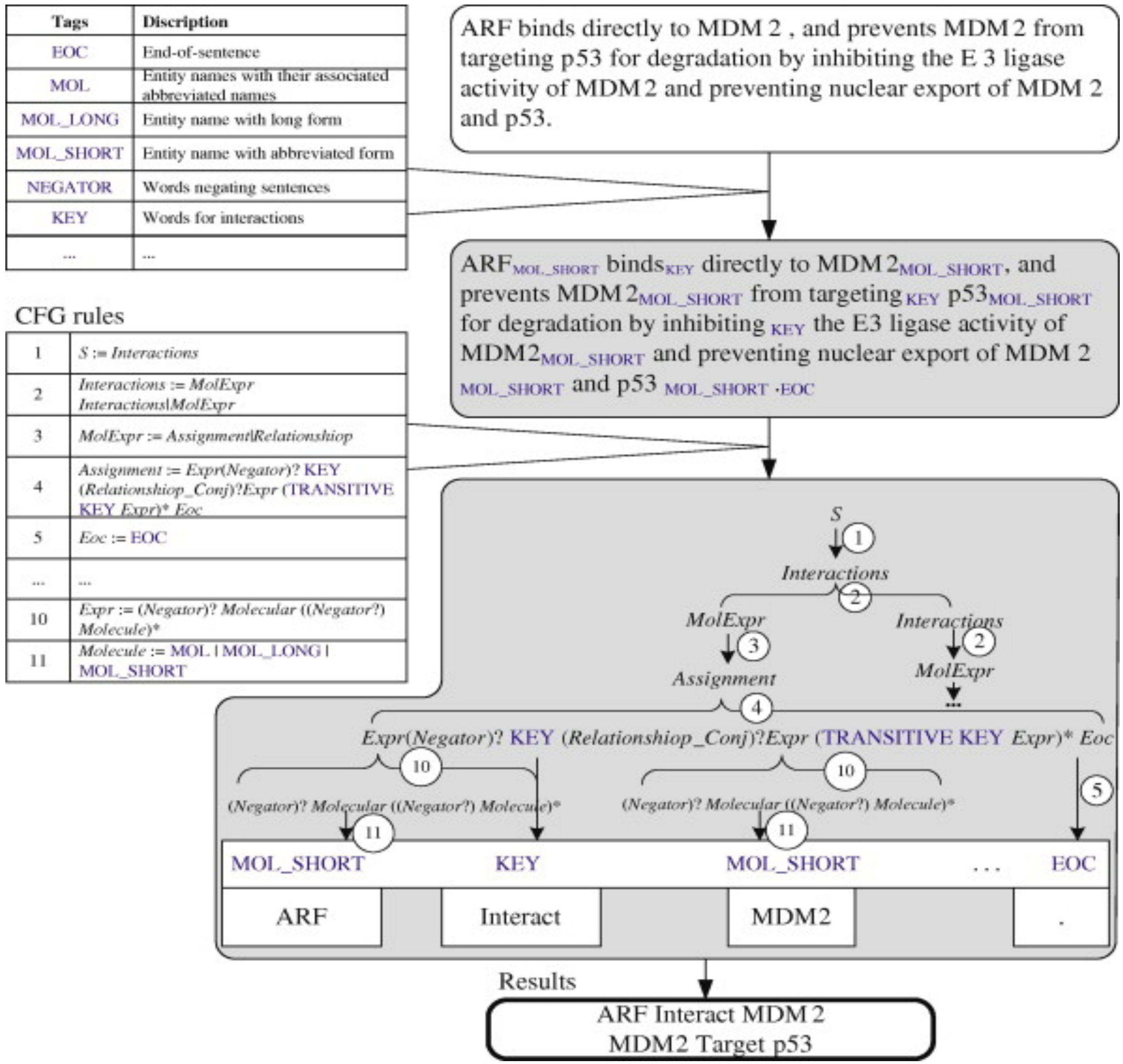
Shallow parsing approaches

- Shallow parsers perform partial decomposition of a sentence structure. They first break sentences into none-overlapping chunks, then extract local dependencies among chunks without reconstructing the structure of an entire sentence.
- For example. shallow parser generate three kinds of tags, such as syntactic, morphological, and boundary tags. Based on the tagging results, subjects and objects were recognized for the most frequently used verbs in a collection of abstracts which were believed to express the interactions between proteins, genes.

Deep parsing approaches

- Systems based on deep parsing deal with the structure of an entire sentence and therefore are potentially more accurate.
- Based on the way of constructing grammars, deep parsing-based approaches can be divided into two types: rationalist methods and empiricist methods.
- Rational methods define grammars by manual efforts
- Empiricist methods automatically generate the grammar by some observations.

An example for deep parsing



Rule-based methods

- A set of rules need to be defined which may be expressed in forms of regular expressions over words or part-of-speech (POS) tags.
- Based on the rules, relations between entities that are relevant to tasks such as proteins, can be recognized.

Rule-based methods: 3 steps

1. **Identification of protein names**
   – Protein names were first identified from sentences based on a predefined biomedical entity dictionary.
2. **Preprocessing compound or complex sentences**
   – Then predefined rules based on the generated POS tags were applied to split those complex sentences.
3. **Recognition of the protein–protein interaction**
   – For example, the defined word patterns could be “A interact with B”, “interaction of A (with—and) B”, “interaction (between | among) A and B” and so on. A and B here indicate protein names.

Machine-learning and statistical approaches

- deducing relationship between two terms based on their co-occurrences in literatures.
- If two proteins frequently appear in the same literature, these two proteins might have an interaction.
- Bayesian classifier, Neuronal work, Support Vector Machine

An Example

1. **Build the training and testing corpora**
    - Training corpus: 260 papers cited by the Database of Interacting Proteins (DIP).
    - Testing data which are denoted as *Yeast MEDLINE* were obtained from MEDLINE
2. **Construct discriminating words**
    - A dictionary was constructed containing the frequencies of the 60,000 most common words used more than three times in the *Yeast MEDLINE* abstracts
3. **Score each abstract in Yeast MEDLINE by its likelihood of discussing protein–protein interaction**

Text-mined PPIs

	Recall (%)	Precision (%)	
Shallow parsing	-	73	34,343 sentences from abstracts retrieved from MEDLINE
	29	69	2,565 unseen abstracts extracted from MEDLINE
	57	90	Training set consists of 500 abstracts from MEDLINE.
Deep parsing	48	80	492 sentences out of 250,000 abstracts on cytosine in MEDLINE
	63.9	70.2	The test corpus consists of 100 randomly selected scientific abstracts from MEDLINE
	26.9	65.6	229 abstracts from MEDLINE correspond to 389 interactions from the DIP database
Rule based	47	70	474 sentences from 50 abstracts retrieved using “E2F1”
	60	87	3343 abstracts were obtained by querying MEDLINE
	80	80	The top 50 biomedical papers were retrieved from the Internet

Online tools

- *Online protein–protein interaction information extraction systems*
  - BioRAT: a search engine and information extraction tool for biological research bioinf.cs.ucl.ac.uk/biorat
  - GeneWays: a system for automatically extracting, analyzing, visualizing and integrating molecular pathway data from the literature. geneways.genomecenter.columbia.edu
  - MedScan: a commercial system based on natural language processing technology for automatic extraction of biological facts from scientific literature such as MEDLINE abstracts, and internal text document www.ariadnegenomics.com/products/medscan.html

Online databases

- *Online tools for biomedical literature mining*
  - CBioC: uses automatic text extraction as a starting point to initialize the interaction database. cbioc.eas.asu.edu
  - Chilibot: a search software for MEDLINE literature database to rapidly identify relationships between genes, proteins, or any keywords that the user might be interested www.chilibot.net
  - GoPubMed: a search engineer that allows users to explore PubMed search results with the Gene Ontology (GO). www.gopubmed.org
  - iHOP; converting the information in MEDLINE into one navigable resource using genes and proteins as hyperlinks between sentences and abstracts. www.ihop-net.org/UniPub/iHOP
  - iProLINK is a resource to facilitate text mining in the area of literature-based database curation, named entity recognition, and protein ontology development. pir.georgetown.edu/iprolink
  - PreBIND: It identifies papers describing interactions using a support vector machine. prebind.bind.ca
  - PubGene is constructed to identify the relationships between genes and proteins, diseases, cell processes, and so on based on their co-occurrences in the abstracts of scientific papers etc. www.pubgene.org
  - Whatizit: a text processing tool that can identify molecular biology terms and linking them to publicly available databases. www.ebi.ac.uk/webservices/whatizit/info.jsf

Outline

- Protein-Protein Interaction Model
- How to get PPI
  - Y2H
  - Bioinformatics
- PPI databases
- PPI network properties
- Analysis method and applications
- Integration with other omic data

Databases that store interaction data

- Database of Interacting Proteins (DIP), http://dip.doe-mbi.ucla.edu/
- Biomolecular Interaction Network Database (BIND) , http://www.bind.ca/
- Molecular Interactions Database (MINT), http://160.80.34.4/mint/
- INTERACT http://www.ebi.ac.uk/intact/index.html
- PIBASE, http://alto.compbio.ucsf.edu/pibase/
- MIPS contains interaction data (both direct and clusters) for yeast
- SCOPPI, http://www.scoppi.org/
- Prolinks, http://mysql5.mbi.ucla.edu/cgi-bin/functionator/pronav

DIP

Database of Interacting Proteins

Jobs
Help
News
Register
Statistics
Satellites
SEARCH
SUBMIT
Software
Services
Articles
Links
Files
MIF

Search by: [protein] [sequence] [motif] [article] [pathBLAST] [Help][LOGIN]

