Master Module
Proteinbiochemistry and Bioinformatics
December 2024

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,
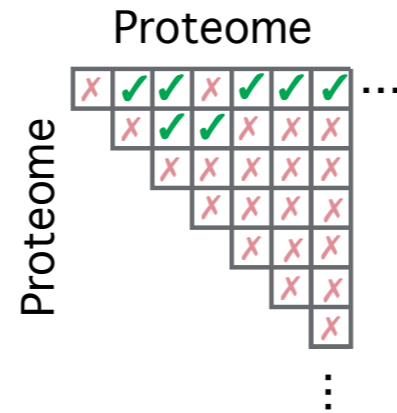does it mean that it is of functional importance?

Resources for protein interactions
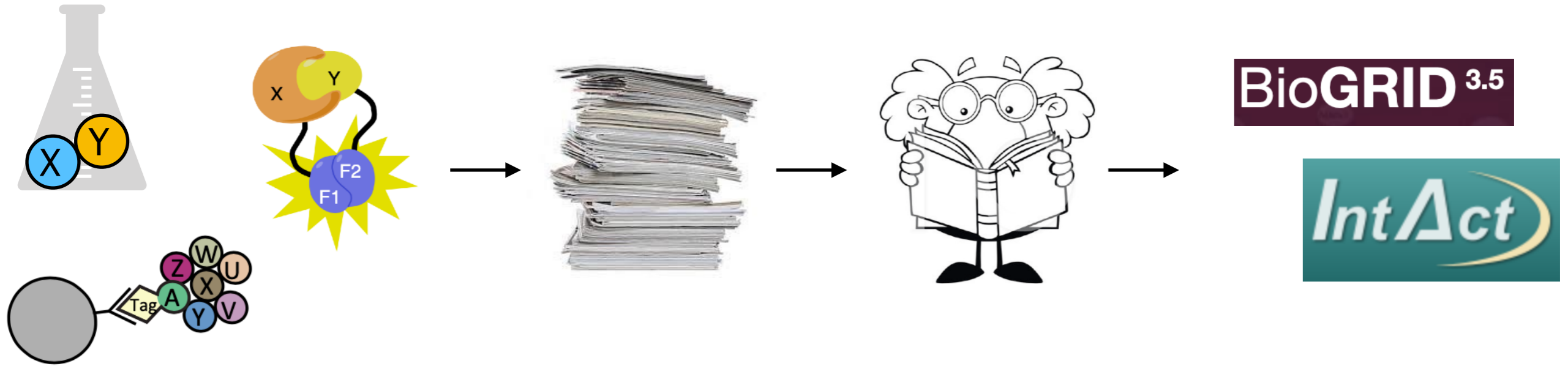
Literature curation

