



蛋白质互作精要：构建和分析互作网络的核心概念

(Protein-Protein Interactions Essentials: Key Concepts to Building and Analyzing Interactome Networks)

小组分工：

翻 译：郭星宇、万紫云、管鹏鹏

ppt制作：管鹏鹏、万紫云

讲 解：万紫云

Author

Javier De Las Rivas

<http://bioinfow.dep.usal.es/>



Javier De Las Rivas



Cancer Research Center (IBMCC, CSIC/USAL), Salamanca, Spain
Bioinformatics, Functional Genomics, Cancer Genomics, Proteomics, Interactomics
在 usal.es 的电子邮件经过验证 - [首页](#)

标题 1-20	引用次数	发表年份
The minimum information required for reporting a molecular interaction experiment (MIMIx) S Orchard, L Salwinski, S Kerrien, L Montecchi-Palazzi, M Oesterheld, ... Nature biotechnology 25 (8), 894-898	224	2007
APID: agile protein interaction DataAnalyzer C Prieto, J De Las Rivas Nucleic acids research 34 (suppl 2), W298-W302	182	2006
Protein-protein interactions essentials: key concepts to building and analyzing interactome networks J De Las Rivas, C Fontanillo PLoS Comput Biol 6 (6), e1000807	174	2010
PSICQUIC and PSIScore: accessing and scoring molecular interactions B Aranda, H Blankenburg, S Kerrien, FSL Brinkman, A Ceol, E Chautard, ... Nature methods 8 (7), 528-529	155	2011
A functional model for the role of cytochrome b559 in the protection against donor and acceptor side photoinhibition J Barber, J De Las Rivas Proceedings of the National Academy of Sciences 90 (23), 10942-10946	155	1993

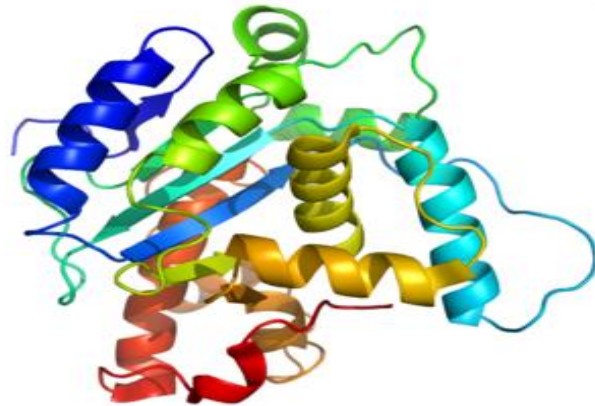
Why protein-protein interactions (PPIs)?

Gene is the basic unit of heredity. Genomes are available.



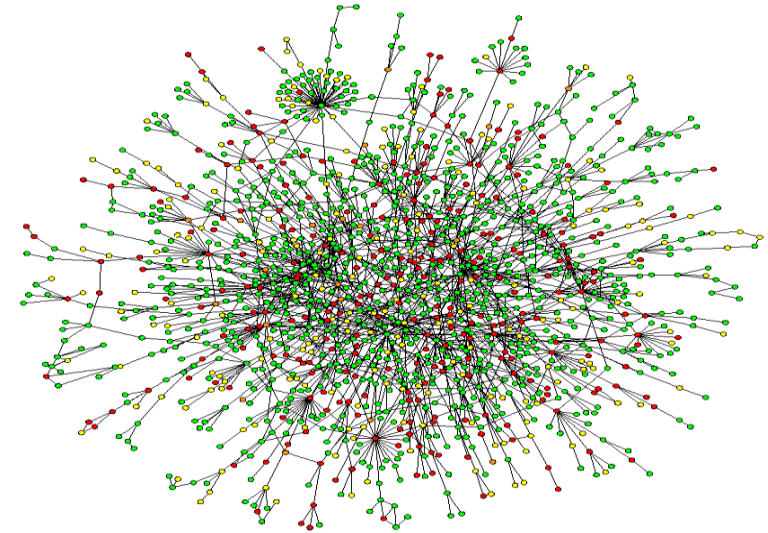
Genome

Proteins, the working molecules of a cell, carry out many biological activities



Proteome (蛋白质组)

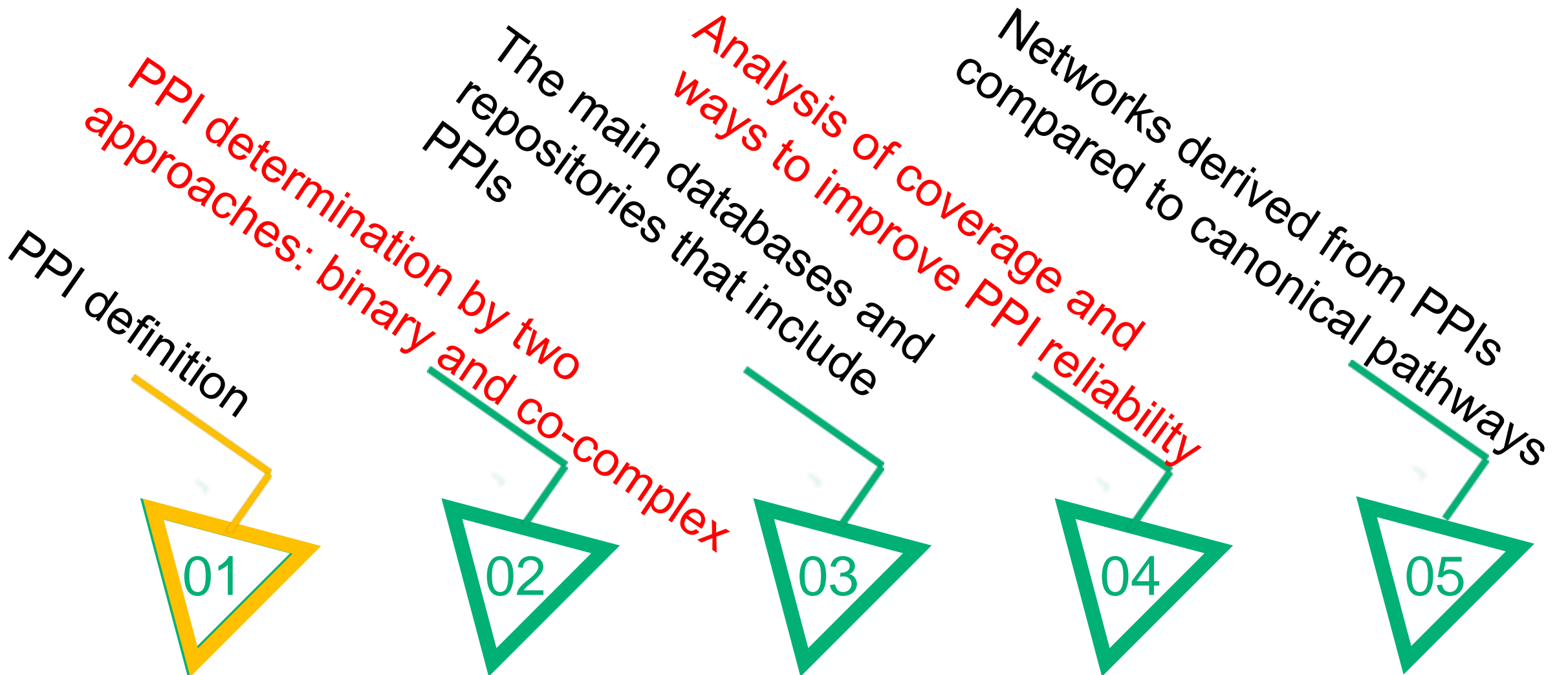
Proteins function by interacting with other proteins.



Interactome (交互组)

- The complete map of protein interactions that can occur in a living organism is called the **interactome**.
- **Efficient large-scale technologies** that measure proteome-wide physical connections between protein pairs are essential for accomplishing a comprehensive knowledge of the protein interactomes.
- To achieve appropriate understanding of PPIs and to design better ways for analyzing and interpreting them, this educational review presents several essential **concepts** and **definitions** intended to facilitate the use of PPI information both by computational and experimental biologists.

