Master Module
Proteinbiochemistry and Bioinformatics
December 2023

Session: Protein interaction networks

3. Resources for protein interactions

How can I use protein interaction data in biological research?

What is the function of my gene of interest?

Is the protein of my interest part of a protein complex?

Can I find new protein complexes?

I found 20 genes in my screen that rescued phenotype X:

- do these genes work in the same biological process?
- are these genes part of the same protein complex?

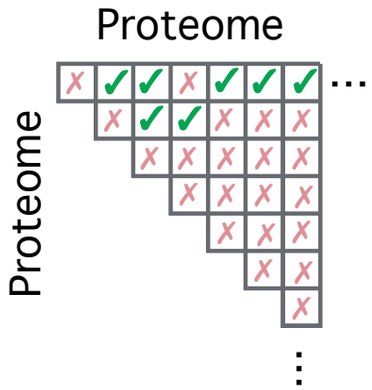
-> do these proteins (tend to) interact with each other?

My protein has many interaction partners,
does it mean that it is of functional importance?

Resources for protein interactions

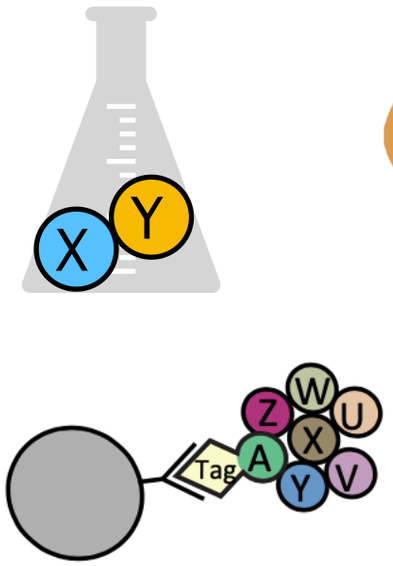
Literature curation

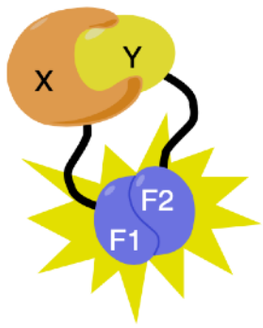
Interactome mapping



Prediction

Literature curation





BioGRID 3.5

IntAct

Which information for a published interaction should be curated?

- organism of interaction partners
- publication
- method (classification of methods?)
- full length proteins or fragments?
- K_D
- fusion constructs used
- identification of proteins (MS, sequencing)
- cellular system (yeast, cell line)
- mutations

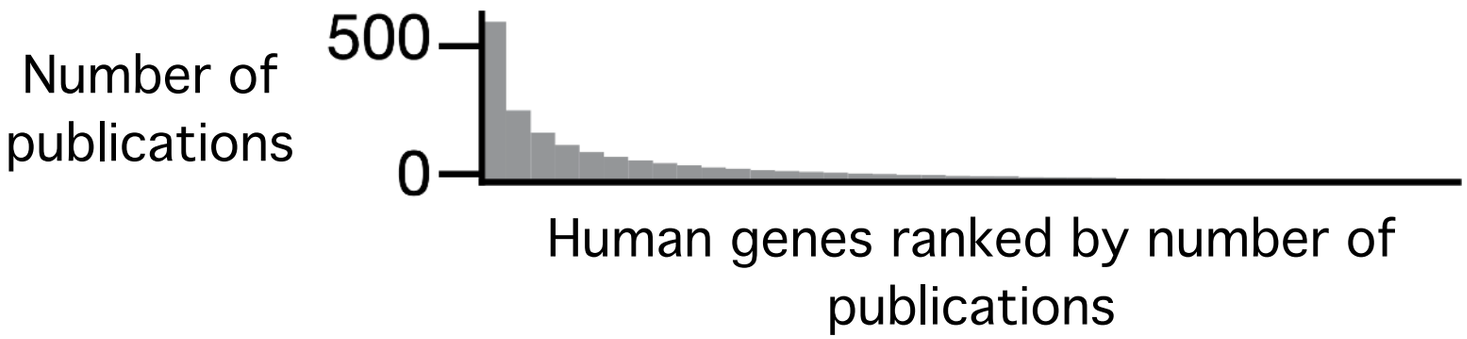
IMEx consortium to standardize curation efforts