DIP 369N **BROWSE LINKS**

Protein: Cellular tumor antigen p53

Binary Complex Functional

DIP			Cross Reference			Protein Name/Description
Interaction	**Interactor(s)**	**Links**	**PIR**	**SWISSPROT**	**GENBANK**	
DIP:88484E	DIP:32548N	•	---	Q60974	---	Nuclear receptor corepressor 1
DIP:40078E	DIP:24169N	•	---	Q64364	gi:6753390	p19ARF tumor suppressor protein
DIP:480E	DIP:1048N	•	TVHUF6	P04049	gi:66762	RAF proto-oncogene serine/threonine-protein kinase
DIP:40079E	DIP:24196N	•	---	P23804	gi:1209699	Ubiquitin-protein ligase E3 Mdm2
DIP:88486E	DIP:46345N	•	---	Q61827	---	Transcription factor MafK
DIP:40141E	DIP:24266N	•	---	Q13625	gi:16197705	(Bbp)
DIP:522E	DIP:1074N	•	TVVPT4	Q9DH70	gi:73275	large T antigen
DIP:88309E	DIP:46342N	•	---	P97302	---	Transcription regulator protein BACH1
DIP:88485E	DIP:31499N	•	---	O09106	---	Histone deacetylase 1
DIP:40140E	DIP:5978N	•	I38604	Q12888	gi:8928568	Tumor suppressor p53-binding protein 1

Tumor suppressor gene P53, PID ID “<DIP:369N>”

DIP Interaction Details

DIP LINK

[-----]

DIP 88484E			
DIP 369N	**PIR** DNMS53	**SwissProt** P02340	**GenBank** gi:2144761
	Name/Description	Cellular tumor antigen p53	
DIP 32548N	**PIR**	**SwissProt** Q60974	**GenBank**
	Name/Description	Nuclear receptor corepressor 1	

Evidence Help

Type	Method	Details	Source	Curation	IMEx
E(d)	anti bait coimmunoprecipitation	•	PMID:19011633	DIP	•
V	SMSC(1)	---			

DIP services

Database of Interacting Proteins

Jobs
Help
News
Register
Statistics
Satellites
SEARCH
SUBMIT
Software
Services
Articles
Links
Files
MIF

[SERVICES:top][EPRI][PVM][DPV] [Help][LOGIN]

DIP SERVICES

Large collections of data, such as the DIP database that gathers information about nearly 11,000 protein-protein interactions, provide a unique opportunity for data anaysis.

The *DIP Services* page provides access to the methods of data analysis that, at their core, utilize the vast amount of information embedded within the DIP database.

Available Services	
EPR Index	Expression Profile Reliability Index (*EPR Index*) evaluates the quality of a large-scale protein-protein interaction data sets by comparing the expression profile of the interacting dataset with that of the high-quality subset of the DIP database.
PVM Score	The Paralogous Verification (*PVM*) method judges an interaction probable if the putatively interacting pair has paralogs that also interact
DPV Score	The Domain Pair Verification (*DPV*) method judges an interaction probable if potential domain-domain interactions between the pair are deemed probable

Expression Profile Reliability (EPR)
Homology methods -Paralogous Verification (PVM)
Domain Pair Verification (DPV)

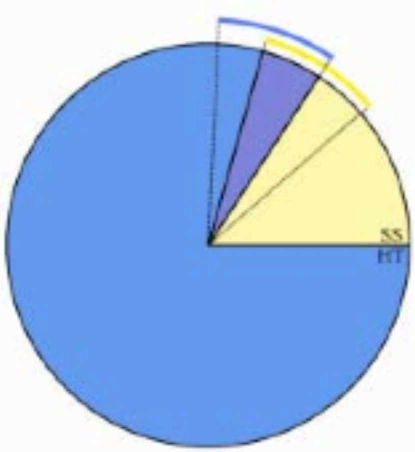
DIP interaction statistics

	All	IMEx DIP	IMEx All
Number of proteins	23201		
Number of organisms	372		
Number of interactions	71276		
Number of distinct experiments describing an interaction	69471	16640	----
Number of data sources (articles)	4607	1602	----

SELECTED ORGANISMS	PROTEINS	INTERACTIONS	EXPERIMENTS	Details
Saccharomyces cerevisiae (baker's yeast)	5051	23860	16444	•
Drosophila melanogaster (fruit fly)	7544	22976	23260	•
Escherichia coli	2949	13688	16742	•
Caenorhabditis elegans	2660	4049	4108	•
Homo sapiens (Human)	2529	3376	4817	•
Helicobacter pylori	714	1424	1443	•
Mus musculus (house mouse)	1003	994	1284	•
Rattus norvegicus (Norway rat)	349	304	425	•
Bos taurus (cow)	129	107	154	•
Arabidopsis thaliana (thale cress)	120	129	168	•

DIP for Yeast

Saccharomyces cerevisiae (baker's yeast)			
PROTEINS	**INTERACTIONS**	**#Exp**	**#Int**
4749	15658	1	13636
		2	1270
		3	402
		4	165
		5	81
		6+	98



Yeast interactions by experiment type:

SS - small-scale experiments

HT - high-throughput experiments

SS/HT overlap - *purple*

Bars mark interactions that were indentified in more than one experiment.

Assessing and filtering interaction data

DIP_CORE is a set of 3,003 interactions considered higher confidence.

DIP_CORE interactions either:

1. Have been observed in a small-scale experiment (2,246)
2. Have been observed in more than one experiment (1,179)
3. Have been confirmed by PVM (1,428)

EXPR DIP 40078E						
DIP 369N	PIR	DNMS53	SwissProt	P53_MOUSE	GenBank	gi:2144761
	Name/Description	cellular tumor antigen p53				
DIP 24169N	PIR		SwissProt	Q64364	GenBank	gi:6753390
	Name/Description	p19ARF tumor suppressor protein				

Evidence Help

Type	Method	Details	Source
E	Immunoprecipitation	---	PMID:9653180
V	SMSC(1)	---	---

verification field indicates that one (1) small-scale experiment supports this interaction

Deane et al. Mol. & Cell. Proteomics (2002) 1.5, 349-356

BIND

- Designed to hold direct interaction, cluster and pathway data 81,000 interactions written in ASN.1 (Abstract Syntax Notation) for computational efficiency

Interaction 13118 Mus musculus Full BIND Record **Launch Viewer:** Select Below

Molecule	Description	Molecular Function	Cellular Component	Biological Process	Experiment(s)	Links	
P53 • Trp53;TP53	Transformation related protein 53. Tumour suppressor protein with DNA binding and transcription factor function. Role in cell cycle; mutations involve [more...]	• DNA binding • transcription factor activity • protein binding	• nucleus • cytoplasm • cytosol	• protein-nucleus import\, translocation • transcription • regulation of transcription\, DNA-dependent • apoptosis • DNA damage response\, signal transduction by p53 class mediator • negative regulation of cell cycle	• Immunoprecipitation	NCBI SeqHound	[2 Pubmed Abstracts] [Other BIND data]
MDM2 • Mdm-2	Transformed Mmouse 3T3 cell Double Minute 2; nuclear phosphoprotein; LocusID:17246	• ubiquitin-protein ligase activity • protein binding • ATP binding • ligase activity	• nucleus	• start control point of mitotic cell cycle • cell growth and/or maintenance • protein ubiquitination • protein catabolism		NCBI SeqHound	

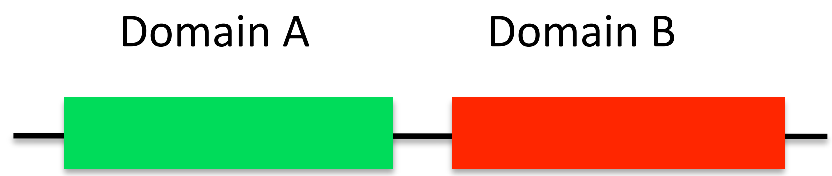
Bader GD, Betel D, Hogue CW. (2003) Nucleic Acids Res. 31(1):248-50

Arabidopsis Databases that store interaction data

- TAIR ftp://ftp.arabidopsis.org/home/tair/Proteins/Interactome2.0/
- http://bioinformatics.psb.ugent.be/supplementary_data/stbod/athPPI/site.php
- AtPIN http://bioinfo.esalq.usp.br/atpin/atpin.pl
- AtPid http://atpid.biosino.org/

Protein Domains

- In protein "language", domains could be considered as "words"
- Analyzing network graph of domains is an effective method to uncover protein functions in genome scale



PDB:1ACO

Domain-Domain interaction Database

- iPfam, http://www.sanger.ac.uk/Software/Pfam/iPfam/
- 3did (domain interactions) http://gatealoy.pcb.ub.es/3did/
- DIMA http://webclu.bio.wzw.tum.de/dima/downloads.jsp

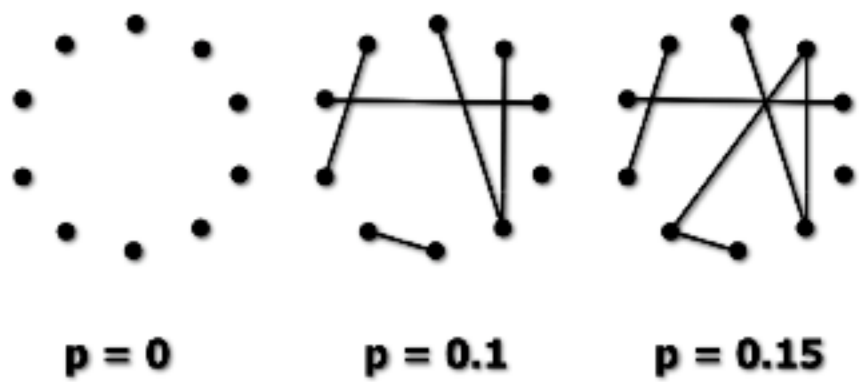
Outline

- Protein-Protein Interaction Model
- How to get PPI
  - Y2H
  - Bioinformatics
- PPI databases
- PPI network properties
- Analysis method and applications
- Integration with other omic data

Random Networks

- Uniformly random network:
  - distributes the edges uniformly among nodes.
- Probabilistic interpretation:
  - There exists a set (ensemble) of networks with given number of nodes and edges. Select a random member of this set.