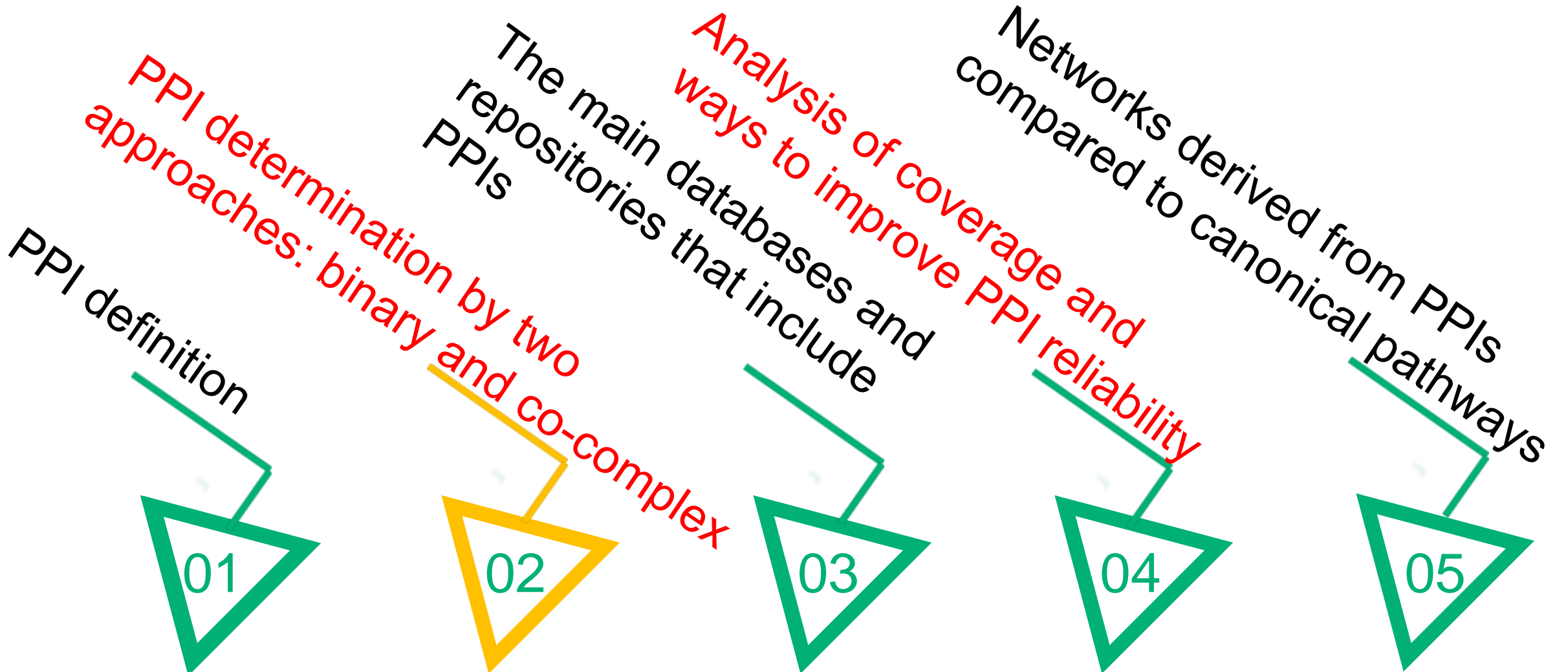
CONTENTS



PPI Definition

- Commonly they are understood as **physical contacts with molecular docking** between proteins that occur in a cell or in a living organism in vivo.
- The issue of whether two proteins share a **“functional contact”** is quite **distinct** from the question of whether the same two proteins **interact directly** with each other.
- The physical contact considered in PPIs should be **specific**, not just all proteins that bump into each other by chance.
- It also should exclude interactions that a protein experiences when it is being made, folded, quality checked, or degraded(i.e. generic interactions).

CONTENTS



Binary——Y2H (酵母双杂实验)

- The goal is to bring together the DNA-binding domain and the activation domain of the yeast Gal4 protein (Gal4p) through the interaction of two proteins, X and Y, to which one or other of the domains is fused.
- This interaction is accompanied by the expression of a reporter gene.

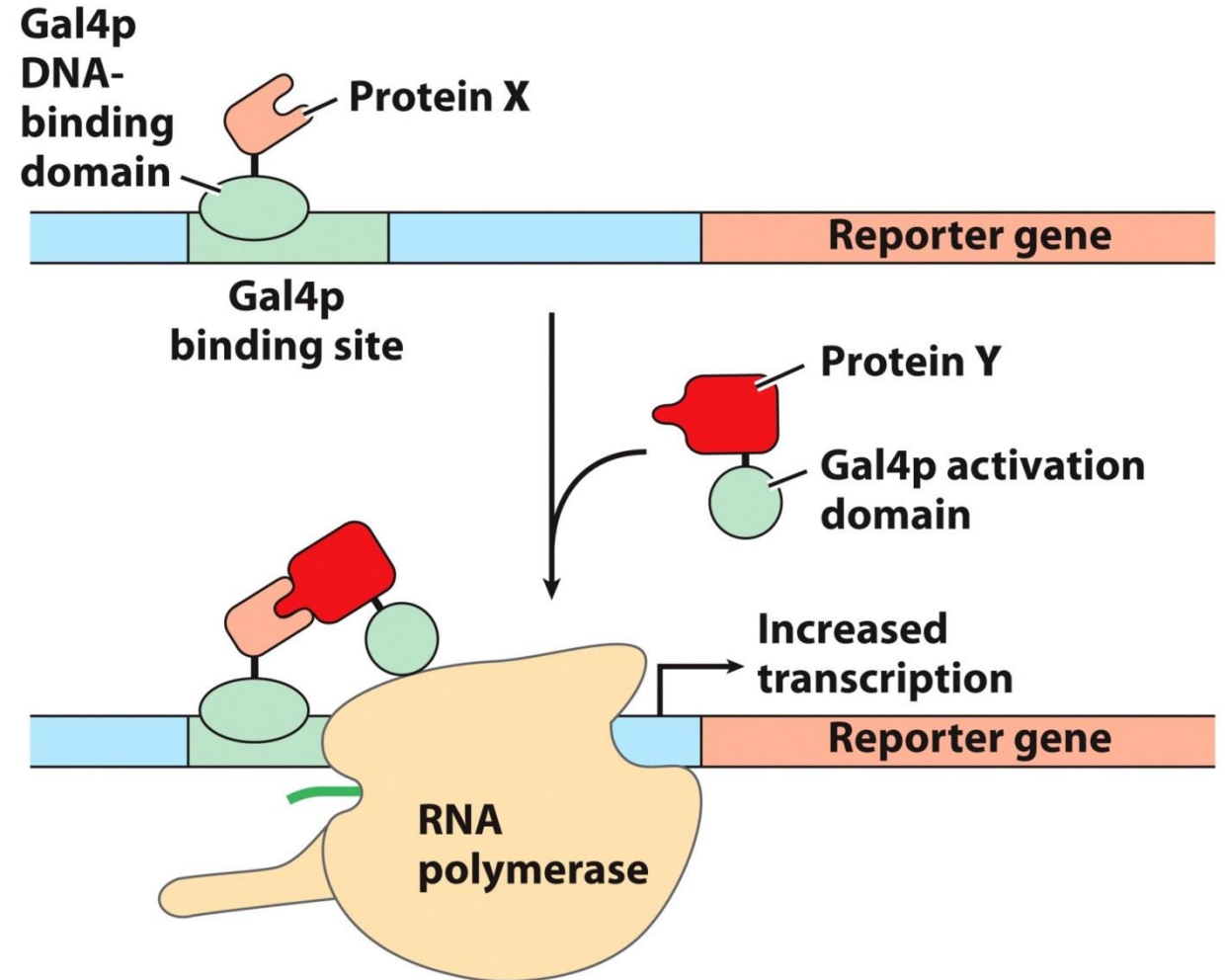


Figure 9-21a
Lehninger Principles of Biochemistry, Sixth Edition
© 2013 W. H. Freeman and Company

Binary——Y2H

(酵母双杂实验)

- The two gene fusions are created in separate yeast strains, which are then mated. The mated mixture is plated on a medium on which the yeast cannot survive unless the reporter gene is expressed.
- Thus, all surviving colonies have interacting fusion proteins.
- Sequencing of the fusion proteins in the survivors reveals which proteins are interacting.