IntAct

Orchard et al *Nature Methods* 2012

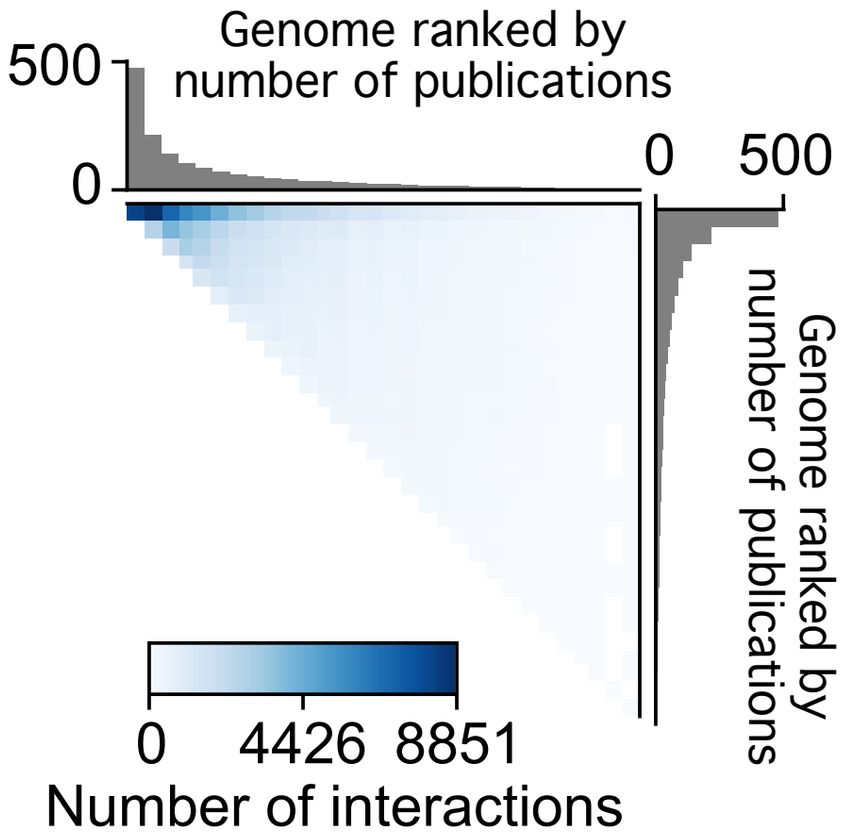
Literature curation

Are human genes/proteins equally well studied?



No!

What does this mean for availability of protein interactions?



Rolland et al *Cell* 2014

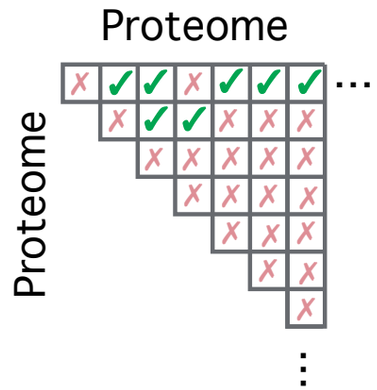
Key facts

- quite comprehensive
- mix of different interaction types
- biased towards well-studied genes