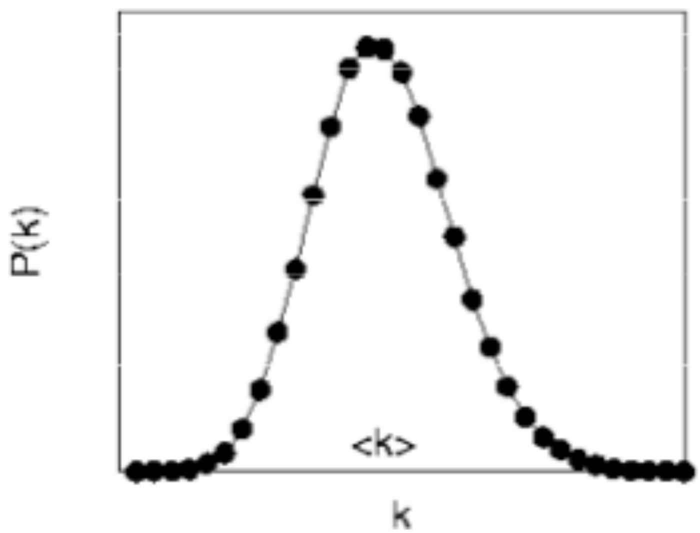
Random Networks



- fixed node number *N*
- connecting pairs of nodes with probability *p*

Expected number of edges: $$E = p\frac{N(N-1)}{2}$$

Node degrees in random graphs



Average degree:

$$\langle k \rangle \approx p|V|$$

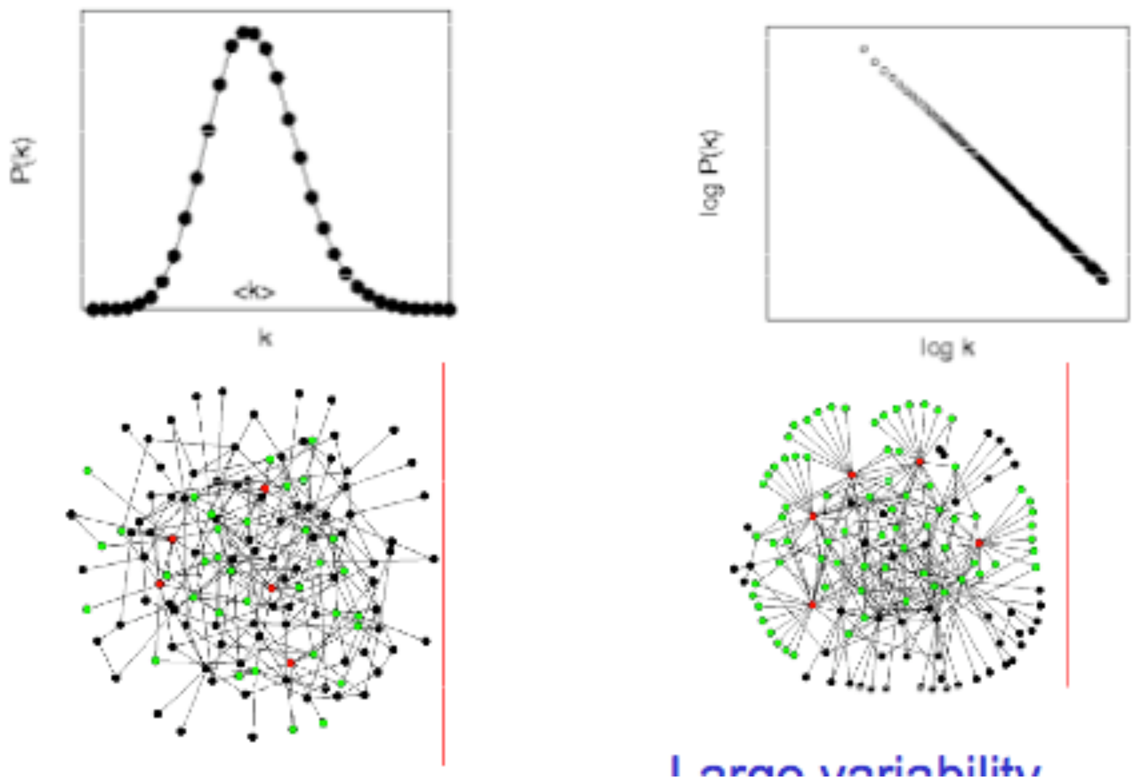
Degree distribution:

$$\mathbf{P}(k) \approx \binom{N-1}{k} p^k (1-p)^{N-1-k}$$

Most of the nodes have approximately the same degree. The probability of very highly connected nodes is exponentially small.

A scale free network

- Power-law degree distributions were found in diverse networks



A scale free network

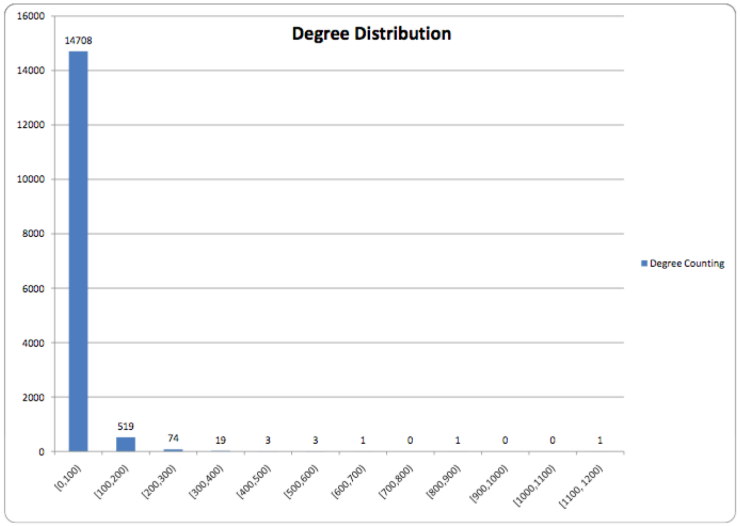
- Power-law degree distributions were found in diverse networks

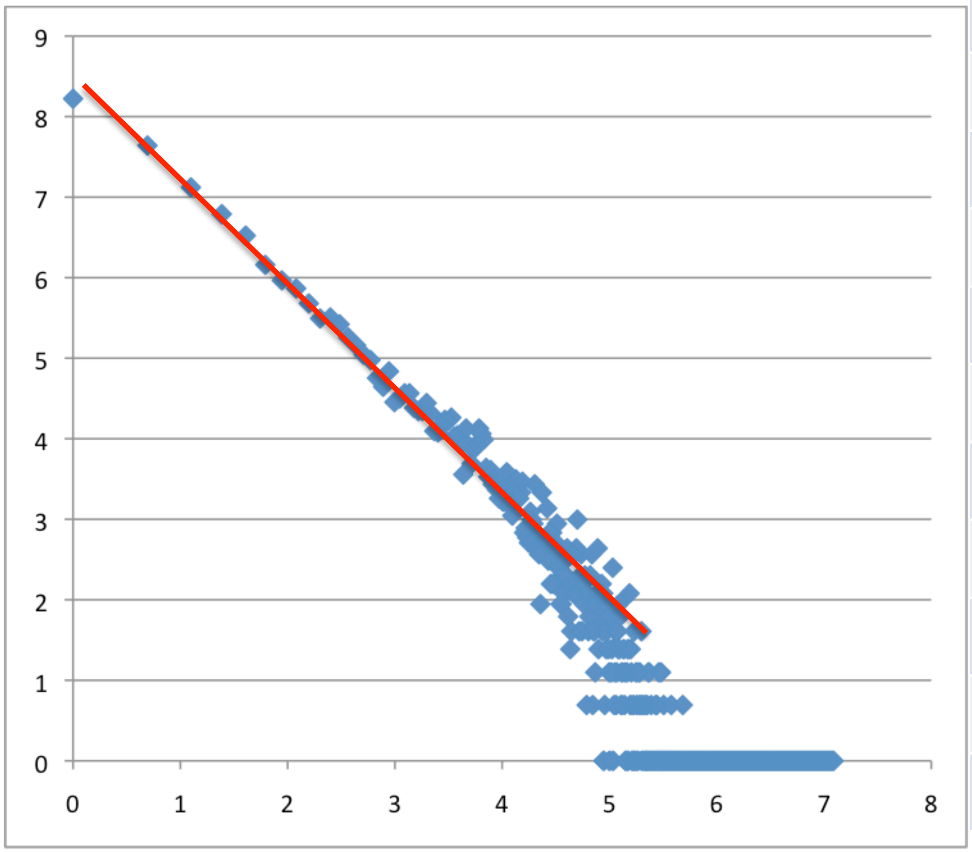
$$\log\left(\mathbf{P}(k)\right) \approx -\gamma \log(k)$$