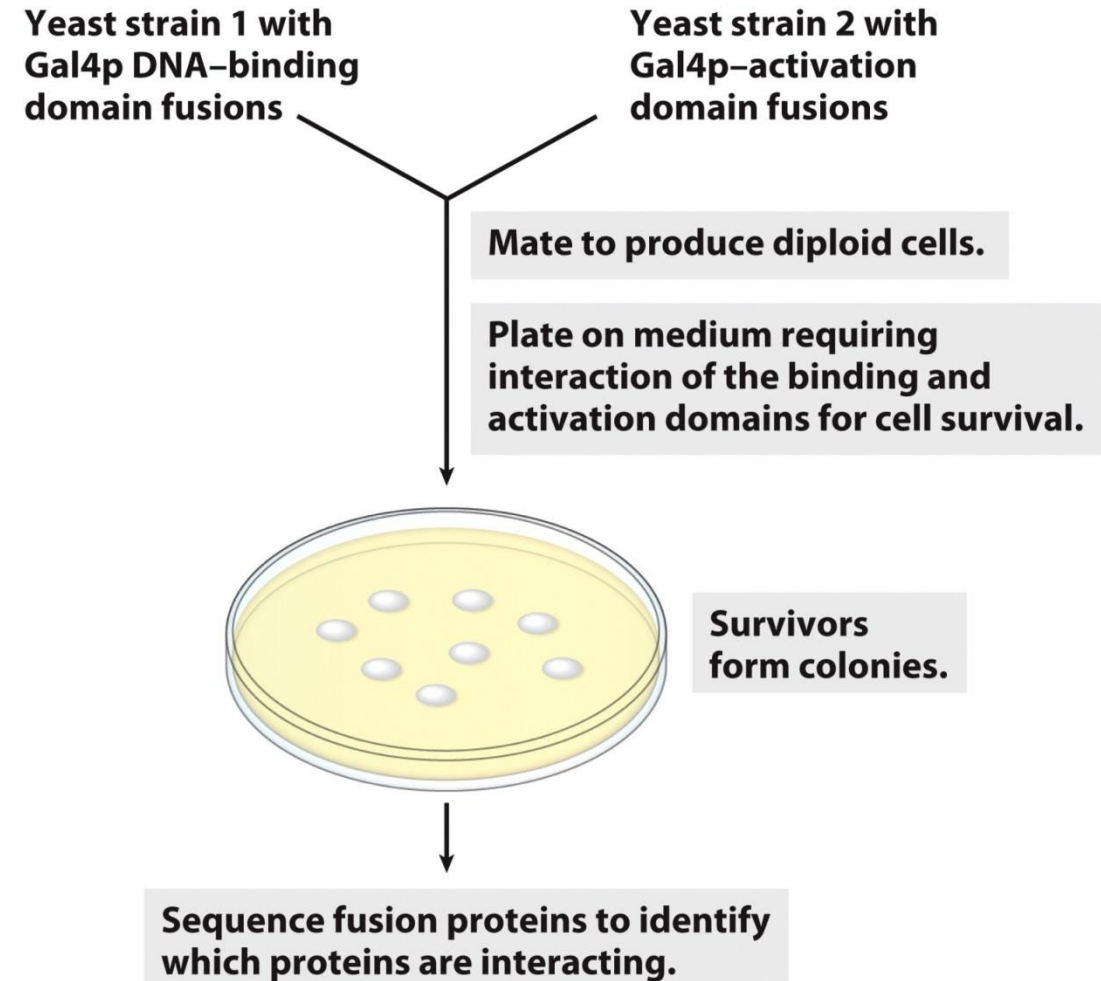


Figure 9-21b
Lehninger Principles of Biochemistry, Sixth Edition
© 2013 W. H. Freeman and Company

Co-complex—TAP-MS (串联亲和纯化-质谱)

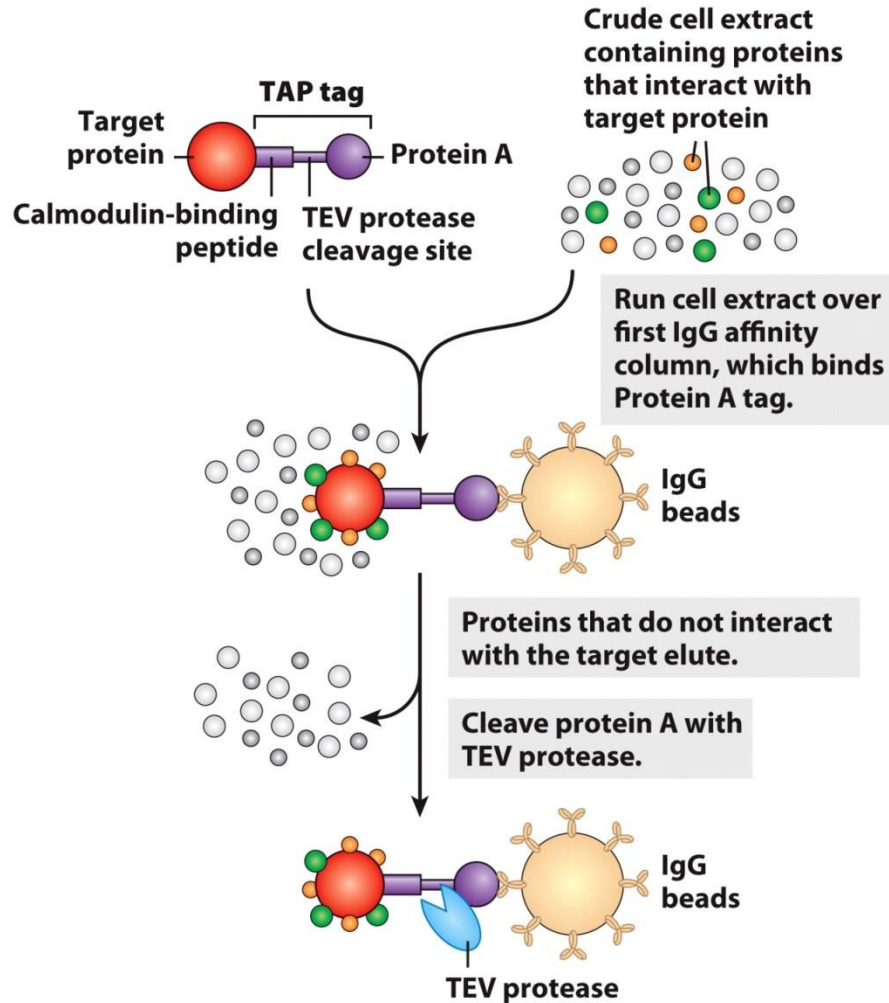


Figure 9-20 part 1
 Lehninger Principles of Biochemistry, Sixth Edition
 © 2013 W. H. Freeman and Company

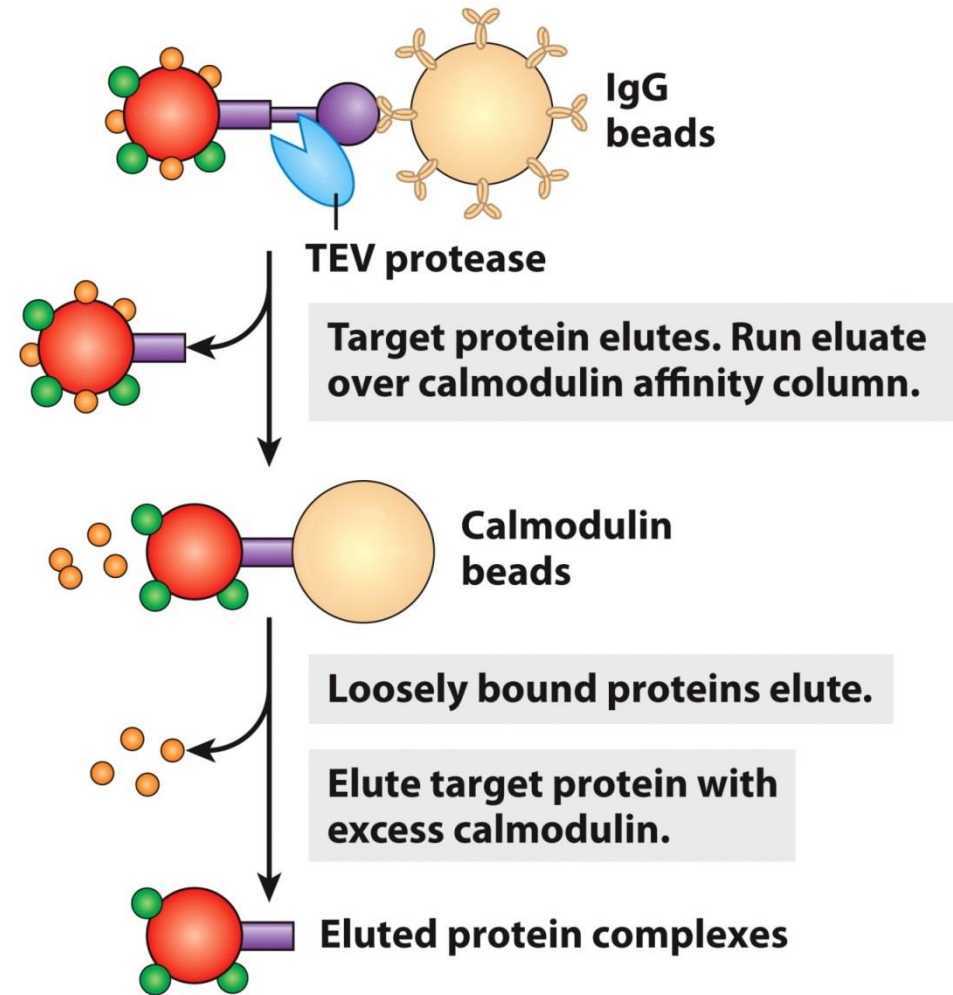
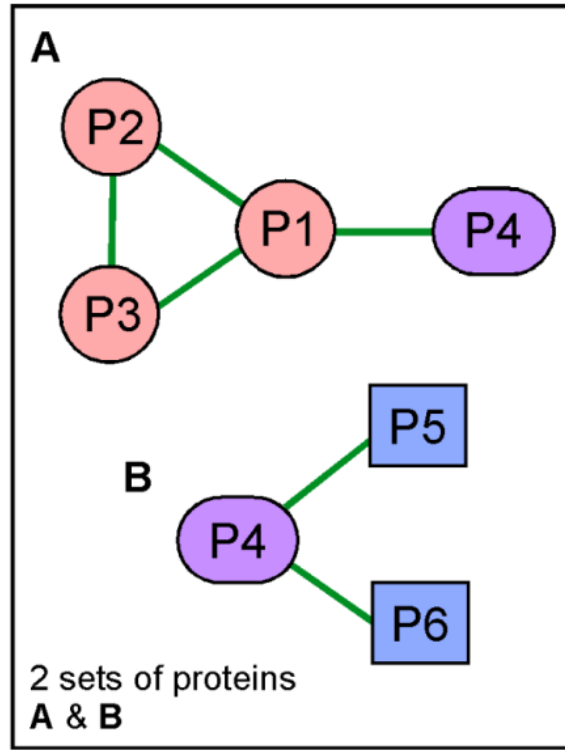


