

Master Module
Proteinbiochemistry and Bioinformatics
March 2022

Session: Protein interaction networks

3. Resources for protein interactions

How can I use protein interaction data in biological research?

How can I use protein interaction data in biological research?

What is the function of my gene of interest?



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?

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?



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?



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?



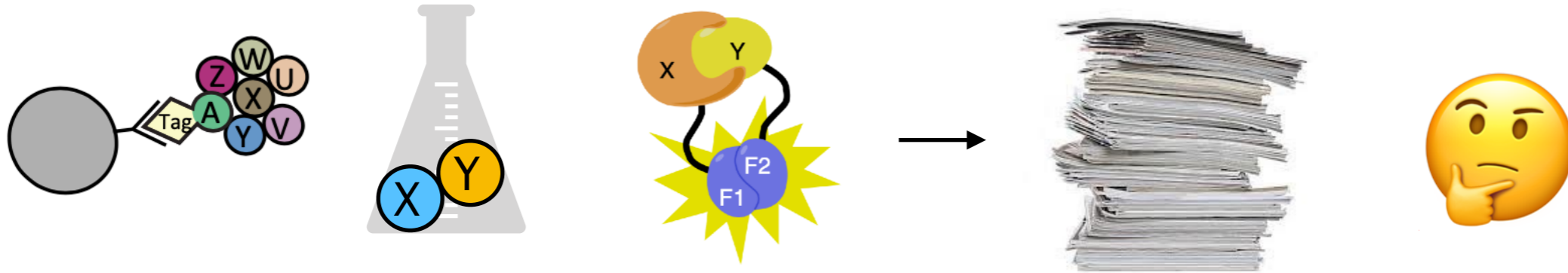
My protein has many interaction partners,
does it mean that it is of functional importance?



Where can I get protein interaction data from?

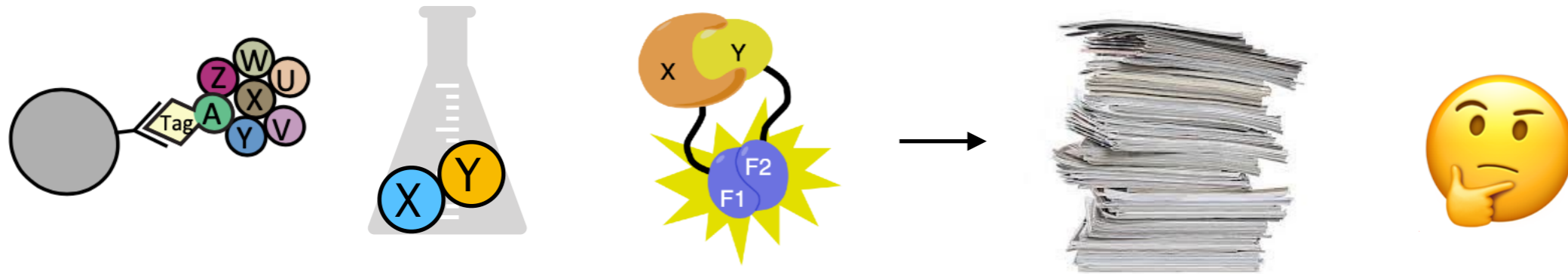
Where can I get protein interaction data from?

Literature curation?

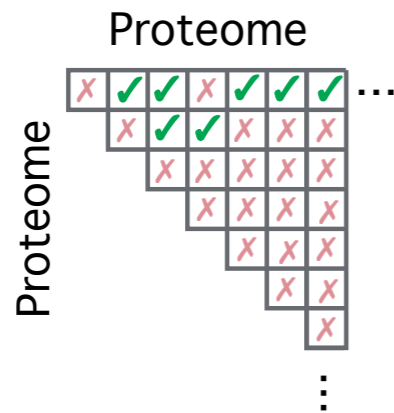


Where can I get protein interaction data from?

Literature curation?



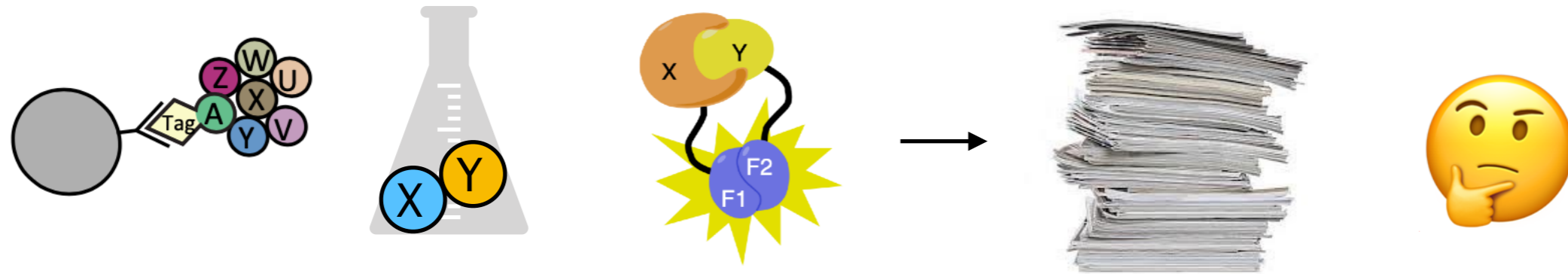
Interactome mapping?



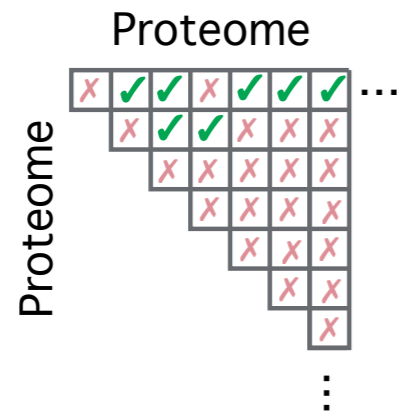
- No PCR for interactions
- Interactions are a state, not a molecule

Where can I get protein interaction data from?

Literature curation?

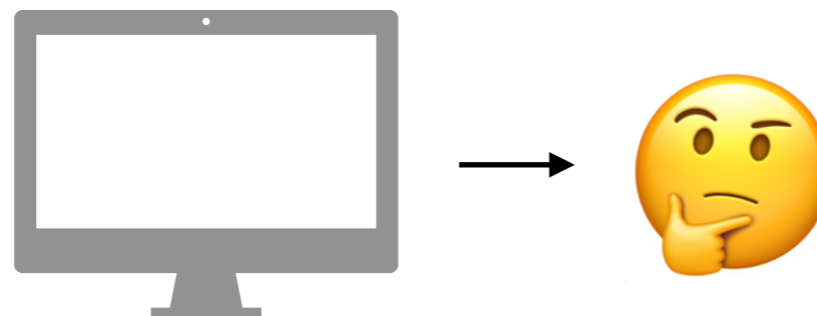


Interactome mapping?



- No PCR for interactions
- Interactions are a state, not a molecule

Prediction?

