

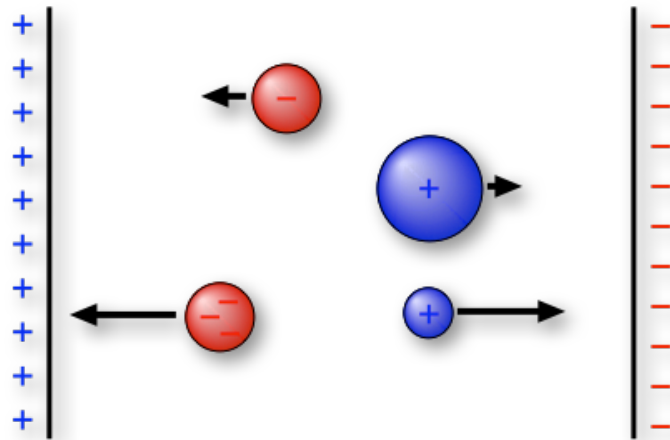
# V 4 – Data for Building Protein Interaction Networks


- Detect PPIs by experimental methods
- Detect (predict) PPIs by computational methods
- Derive condition-specific PPIs by data integration

Tue, Oct. 29, 2019

# Identification of proteins / components of a complex (1): gel electrophoresis

Electrophoresis: directed diffusion of charged particles in an electric field

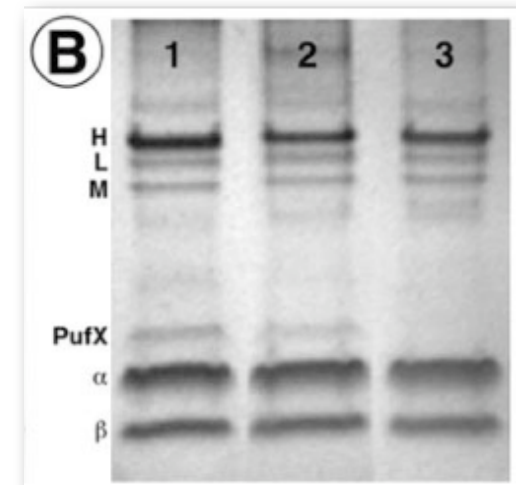


**faster diffusion**  
higher charge, smaller  
  
lower charge, larger  
**slower diffusion**

Put proteins in a spot on a gel-like matrix,  
apply electric field

- separation according to size (mass) and charge
- identify constituents of a complex

Nasty details: protein charge vs. pH, cloud of counter ions,  
protein shape, denaturation, ...



# SDS-PAGE

For better control: denature proteins with detergent

Often used: sodium dodecyl sulfate (**SDS**)

→ denatures and coats the proteins with a negative charge

→ charge proportional to mass

→ traveled distance  $x$  per time

$$x \propto \frac{1}{\log(M)}$$

→ **SDS-polyacrylamide gel electrophoresis**

After the run: **staining** to make proteins visible

For "quantitative" analysis: compare to **marker**  
(set of proteins with known masses)

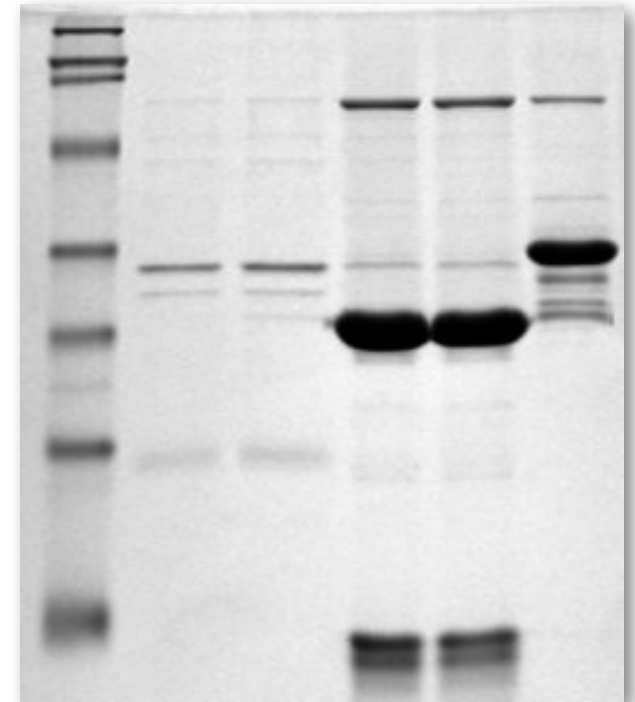


Image from Wikipedia, marker on the left lane

# Protein Charge?

*Protein charge at pH=7*

$$\cong \sum Lys + \sum Arg - \sum Asp - \sum Glu + \sum co - factors$$

Main source for charge differences: pH-dependent protonation states

<=> Equilibrium between

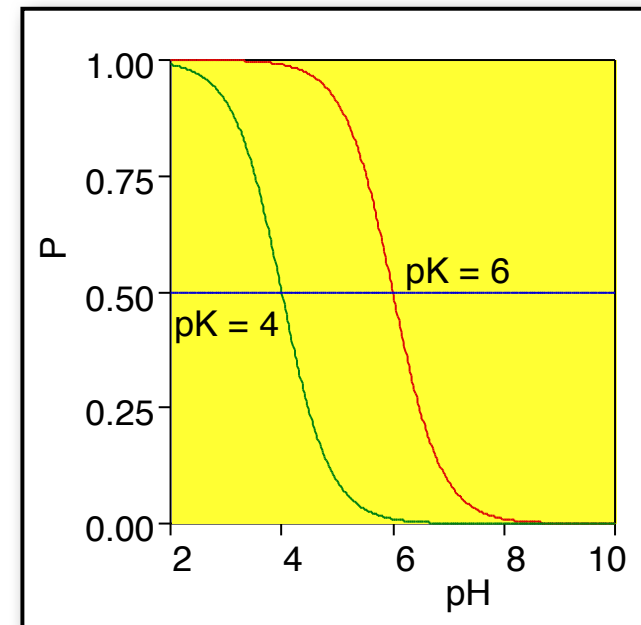
- density (pH) dependent H<sup>+</sup>-binding and
- density independent H<sup>+</sup>-dissociation

Probability to have a proton:

$$P = \frac{1}{1 + 10^{pH-pK}}$$

pKa = pH value for 50% protonation

Asp 3.7–4.0 ... His 6.7–7.1 ... Lys 9.3-9.5



Each H<sup>+</sup> has a +1e charge

→ **Isoelectric point:** pH at which the protein is **uncharged**

→ protonation state cancels permanent charges

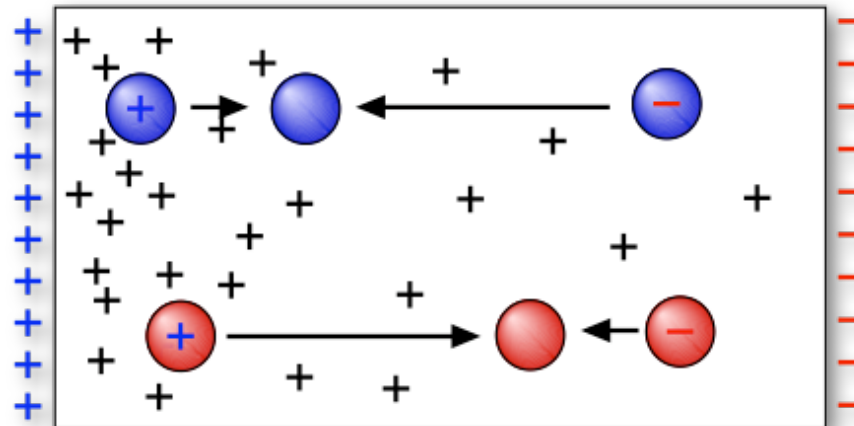


# 2D Gel Electrophoresis

**Two steps:** i) separation **by isoelectric** point via pH-gradient  
ii) separation **by mass** with SDS-PAGE

Step 1:

low pH high pH  
↓  
protonated unprotonated  
=> pos. charge => neg. charge



Nowadays, “blue native gels” are used to detect protein complexes of membrane proteins.

Step 2:

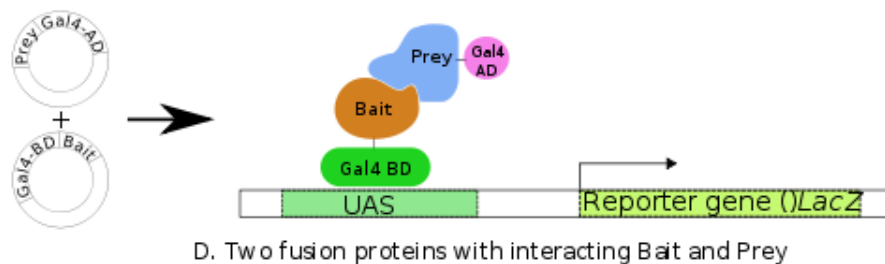
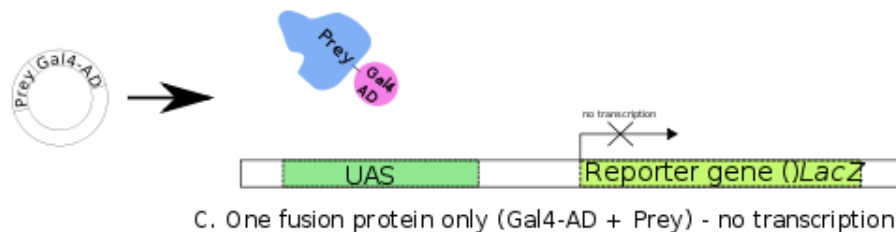
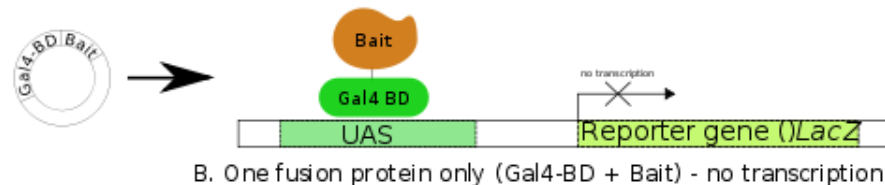
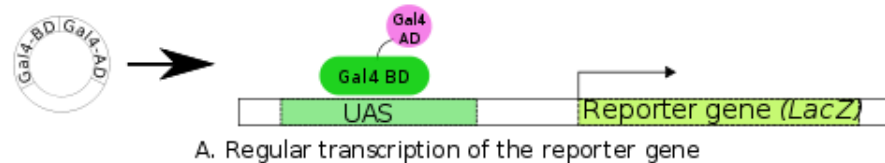
SDS-Page



→ Most proteins differ in mass and isoelectric point (pI)

# Detect interactions: Yeast Two-Hybrid method

Discover binary protein-protein interactions (bait/prey) via physical interaction



Transcription factor consisting of binding domain (BD) + activator domain (AD) induces expression of reporter gene (LacZ or GFP)

Disrupt BD-AD protein;  
fuse bait to BD, prey to AD

→ expression only when  
bait:prey-complex formed

Reporter gene may be fused to green fluorescent protein.


[www.wikipedia.org](http://www.wikipedia.org)

# Pros and Cons of Y2H

## Advantages:

- *in vivo* test for interactions
- cheap + robust → large scale (genome-wide) tests possible

## Problems:

- investigates the interaction between
    - (i) overexpressed
    - (ii) fusion proteins in the
    - (iii) yeast
    - (iv) nucleus
  - spurious interactions via third protein
- 
- many false positives  
(up to 50% errors)

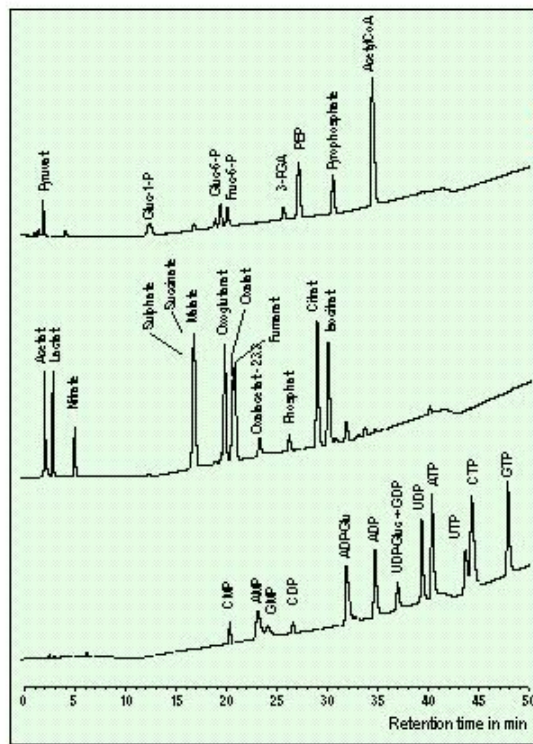
# Identify fragments of proteins / components of a complex (2): Mass Spectrometry

HPLC: high pressure liquid chromatography (first purification step)

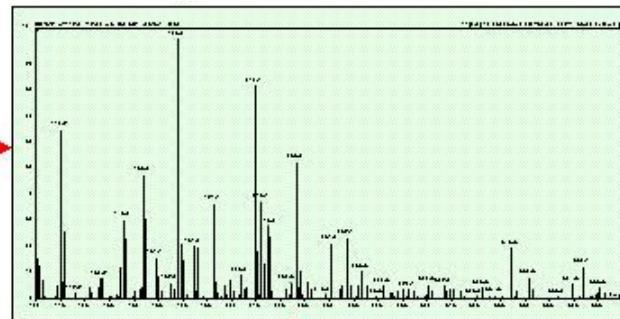
Then identify constituents of a (fragmented) complex by MS via their mass/charge patterns  $m/z$

## Overview LC-MS

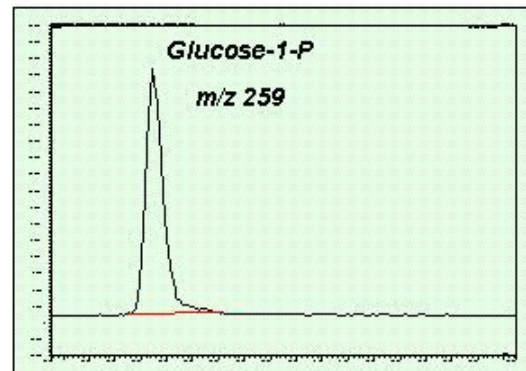
1) Metabolite separation via IC/HPLC



2) Mass detection



3) Extraction of specific masses



[http://gene-exp.ipk-gatersleben.de/body\\_methods.html](http://gene-exp.ipk-gatersleben.de/body_methods.html)

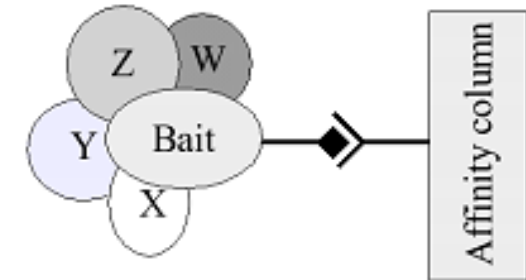
# Detect interactions: Tandem affinity purification (also „pull-down“)

Yeast 2-Hybrid-method can only identify binary complexes.