Figure 9-20 part 2
 Lehninger Principles of Biochemistry, Sixth Edition
 © 2013 W. H. Freeman and Company

Comparison of Two Approaches

True interactions (PPIs)
physical topology *in vivo*

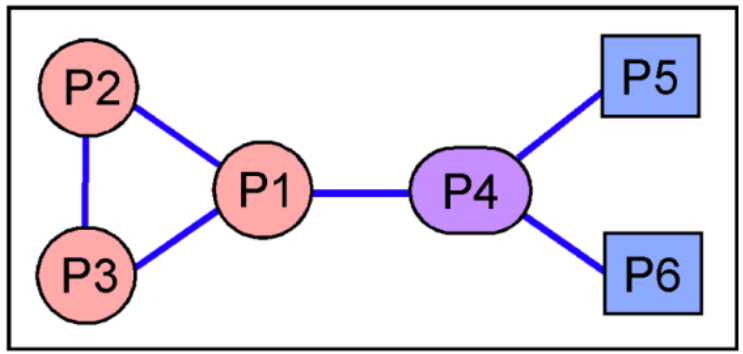


Binary methods
measure physical direct PPIs

e.g. Y2H

P1-P2	P2-P3	--- direct assignment -->
P1-P3	P4-P5	
P1-P4	P4-P6	

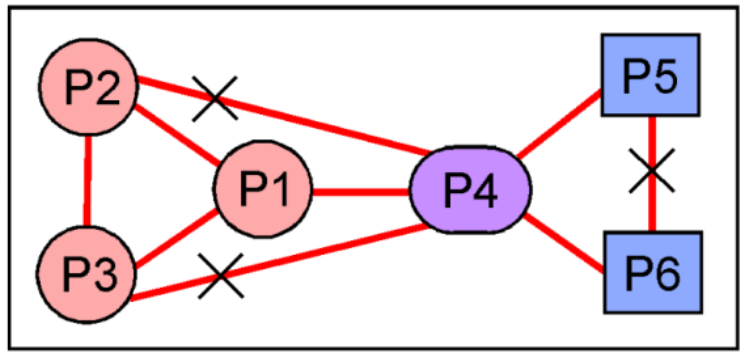
Experimental interactions (PPIs)
obtained from binary or co-complex methods



Two different PPI networks derived from two types of experimental data
(the X below indicate interactions that do not occur, i.e. they will be false positives)

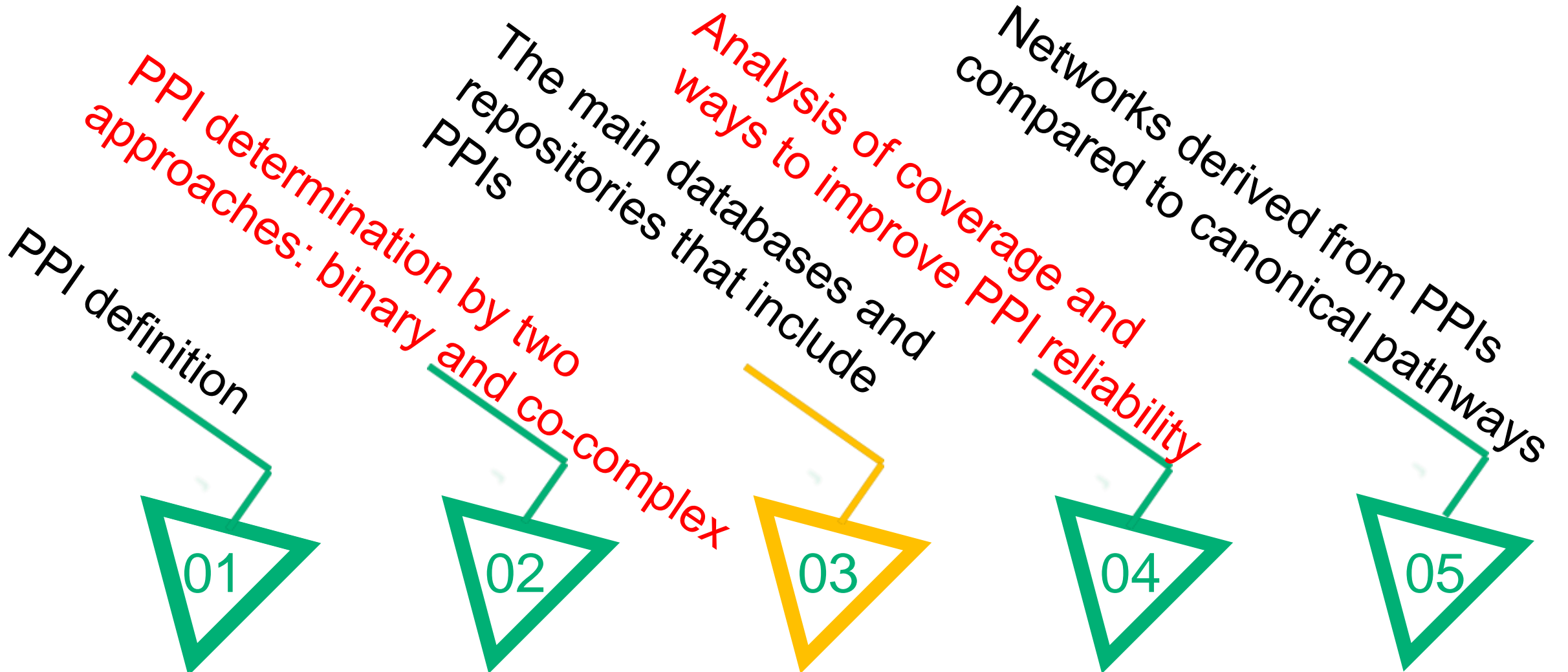
e.g. TAP-MS
CoIP

P1=P2,P3,P4	--- assignment with spoke model -->
P2=P1,P3,P4	
P3=P1,P2,P4	
P4=P1,P2,P3,P5,P6	
P5=P4,P6	
P6=P4,P5	



Co-complex methods
measure physical PPIs (direct & indirect)

CONTENTS



Primary Databases

Acronym	Database Full Name and URL	PPI Sources	Type of MI	Species	<i>n</i> Proteins (Dec. 2009)	<i>n</i> Interactions (Dec. 2009)
Primary Databases: PPI experimental data (curated from specific SSc & LSc published studies)						
BIND	Biomolecular Interaction Network Database, http://bond.unleashedinformatics.com/	Ssc & Lsc published studies (literature-curated)	PPIs & others	All	[31,972]	[58,266]
BioGRID	Biological General Repository for Interaction Datasets, http://www.thebiogrid.org/	Ssc & Lsc published studies (literature-curated)	PPIs & others	All	[28,717]	[108,691]
DIP	Database of Interacting Proteins, http://dip.doe-mbi.ucla.edu/dip/	Ssc & Lsc published studies (literature-curated)	Only PPIs	All	20,728	57,683
HPRD	Human Protein Reference Database, http://www.hprd.org/	Ssc & Lsc published studies (literature-curated)	Only PPIs	Human	27,081	38,806
IntAct	IntAct Molecular Interaction Database, http://www.ebi.ac.uk/intact/	Ssc & Lsc published studies (literature-curated)	PPIs & others	All	[60,504]	[202,826]
MINT	Molecular INTeraction database, http://mint.bio.uniroma2.it/mint/	Ssc & Lsc published studies (literature-curated)	Only PPIs	All	30,089	83,744
MIPS-MPact	MIPS protein interaction resource on yeast, http://mips.gsf.de/genre/proj/mpact/	Derived from CYGD	Only PPIs	Yeast	1,500	4,300
MIPS-MPPI	MIPS Mammalian Protein-Protein Interaction Database, http://mips.gsf.de/proj/ppi	Ssc published studies (literature-curated)	Only PPIs	Mammalian	982	937



Meta-Databases

Meta-Databases: PPI experimental data (integrated and unified from different public repositories)