Resources for protein interactions

Literature curation

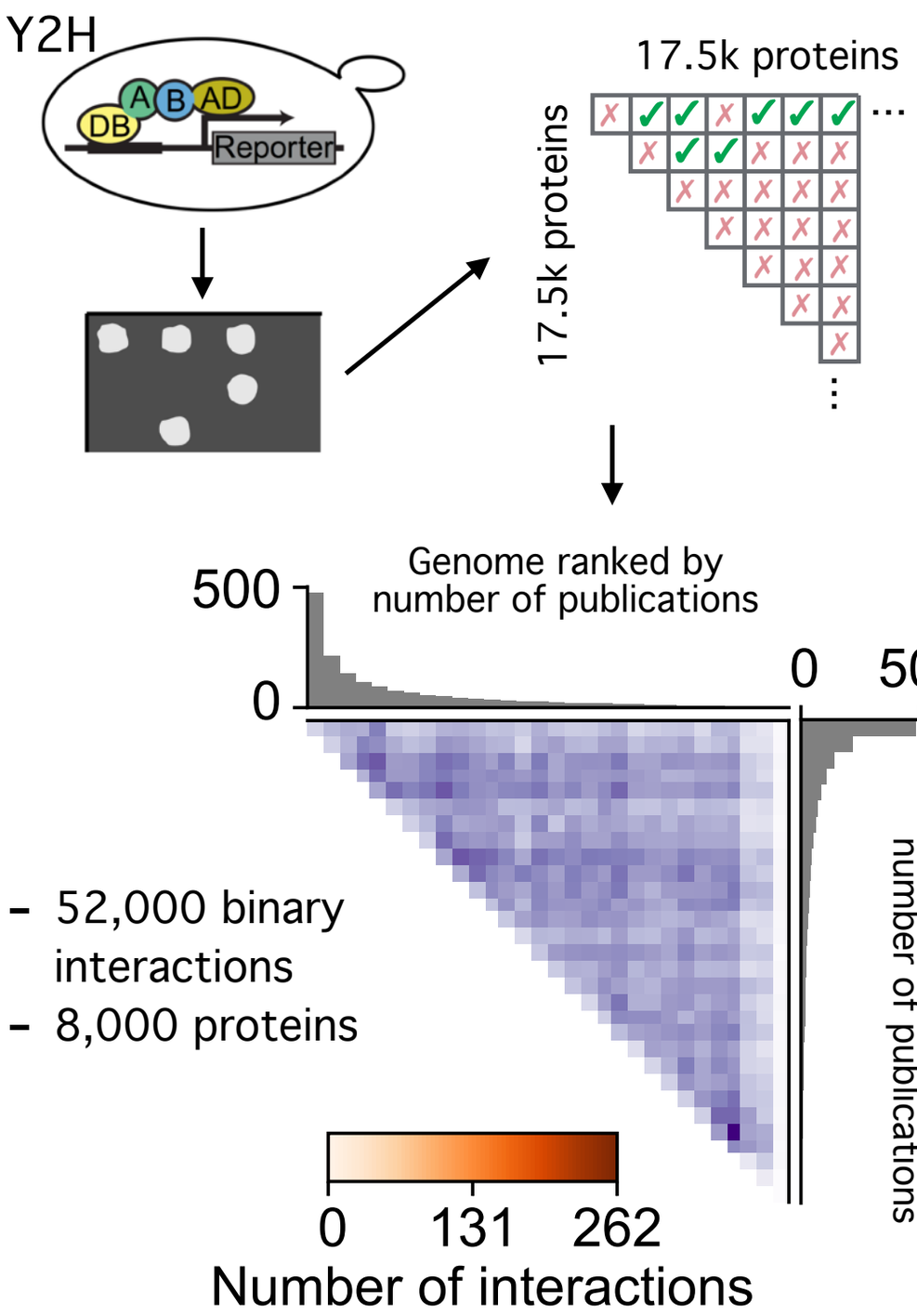
Interactome mapping



Prediction

Systematic protein interactome mapping

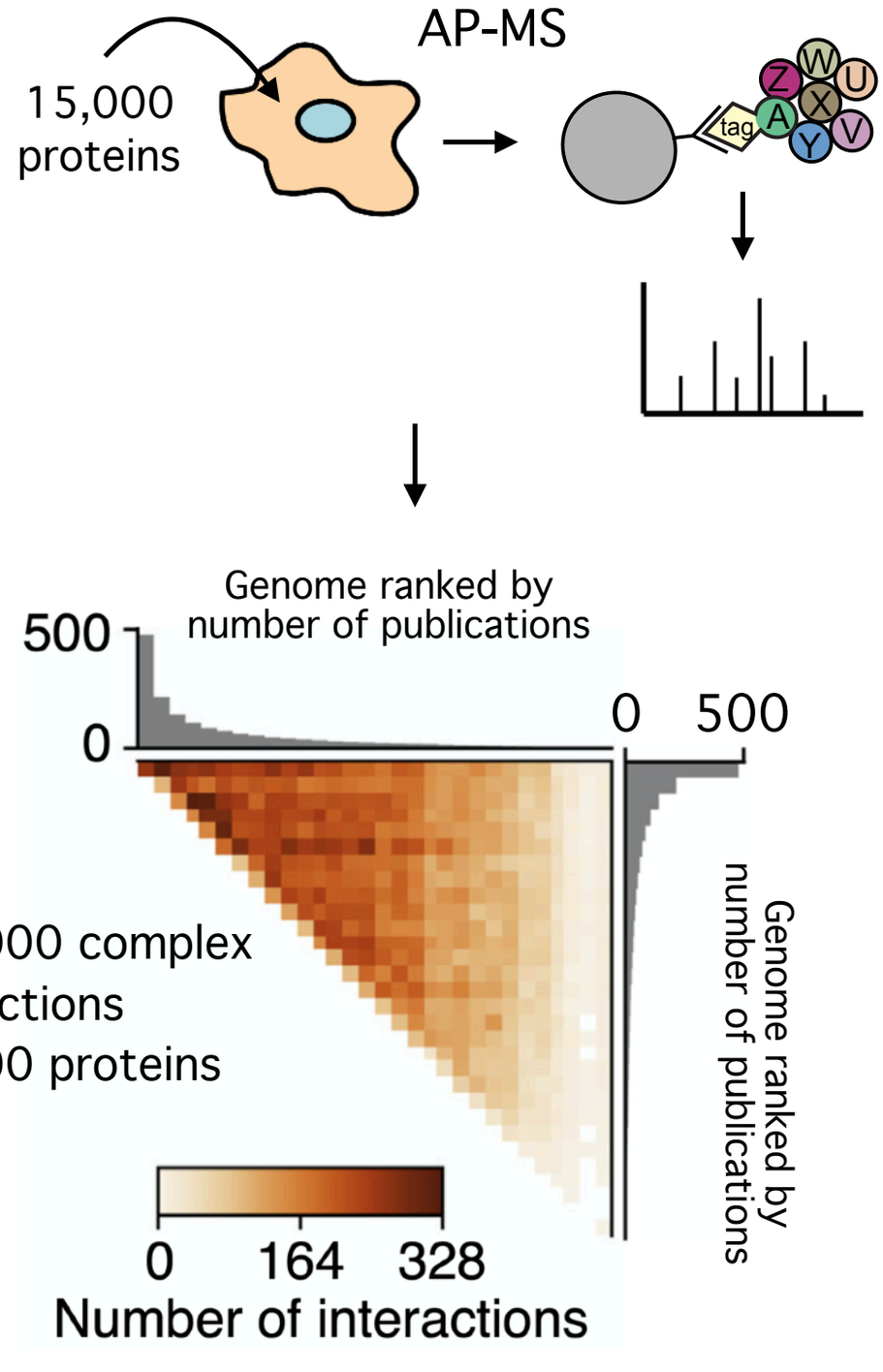
HuRI (binary)



- 52,000 binary interactions
- 8,000 proteins

Luck et al *Nature* 2020

BioPlex (co-complex)



- 118,000 complex interactions
- 15,000 proteins

Huttlin et al *Cell* 2021

Systematic protein interaction mapping

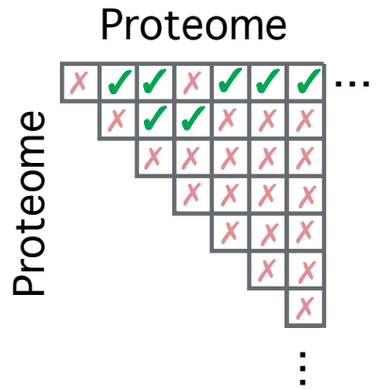
Key facts

- systematic -> not biased towards highly studied genes
- highly controlled experiments
- well documented
- not as comprehensive as curated protein interaction resources