Interactome mapping
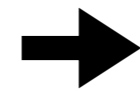
Proteome

Proteome

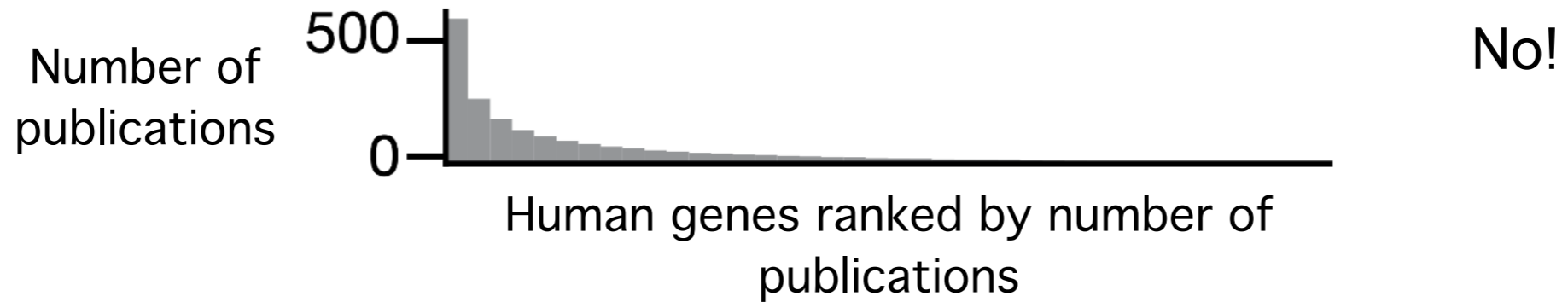## Prediction

Literature curation

Which information for a published interaction should be curated?
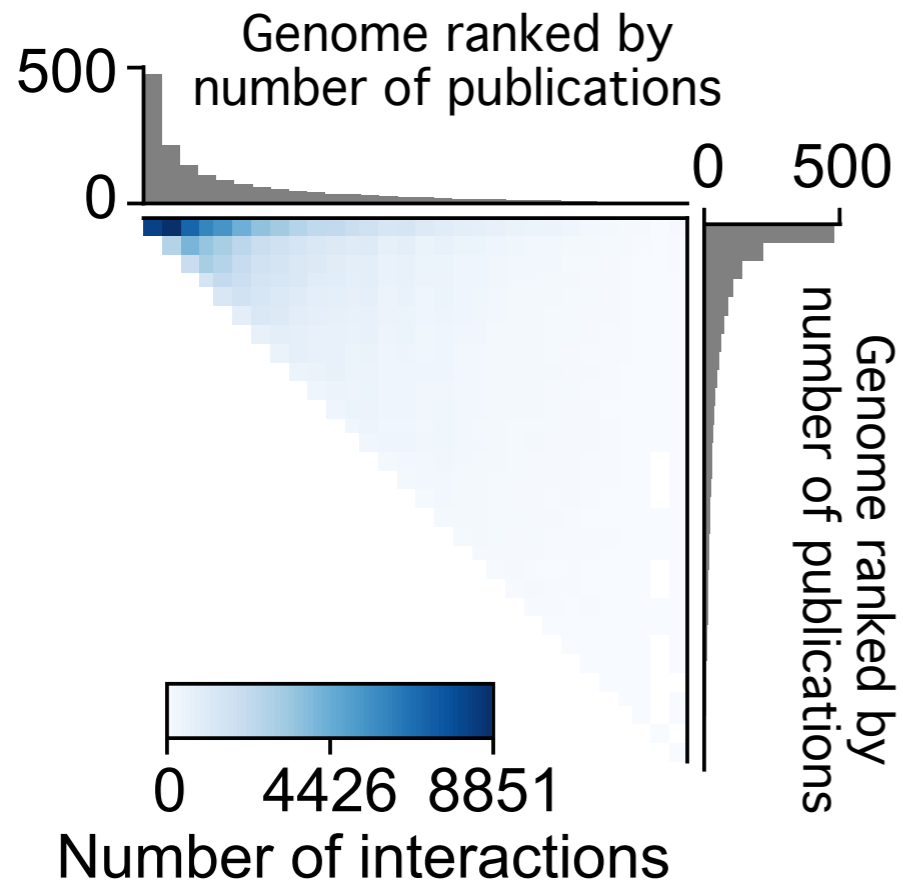
IMEx consortium to standardize curation efforts

Literature curation

Are human genes/proteins equally well studied?

No!

Number of publications

500 —

0 —

Human genes ranked by number of publications

What does this mean for availability of protein interactions?

Genome ranked by number of publications

500 —

0 —

0   500

Genome ranked by number of publications

Key facts

- quite comprehensive

- mix of different interaction types

- biased towards well-studied genes

0   4426 8851

Number of interactions

Rolland et al *Cell* 2014

Resources for protein interactions

Literature curation

Interactome mapping

Prediction

Proteome

Proteome
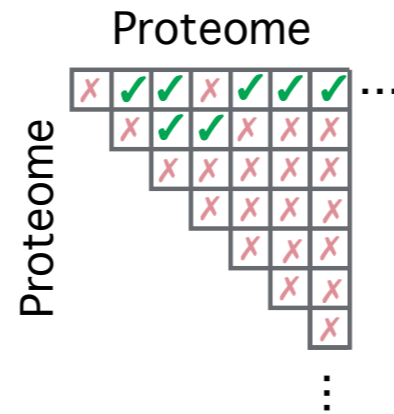
Systematic protein interactome mapping

HuRI (binary)

BioPlex (co-complex)