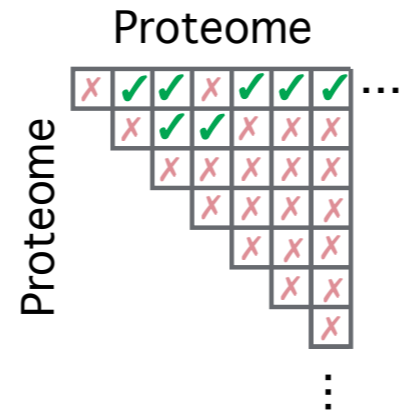


Resources for protein interactions

Literature curation



Interactome mapping



Prediction

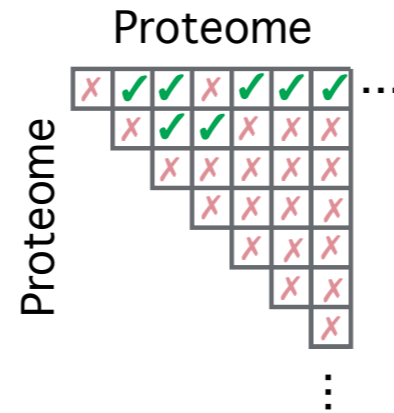


Resources for protein interactions

Literature curation



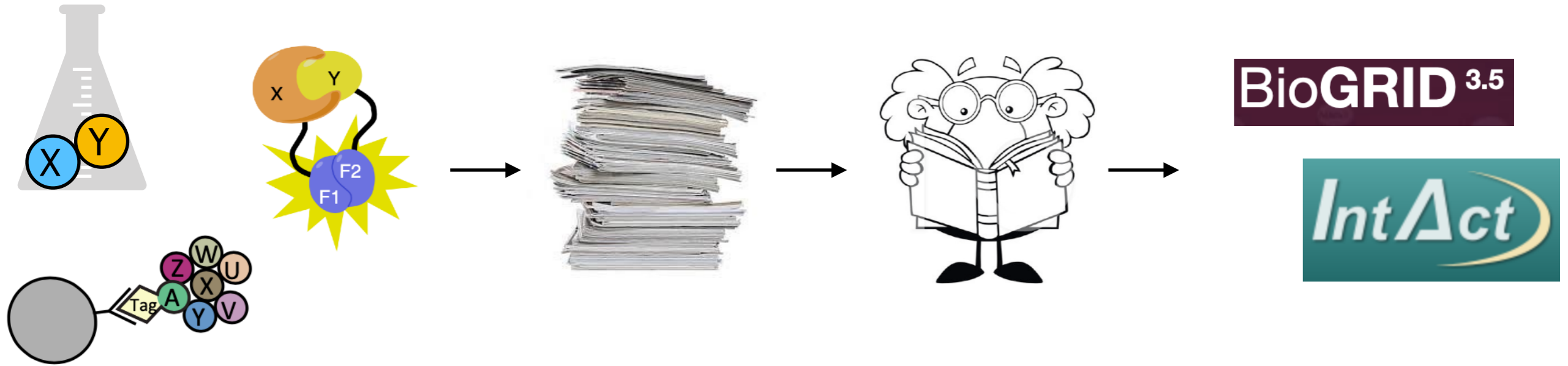
Interactome mapping



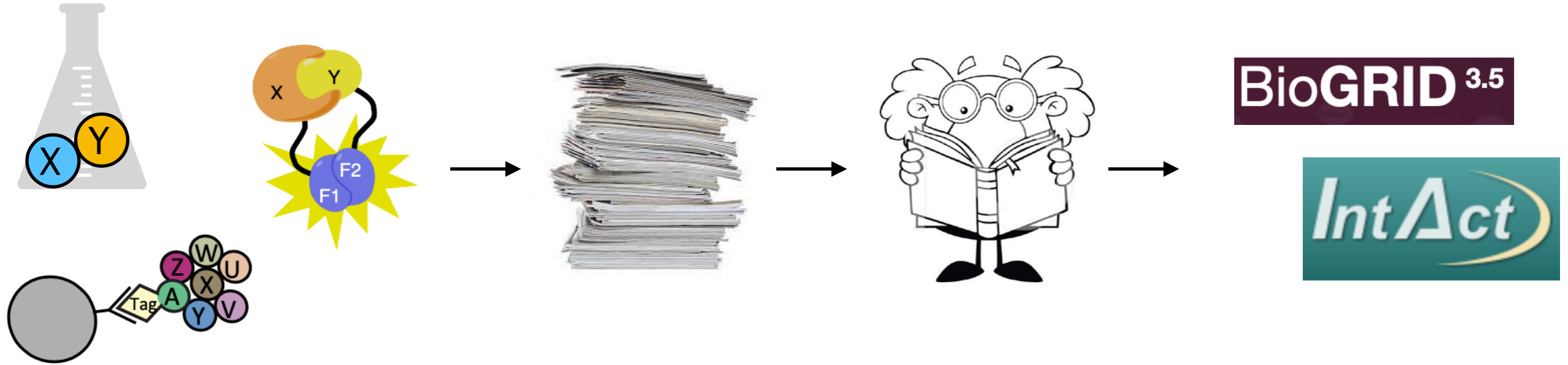
Prediction



Literature curation

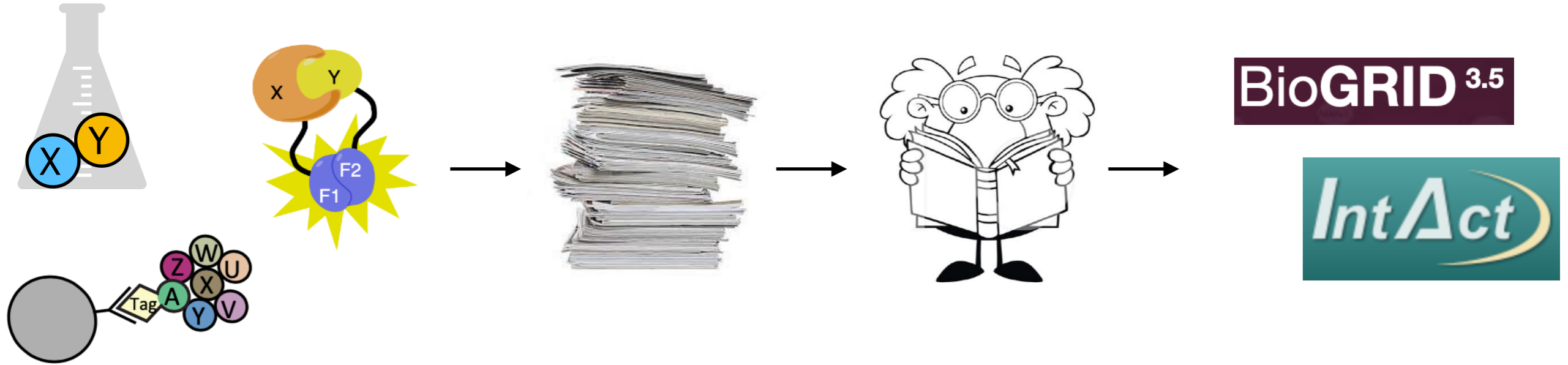


Literature curation



Which information for a published interaction should be curated?

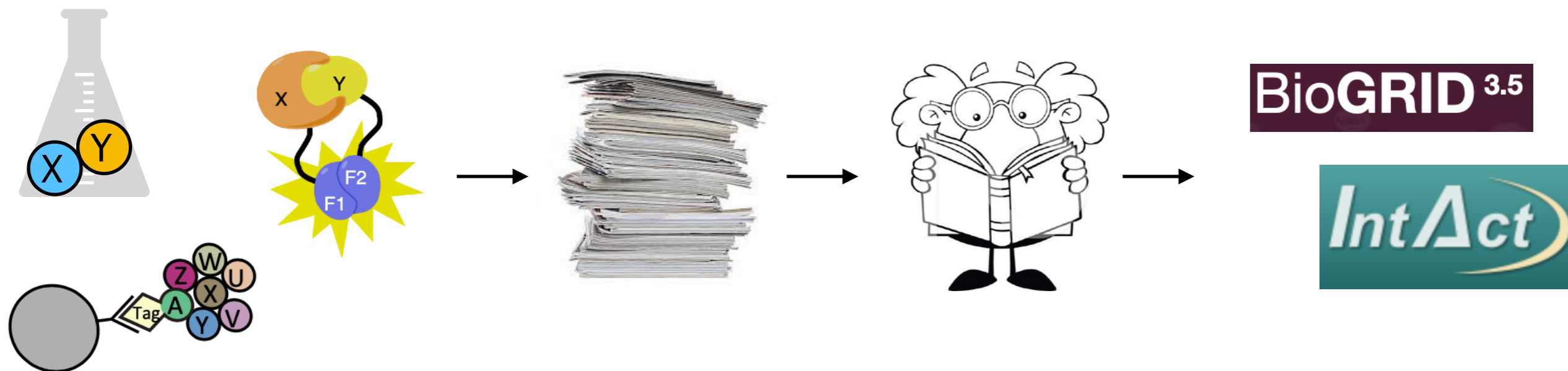
Literature curation



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

Literature curation



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



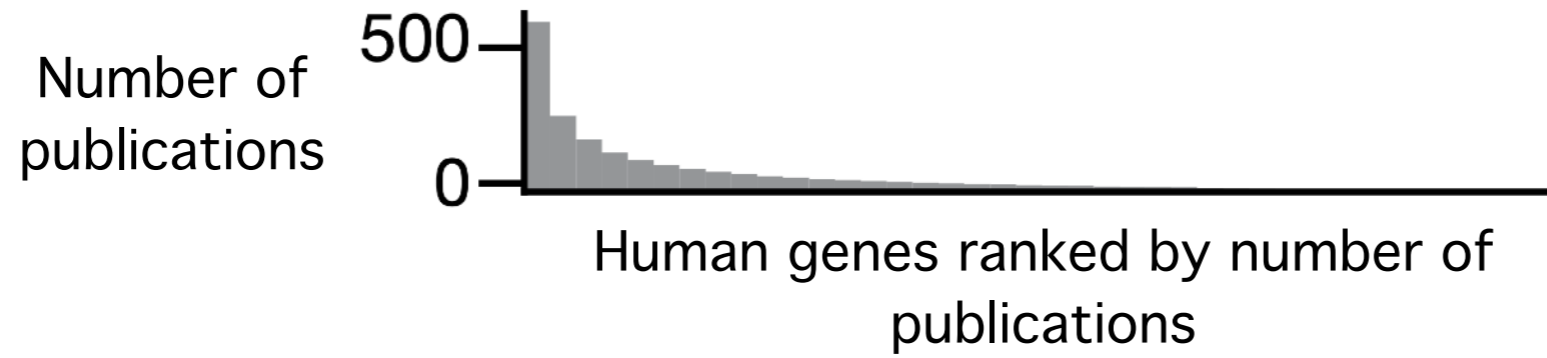
Literature curation

Literature curation

Are human genes/proteins equally well studied?

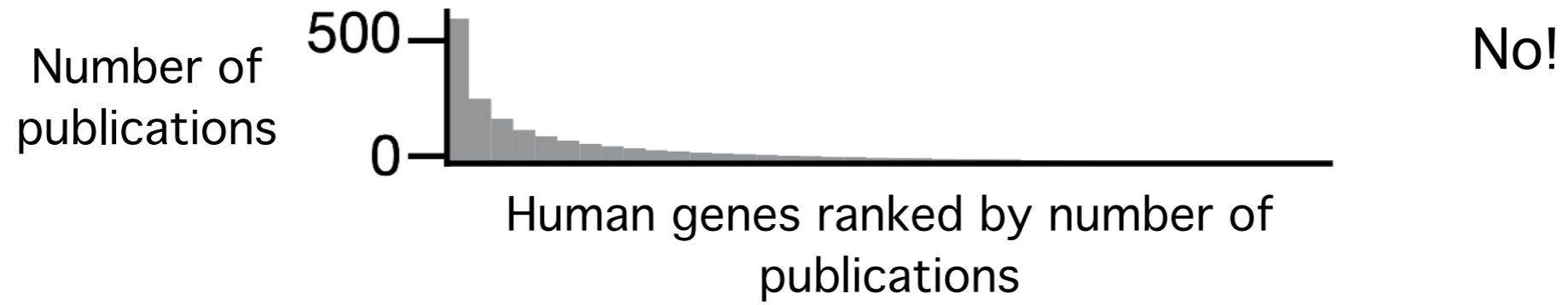
Literature curation

Are human genes/proteins equally well studied?