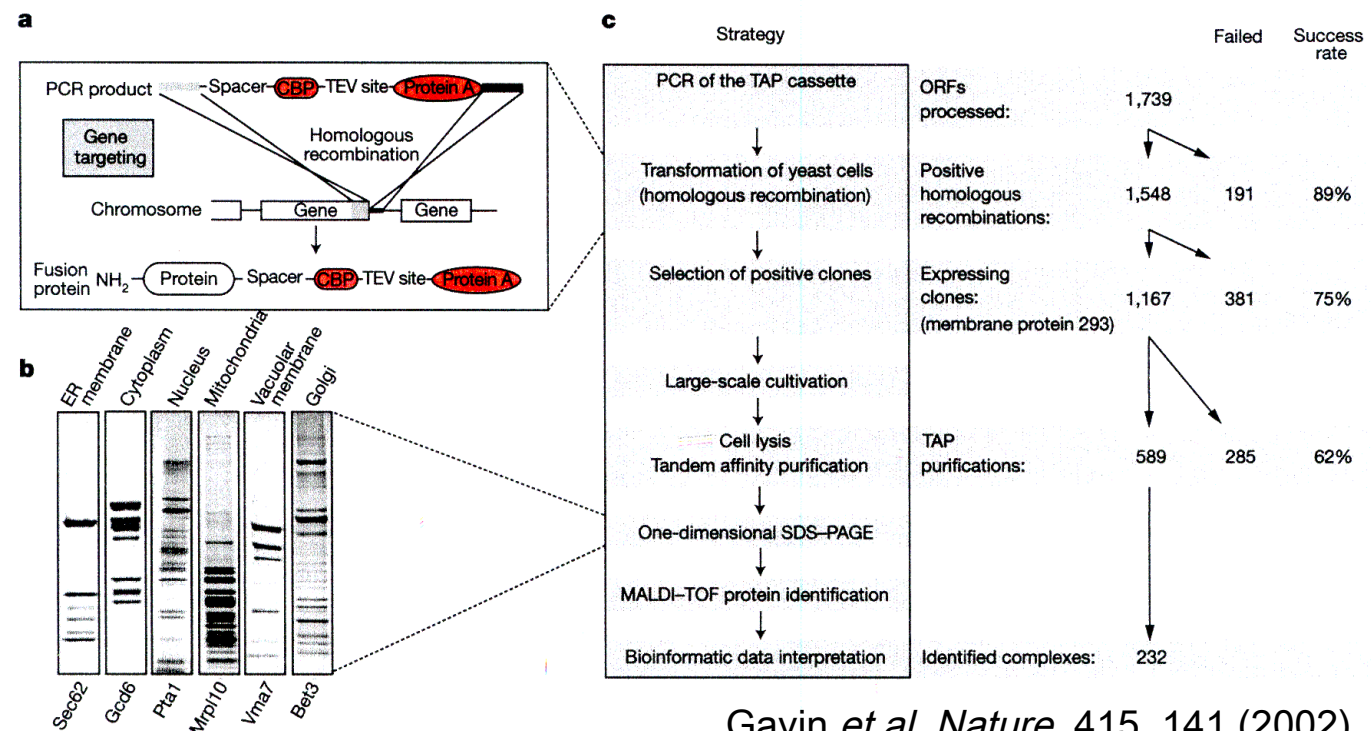
In **affinity purification**, a protein of interest (bait) is tagged with a molecular label (dark route in the middle of the figure) to allow easy purification.

The tagged protein is then co-purified together with its interacting partners (W–Z).

This strategy can be applied on a genome scale (as Y2H).



Identify proteins by  
mass spectrometry and  
scanning of yeast  
protein database for  
protein composed of  
fragments of suitable  
mass.

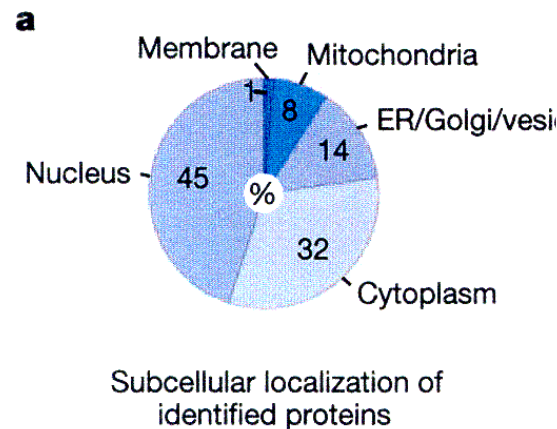


Gavin *et al.* *Nature* 415, 141 (2002)

# TAP analysis of yeast PP complexes

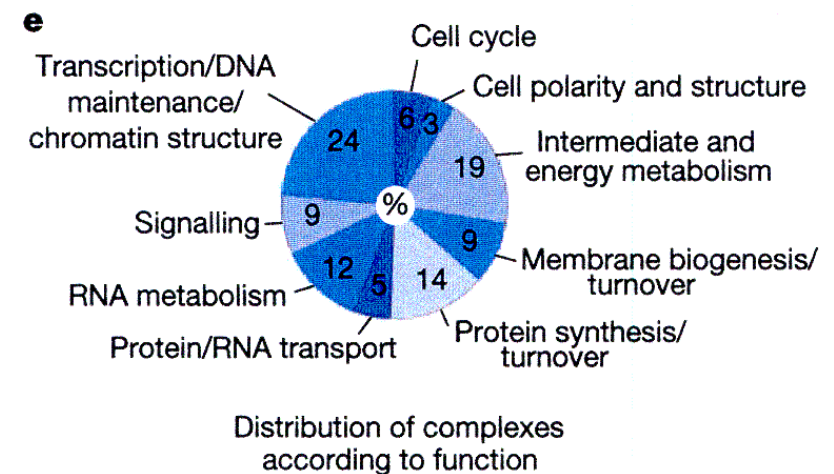
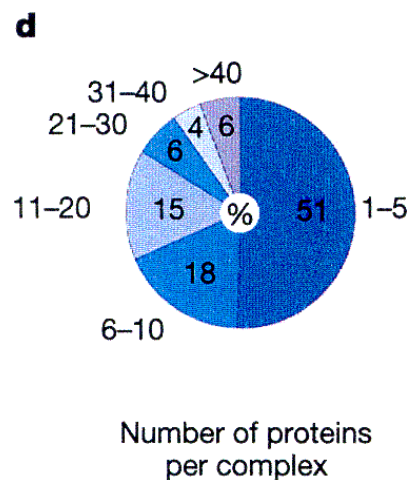
(a) Localization of identified proteins

→ no apparent bias for one compartment, but very few membrane proteins (should be ca. 25%)



(e) complexes are involved in practically all cellular processes

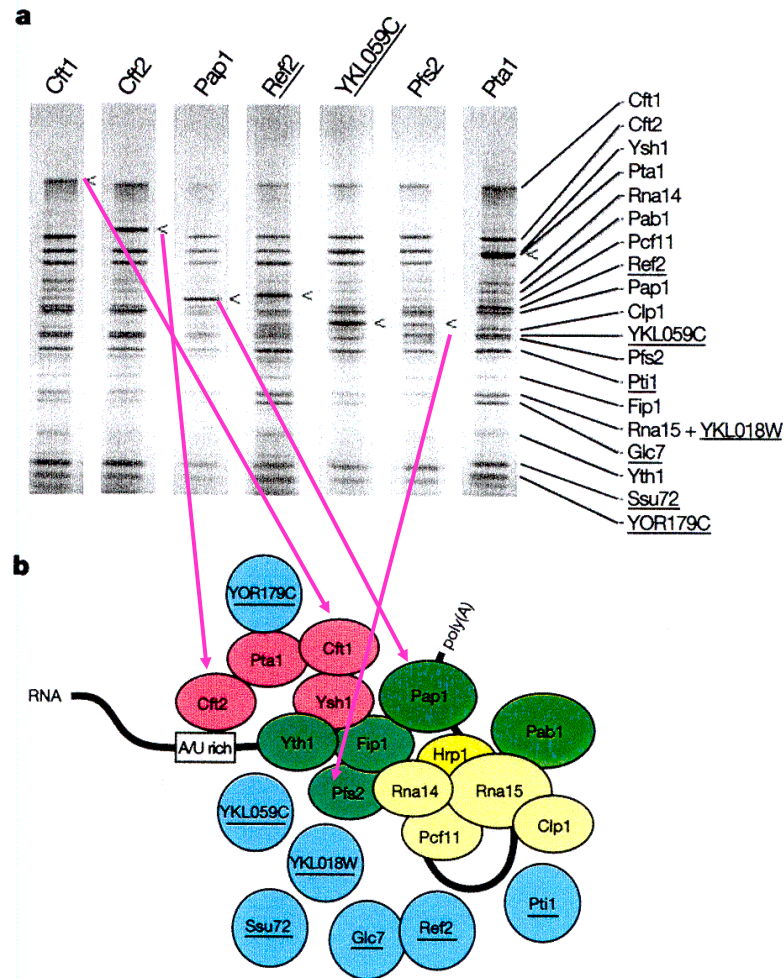
(d) number of proteins per complex → half of all PP complexes have 1-5 members, the other half is larger



Gavin *et al.* *Nature* 415, 141 (2002)



# Validation of TAP methodology



Check of the method:

can the same complex be obtained for different choices of the attachment point (tag protein is attached to different components of complex shown in (b))?

Yes, more or less (see gel in (a)).

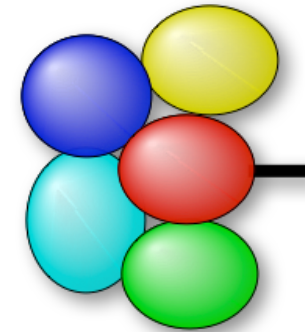
< signs mark tag proteins in the gel lane

Gavin *et al.* *Nature* 415, 141 (2002)

# Pros and Cons of TAP-MS

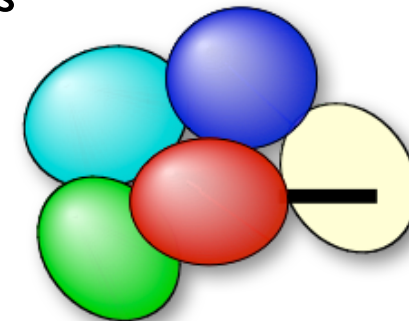
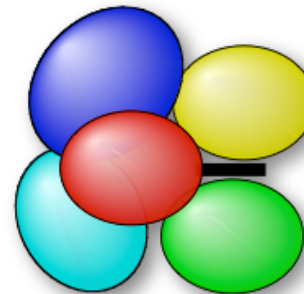
## Advantages:

- **quantitative** determination of complex partners *in vivo* without prior knowledge
- simple method, high yield, high throughput



## Difficulties:

- tag may **prevent** binding of the interaction partners
- tag may change (relative) **expression** levels
- tag may be **buried** between interaction partners  
→ no binding to beads