$$\mathbf{P}(k) \approx c\, k^{-\gamma}$$

Power-law degree distributions

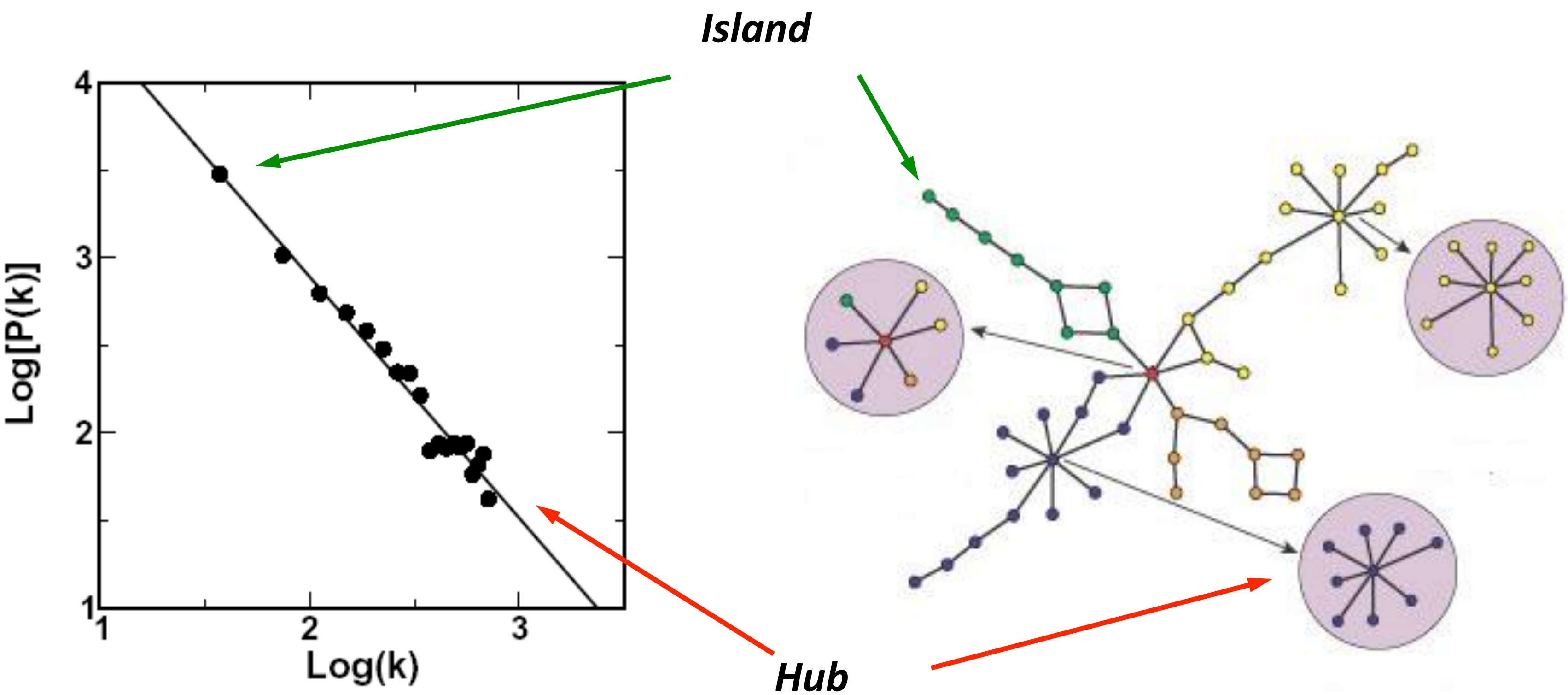
ATH PPI





k	log(k)	P(k)	log(P(k))
1	0	3721	8.221
2	0.693	2082	7.641
3	1.098	1238	7.121
4	1.386	888	6.788
5	1.609	680	6.522
6	1.791	473	6.159
7	1.945	390	5.966
8	2.079	353	5.866
9	2.197	293	5.680
10	2.302	243	5.493
11	2.397	246	5.505
12	2.484	226	5.4205
13	2.564	192	5.257
14	2.639	174	5.159
15	2.708	155	5.043
16	2.772	145	4.9767
17	2.833	116	4.753

Scale Free

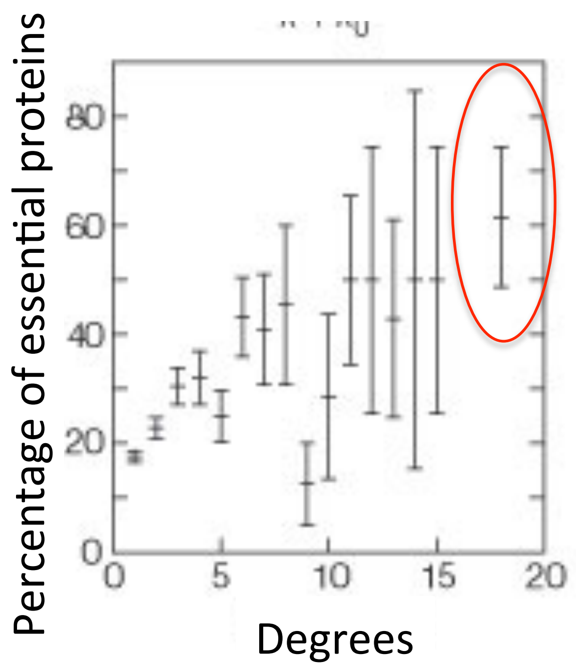


$$P(k) \sim k^{-\gamma}$$

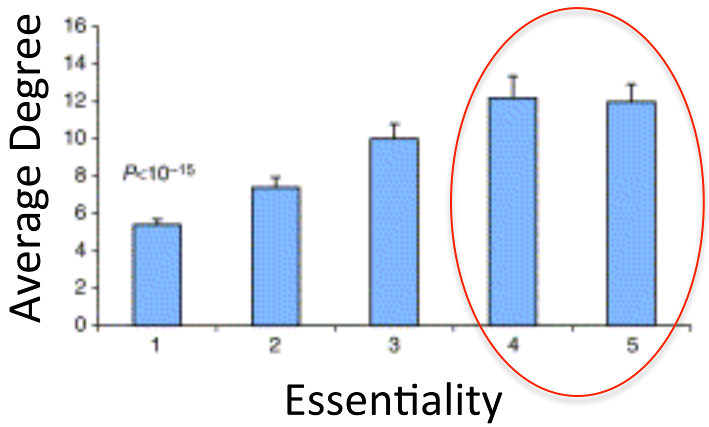
Han *et al.* Nature, 2004

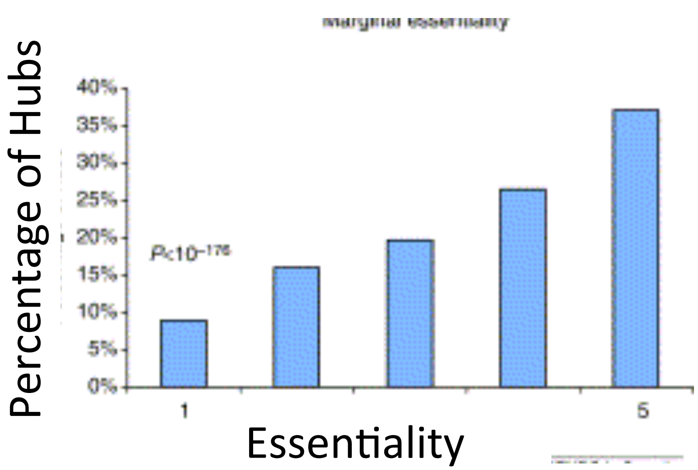
Hub proteins=Essential proteins

- An essential gene is one that, when knocked out, renders the cell unviable.
- Hub proteins are significantly enriched for essential proteins. (Jeong et al. 2001, Nature 411,41)



Essential proteins





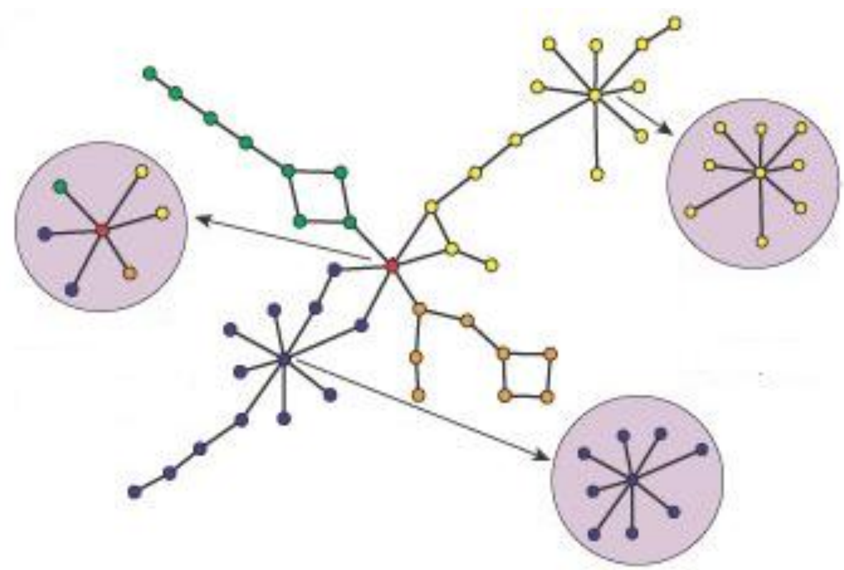
Hubs have high degrees

Essential genes have high essentiality.