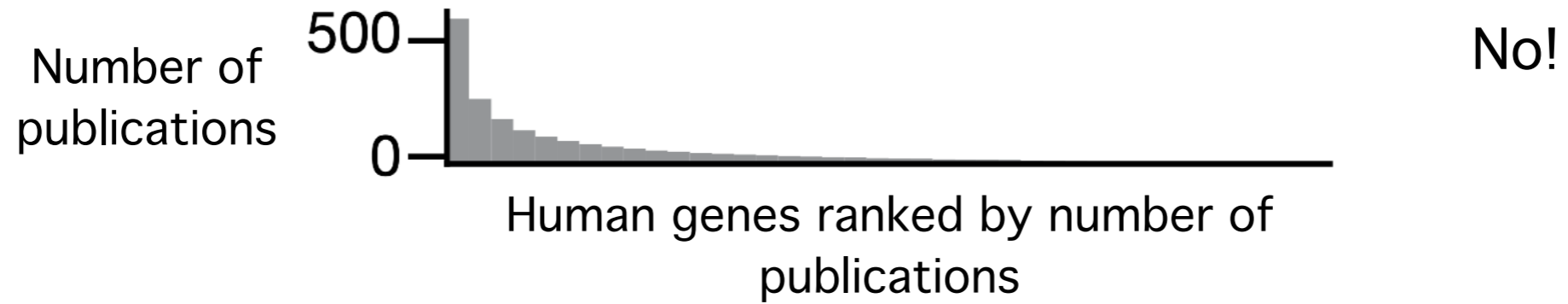
Literature curation

Are human genes/proteins equally well studied?



Literature curation

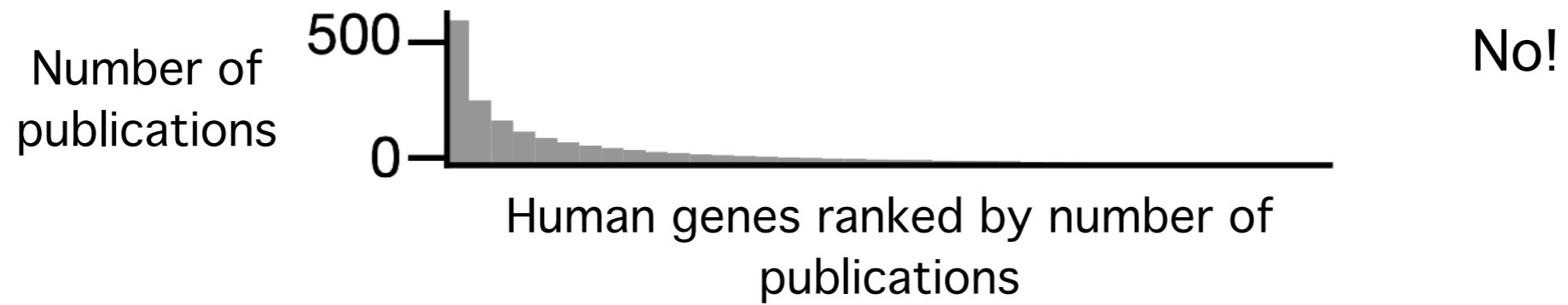
Are human genes/proteins equally well studied?



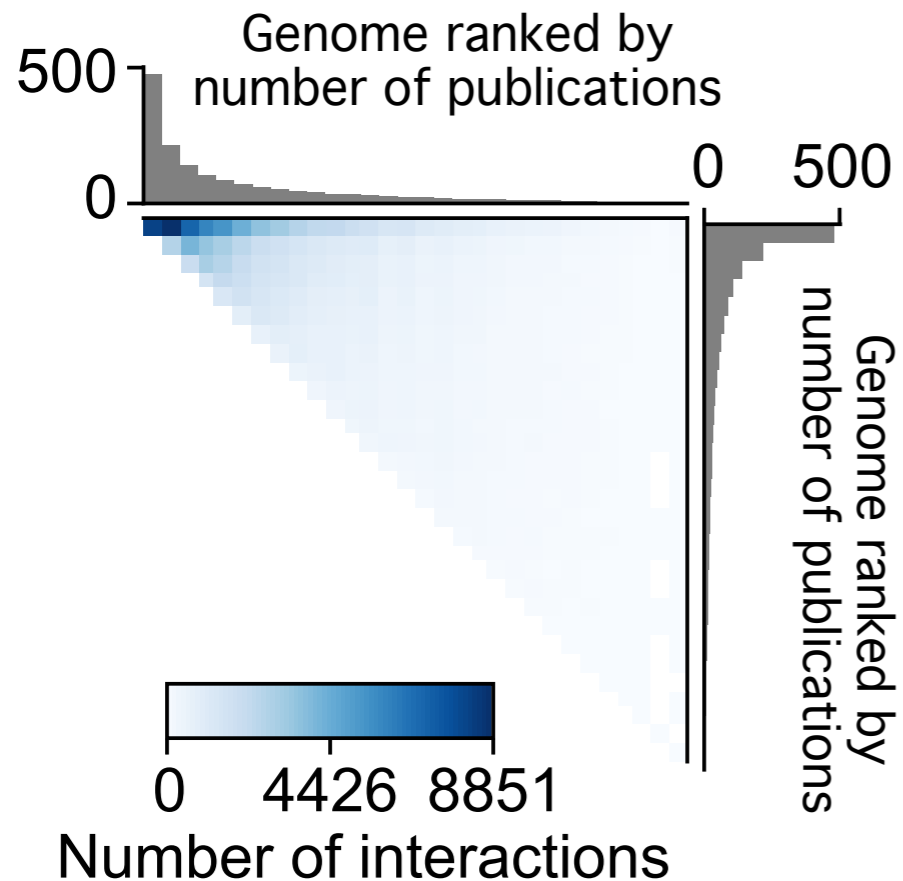
What does this mean for availability of protein interactions?

Literature curation

Are human genes/proteins equally well studied?

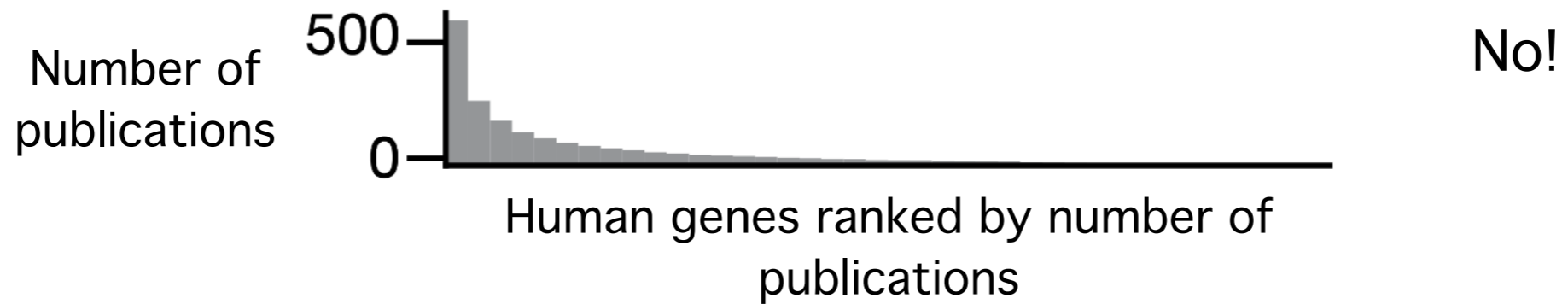


What does this mean for availability of protein interactions?

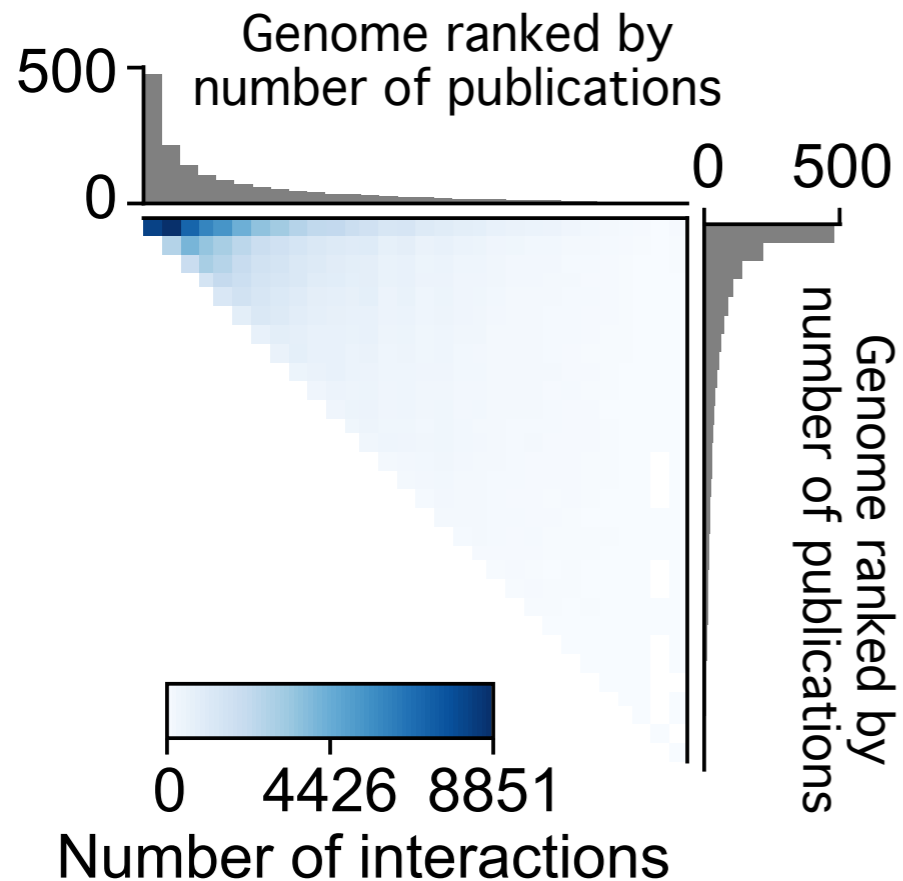


Literature curation

Are human genes/proteins equally well studied?