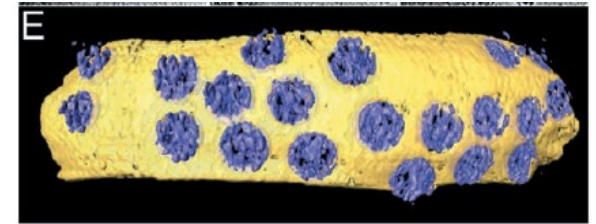




# Protein interactions in nuclear pore complex

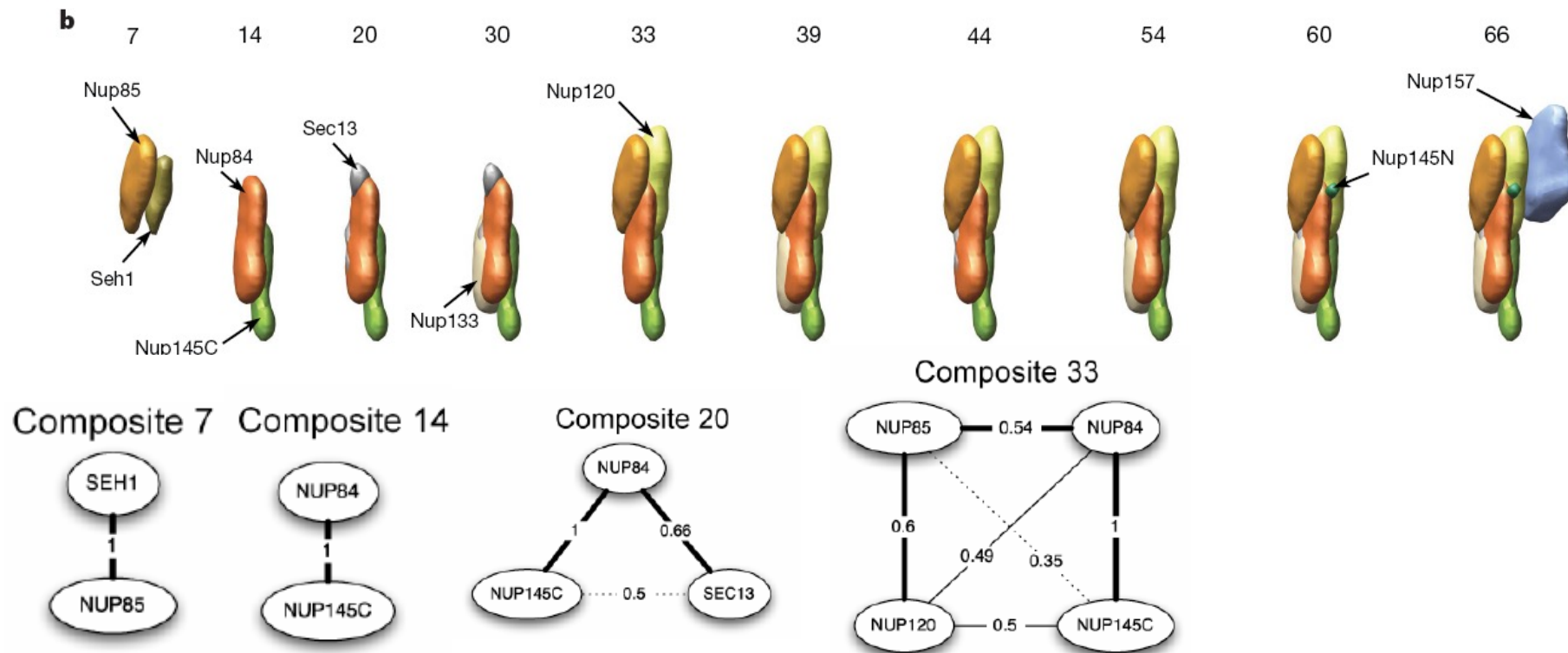
Figure (right) shows 20 NPCs (blue) in a slice of a nucleus.

**Aim:** identify individual PPIs in Nuclear Pore Complex.



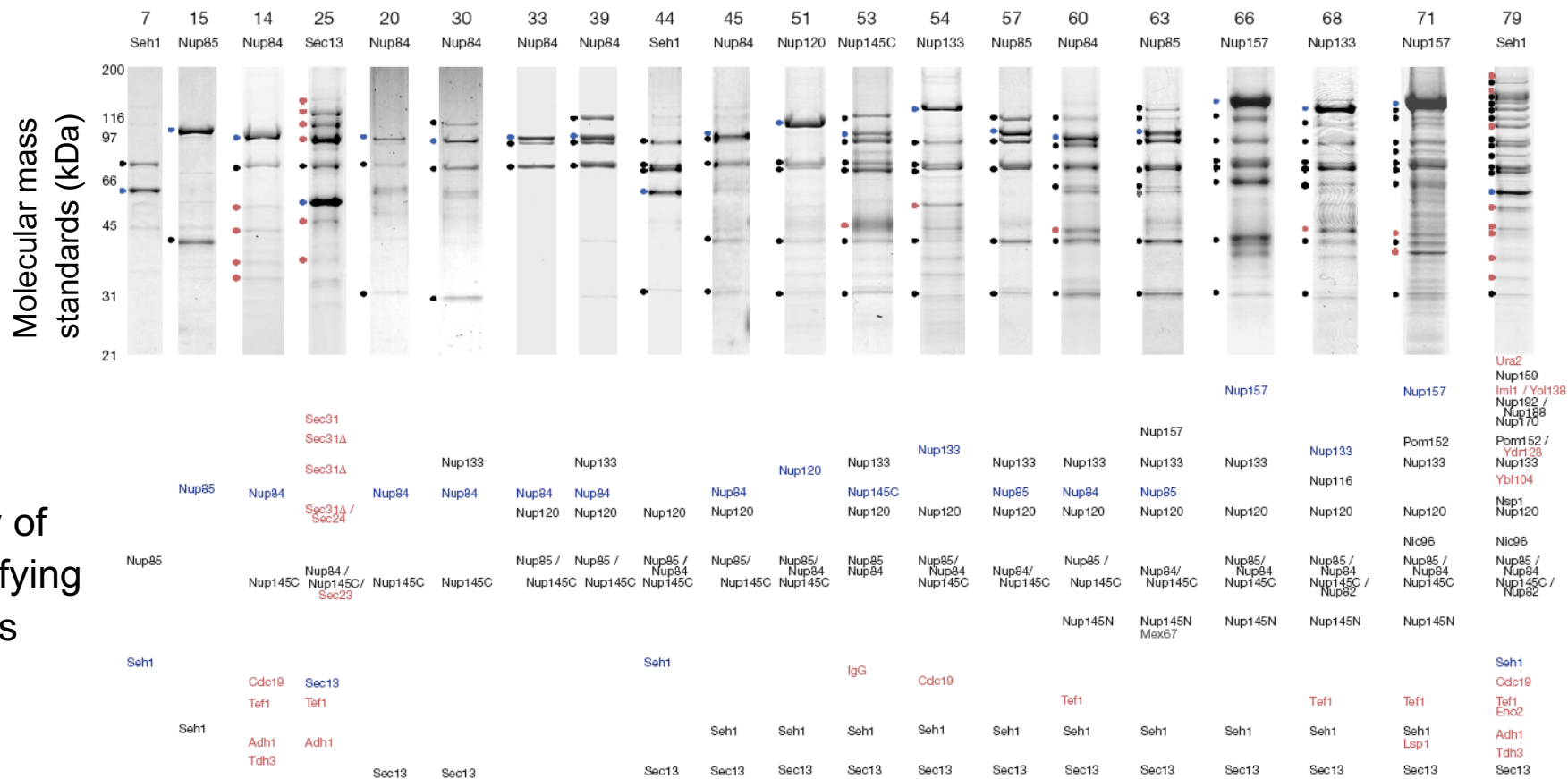
Below : mutual arrangement of Nup84-complex-associated proteins as visualized by their localization volumes in the final NPC structure.

Nup84 protein shown in **light brown**.



# SDS + MS:Composites involving Nup84

**a** above lanes: name of ProteinA-tagged protein and identification number for composite



**Blue:** PrA-tagged proteins,  
**Black:** co-purifying nucleoporins,  
**Grey:** NPC-associated proteins,  
**Red:** and other proteins (e.g. **contaminants**)

Affinity-purified PrA-tagged proteins and interacting proteins were resolved by **SDS-PAGE** and visualized with Coomassie blue. The bands marked by filled circles at the left of the gel lanes were identified by **mass spectrometry** (cut out band from the gel and use as input for MS).

# Indirect Evidence on PPIs: Synthetic Lethality

Apply two mutations that are viable on their own, but lethal when combined.

In cancer therapy, this effect implies that inhibiting one of these genes in a context where the other is defective should be selectively lethal to the tumor cells but not toxic to the normal cells, potentially leading to a large therapeutic window.

Gene X	Gene Y	
+	+	No effect
—	+	No effect
+	—	No effect
—	—	Death

<http://jco.ascopubs.org/>

Synthetic lethality may point either to:

- physical interaction of proteins (they are building blocks of a complex)
- both proteins belong to the same pathway
- both proteins have the same function (redundancy)

# Indirect Evidence on PPIs: Gene Coexpression

All constituents of a PP complex should be present  
be present

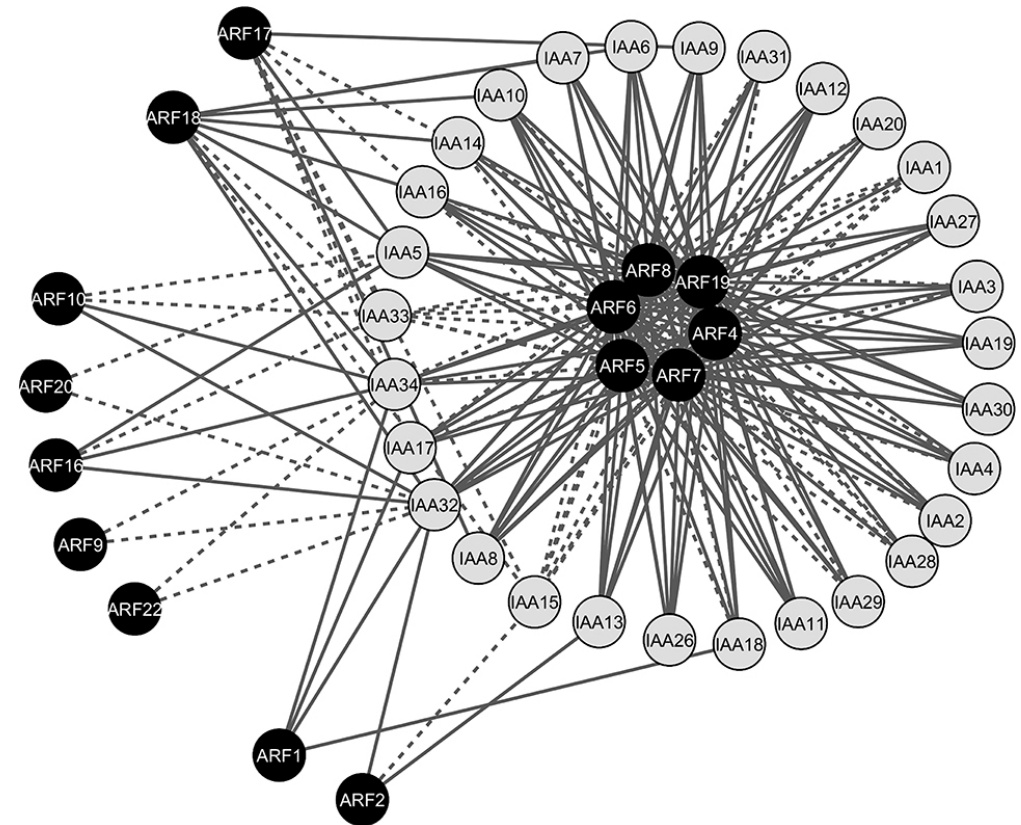
→ test for correlated expression

## Gene co-expression network of interacting ARF-Aux/IAA proteins in *A. thaliana*.

Network contains 44 nodes and 213 edges (interacting combinations based on Y2H/BiFC).

Continuous edges: protein pairs with significantly correlated expression profiles in at least one tissue.

Dotted edges: protein pairs without significantly correlated expression profiles.



Note that co-expression is no direct indication for formation of complexes (there are too many co-regulated genes), but it is a useful "filter"-criterion.

Piya S et al. (2014) *Front. Plant Sci.* **5**:744.

# Interaction Databases

Bioinformatics: make experimental data available in databases

## 3.2 Experimental High-Throughput Methods for Detecting Protein-Protein Interactions | 4

**Table 3.1** Some public databases compiling data related to protein interactions: (P) and (D) stand for proteins and domains (the number of interactions reflects the status of June 2007).

	URL	Number of interactions	Type	Proteins /domains
MIPS	mips.gsf.de/genre/proj/mpact	4300	curated	
BIND	bond.unleashedinformatics.com	200000	curated	P
MINT	160.80.34.4/mint/	103800	curated	P
DIP	dip.doe-mbi.ucla.edu	56000	curated	P
PDB	www.rcsb.org/pdb	800 complexes	curated	
HPRD	www.hprd.org	37500	curated	P, D
Scoppi	www.scoppi.org	102000	automatic	D
UniHI	theoderich.fb3.mdc-berlin.de:8080/unihi/home	209000	integrated data	P
STRING	string.embl.de	interactions of 1500000 proteins	integrated data from genomic context, high-throughput experiments, coexpression, previous knowledge	P
iPfam	www.sanger.ac.uk/Software/Pfam/iPfam	3019	data extracted from PDB	D
YEAST protein complex database	yeast.cellzome.com	232 complexes	experimental	P
ABC	service.bioinformatik.uni-saarland.de/abc	13000 complexes	semiautomatic	P

# Initially low overlap of results

For **yeast**: ~ 6000 proteins => ~18 million potential interactions  
rough estimates: ≤ 100000 interactions occur

→ 1 true positive for 200 potential candidates = **0.5%**

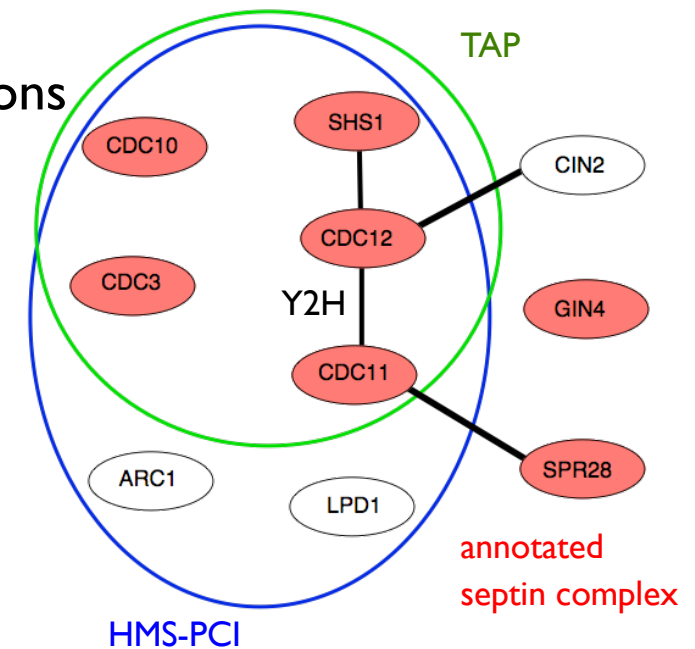
→ **decisive** experiment must have **accuracy** << 0.5% false positives

**Different experiments** detect different interactions

For yeast: 80000 interactions known in 2002  
only 2400 were found by ≥ 2 experiments

Problems with experiments:

- i) incomplete coverage
- ii) (many) false positives
- iii) selective to type of interaction  
and/or compartment



von Mering (2002)

Y2H: yeast two hybrid screen

TAP: tandem affinity purification

HMS-PCI: protein complex identification by MS



# Criteria for reliability of detected PPIs

Guiding principles to judge experimental results on PPIs (incomplete list!):

1) check **mRNA abundance** of detected PPIs:

most experimental techniques are biased towards high-abundance proteins.