Resources for protein interactions

Literature curation

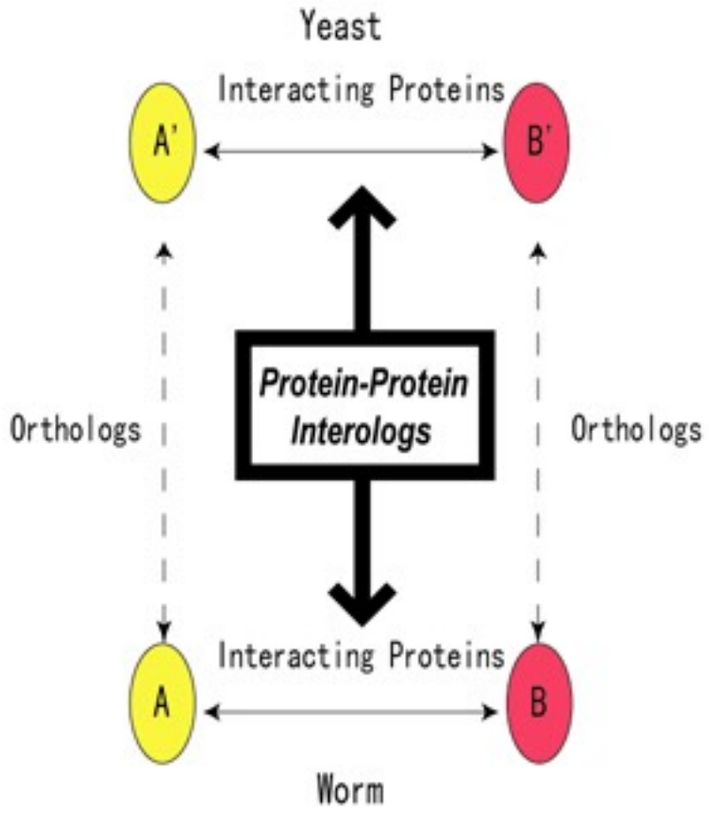
Systematic mapping



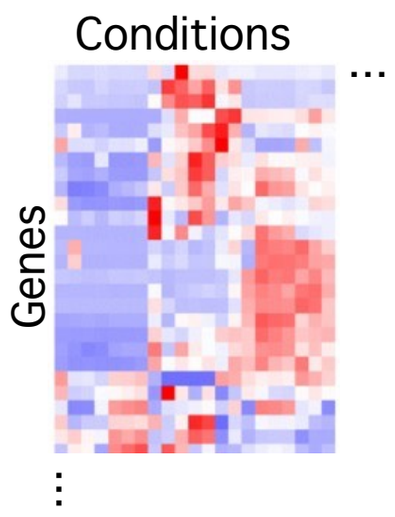
Prediction

Prediction of protein interactions

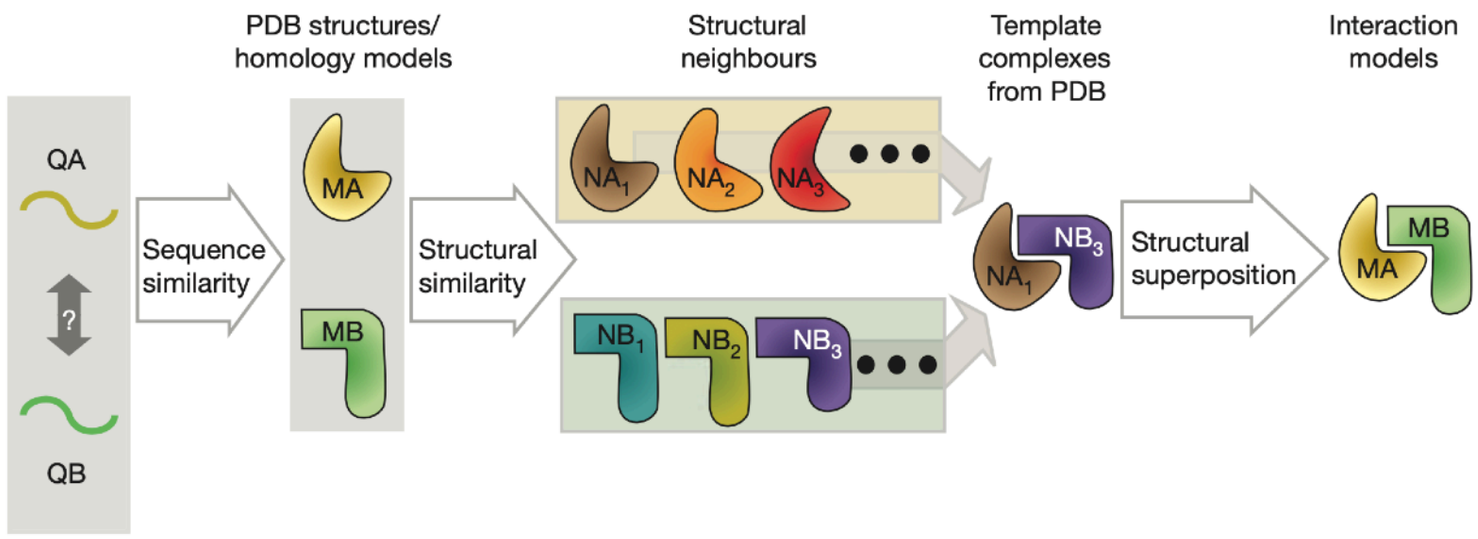
- Identification of interologs



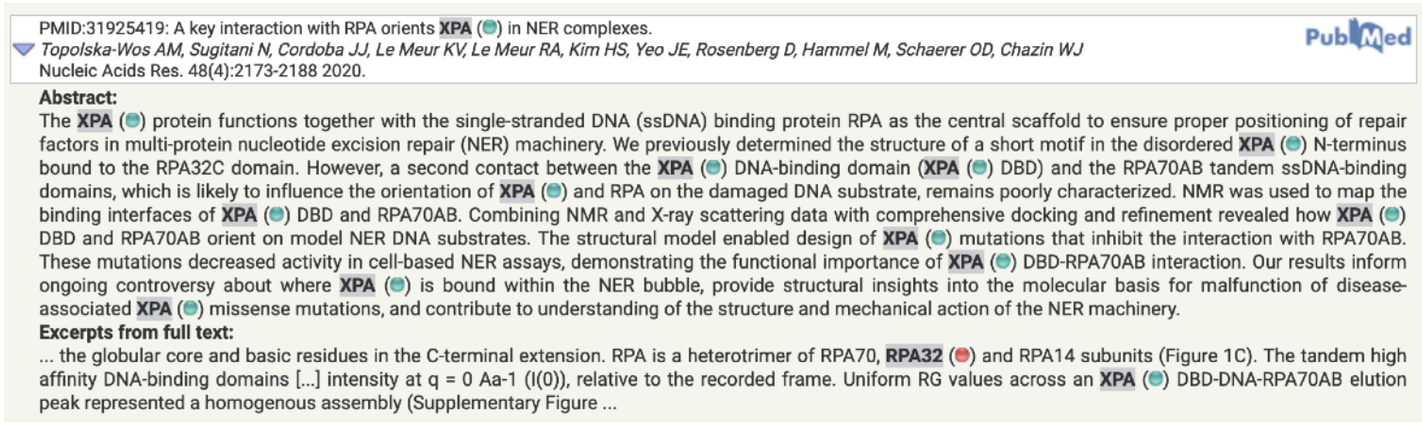
- Co-evolution, co-regulation, co-occurrence



- Structure-based modeling



- Textmining



PMID:31925419: A key interaction with RPA orients **XPA** (🟢) in NER complexes.
Topolska-Wos AM, Sugitani N, Cordoba JJ, Le Meur KV, Le Meur RA, Kim HS, Yeo JE, Rosenberg D, Hammel M, Schaerer OD, Chazin WJ
Nucleic Acids Res. 48(4):2173-2188 2020.