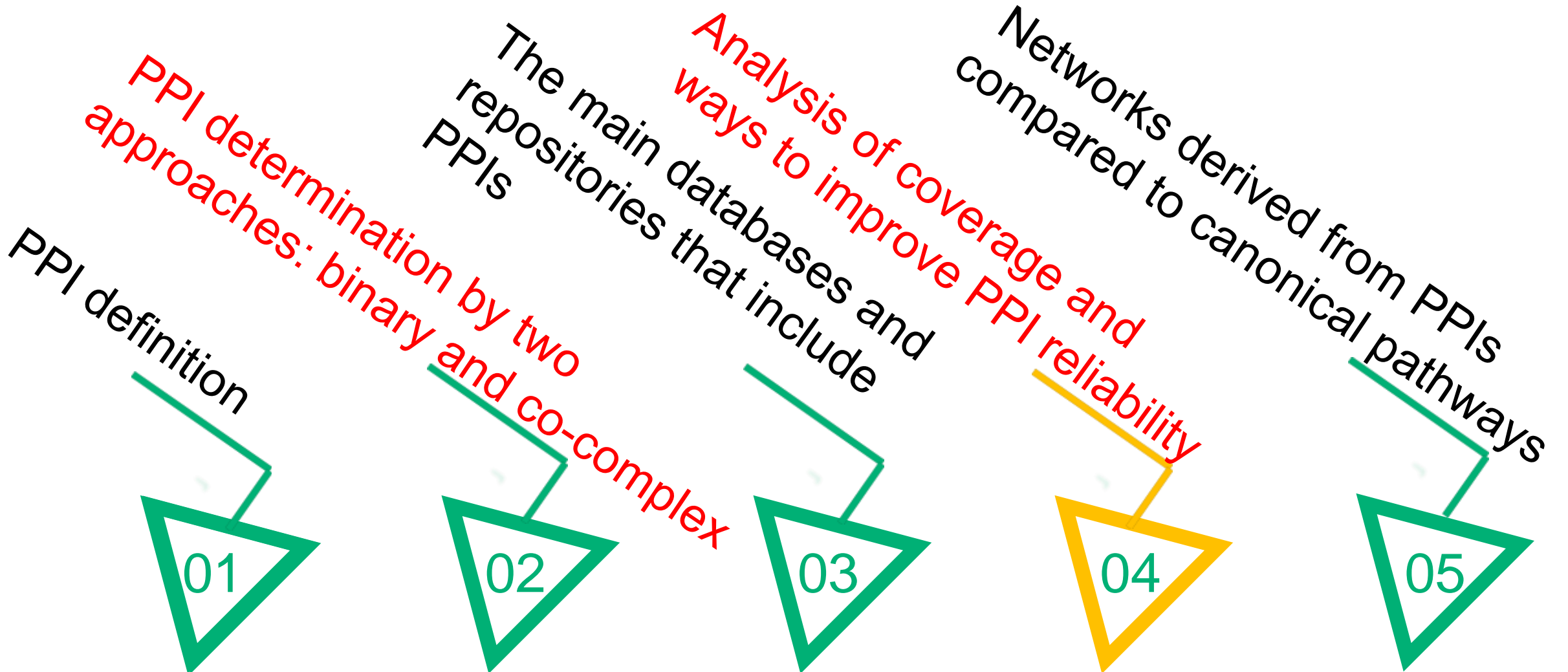
APID	Agile Protein Interaction DataAnalyzer, http://bioinfow.dep.usal.es/apid/	BIND, BioGRID, DIP, HPRD, IntAct, MINT	Only PPIs	All	56,460	322,579
MPIDB	The Microbial Protein Interaction Database, http://www.jcvi.org/mpidb/	BIND, DIP, IntAct, MINT, other sets (exp & lit.-curated)	Only PPIs	Microbial	7,810	24,295
PINA	Protein Interaction Network Analysis platform, http://csbi.itdk.helsinki.fi/pina/	BioGRID, DIP, HPRD, IntAct, MINT, MPact	Only PPIs	All	[?]	188,823

Prediction Databases

Prediction Databases: PPI experimental and predicted data (“functional interactions”, i.e., interactions *lato sensu* derived from different types of data)

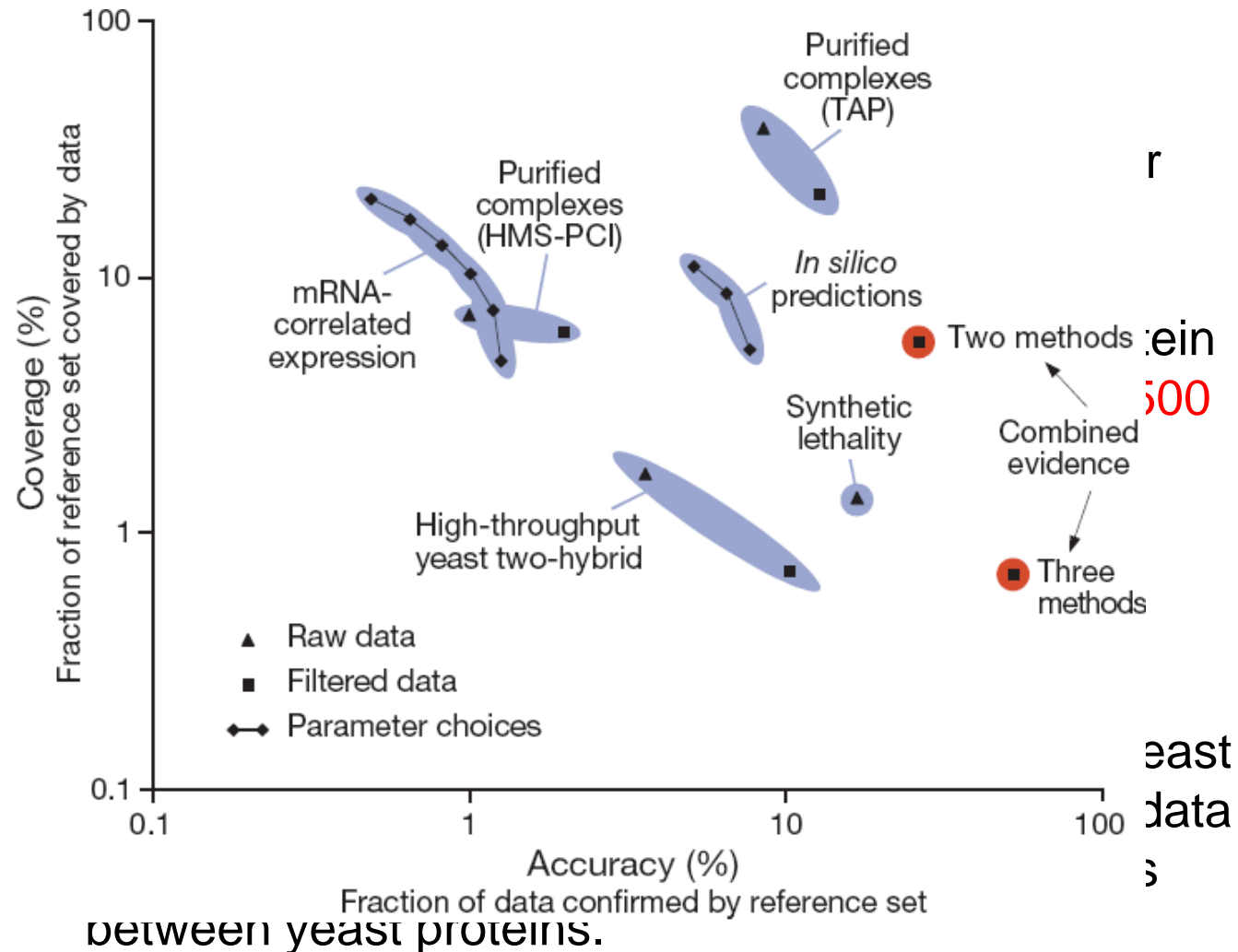
MiMI	Michigan Molecular Interactions, http://mimi.ncibi.org/MimiWeb/	BIND, BioGRID, DIP, HPRD, IntAct, & nonPPI data	PPIs & others	All	[45,452]	[391,386]
PIPs	Human PPI Prediction database, http://www.compbio.dundee.ac.uk/www-pips/	BIND, DIP, HPRD, OPHID, & nonPPI data	PPIs & others	Human	[?]	[37,606]
OPHID	Online Predicted Human Interaction Database, http://ophid.utoronto.ca/	BIND, BioGRID, HPRD, IntAct, MINT, MPact, & nonPPI data	PPIs & others	Human	[?]	[424,066]
STRING	Known and Predicted Protein-Protein Interactions, http://string.embl.de/	BIND, BioGRID, DIP, HPRD, IntAct, MINT, & nonPPI data	PPIs & others	All	[2,590,259]	[88,633,860]
UniHI	Unified Human Interactome, http://www.mdc-berlin.de/unihi/	BIND, BioGRID, DIP, HPRD, IntAct, MINT, & nonPPI data	PPIs & others	Human	[22,307]	[200,473]