What does this mean for availability of protein interactions?



Key facts

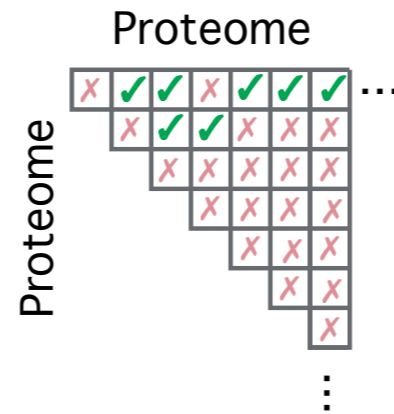
- quite comprehensive
- mix of different interaction types
- biased towards well-studied genes

Resources for protein interactions

Literature curation



Interactome mapping



Prediction



Systematic protein interactome mapping

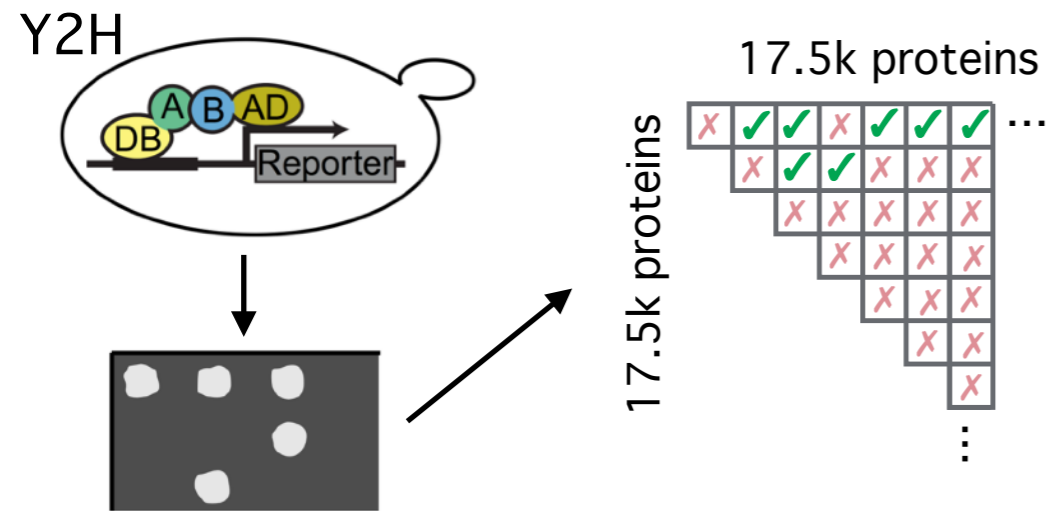
HuRI (binary)

BioPlex (co-complex)

Systematic protein interactome mapping

HuRI (binary)

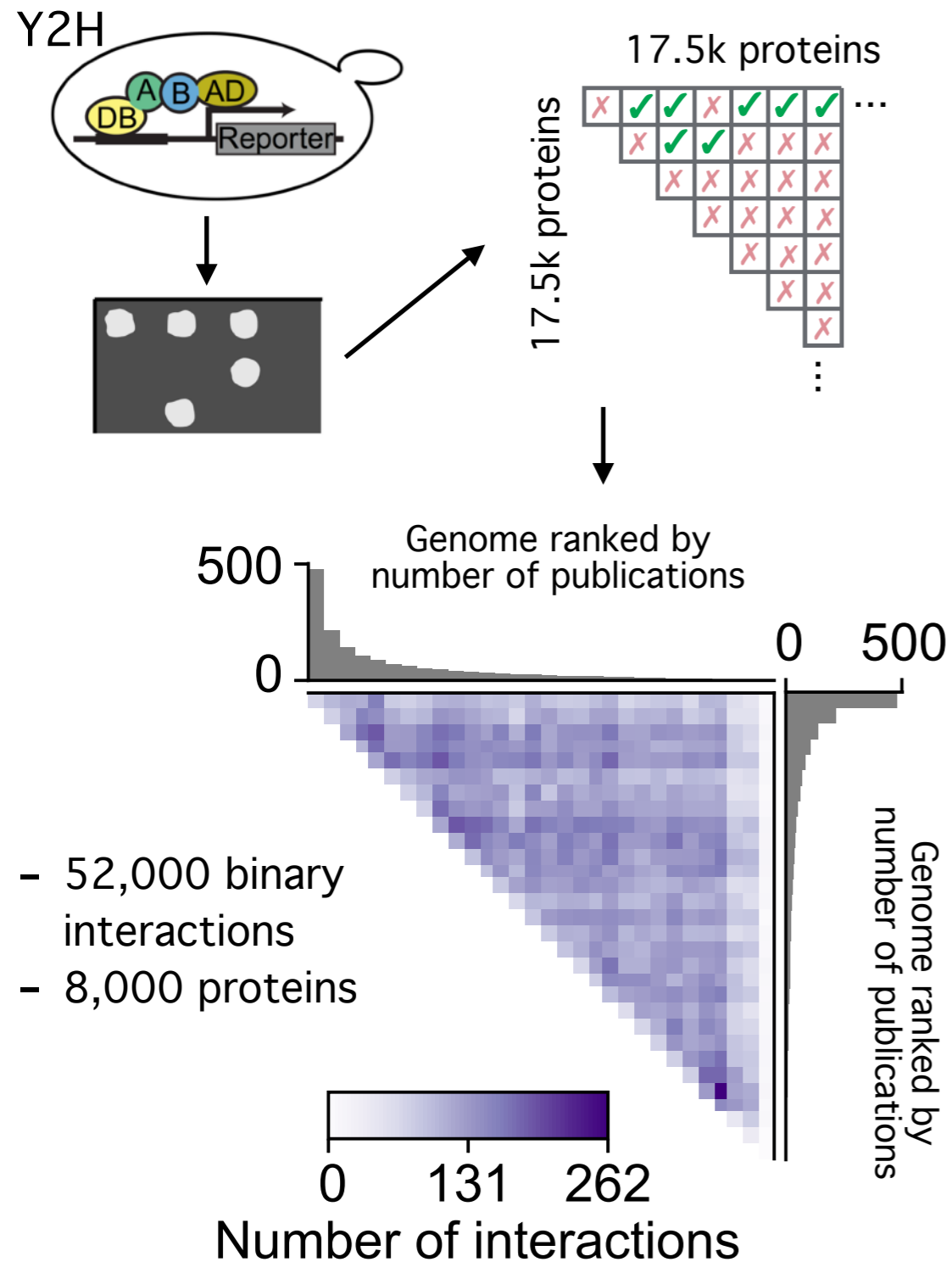
BioPlex (co-complex)