Y2H

17.5k proteins

17.5k proteins

AP-MS

15,000 proteins

Genome ranked by number of publications

500

0   500

0

Genome ranked by number of publications

- 52,000 binary interactions
- 8,000 proteins

0   131   262
Number of interactions

Genome ranked by number of publications

500

0   500

0

Genome ranked by number of publications

- 118,000 complex interactions
- 15,000 proteins

0   164   328
Number of interactions

Luck et al *Nature* 2020

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
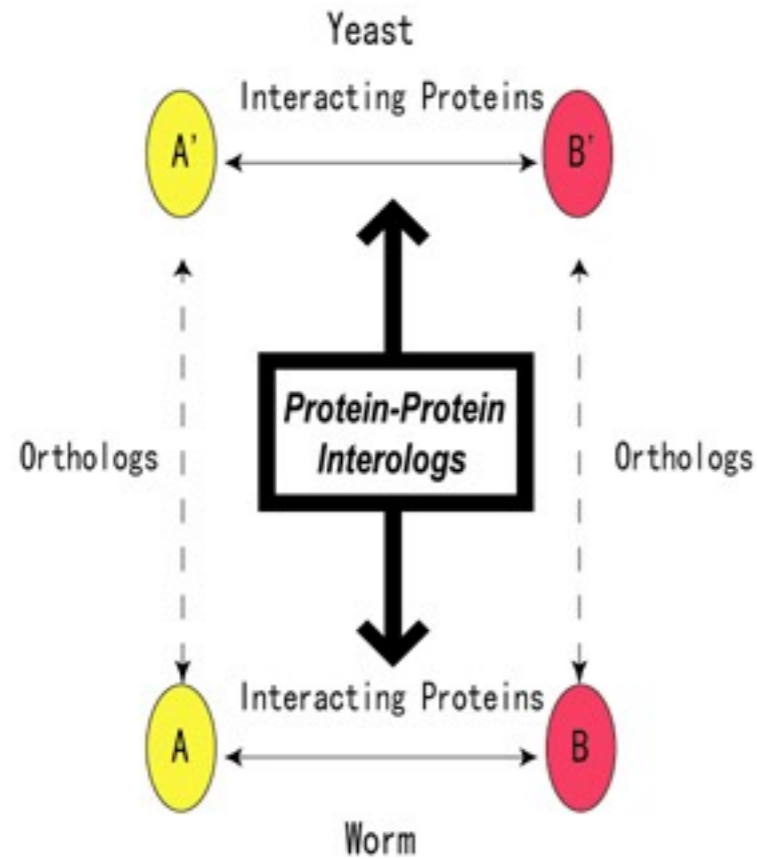
Literature curation
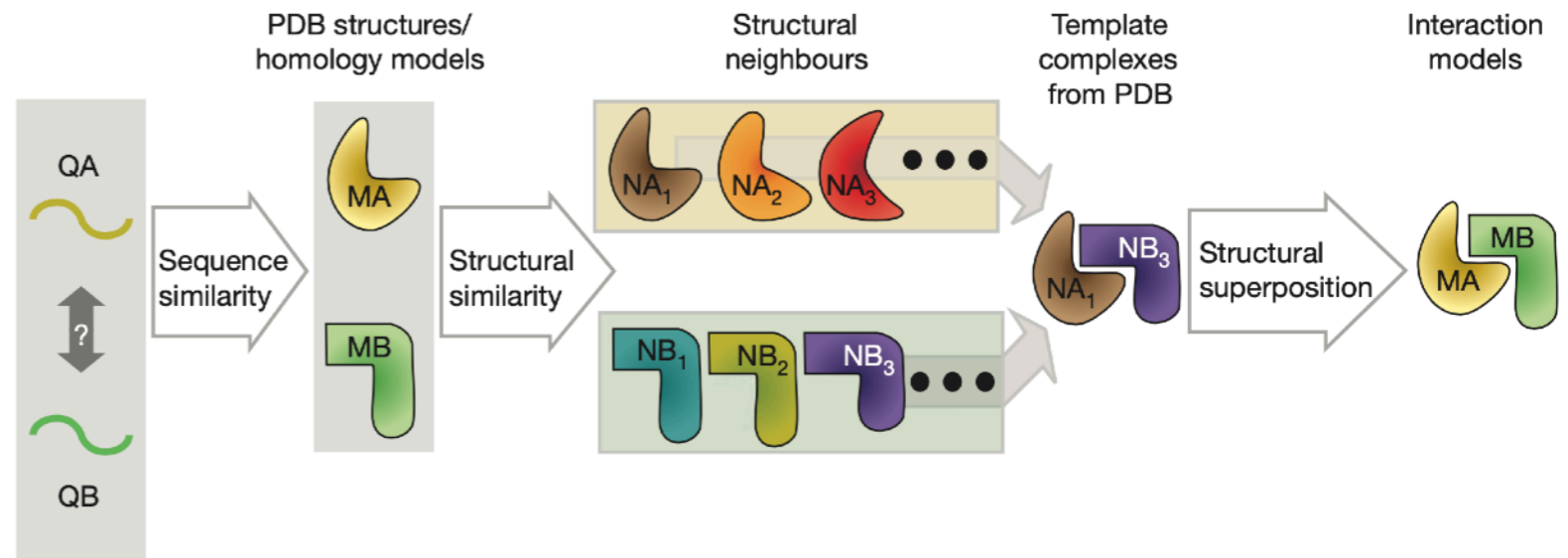
Systematic mapping

Prediction

Proteome

Proteome
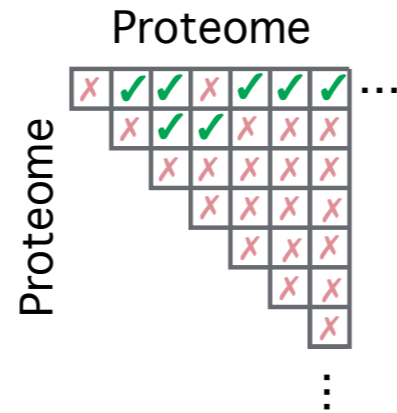
Prediction of protein interactions

- Identification of interologs

- Structure-based modeling

– Textmining

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

Prediction of protein interactions
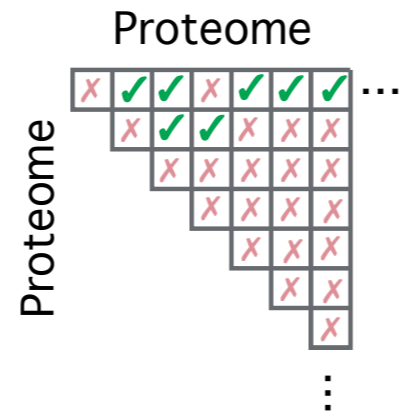
Key facts

- 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

Proteome

Proteome

Exercise: Explore STRING DB

40 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

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?

20 min

2nd part:

Discussion of results with everyone

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