If this is the case, results for low-abundance proteins are not reliable.

2) Check localization to cellular **compartments**:

- most methods have their "preferred compartment"
- if interacting proteins belong to the same compartment  
=> results are more reliable

3) **co-functionality**

it is realistic to assume that members of a protein complex should have closely related biological functions -> check whether interaction proteins have overlapping annotations with terms from Genome Ontology (GO)

# In-Silico Prediction Methods

## Sequence-based:

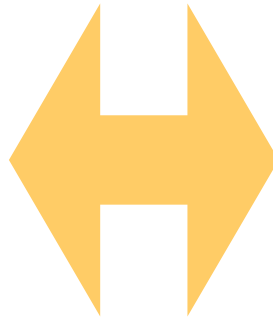
- gene clustering
- gene neighborhood
- Rosetta stone
- phylogenetic profiling
- coevolution



"Work on the parts list"

- fast
- unspecific
- high-throughput methods for pre-sorting

Will be covered today



## Structure-based:

- interface propensities
- protein-protein docking
- spatial simulations (e.g. MD)



"Work on the parts"

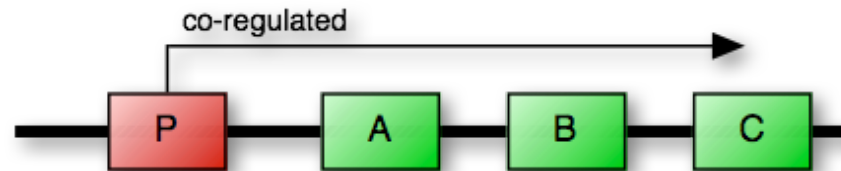
- specific, detailed
- expensive
- accurate

Not subject of this lecture



# Gene Clustering

**Idea:** functionally **related** proteins or parts of a complex are expressed **simultaneously**



Search for genes with a **common promoter**

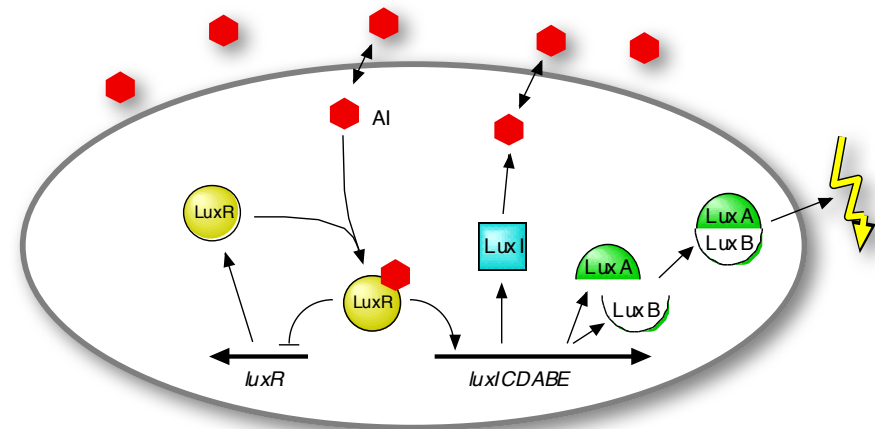
→ when activated, all are transcribed together as one *operon*

## Example:

bioluminescence in *V. fischeri* is regulated via quorum sensing

→ three proteins: I, AB, CDE are responsible for this.

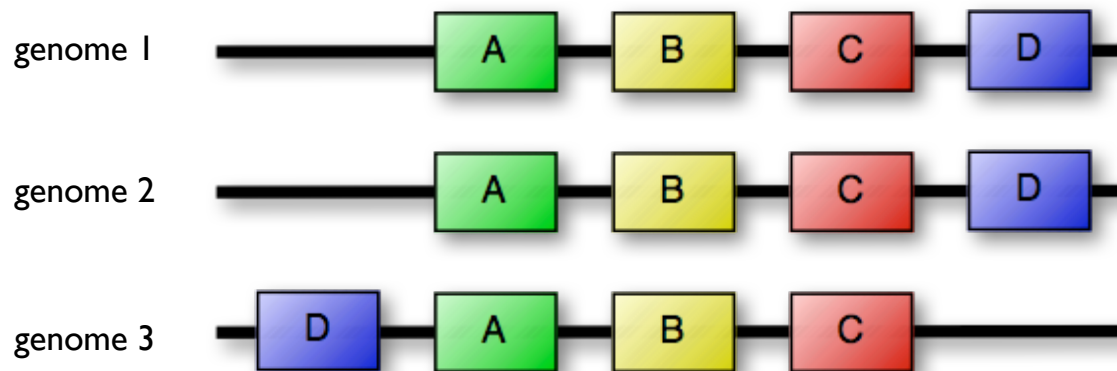
They are organized as 1 operon named *luxICDABE*.



# Gene Neighborhood

**Hypothesis** again: functionally **related** genes are expressed **together**

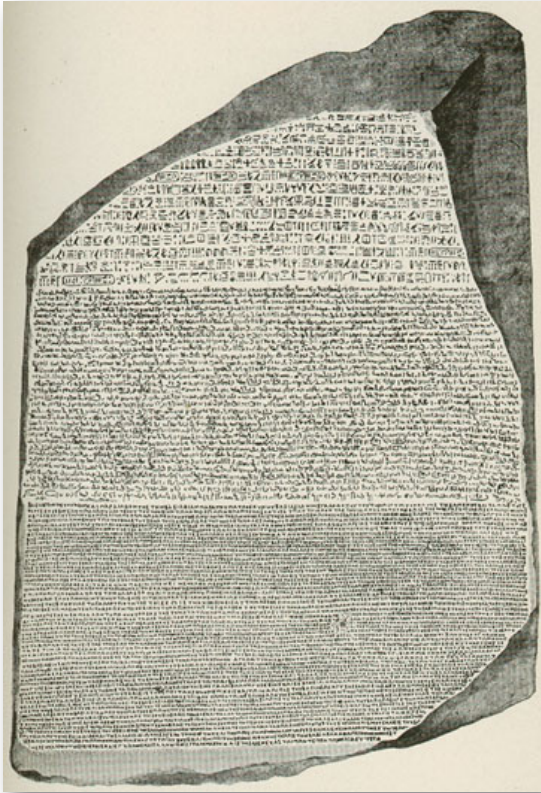
"functionally related" means same {complex | pathway | function | ...}



→ Search for **similar arrangement** of related genes in **different organisms**

(<=> Gene clustering: done in one species, need to know promoters)

# Rosetta Stone Method



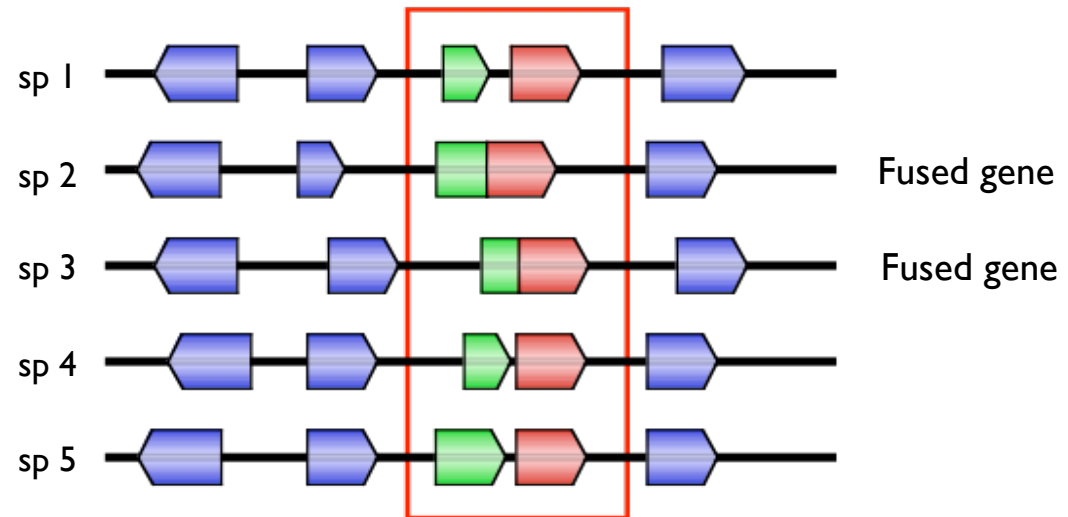
Multi-lingual stele from 196 BC,  
found by the French in 1799

The same decree is inscribed on the stone  
3 times, in hieroglyphic, demotic, and greek.  
→ key to deciphering meaning of  
hieroglyphs

**Idea:** find homologous genes ("words") in genomes of  
different organisms ("texts")

- check if *fused gene pair* exists in one organism

→ May indicate that these 2 proteins form a complex

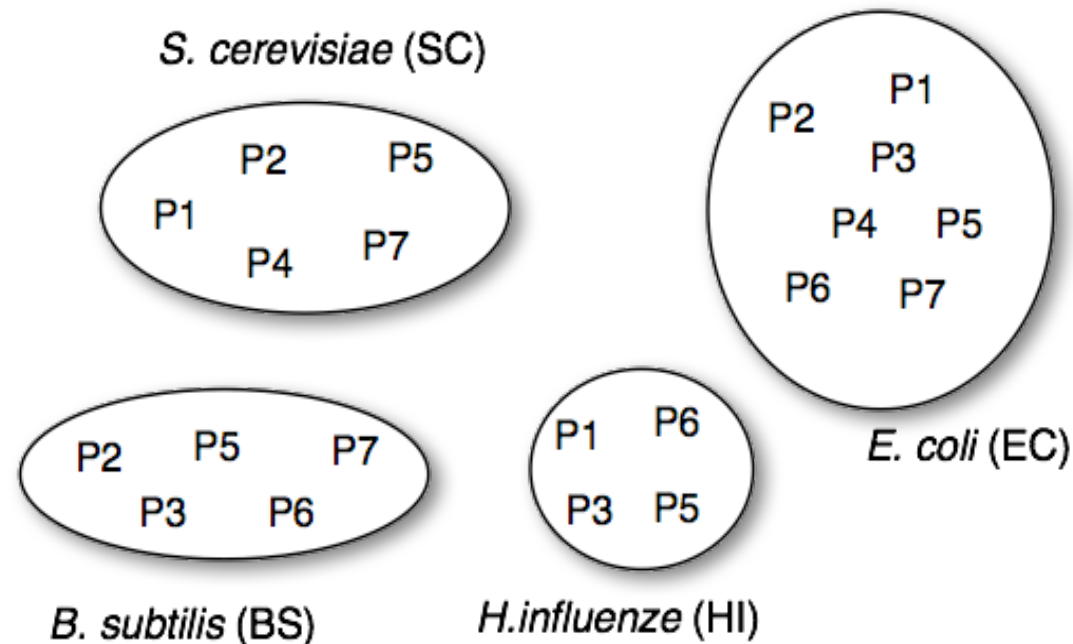


Enright, Ouzounis (2001):  
40000 predicted pair-wise interactions  
from search across 23 species

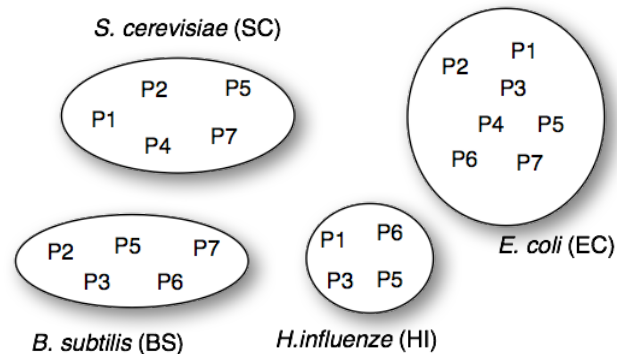
# Phylogenetic Profiling

**Idea:** either **all** or **none** of the proteins of a complex should be **present** in an organism

→ compare presence of protein homologs across species  
(e.g., via sequence alignment)



# Distances in Phylogenetic Profiling

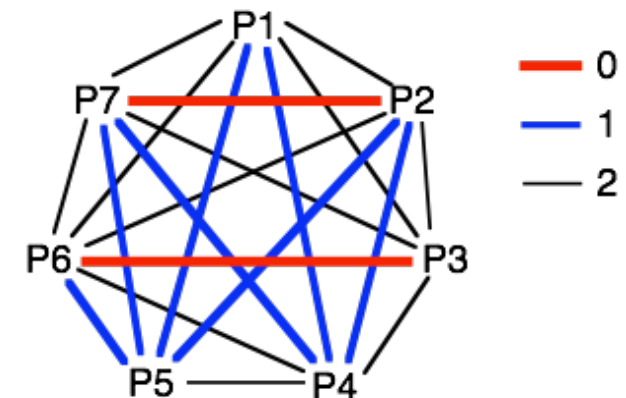


Decode presence/absence

	EC	SC	BS	HI
P1			0	
P2				0
P3		0		
P4			0	0
P5				
P6		0		
P7				0

**Hamming** distance between species: number of different protein occurrences

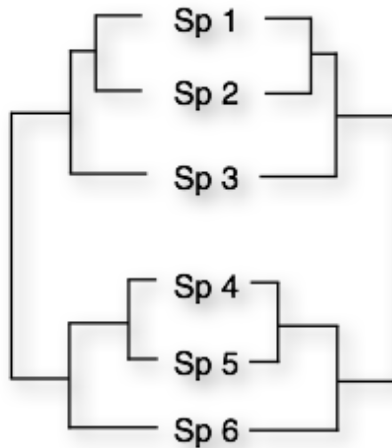
	P1	P2	P3	P4	P5	P6	P7
P1	0	2	2	1	1	2	2
P2		0	2	1	1	2	<b>0</b>
P3			0	3	1	<b>0</b>	2
P4				0	2	3	1
P5					0	1	1
P6						0	2
P7							0



Two pairs with similar occurrence: P2-P7 and P3-P6

These are candidates to interact with each other.