Systematic protein interactome mapping

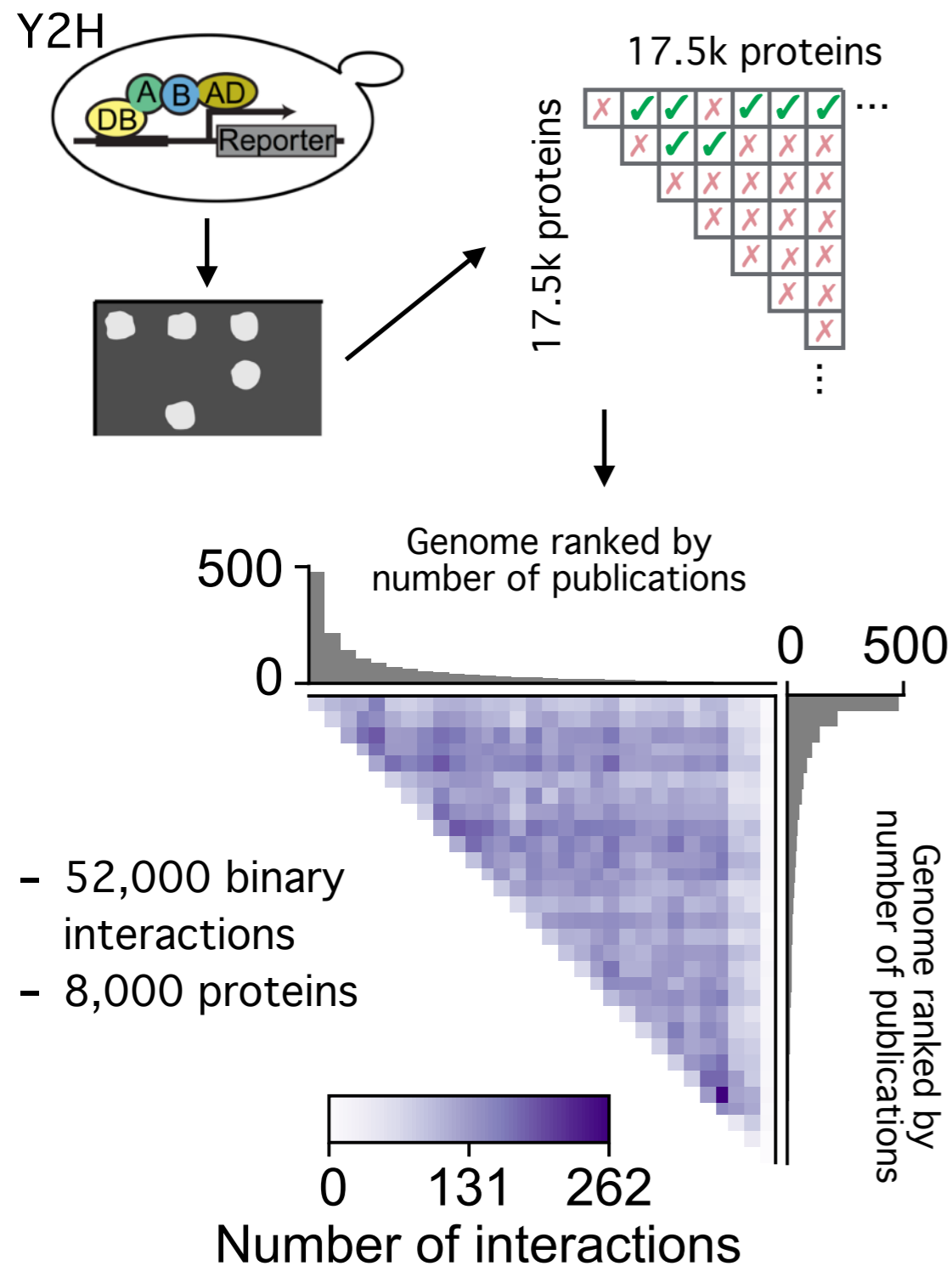
HuRI (binary)

BioPlex (co-complex)

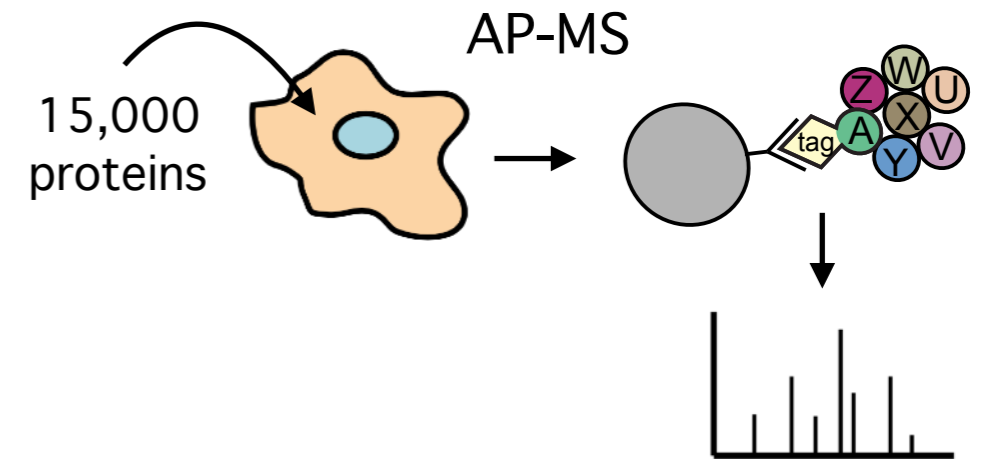


Systematic protein interactome mapping

HuRI (binary)

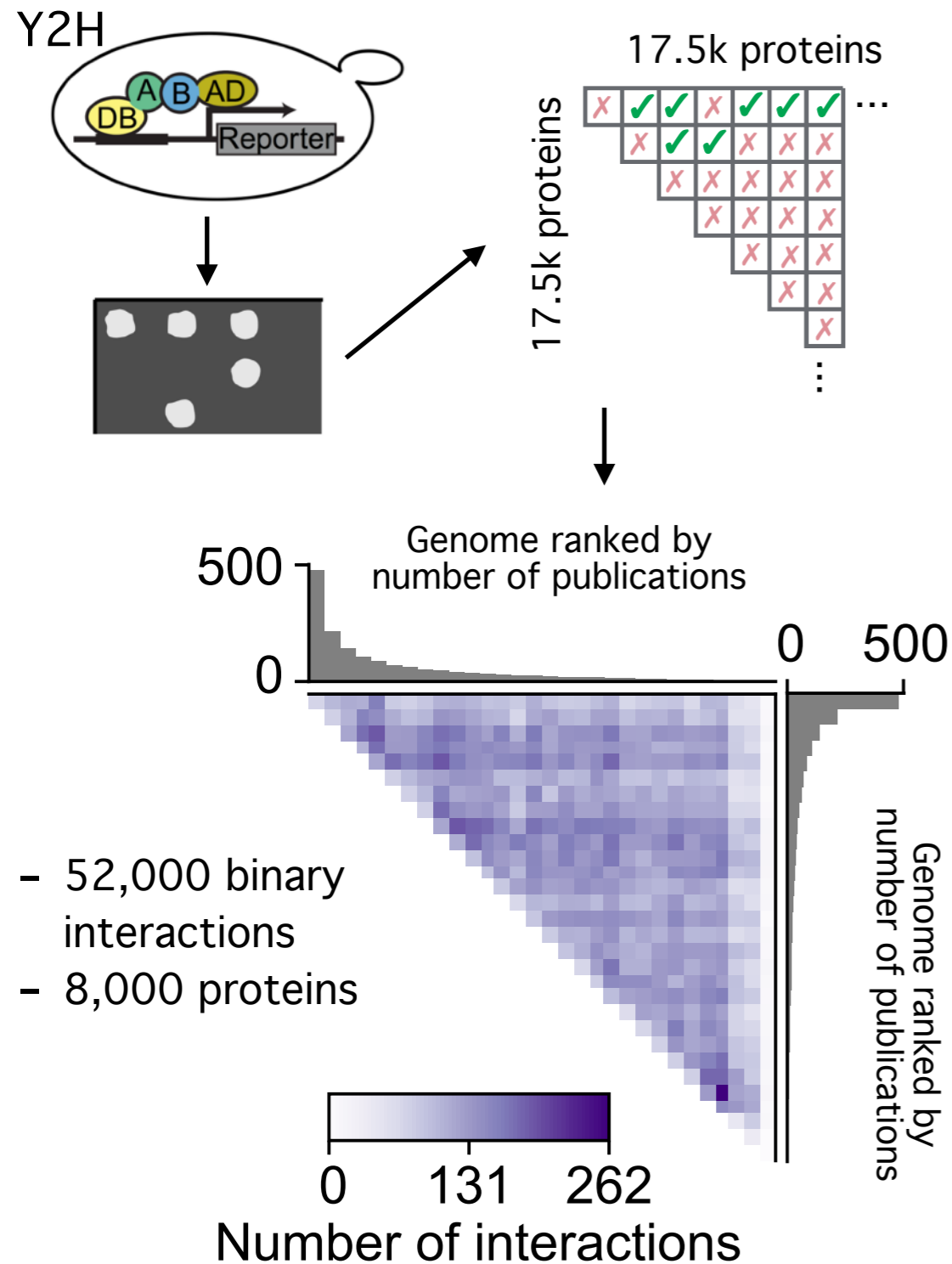


BioPlex (co-complex)