Abstract:
The **XPA** (🟢) protein functions together with the single-stranded DNA (ssDNA) binding protein RPA as the central scaffold to ensure proper positioning of repair factors in multi-protein nucleotide excision repair (NER) machinery. We previously determined the structure of a short motif in the disordered **XPA** (🟢) N-terminus bound to the RPA32C domain. However, a second contact between the **XPA** (🟢) DNA-binding domain (**XPA** (🟢) DBD) and the RPA70AB tandem ssDNA-binding domains, which is likely to influence the orientation of **XPA** (🟢) and RPA on the damaged DNA substrate, remains poorly characterized. NMR was used to map the binding interfaces of **XPA** (🟢) DBD and RPA70AB. Combining NMR and X-ray scattering data with comprehensive docking and refinement revealed how **XPA** (🟢) DBD and RPA70AB orient on model NER DNA substrates. The structural model enabled design of **XPA** (🟢) mutations that inhibit the interaction with RPA70AB. These mutations decreased activity in cell-based NER assays, demonstrating the functional importance of **XPA** (🟢) DBD-RPA70AB interaction. Our results inform ongoing controversy about where **XPA** (🟢) is bound within the NER bubble, provide structural insights into the molecular basis for malfunction of disease-associated **XPA** (🟢) missense mutations, and contribute to understanding of the structure and mechanical action of the NER machinery.
Excerpts from full text:
... the globular core and basic residues in the C-terminal extension. RPA is a heterotrimer of RPA70, **RPA32** (🔴) and RPA14 subunits (Figure 1C). The tandem high affinity DNA-binding domains [...] intensity at q = 0 Aa-1 (I(0)), relative to the recorded frame. Uniform RG values across an **XPA** (🟢) DBD-DNA-RPA70AB elution peak represented a homogenous assembly (Supplementary Figure ...

genecensus.org, Zhang et al *Nature* 2012, Szklarczyk et al *NAR* 2023

Prediction of protein interactions

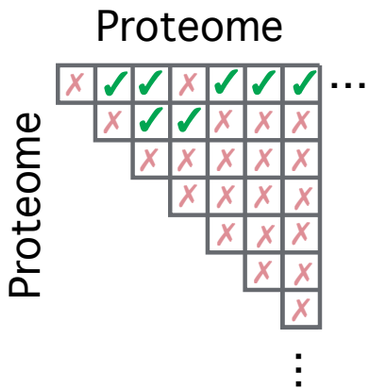
Key facts

- high false positive rate
- highly biased (orthologs, structures available)
- for some species only way to get protein interaction data

Resources for protein interactions

Literature curation

Systematic mapping



Prediction

STRING

BioGRID 3.4

IntAct

HIPPIE

IMEx

mentha
the interactome browser

CORUM

3D Interactome3D

Exercise: Explore STRING DB

30 min

1st part:

- Explore the STRING DB (string-db.org) with the help of the questions (STRING_questions.txt) and input list of proteins (STRING_input_28_genes.txt) provided
- Take notes and/or screenshots of your observations

30 min

2nd part:

Discussion of results with everyone

I found 28 genes in my screen that are likely associated with Neurodevelopmental disorders:

- do these genes work in the same biological process?
- are these genes part of the same protein complex?

-> do these proteins (tend to) interact with each other?

Exercise: Explore protein interaction databases

Take home messages

- STRING contains predicted and experimentally based protein associations -> only a small fraction corresponds to actual protein interactions
- You can filter your search results based on your question/interest -> make use of it to get a meaningful output
- STRING provides many tools to analyse and explore your network
- Make sure you understand the content of a bioinformatic resource before using it