# Co-evolution



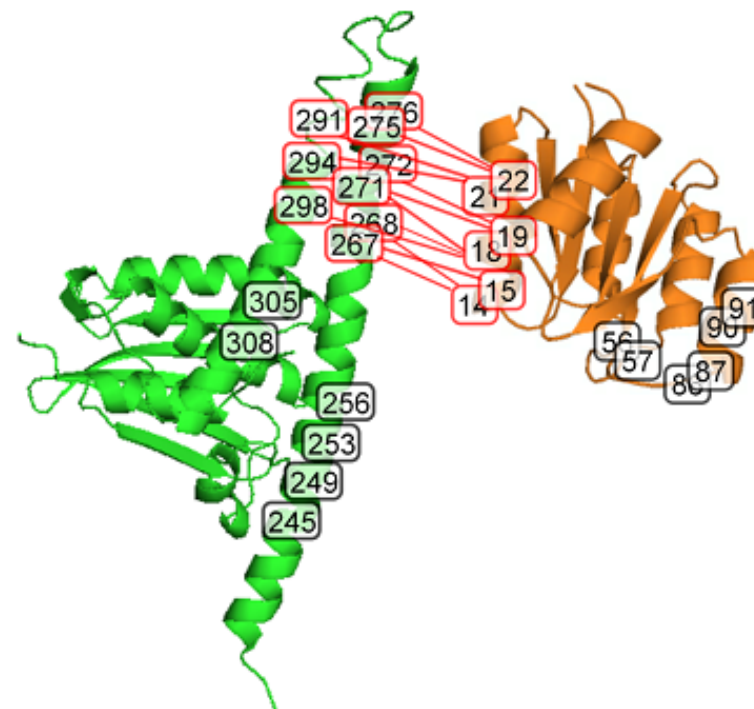
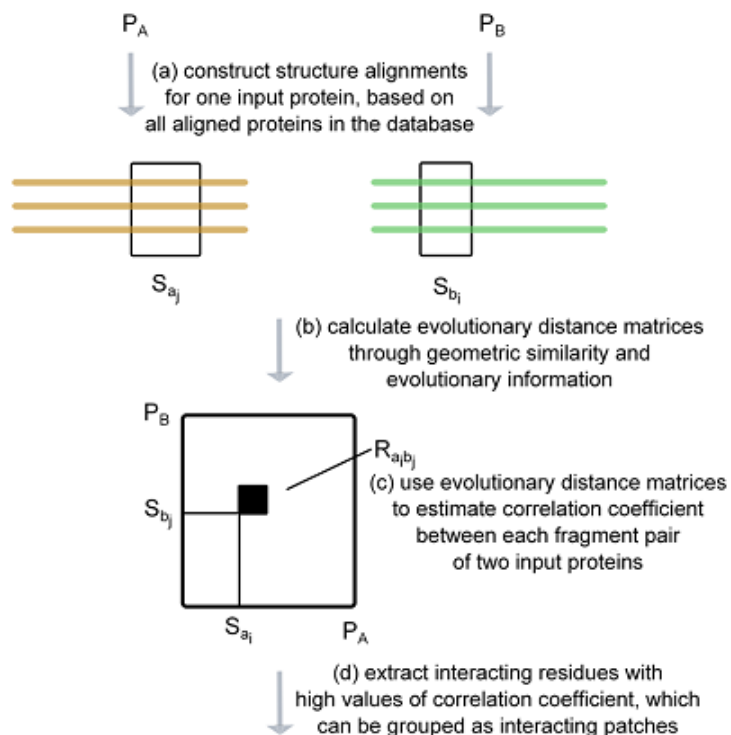
Binding interfaces of complexes are often **better conserved** in evolution than the rest of the protein surfaces.

**Idea** of Pazos & Valencia (1997):  
if a mutation occurs at one interface that changes the character of this residue (e.g. polar → hydrophobic), a corresponding mutation could occur at the other interface at one of the residues that is in contact with the first residue.

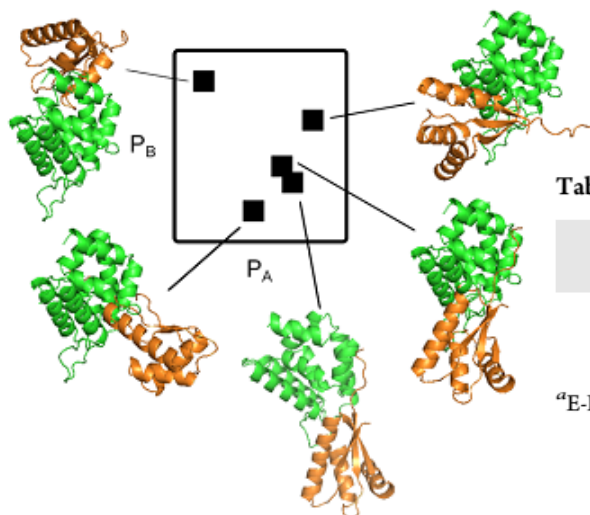
Detecting such correlated mutations could help in identifying binding candidates.



# Correlated mutations



**Figure 3.** Our method detects correlated residues on SK/RR interaction. Interface residues are indicated in red boxes, and non-interface residues are indicated in black boxes.



**Table 6.** Comparison to metaPPI, meta-PPISP, and PPI-Pred

type	our method		metaPPI		meta-PPISP		PPI-Pred	
	Acc	Cov	Acc	Cov	Acc	Cov	Acc	Cov
E-I <sup>a</sup>	73%	65%	61%	37%	56%	55%	46%	47%
others	55%	57%	41%	22%	39%	26%	29%	31%
overall	63%	61%	49%	28%	46%	38%	36%	38%

<sup>a</sup>E-I is type of enzyme–inhibitor.

Guo et al. J. Chem. Inf. Model. 2015, 55, 2042–2049



# Correlated mutations (Gremlin)

Detect positional correlations in paired multiple sequence alignments of thousands of protein sequences.

Gremlin constructs a global statistical model of the alignment of the protein family pair A and B by assigning a probability to every amino acid sequence in the paired alignment:

$$p(X_1, X_2, \dots, X_p; X_{p+1}, \dots, X_{p+q}) = \frac{1}{Z} \exp \left( \sum_{i=1}^{p+q} \left[ v_i(X_i) + \sum_{j=1}^{p+q} w_{ij}(X_i, X_j) \right] \right)$$

$X_i$  : amino acid composition at position  $i$ ,

$v_i$  : vectors encoding position-specific amino acid propensities

$w_{ij}$  : matrices encoding amino acid coupling between positions  $i$  and  $j$ .

$Z$  : partition function, normalizes sum of probabilities to 1.

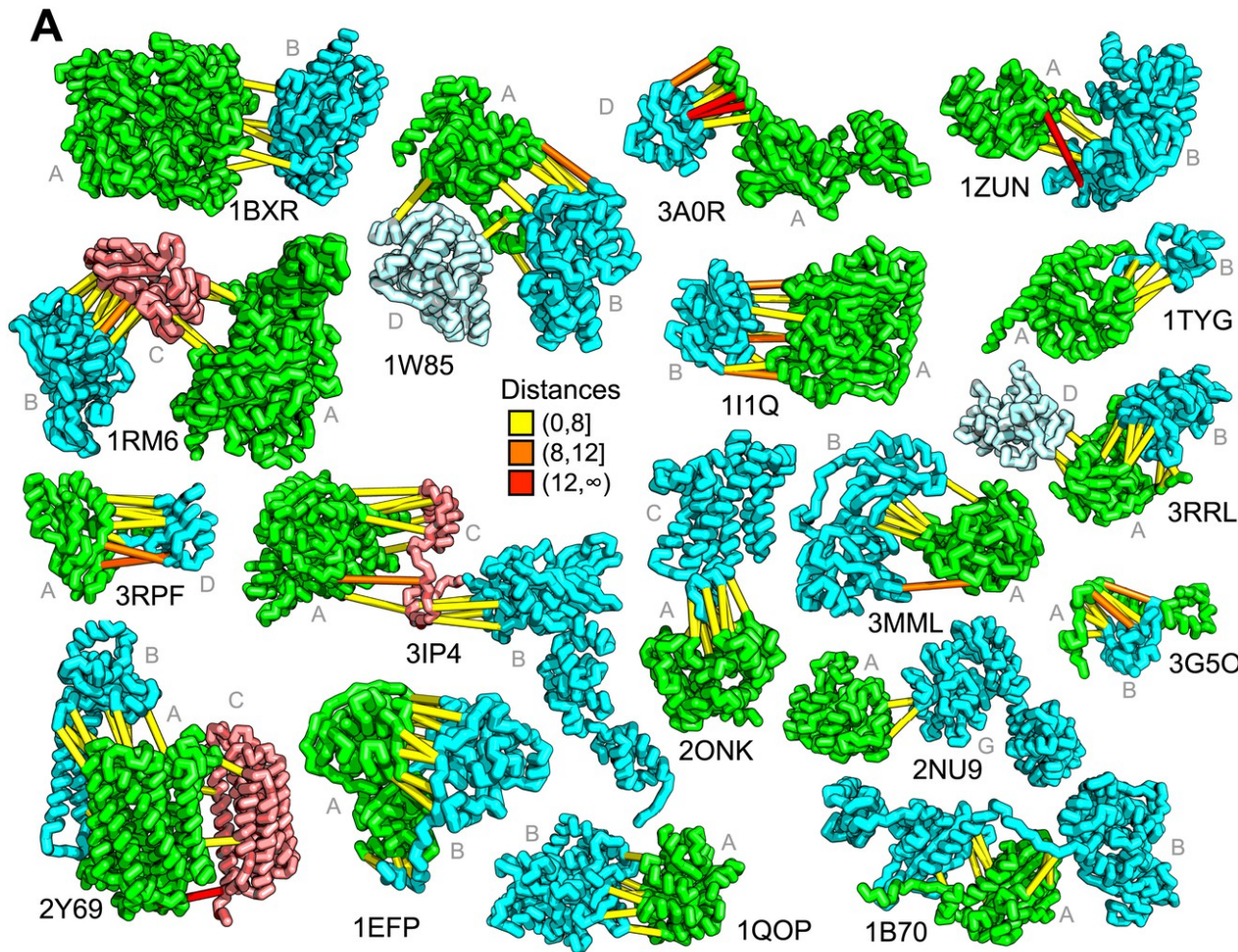
$v_i$  and  $w_{ij}$  are obtained from the aligned sequences by a maximum likelihood approach.

The derived coupling strengths  $w_{ij}$  are then normalized and converted into distance restraints that can be used e.g. in scoring protein-protein docking models.

Ovchinnikov, Kamisetty,  
Baker (2014) eLife 3:e02030



# Correlated mutations



Residue-pairs across protein chains with high GREMLIN scores almost always make contact across protein interfaces in experimentally determined complex structures.

Shown are all contacts with GREMLIN scores greater than 0.6.

Yellow: residue pairs within a distance  $< 8$  Å,  
orange:  $8 - 12$  Å,  
red:  $> 12$  Å.