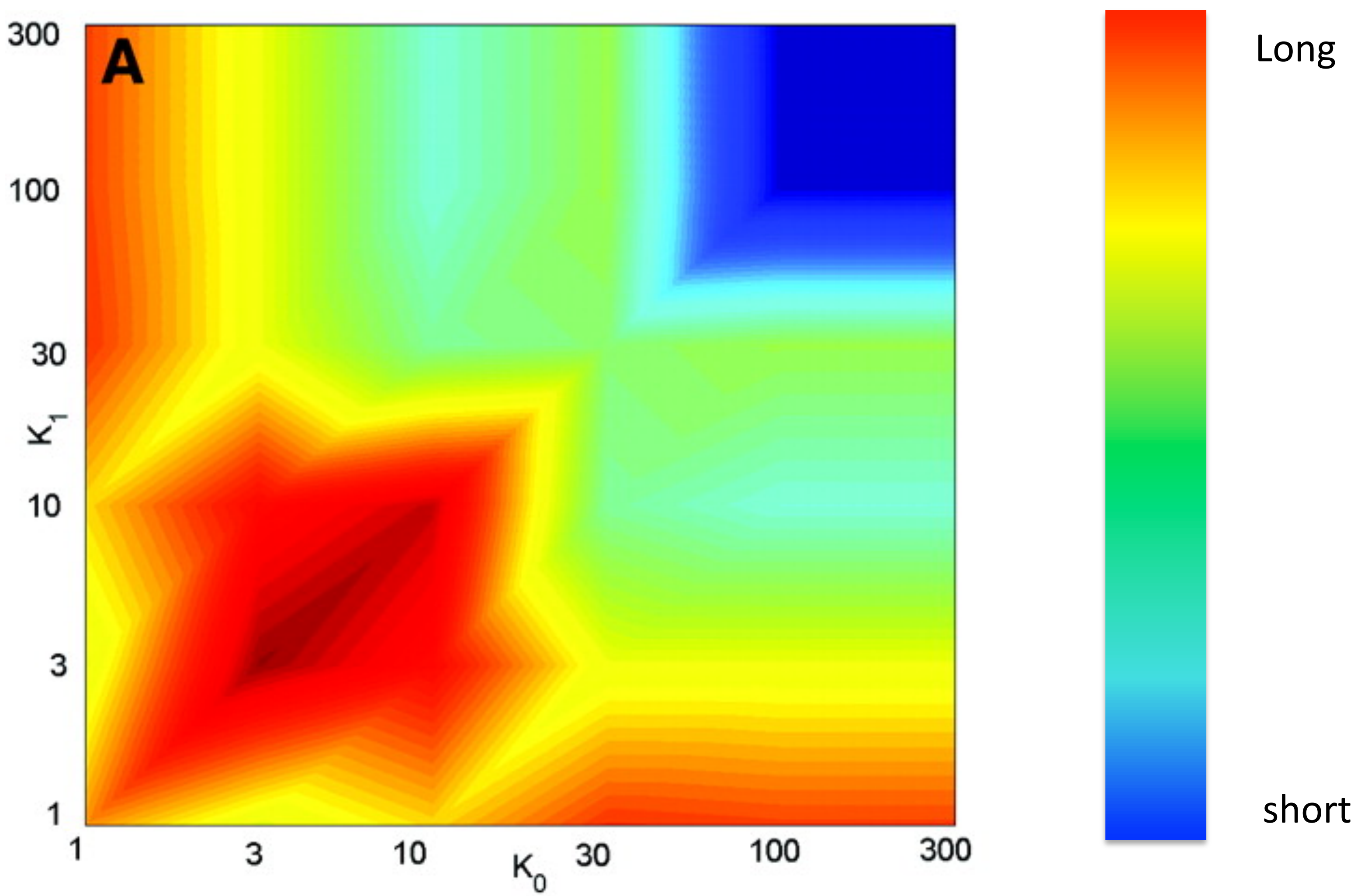
Yu (2004) Trends in Genetics, 20(6), 227

Hub proteins close to each other

- Hub proteins have lower average length of shortest path among themselves than non-hub proteins. (Moslov et al. 2002 *Science 296, 910*)

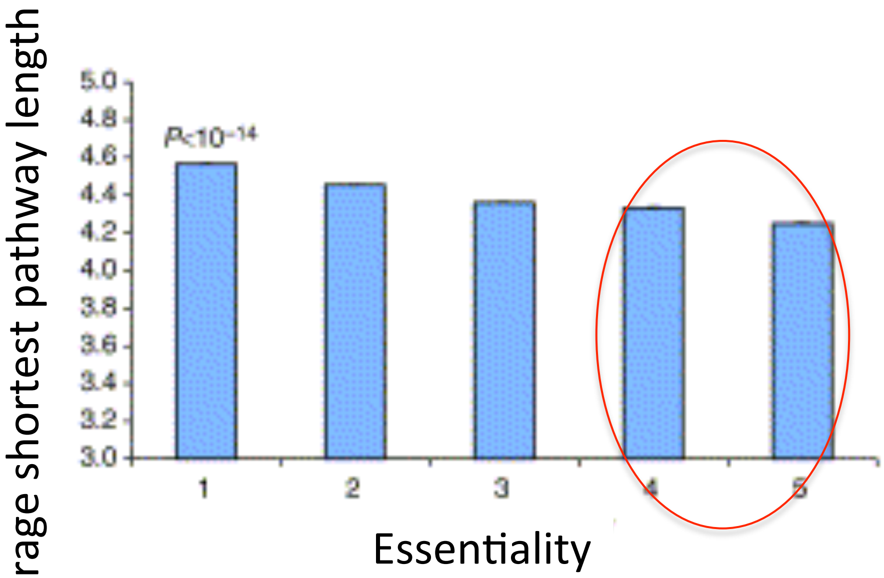


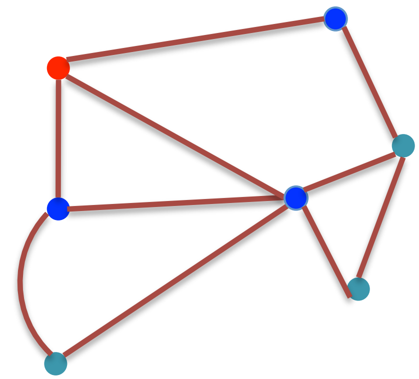
Length of shortest path

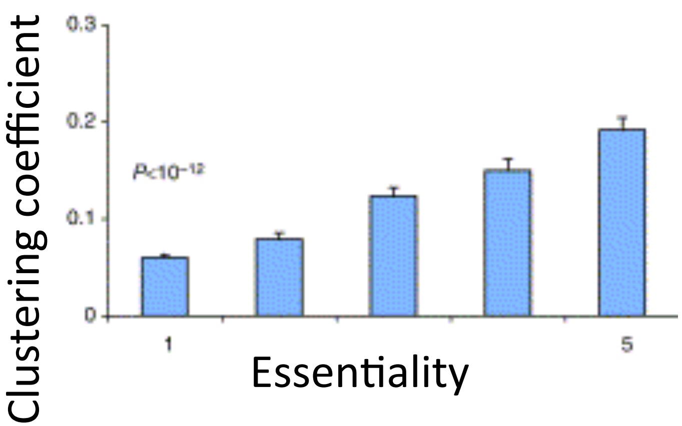


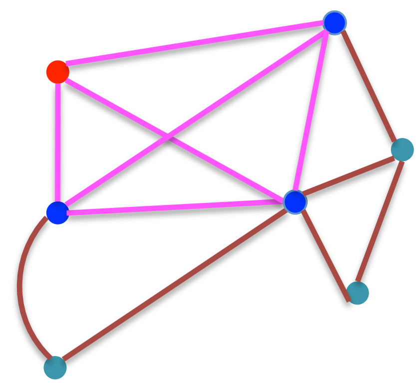
Moslov et al. 2002 *Science 296, 910*

Essential proteins





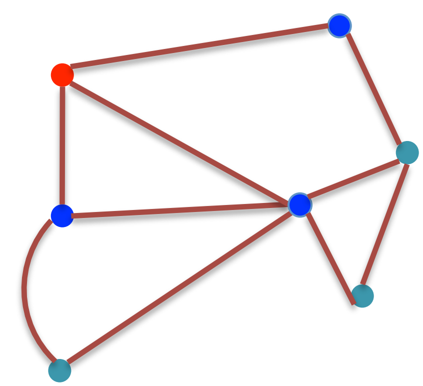




Clustering coefficient

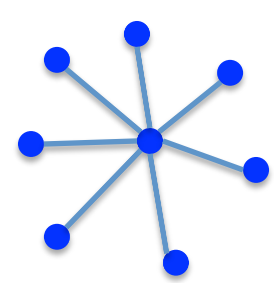
- Local clustering coefficient C_i for a vertex v_i is given by the proportion of links between the vertices within its neighborhood divided by the number of links that could possibly exist between them.

$$C_i = \frac{|e_{ij}|}{\mathbf{V}(\mathbf{V}-1)/2}$$

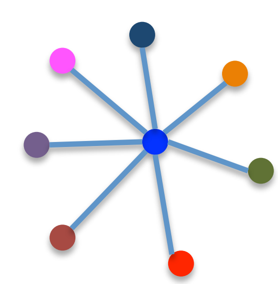


Static or Dynamic

- Combined PPI with gene expression profiles.
- Calculate co-express correlation between hubs and their neighbors.
- Two types of hubs:



Party Hub



Date Hub

Han et al. (2004) Nature 430(6995):88-93

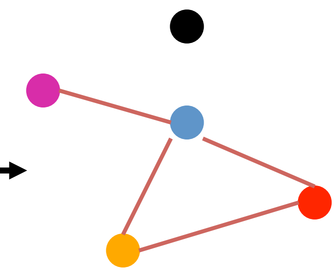
Gene Co-expression correlation

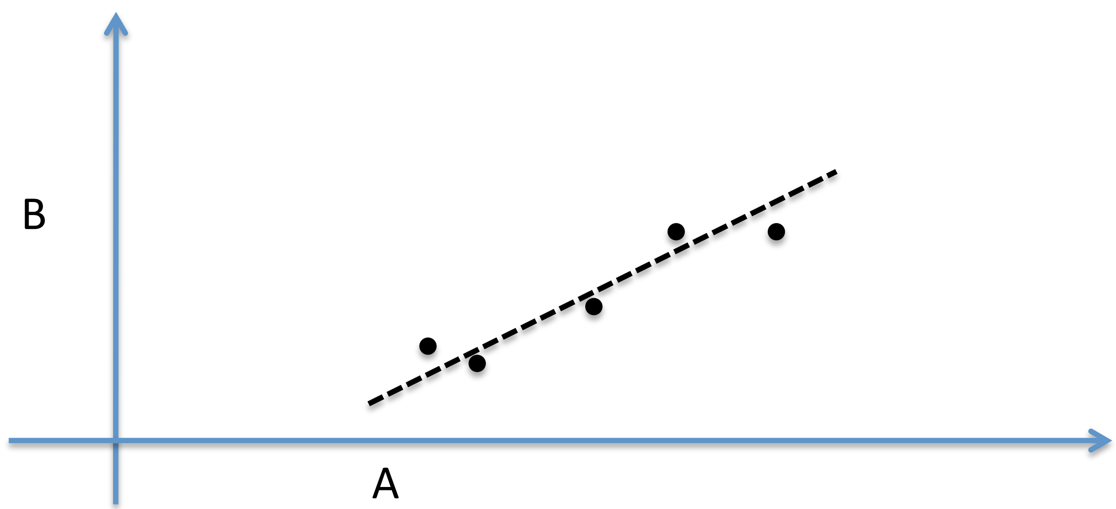
	T1	T2	T3	T4	T5
A	2.5	2.8	3.7	4.6	1.5
B	0.2	0.8	0.3	1.5	0.6
C	1.9	1.3	0.2	0.8	1.6
D	0.8	1.4	0.7	1.6	1.7
E	1.5	1.8	0.3	0.5	1.9

pair-wise → correlation

	C.C
A-B	0.76
A-C	0.90
A-D	0.50
…	0.83
D-E	0.42

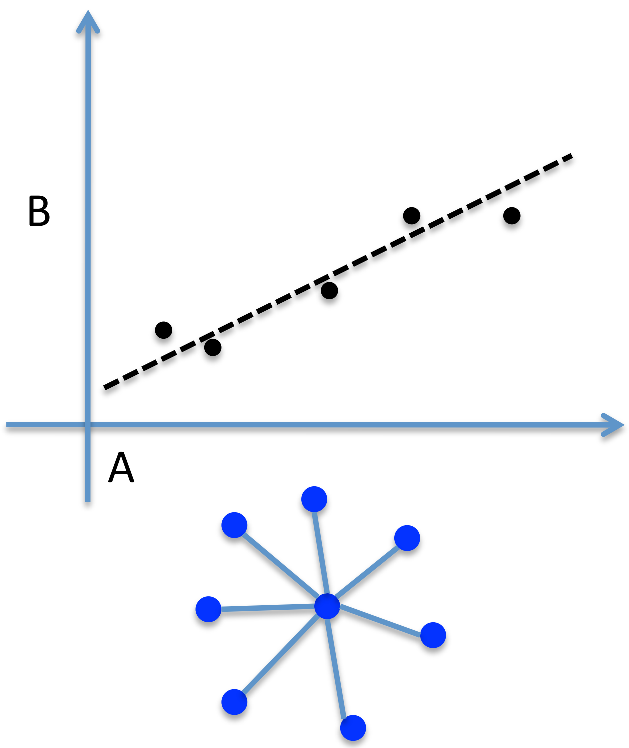
cutoff → >= 0.6



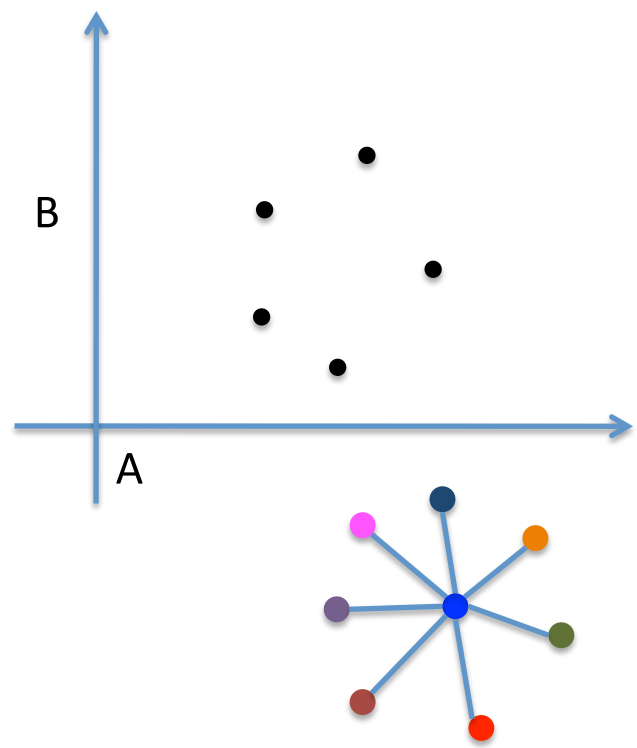


Hub Co-expression correlation

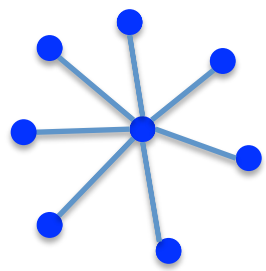
	T1	T2	T3	T4	T5
A	2.5	2.8	3.7	4.6	1.5
B	2.4	2.8	3.6	4.7	1.6
C	1.9	2.0	3.2	4.2	1.3
D	2.8	3.0	4.1	5.0	2.5
E	1.5	1.8	3.0	4.0	1.2



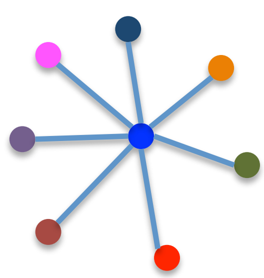
	T1	T2	T3	T4	T5
A	2.5	2.8	3.7	4.6	1.5
B	5.4	0.8	1.6	4.7	3.6
C	1.0	5.0	1.2	2.2	3.3
D	4.8	0.3	0.1	6.0	1.5
E	1.0	2.8	3.4	0.0	1.2



Date or Party Hubs

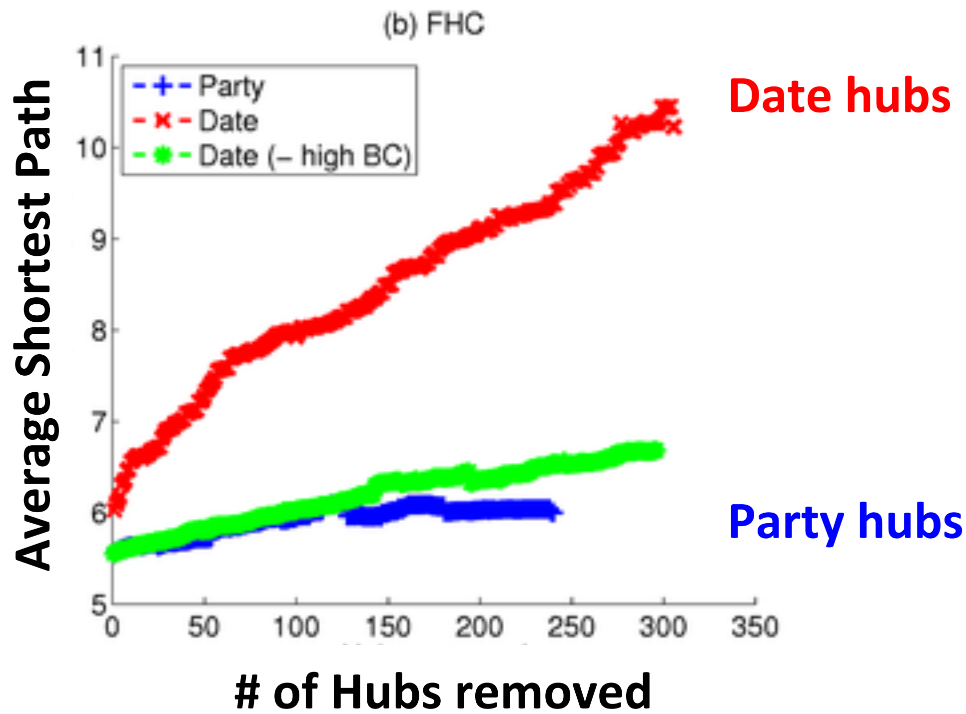


Party Hubs are expressed with their connection partners at same time. They will form a large protein complex. They are more essential. Most of them are house keeping genes.



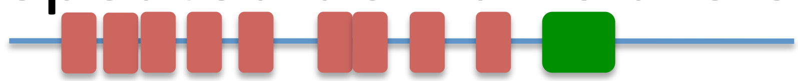
Date Hubs bind with their different connection partners at different time. They have many different binding sites. They have more disorder regions.