CONTENTS



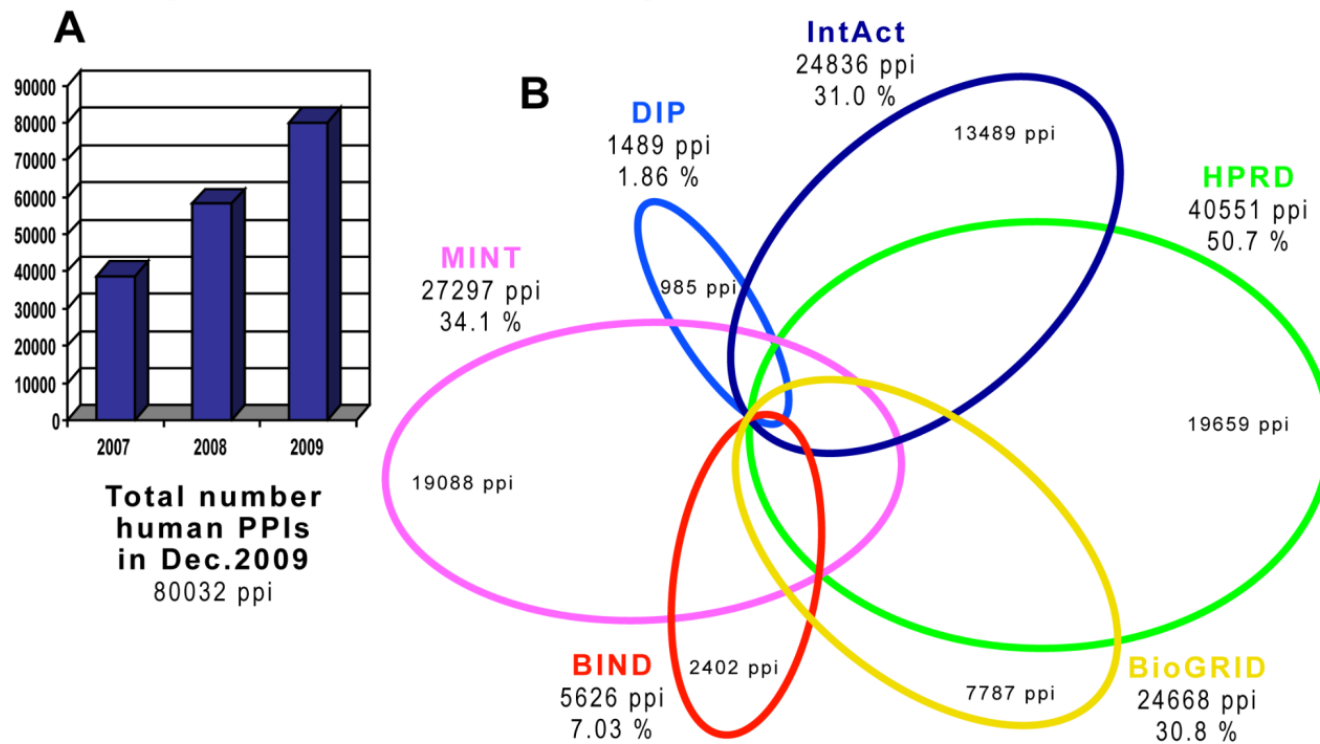


High error rate (False Positives)



Human Interactome

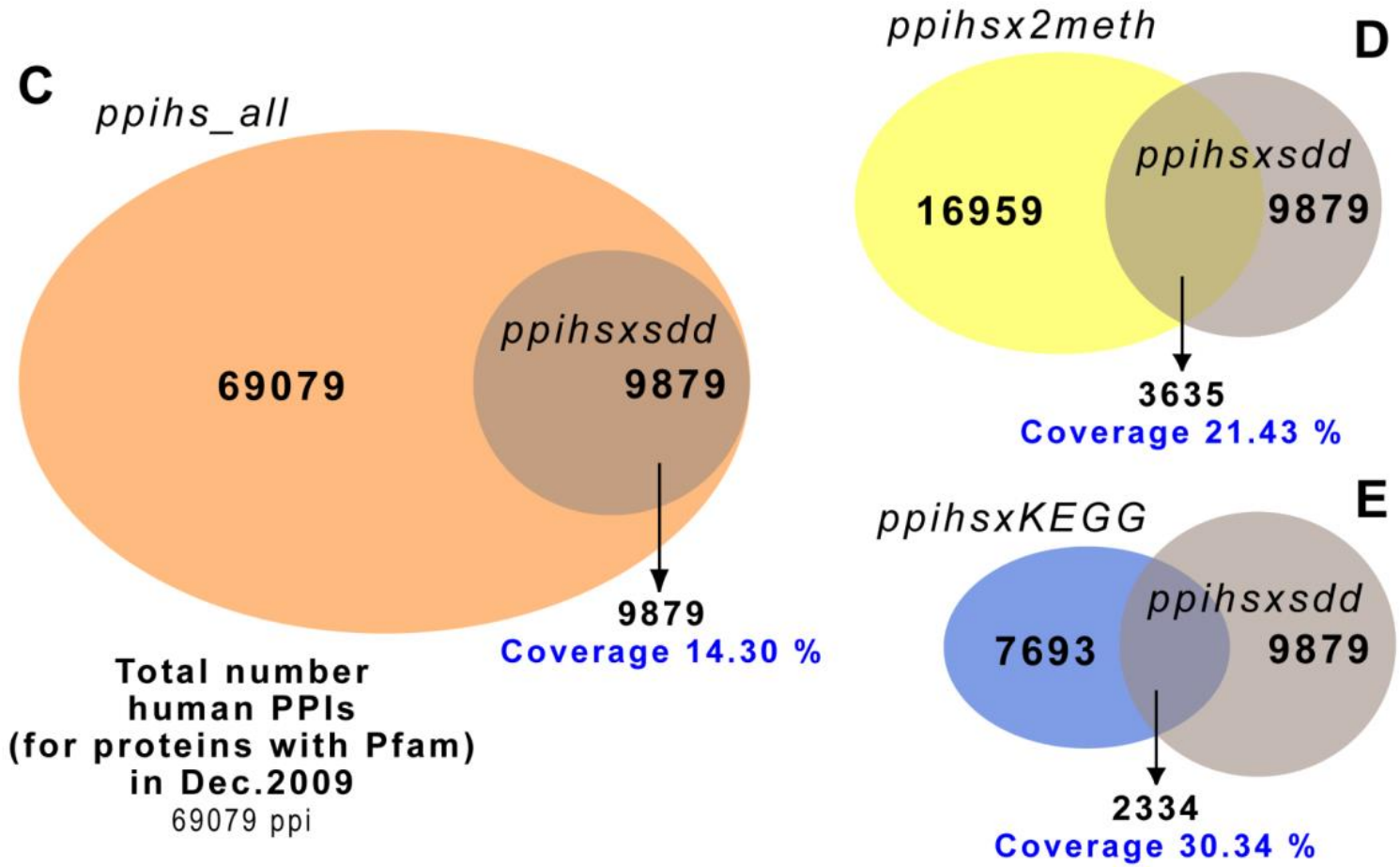
Coverage of human PPIs on major public repositories