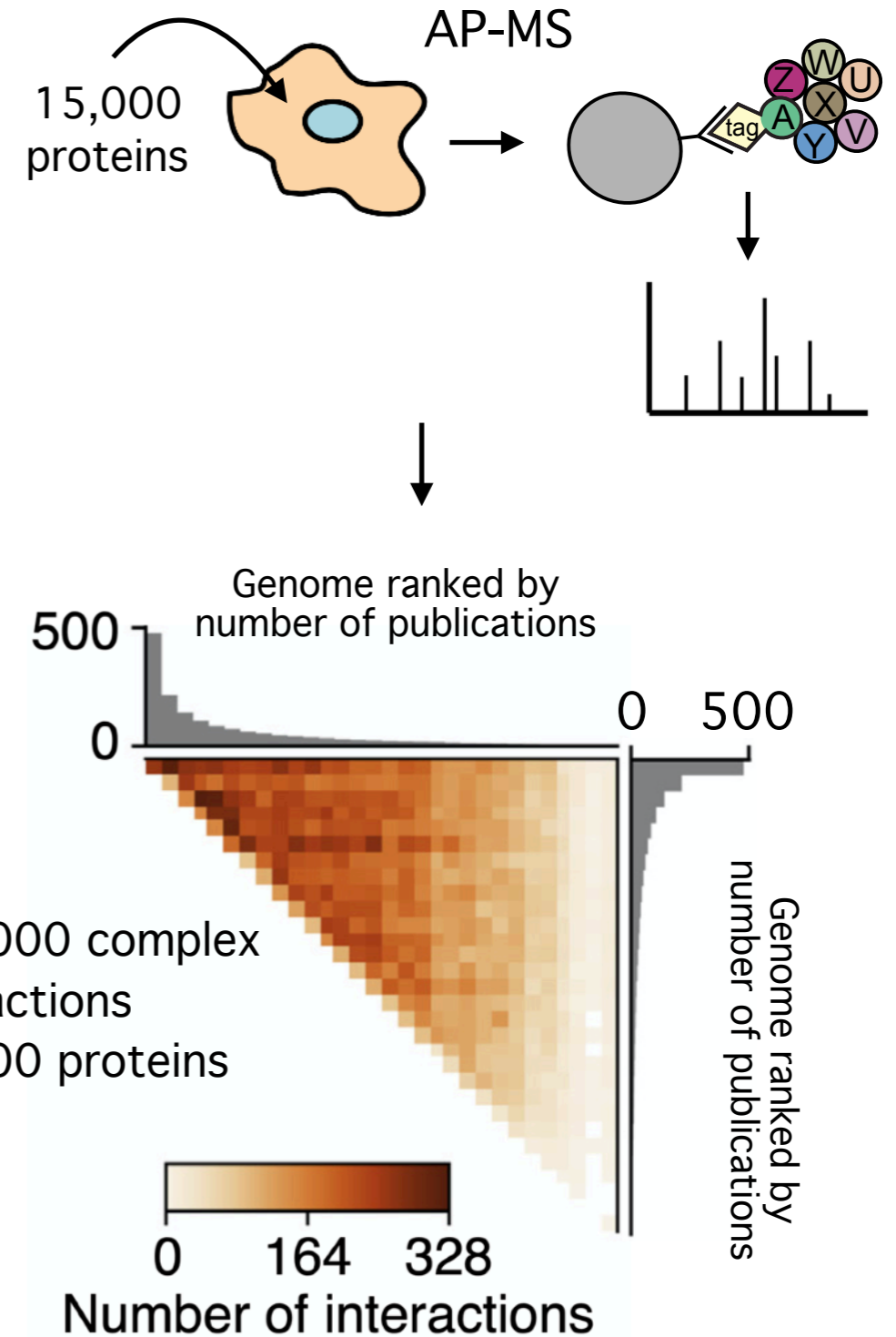


Systematic protein interactome mapping

HuRI (binary)



BioPlex (co-complex)