Network topology of hubs

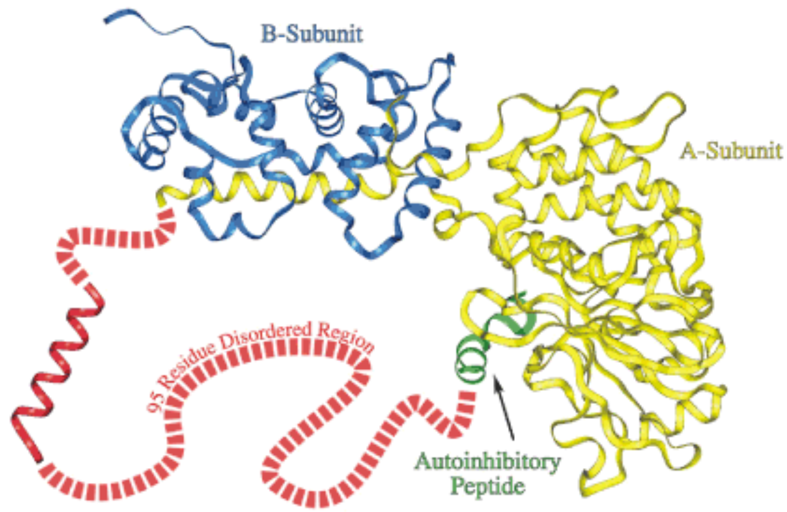


Hub proteins

- Multiple and repeated domains are enriched in hub proteins



- Long disordered regions are common in hubs.

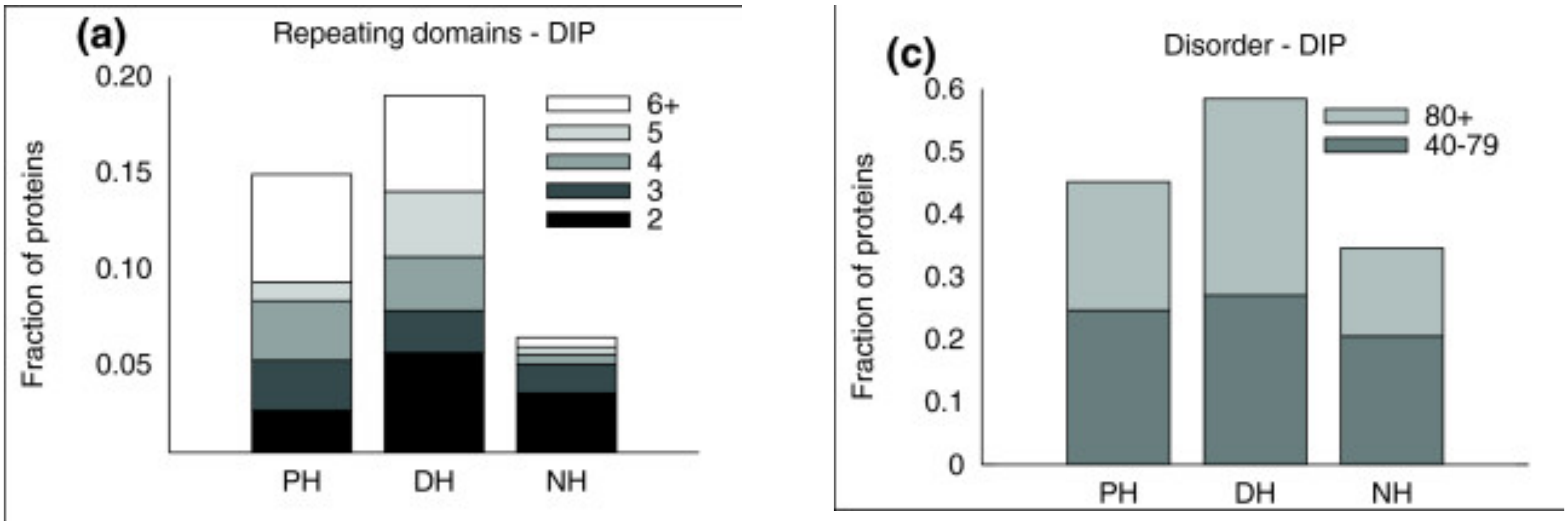


(Image adapted from: Kissinger CR, et al. 1995. "Crystal structures of human calcineurin and the human FKBP12-FK506-calcineurin complex." Nature 378:641-4.)

disordered regions are typically involved in regulation, signaling and control pathways in which interactions with multiple partners and high-specificity/low-affinity interactions are often requisite.

(Ekman et al. 2006 Genome Biol. 7(6): R45)

Hub proteins



PH: Party Hubs
DH: Date Hubs
NH: Non-hubs

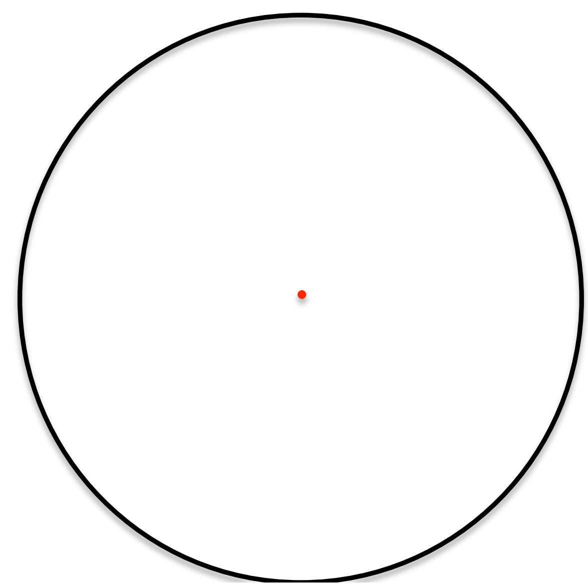
Centrality of PPI

- Compared yeast, worm, and fly PPI
- the number of degrees and the centrality of proteins in the networks have similar distributions.
- Essential proteins have significant centrality.
- Proteins that have a more central position in all three networks, regardless of the number of direct interactors, evolve more slowly and are more likely to be essential for survival.

Hahn et al. (2004) Molecular Biology and Evolution, 22(4) 803.

Centrality

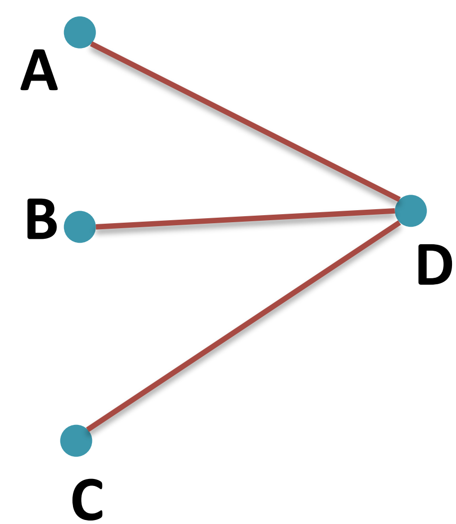
- Measure of the **centrality** of a vertex within a graph that determine the relative importance of a vertex within the graph.
  - Closeness centrality
  - Betweenness centrality



Closeness centrality

- It is defined as the average distance between a vertex *v* and all other vertices reachable from it.
- For a graph *G*: = (*V*,*E*) with *n* vertices, the degree centrality $C_c(v)$ for vertex *v* is

$$C_c = \frac{\sum_i \mathrm{dis}(vi)}{n-1}$$



A node is important if it has a small closeness centrality, because it is close to any other node.

Betweenness centrality

- Vertices that occur on many shortest paths between other vertices have higher betweenness than those that do not.
- For all node pairs (i, j), find the number of shortest paths between them, $\sigma(i,j)$, and determine how many of these pass through node k - $\sigma_k(i,j)$

$$C_k = \sum_{i,j} \frac{\sigma_k(i,j)}{\sigma(i,j)}$$

A node is important if it has a large Betweenness centrality, because many shortest paths pass it.

Essentiality and Centrality

		Yeast	Worm	Fly
Betweeness Centrality	Essential	0.0009	0.0017	0.0007
	Non-Essential	0.0007	0.0009	0.0004
1/ Closeness Centrality	Essential	0.244	0.183	0.238
	Non-Essential	0.239	0.175	0.221
Degrees	Essential	19.3	8.2	9.8
	Non-Essential	15.8	5.6	5.7

Hahn et al. (2004) Molecular Biology and Evolution, 22(4) 803.

Essentiality, Centrality, slow evolution rate

correlation	Yeast	Worm	Fly
D_n - Betweeness	-0.174	-0.118	-0.071
D_n - Closeness	-0.085	-0.114	-0.064
D_n - Degrees	-0.161	-0.027	-0.053

- Identified orthologs of the proteins in the yeast, worm, and fly networks in the related species *S. paradoxus*, *C. briggsae*, and *D. pseudoobscura*, respectively.
- D_n = the number of nonsynonymous differences per nonsynonymous site. (that changes amino acid). This is proportional to the evolution rate.
- Essential genes are house-keeping genes, have slow evolution rate.

Hahn et al. (2004) Molecular Biology and Evolution, 22(4) 803.

Evolution Rates of party or date hubs

	Date Hubs	Party Hubs
Dn	0.7597	0.5652
Ds	2.3133	2.4254
Dn/Ds	0.3631	0.2627

- The lowering of evolutionary rate of the party hub proteins than the date hub proteins.
- Party hubs form a big protein complex; they are more essential.