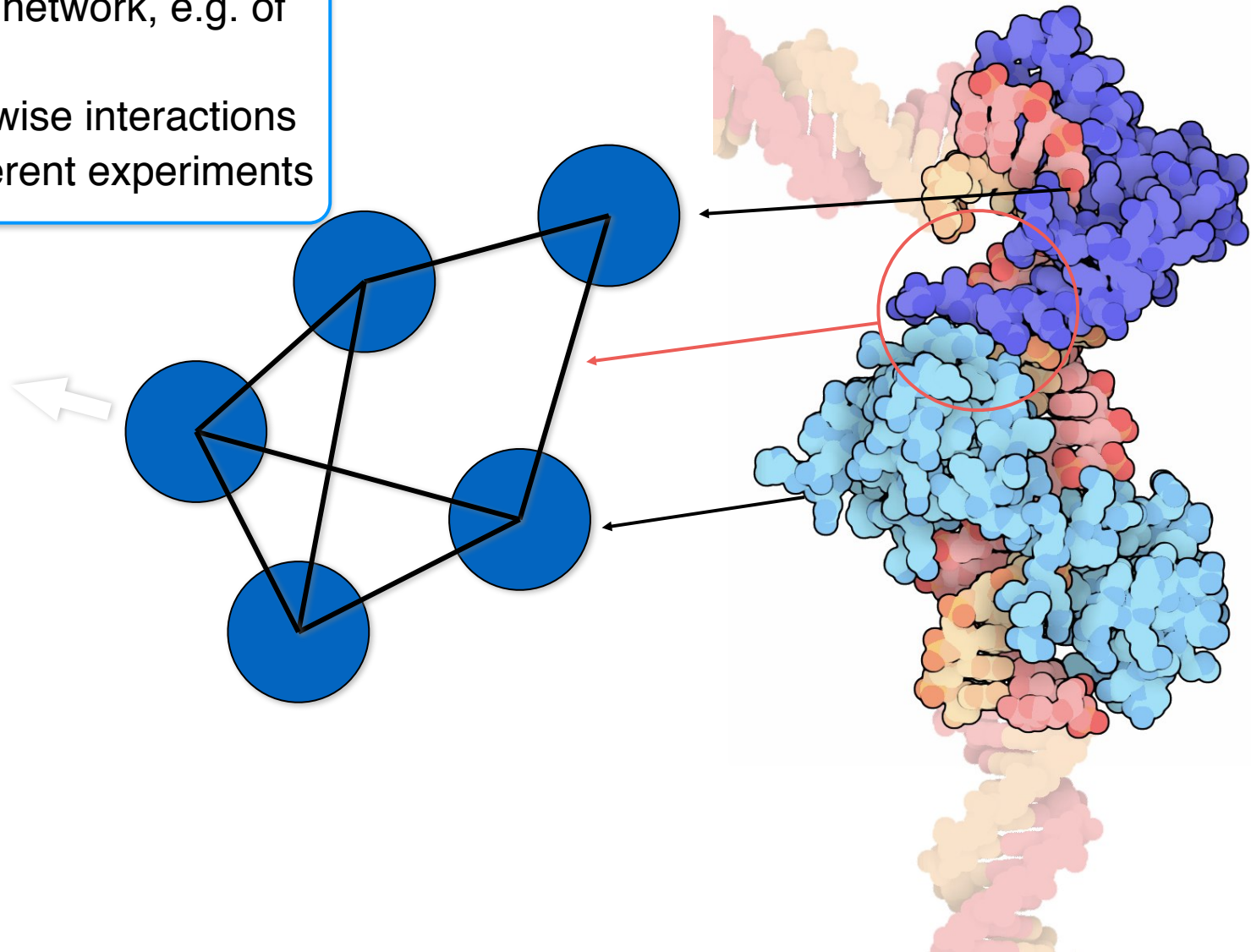
The structures are pulled apart for clarity.

Ovchinnikov, Kamisetty,  
Baker (2014) eLife 3:e02030

# Toward condition-specific protein interaction networks

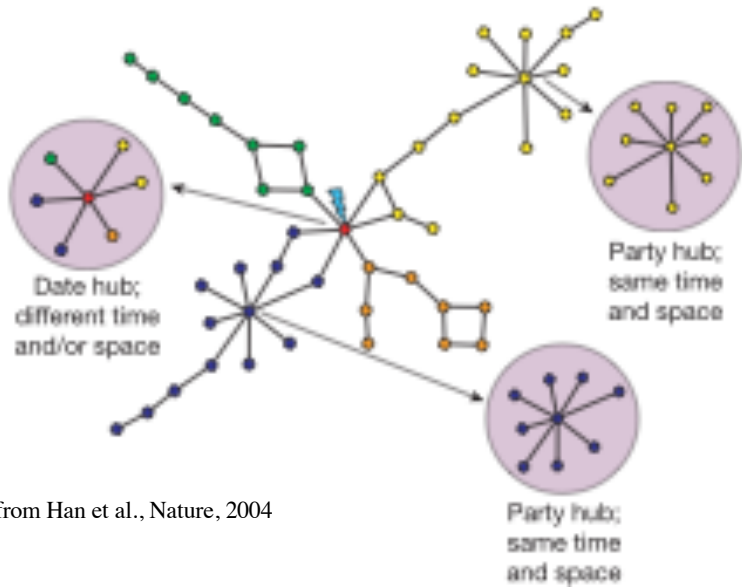
Full interaction PP network, e.g. of human  
= collection of pairwise interactions  
compiled from different experiments

broad range of  
applications



# But protein interactions can be ...

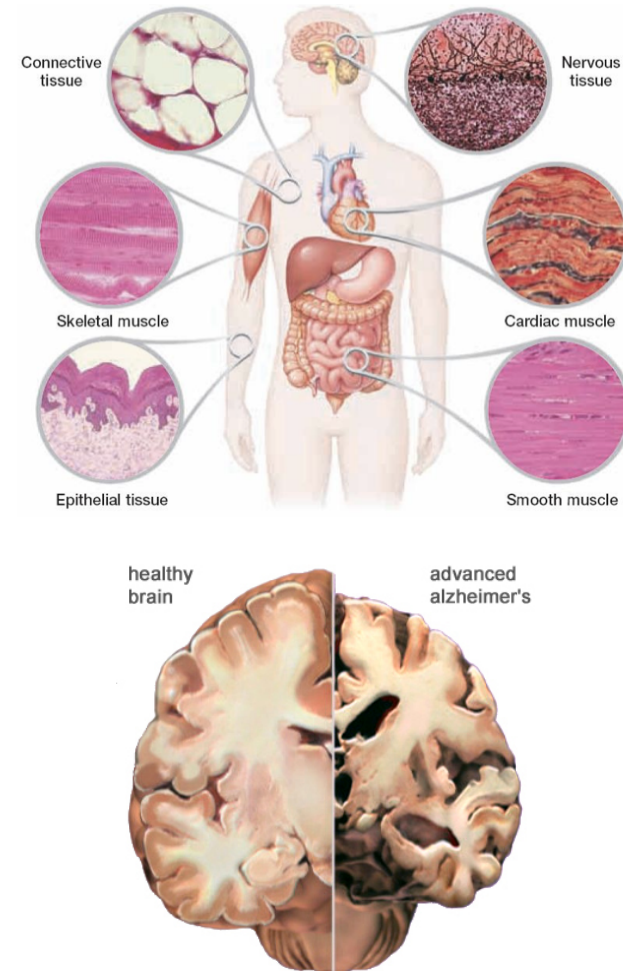
dynamic in time and space



same color = similar expression profiles

interaction data itself  
generally **static**

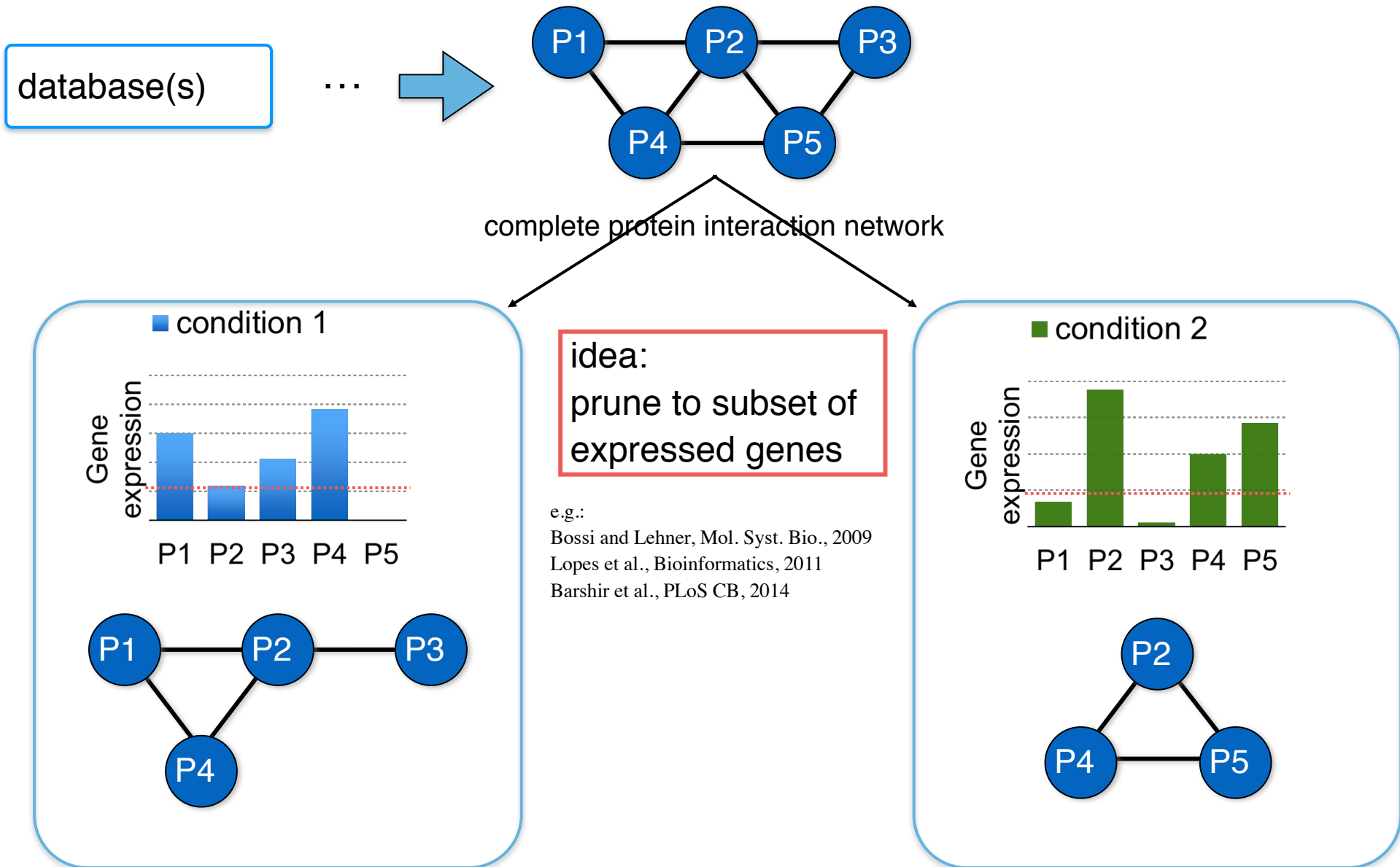
condition-specific  
protein composition



Human tissues from [www.pharmaworld.pk](http://www.pharmaworld.pk)  
Alzheimer from [www.alz.org](http://www.alz.org)

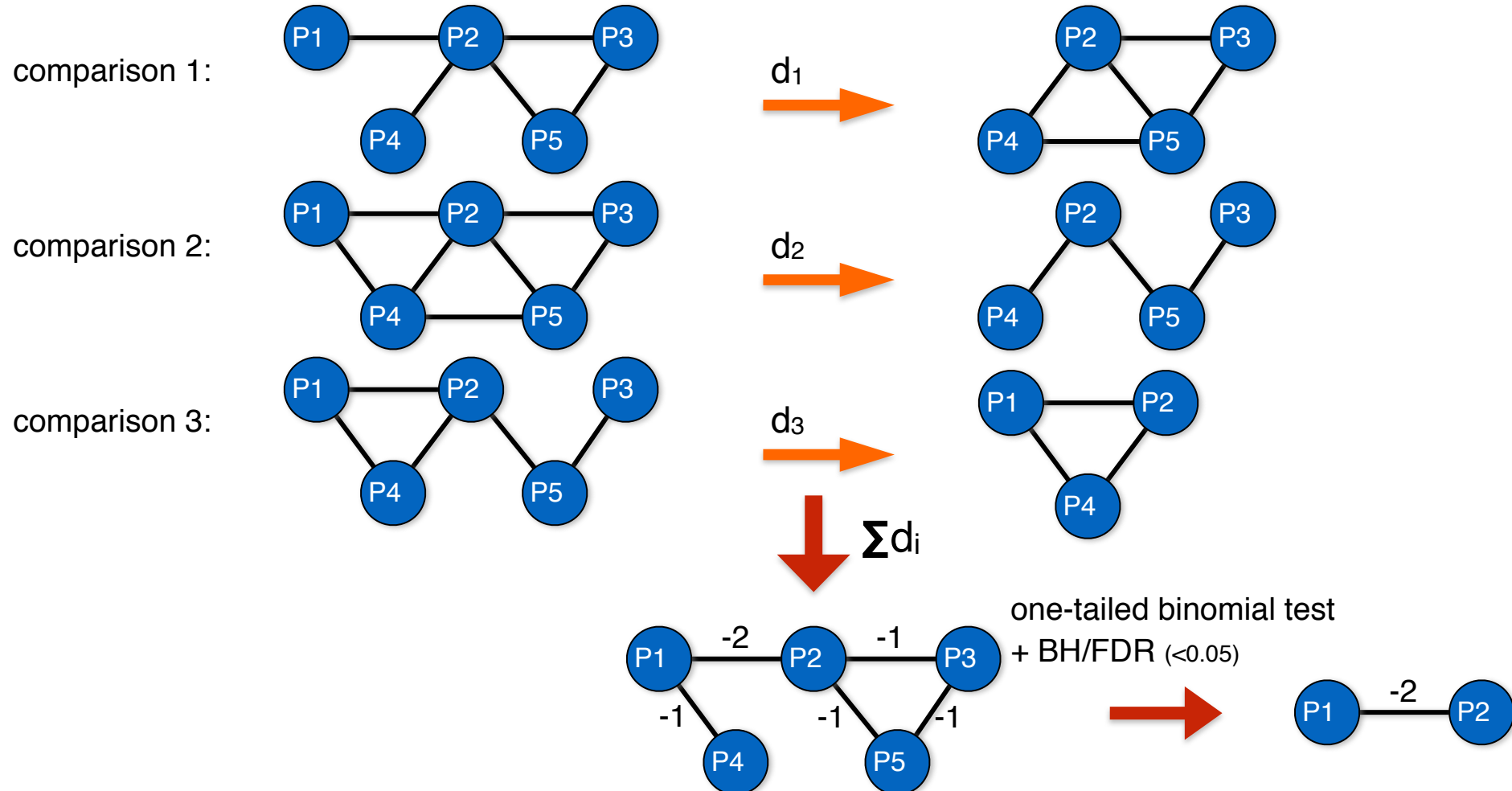


# Simple condition-specific PPI networks



# Differential PPI wiring analysis

112 matched normal tissues (TCGA)    112 breast cancer tissues (TCGA)



Check whether rewiring of a particular PP interaction occurs in a significantly large number of patients compared to what is expected by chance rewiring events.