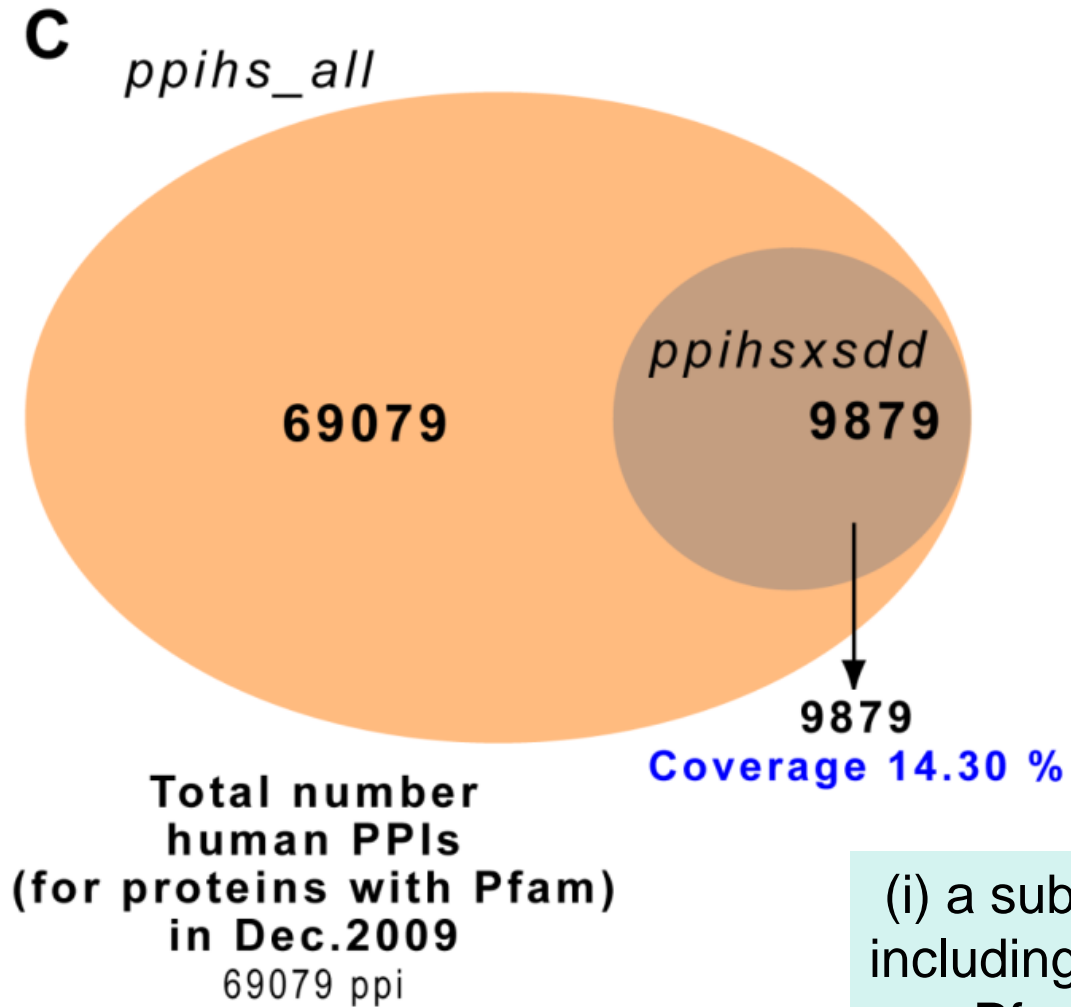


- The graph in Figure 2A shows the observed growth of human PPIs in the past 3 years.
- HPRD and MINT are the primary databases that include the most human PPIs: 50.7% and 34.1%, respectively.

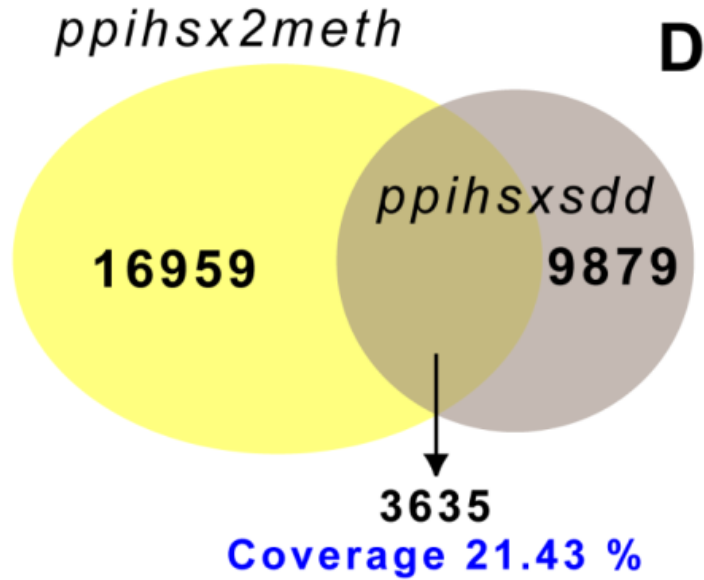
Incorporate 3-D Structural Information

Coverage of human PPIs with 3D structure (only proteins with Pfam)

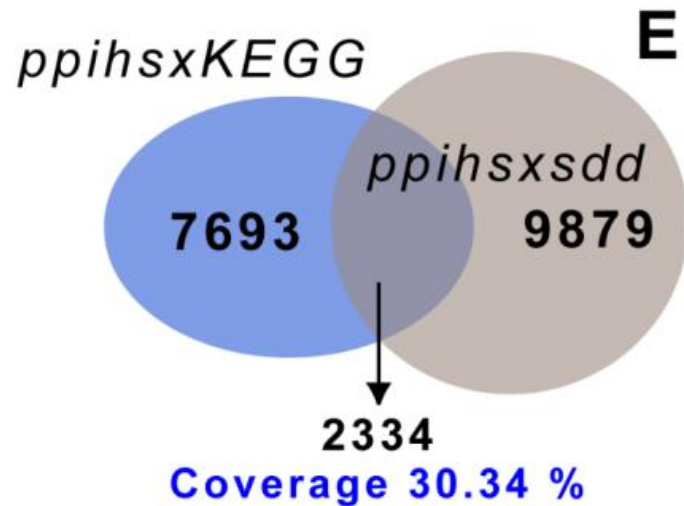




(i) a subset of the complete human PPI data including only the proteins that have at least one Pfam domain assigned: 69,079 interactions, called *ppihs_all* (Figure 2C)

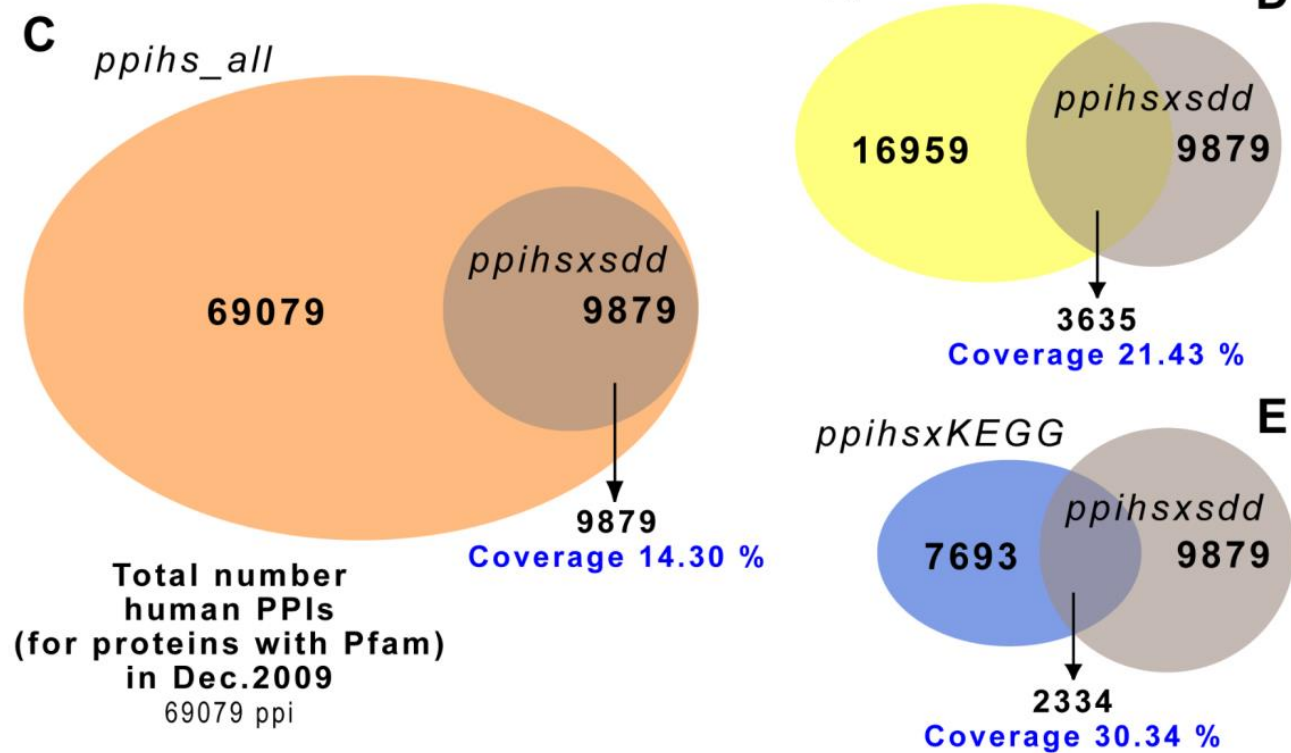


(ii) a subset of *ppihs_all* with only the interactions that have been validated by at least two experimental methods that demonstrate the interaction or by the same experimental method reported in at least two independently published articles: 16,959 interactions, called *ppihsx2meth* (Figure 2D)