Dn: non-synonymous distance (changes amino acid)
Ds: Pairwise synonymous (do not change amino acid)

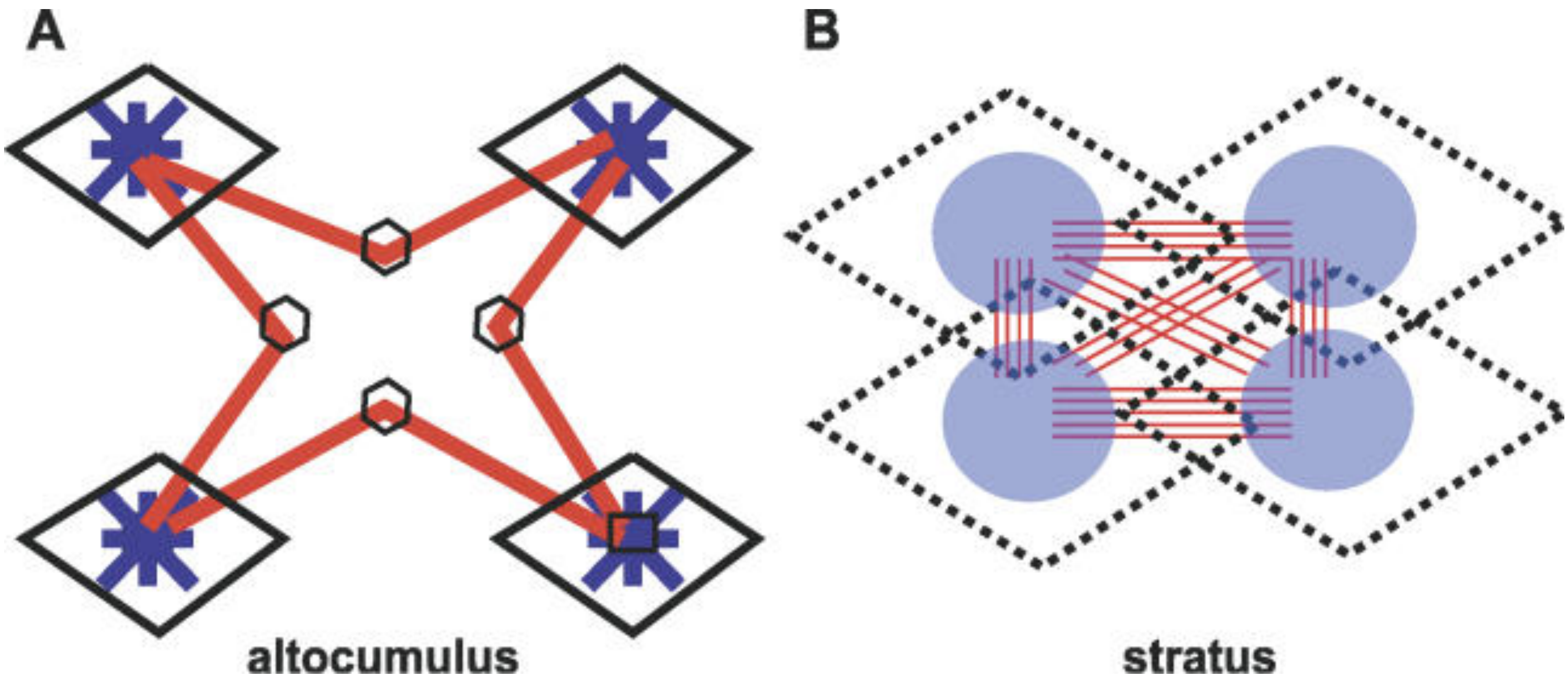
Kahali Et al (2009) Gene, 429, 18

PPI Network topology

- Global protein interaction network is highly interconnected and hence interdependent, more like the continuous dense aggregations of stratus clouds than the segregated configuration of altocumulus clouds.

Batada et al. (2006) PloS Biology, 4(10), e137

Altocumulus or Stratus



highly interconnected and
hence interdependent

Fault tolerance of PPI Networks

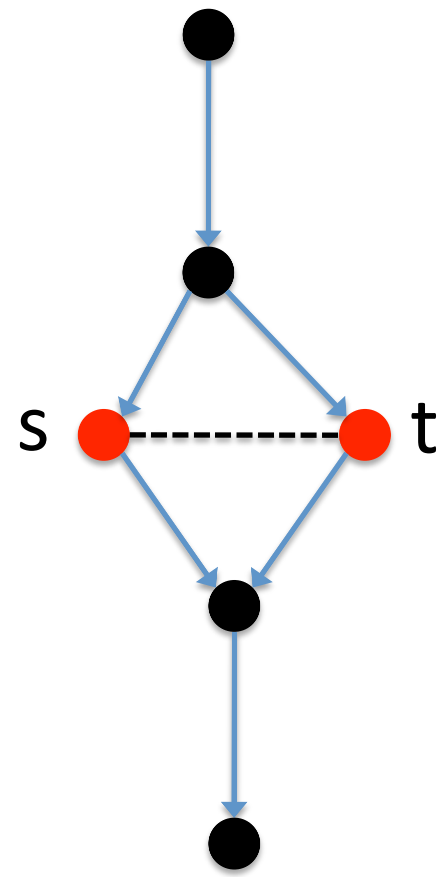
- Whether there exist alternative pathways that can perform some required function if a gene essential to the main mechanism is defective, absent or suppressed.
- Redundant pathways is the BPM (between-pathway model) motif

http://www.ncbi.nlm.nih.gov/pubmed/19399174

Brady et al. (2009) Plos One, 4(4) e5364

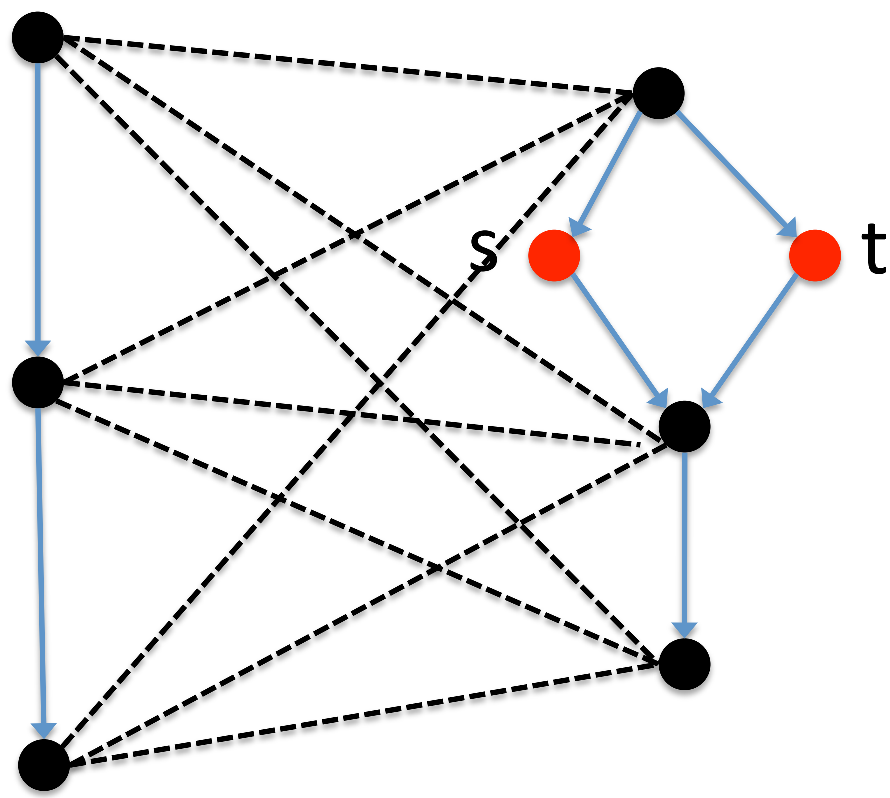
BMP motif

“synthetic-lethality” interaction: both genes are nonessential, but their simultaneous deletion destroys the viability of the cell.



Redundant pathways

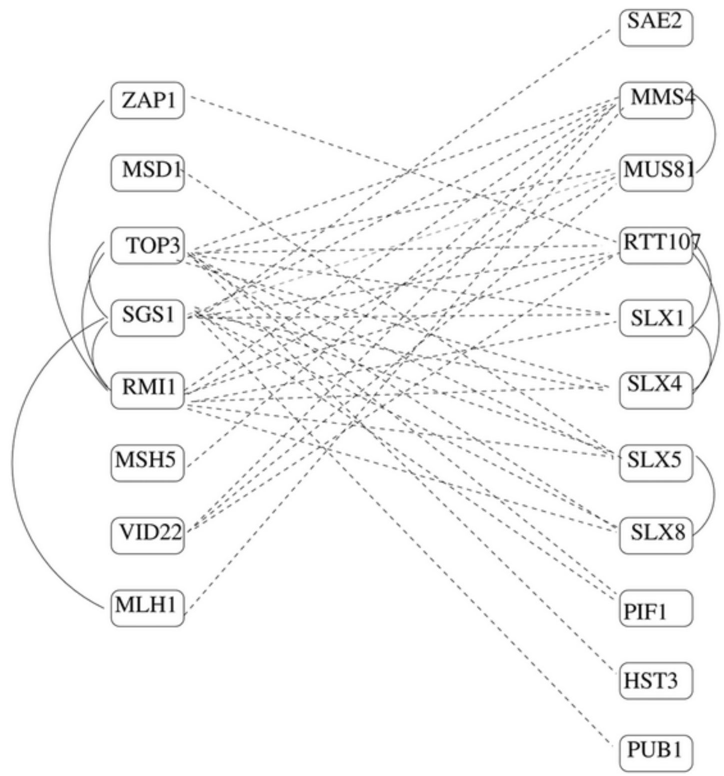
The BPM motif reduces the number of synthetic-lethal interactions, and increase the fault-tolerance for a cell.



This is not a bipartite network

Brady et al. (2009) Plos One, 4(4) e5364

Redundant pathways



This is not a bipartite network

Brady et al. (2009) Plos One, 4(4) e5364

Outline

- Protein-Protein Interaction Model
- How to get PPI
  - Experimental methods
  - Bioinformatic methods
- PPI databases
- Network properties
- Analysis method and applications