Systematic protein interaction mapping

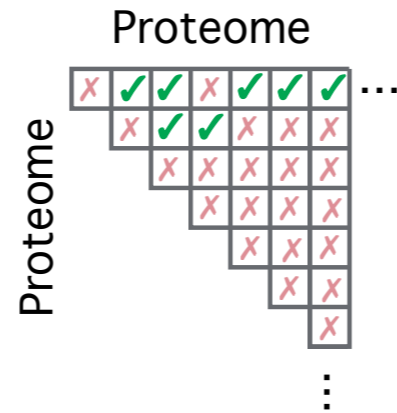
Key facts

- 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

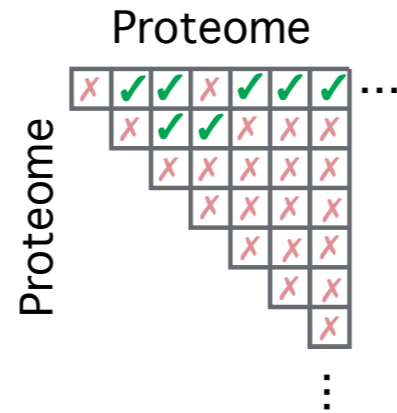


Resources for protein interactions

Literature curation



Systematic mapping



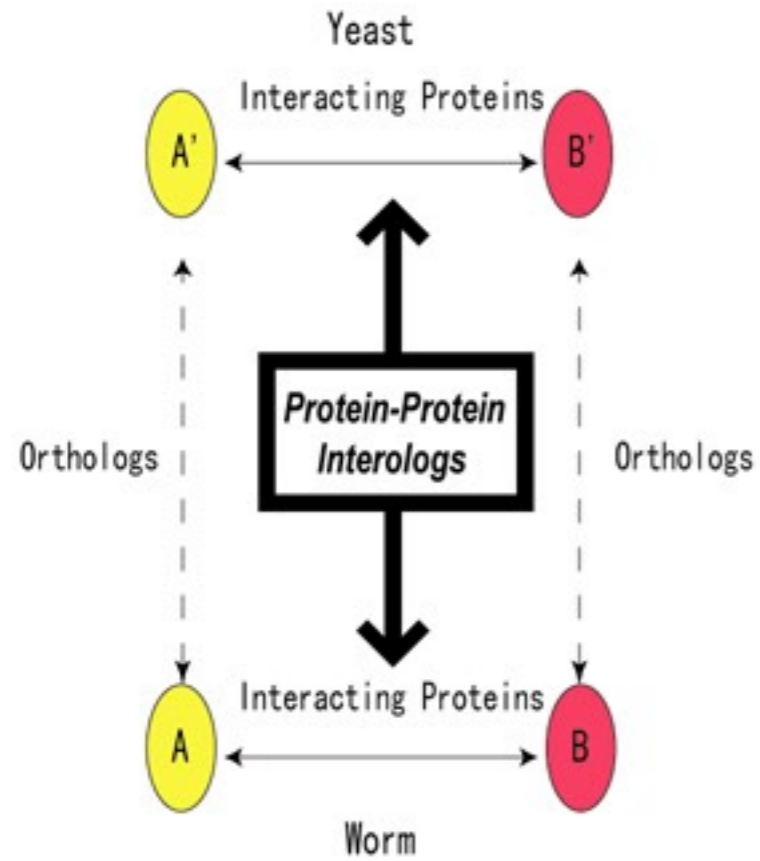
Prediction



Prediction of protein interactions

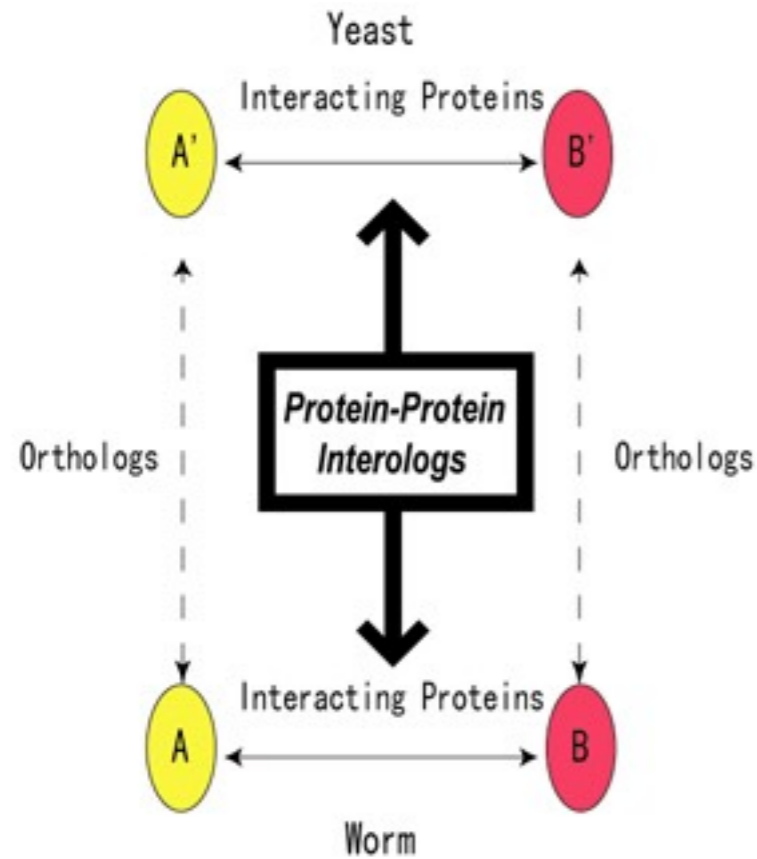
Prediction of protein interactions

- Identification of interologs

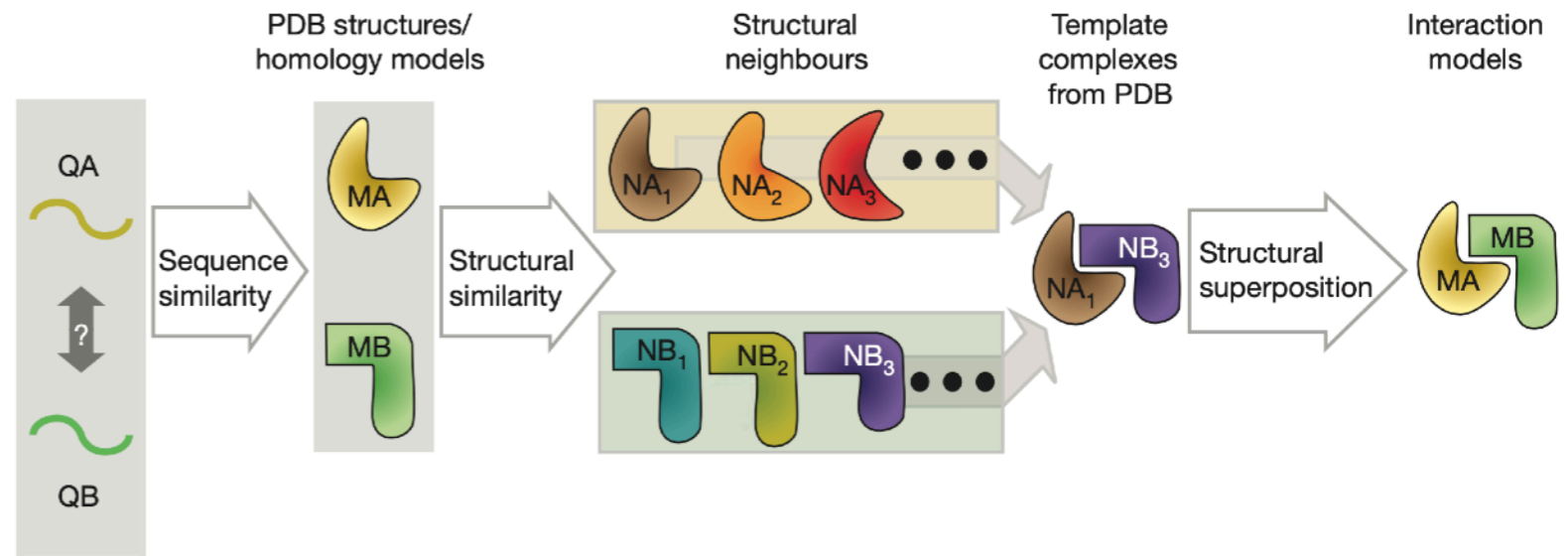


Prediction of protein interactions

- Identification of interologs

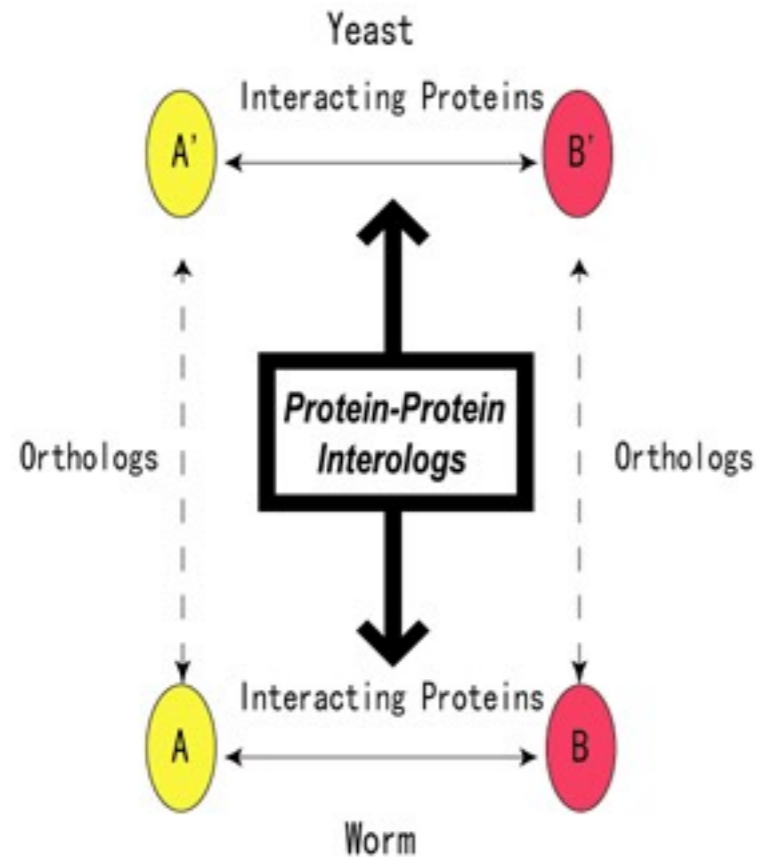


- Structure-based modeling

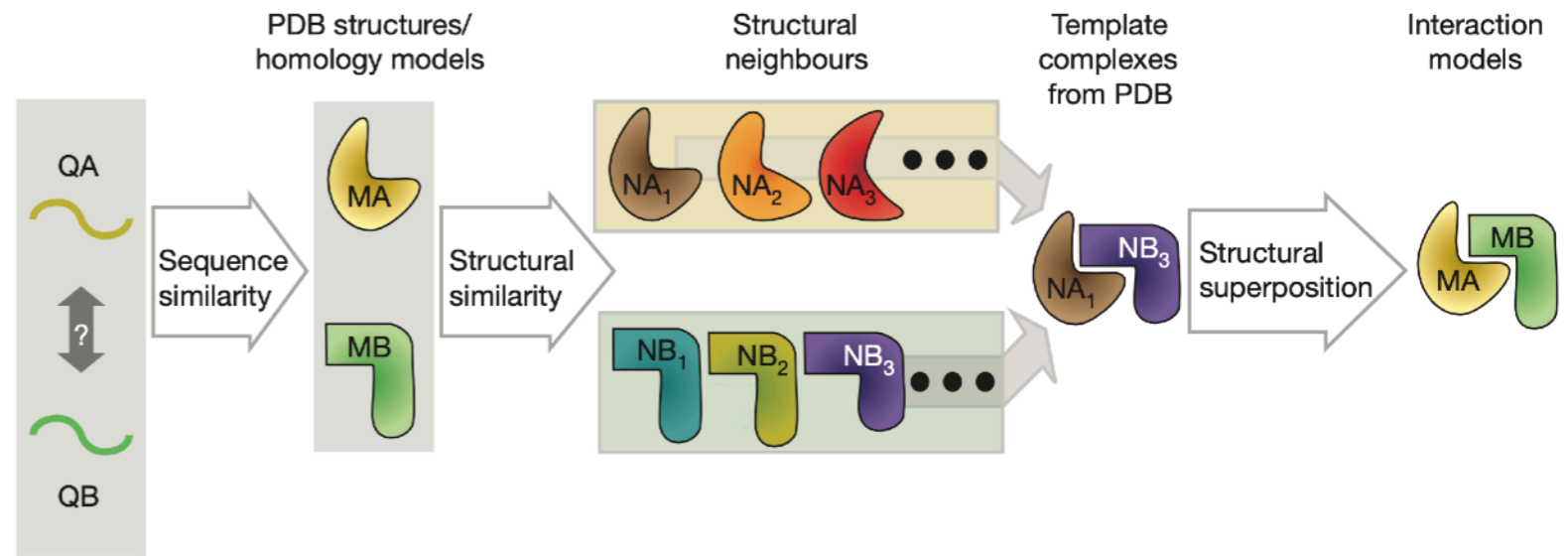


Prediction of protein interactions

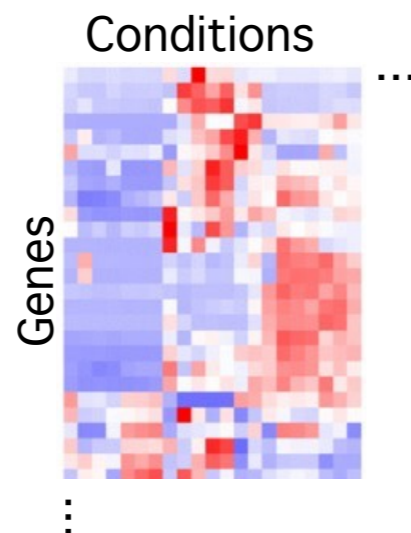
- Identification of interologs



- Structure-based modeling

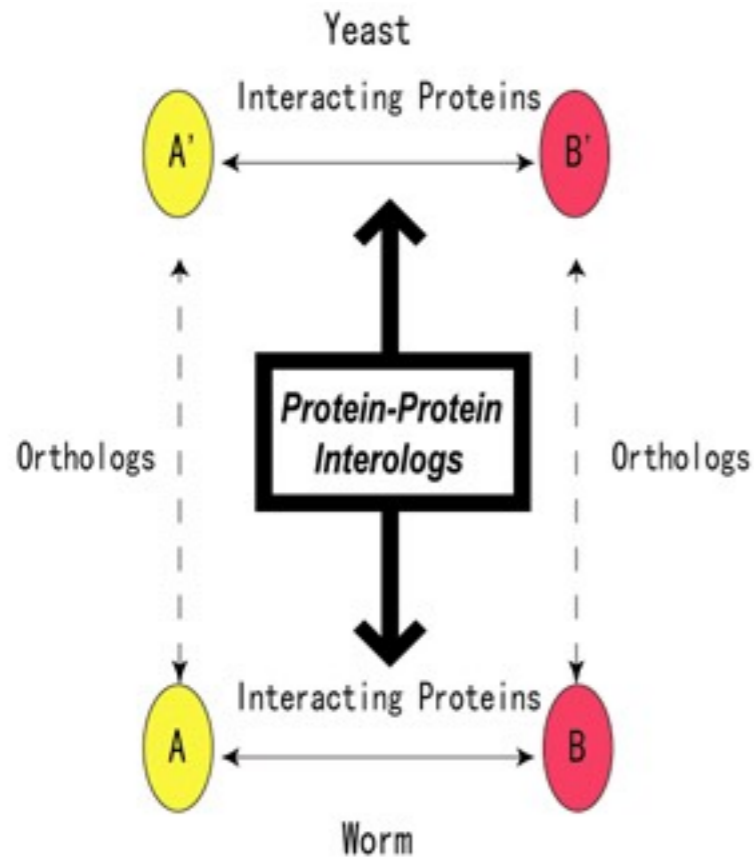


- Co-evolution,
co-regulation,
co-occurrence

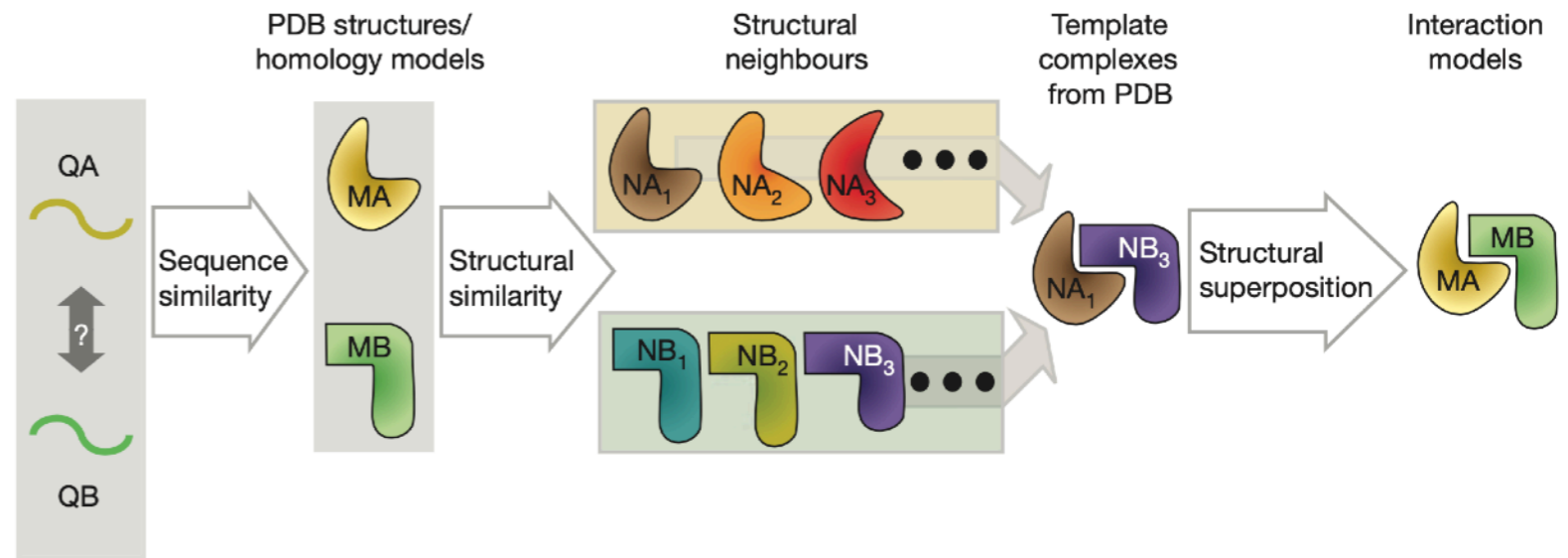


Prediction of protein interactions

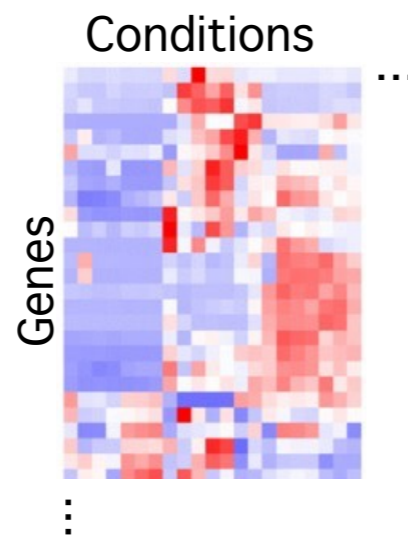
- Identification of interologs



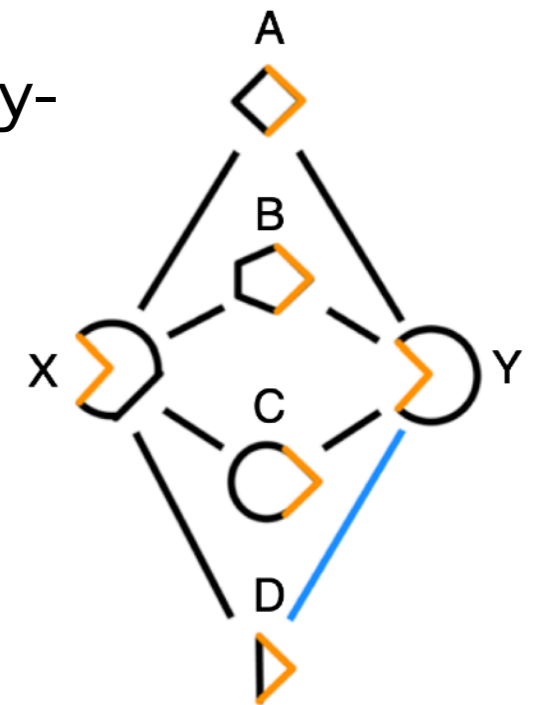
- Structure-based modeling



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



- Network topology-based prediction



Prediction of protein interactions

Key facts

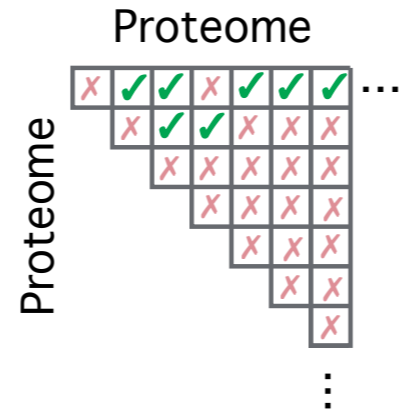
- high false positive rate
- highly biased (orthologs, structures available)
- for some species only way to get protein interaction data

Resources for protein interactions

Literature curation



Systematic mapping



Prediction

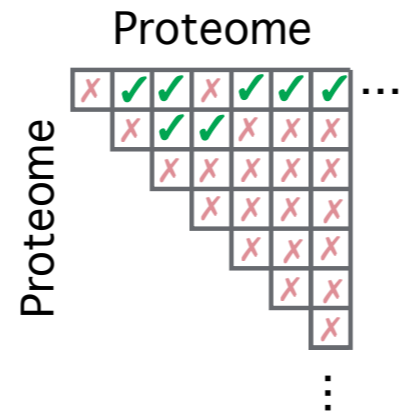


Resources for protein interactions

Literature curation



Systematic mapping



Prediction



Exercise: Explore protein interaction databases

Exercise: Explore protein interaction databases

Student - database assignment:

Last name	First Name	DB	Question
Bächle	Natalie	STRING	1
Bahn	Anja	STRING	2
Bayer	Marie Elisabeth	STRING	3
Brühl	Franziska	STRING	4
Buwen	Alena	STRING	5
Cañon Duque	Santiago Luis	STRING	6
Coelle	Mara	STRING	7
Maria	De la Cal Gonzalez	STRING	8