# How much rewiring of PPIs exists?

Standard deviations reflect differences between patients.

About 10.000 out of 133.000 protein-protein interactions are significantly rewired between normal and cancer samples.

	GENE
avg. number of proteins (normal)	12,678 ± 223
avg. number of proteins (tumor)	12,528 ± 206
avg. number of interactions (normal)	134,348 ± 2,387
avg. number of interactions (tumor)	133,128 ± 2,144
$P_{\text{rew}}$	0.067 ± 0.016
significantly rewired interactions	9,754

*Table S7: Results obtained using the BioGRID interaction data and using either gene- or various transcript-based network construction approaches. The given numbers denote the sizes of the constructed networks. For all deterministic approaches the standard deviation across all 112 matched samples is shown, for the randomized approach the deviation shown is the average of standard deviations per run. A part of the results for  $P_{\text{rew}}$  and significantly rewired interactions are also shown in the upper half of Table 3 in the main text. Both net loss of proteins and interactions from normal to tumor were significant according to a two-sided Wilcoxon signed-rank test applied to the matched pairs of samples. For the*

# Rewired PPIs are associated with hallmarks

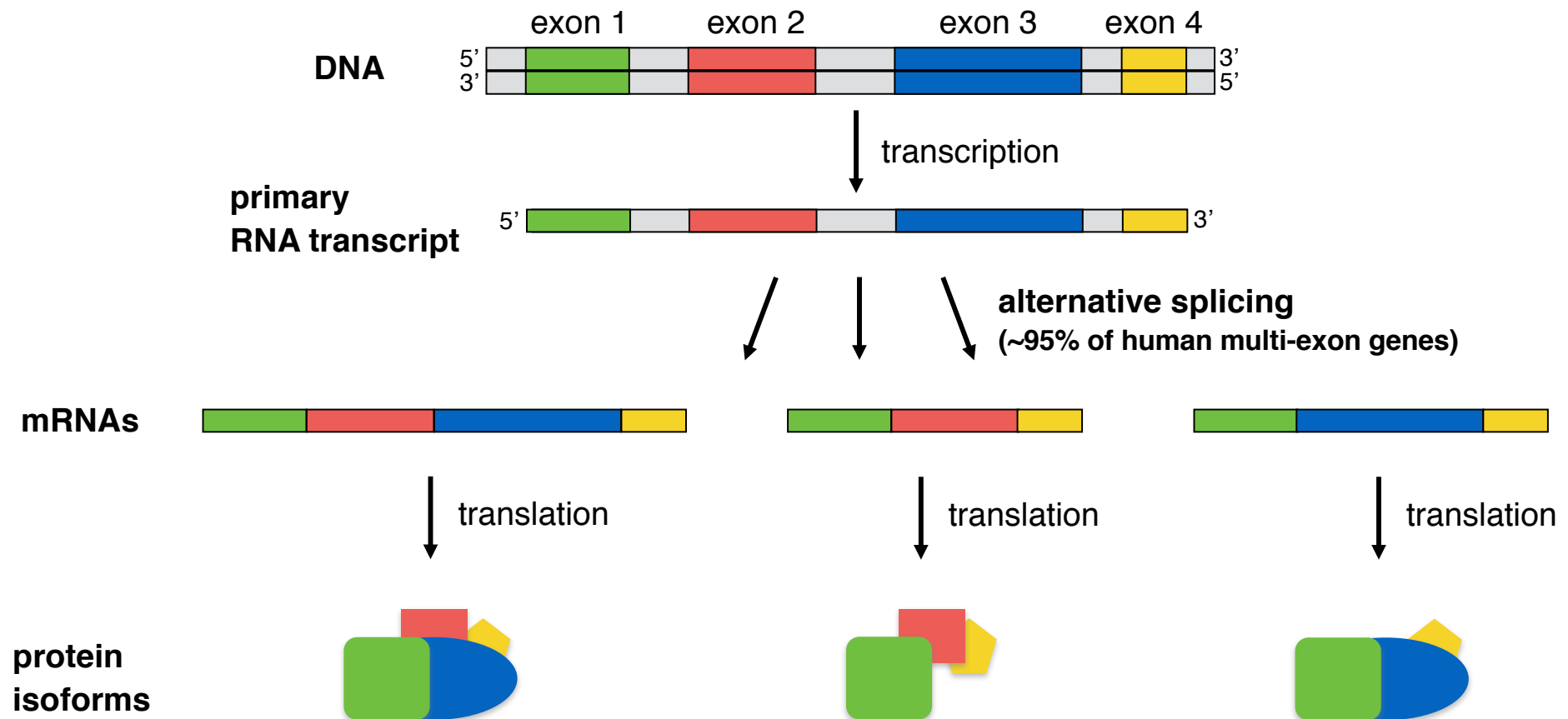
	GENE
rewired interactions	9,754
participation in any hallmark term	7,028
fraction in any hallmark term	0.721
Resisting Cell Death	4,064 (0.417)
Activating Invasion and Metastasis	2,244 (0.230)
Sustaining Proliferative Signaling	3,964 (0.406)
Inducing Angiogenesis	169 (0.017)
Tumor-Promoting Inflammation	516 (0.053)
Genome Instability and Mutation	1,362 (0.140)
Enabling Replicative Immortality	232 (0.024)
Evading Growth Suppressors	3,362 (0.345)
Avoiding Immune Destruction	752 (0.077)
Deregulating Cellular Energetics	821 (0.084)
avg.	1,749 (0.179)

A large fraction (72%) of the rewired interactions affects genes that are associated with „hallmark of cancer“ terms.

*Table S10: Results for the rewiring analysis of the BioGRID network in terms of rewired interactions that affect proteins associated with hallmarks of cancer as defined by [1]. A protein interaction was considered relevant regarding a hallmark term if at least one of its associated proteins was part of the corresponding set of hallmark proteins. The results for individual hallmark terms are reported as the absolute quantity of matches (left number) and as fraction of the total number of rewired interactions listed in the first row (in brackets).*



# Not considered yet: alternative splicing

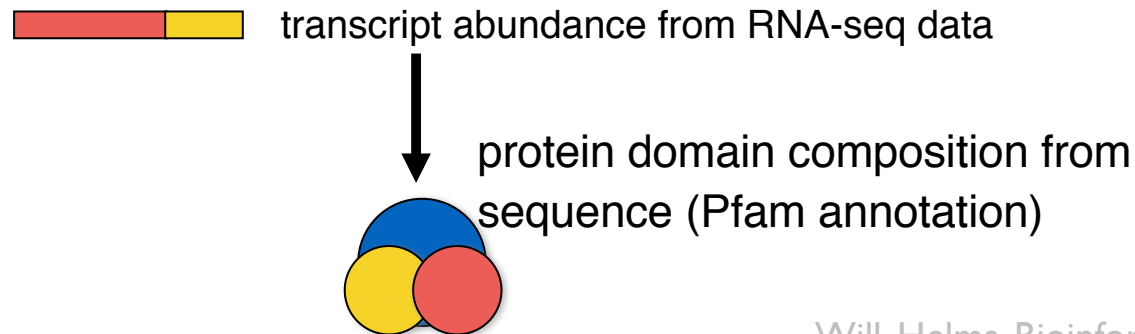


AS affects ability of proteins to interact with other proteins

# PPIXpress uses domain information

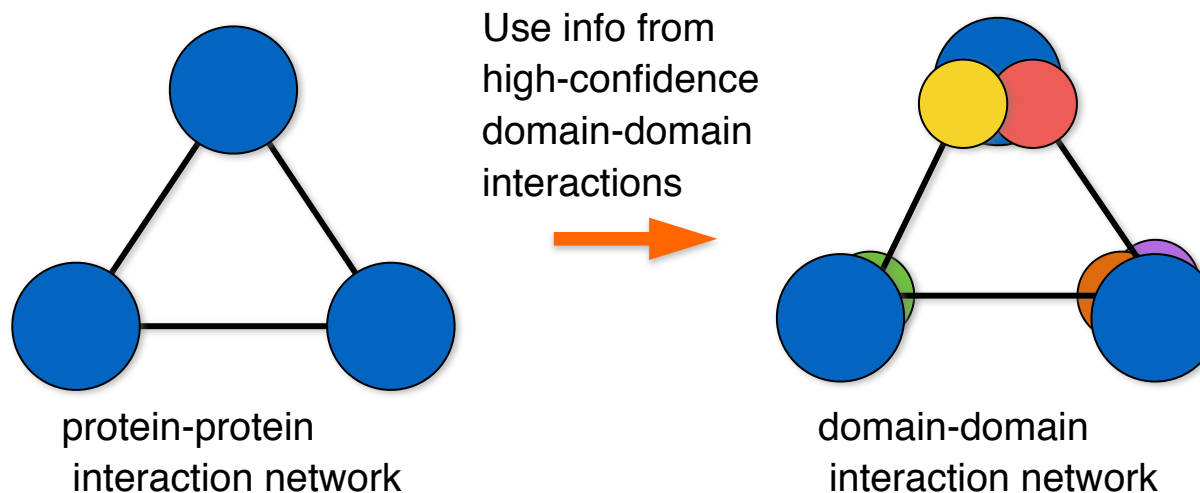
see <http://sourceforge.net/projects/ppixpress>

## I. Determine “building blocks” for all proteins



Will, Helms, Bioinformatics, 47, 219 (2015)  
doi: 10.1093/bioinformatics/btv620

## II. Connect them on the domain-level



# Coverage of PPIs with domain information

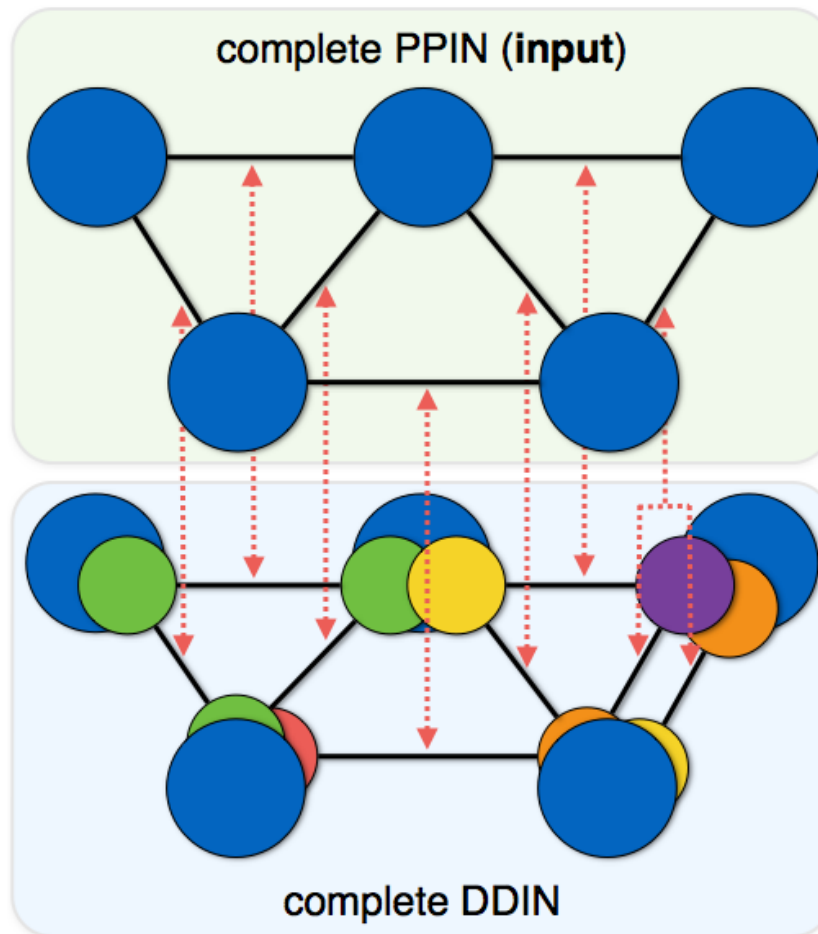
protein set	size of set	fraction of	
		matched PPIs	contributing proteins
complete network*	15086	0.264	0.517
all HM	4407	0.280	0.684
non HM	10679	0.227	0.449

Domain information is currently available for 51.7% of the proteins of the PP interaction network.

This means that domain information supports about one quarter (26.7%) of all PPIs.

All other PPIs were connected by us via artificially added domains (1 protein = 1 domain).