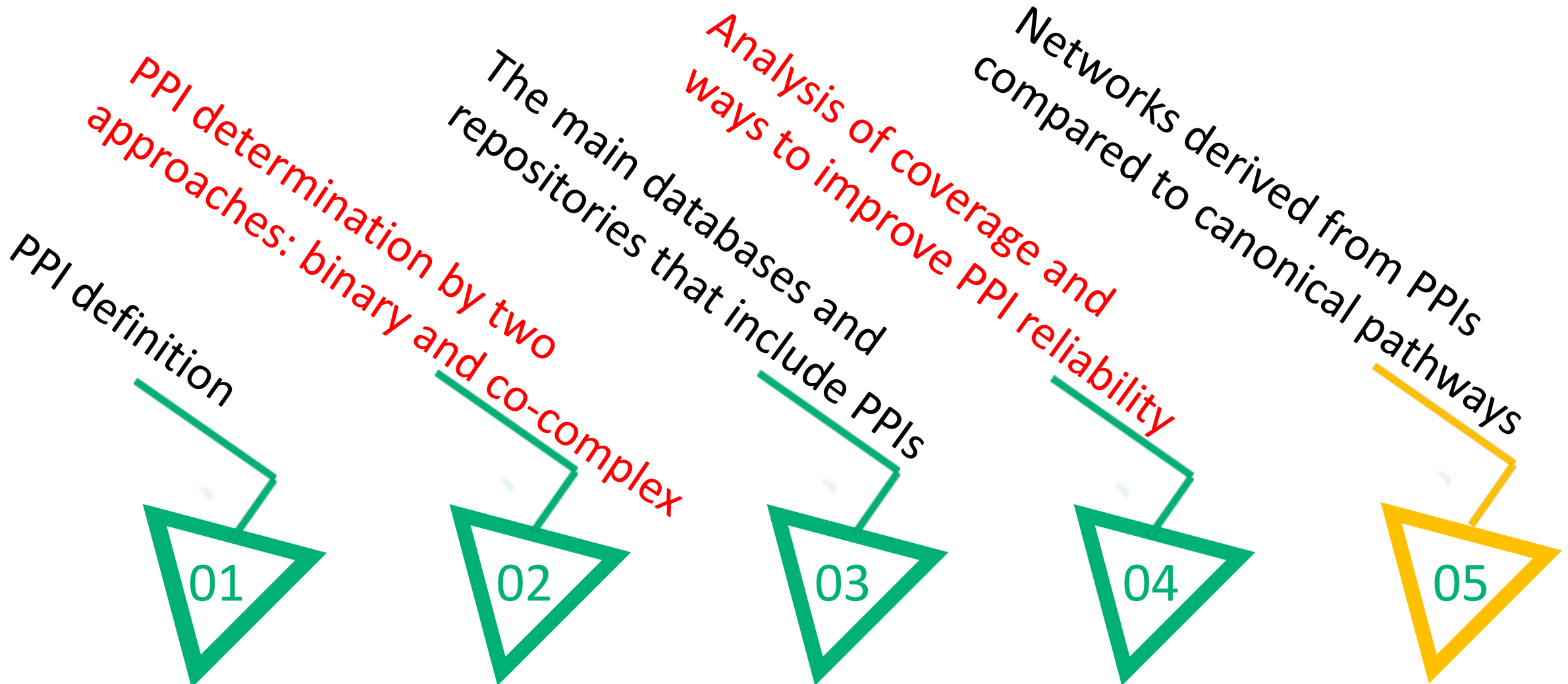
(iii) a subset of *ppihs_all* with only the interactions corresponding to proteins that work together in the same KEGG biological pathway: 7,693 interactions, called *ppihsxKEGG* (Figure 2E)

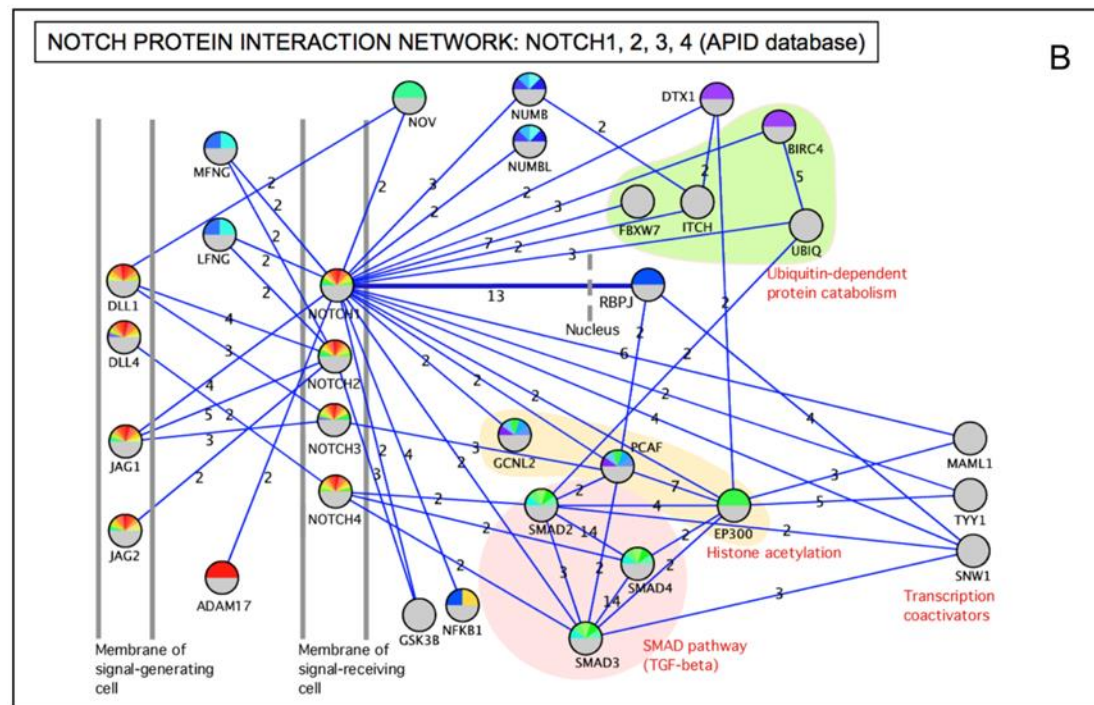
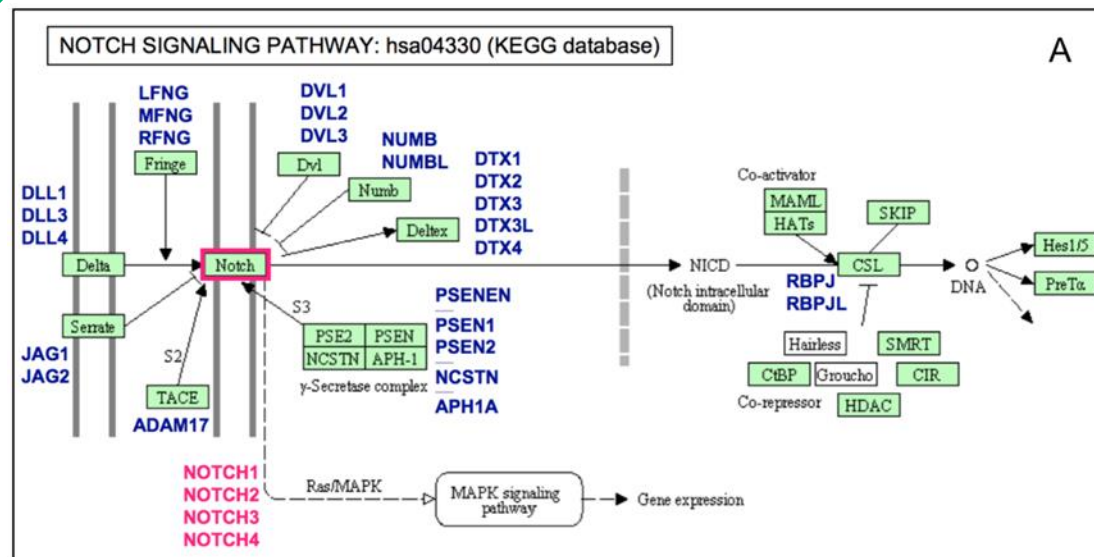
Coverage of human PPIs with 3D structure (only proteins with Pfam)



The Venn diagrams (Figure 2C–2E) indicate that the coverage of structural data increases from 14.3% to 21.4% and 30.3%, following the increase in “stringency” of the interactome datasets. Therefore, the structural validation can help to increase reliability of PPI data, as shown by the larger percentage (21.4%) of sddis getting included in the interactome proven by two methods (*ppihsx2meth*).

CONTENTS





- ◆ The KEGG pathway representation does not distinguish the relations between the four **NOTCH paralogous proteins**, while the PPI network separates the links proven for each NOTCH paralogous protein.
- ◆ The KEGG pathway representation distinguishes the **direction and properties** of the links, while the PPI network does not include such directional information.
- ◆ The γ -secretase complex is not included in the PPI network, while the interaction of NOTCH with the SMAD pathway is not present in the KEGG network.

Looking forward

Two main challenges remain for the field and for database providers:

(i) a better **filtering of false positives** in PPI collections and

(ii) an adequate distinction of the biological context that specifies and determines the existence or not of a given PPI at a given biological situation.

Main Reference

- A. Panchenko, T. Przytycka (eds.), Protein-protein Interactions and Networks, DOI: 10.1007/978-1-84800-125-1_1, © Springer-Verlag London Limited 2008
- Hakes L, Robertson DL, Oliver SG, Lovell SC (2007) Protein interactions from complexes: a structural perspective. *Comp Funct Genomics*.49356 p
- Lehninger Principles of Biochemistry, 6th edition, 2013

Q&A