# PPIXpress method



## mapping:

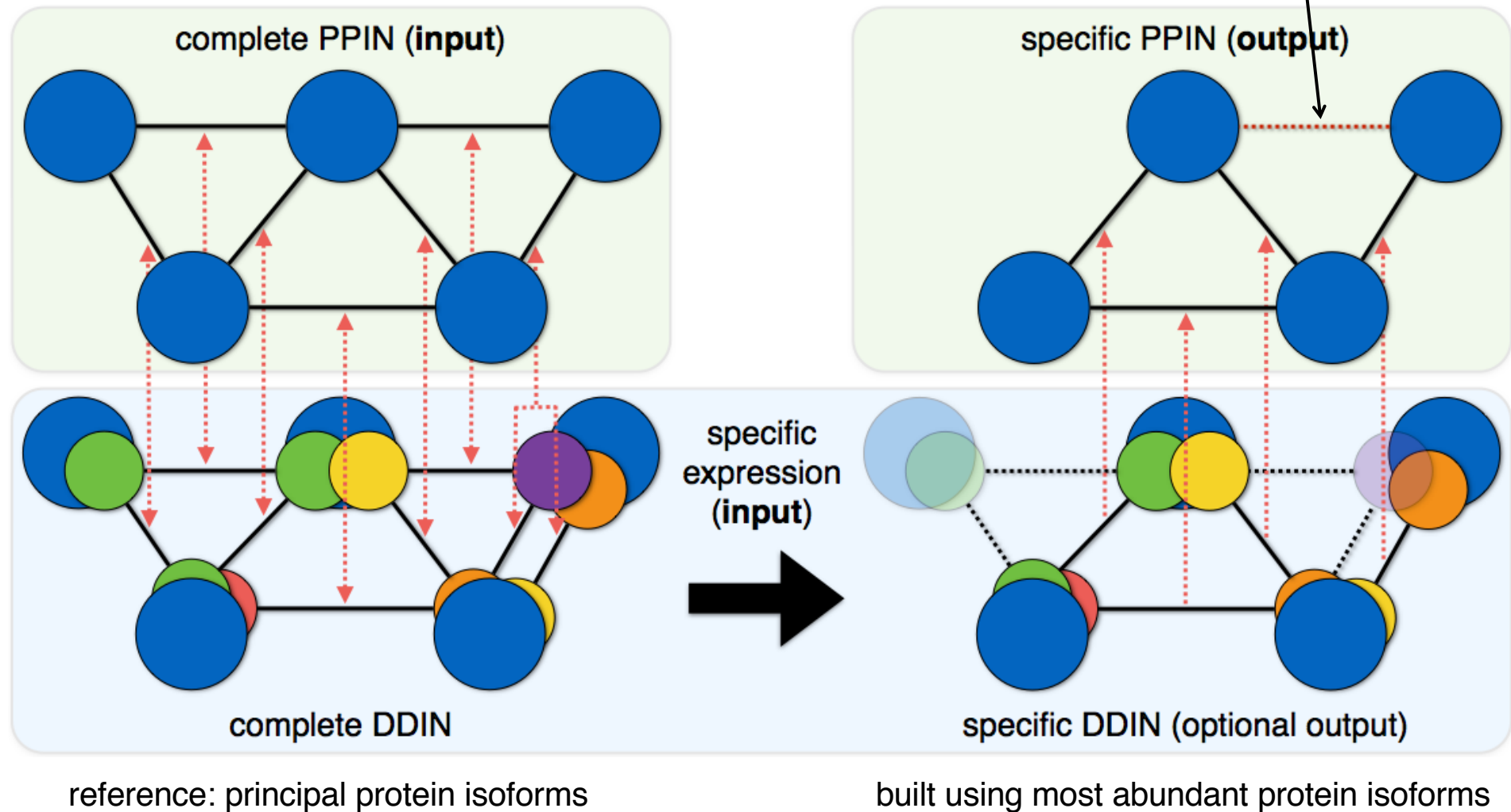
protein-protein interaction

establish  
one-to-at-least-one  
relationship

domain-domain interaction

reference: principal protein isoforms = longest coding transcript

# PPIXpress method

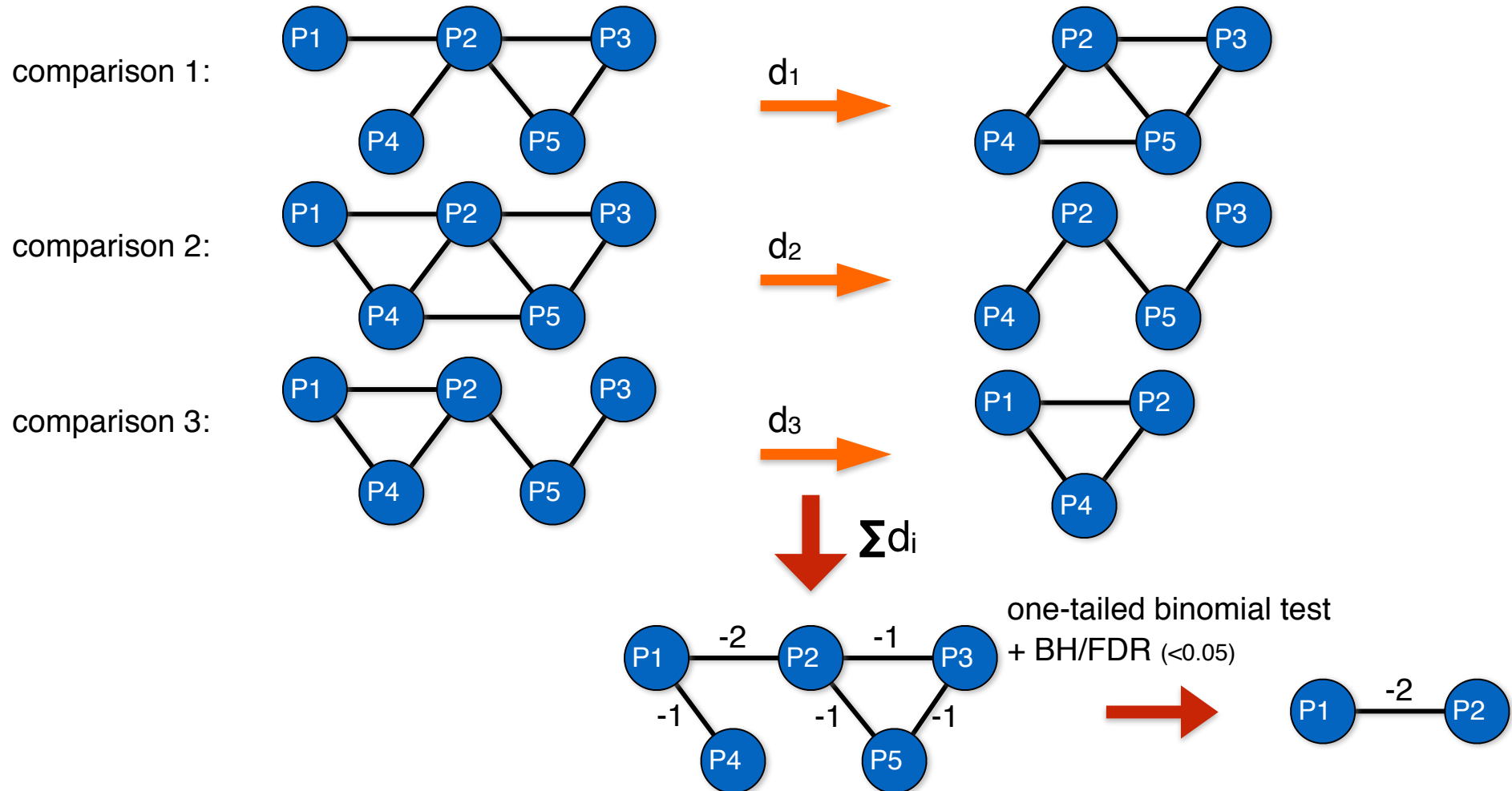


## I. mapping

## II. instantiation

# Differential PPI wiring analysis at domain level

112 matched normal tissues (TCGA)    112 breast cancer tissues (TCGA)



# Rewired PPIs are associated with hallmarks

	GENE	ALL-DDI
rewired interactions	9,754	10,111
participation in any hallmark term	7,028	7,343
fraction in any hallmark term	0.721	0.726
Resisting Cell Death	4,064 (0.417)	4,316 (0.427)
Activating Invasion and Metastasis	2,244 (0.230)	2,285 (0.226)
Sustaining Proliferative Signaling	3,964 (0.406)	4,142 (0.410)
Inducing Angiogenesis	169 (0.017)	172 (0.017)
Tumor-Promoting Inflammation	516 (0.053)	537 (0.053)
Genome Instability and Mutation	1,362 (0.140)	1,419 (0.140)
Enabling Replicative Immortality	232 (0.024)	360 (0.036)
Evading Growth Suppressors	3,362 (0.345)	3,557 (0.352)
Avoiding Immune Destruction	752 (0.077)	772 (0.076)
Deregulating Cellular Energetics	821 (0.084)	850 (0.084)
avg.	1,749 (0.179)	1,841 (0.182)

The construction at transcript-level found a larger fraction (72.6 vs 72.1%) of differential interactions that can be associated with hallmark terms than the gene-level based approach.

*Table S10: Results for the rewiring analysis of the BioGRID network in terms of rewired interactions that affect proteins associated with hallmarks of cancer as defined by [1]. A protein interaction was considered relevant regarding a hallmark term if at least one of its associated proteins was part of the corresponding set of hallmark proteins. The results for individual hallmark terms are reported as the absolute quantity of matches (left number) and as fraction of the total number of rewired interactions listed in the first row (in brackets).*



# Enriched KEGG and GO-BP terms in gene-level \ transcript-level set

GENE			ALL-DDI	
	term	p	term	p
KEGG	hsa04012:ErbB signaling pathway	0.0013	hsa05200:Pathways in cancer	$1.5 * 10^{-17}$
	hsa05212:Pancreatic cancer	0.0491	hsa04110:Cell cycle	$1.8 * 10^{-15}$
			hsa05220:Chronic myeloid leukemia	$3.5 * 10^{-15}$
			hsa05212:Pancreatic cancer	$1.4 * 10^{-8}$
			hsa05223:Non-small cell lung cancer	$4.3 * 10^{-8}$
GO BP	GO:0007242 intracellular signaling cascade	$6.9 * 10^{-5}$	GO:0010604 positive regulation of macromolecule metabolic process	$4.3 * 10^{-16}$
	GO:0043065 positive regulation of apoptosis	0.0252	GO:0042981 regulation of apoptosis	$3.6 * 10^{-15}$
	GO:0043068 positive regulation of programmed cell death	0.0272	GO:0043067 regulation of programmed cell death	$6.1 * 10^{-15}$
	GO:0010942 positive regulation of cell death	0.0287	GO:0010941 regulation of cell death	$7.7 * 10^{-15}$
	GO:0051329 interphase of mitotic cell cycle	0.0409	GO:0007049 cell cycle	$1.7 * 10^{-14}$

*Table S16: Comparison of rewiring results between the gene-based construction and a transcript-based construction method for the BioGRID network. Here, the top five enriched terms and their p-values are shown for the proteins affected by interactions exclusively found by the transcript-based method using the ALL-DDI dataset or the gene-based approach, respectively. Enrichment in KEGG pathways and GO biological processes was determined using DAVID [2] where we used the proteins included in the corresponding input network as the background. Enrichment was defined as  $p < 0.05$  (Bonferroni-adjusted).*

The enriched terms that are exclusively found by the transcript-level method (right) are closely linked to carcinogenetic processes.

Hardly any significant terms are exclusively found at the gene level (left).

# Conclusion (PPIXpress)

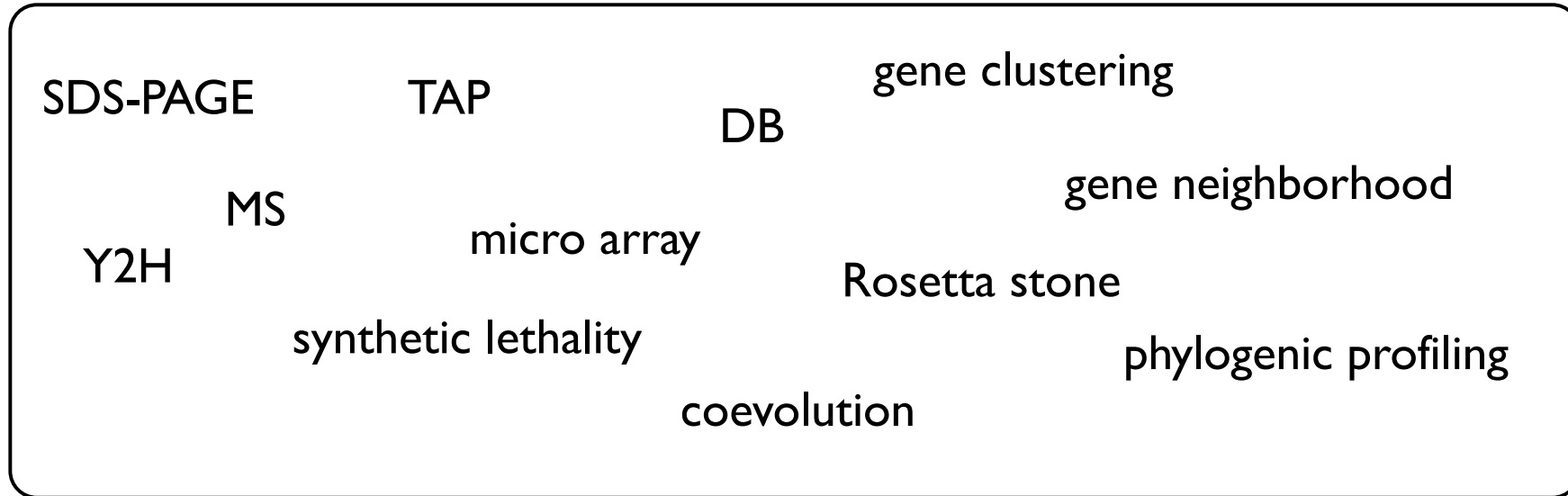
About 10.000 out of 130.000 PP interactions are **rewired** in cancer tissue compared to matched normal tissue due to **altered gene expression**.

The method PPIXpress exploits domain interaction data to adapt protein interaction networks to specific cellular conditions at transcript-level detail.

For the example of protein interactions in breast cancer this increase in granularity positively affected the performance of the network construction compared to a method that only makes use of gene expression data.

# Summary

What you learned **today**: how to get some data on PP interactions



type of interaction? — reliability? — sensitivity? — coverage? — ...

## Next lecture:

- combining weak indicators: Bayesian analysis
- identifying communities in networks