Duchmann	Tamara	STRING	9
Efinger	Leila	STRING	9
Ehlers	Yannick	STRING	3
Hosseini	Mersedeh	STRING	5
Jager	Loreen Maria	IntAct	1
Kaya	Can	IntAct	2
Längst	Vivian Teresa Christin	IntAct	3
Lösch	Neele	IntAct	4
Maurer	Svenja	IntAct	5
Neuroth	Sarah	IntAct	6
Rein	Ann Sophie	IntAct	7
Sarpong	Benenita Cindy	IntAct	8
Schäfer	Helena Sophie	IntAct	9
Schapelhouman	Andy	IntAct	9
Scherer	Lena	IntAct	5
Schorsch	Eva-Maria	BioGRID	1
Schulz	Dominique	BioGRID	2
Sengouni	Maria-Vasiliki	BioGRID	3
Simon	Niaef Beate	BioGRID	4
Simon Sanchez	Claudia	BioGRID	5
Solbach	Julian	BioGRID	6
Stastny	Anna	BioGRID	7
Sturm	Maximilian	BioGRID	8
Torrisi	Isabella Maria	BioGRID	9
Urschel	Jule	BioGRID	9
Wirtz Martin	Katja Montserrat	BioGRID	3
Zamor	Julie Catherine	BioGRID	7

Exercise: Explore protein interaction databases

Student - database assignment:

Last name	First Name	DB	Question
Bächle	Natalie	STRING	1
Bahn	Anja	STRING	2
Bayer	Marie Elisabeth	STRING	3
Brühl	Franziska	STRING	4
Buwen	Alena	STRING	5
Cañon Duque	Santiago Luis	STRING	6
Coelle	Mara	STRING	7
Maria	De la Cal Gonzalez	STRING	8
Duchmann	Tamara	STRING	9
Efinger	Leila	STRING	9
Ehlers	Yannick	STRING	3
Hosseini	Mersedeh	STRING	5
Jager	Loreen Maria	IntAct	1
Kaya	Can	IntAct	2
Längst	Vivian Teresa Christin	IntAct	3
Lösch	Neele	IntAct	4
Maurer	Svenja	IntAct	5
Neuroth	Sarah	IntAct	6
Rein	Ann Sophie	IntAct	7
Sarpong	Benenita Cindy	IntAct	8
Schäfer	Helena Sophie	IntAct	9
Schapelhouman	Andy	IntAct	9
Scherer	Lena	IntAct	5
Schorsch	Eva-Maria	BioGRID	1
Schulz	Dominique	BioGRID	2
Sengouni	Maria-Vasiliki	BioGRID	3
Simon	Niaef Beate	BioGRID	4
Simon Sanchez	Claudia	BioGRID	5
Solbach	Julian	BioGRID	6
Stastny	Anna	BioGRID	7
Sturm	Maximilian	BioGRID	8
Torrisi	Isabella Maria	BioGRID	9
Urschel	Jule	BioGRID	9
Wirtz Martin	Katja Montserrat	BioGRID	3
Zamor	Julie Catherine	BioGRID	7

1st part:

- Understand content and usage of database
- Fill google sheet with answers:
<https://tinyurl.com/53azz5yu>

Exercise: Explore protein interaction databases

Student - database assignment:

Last name	First Name	DB	Question
Bächle	Natalie	STRING	1
Bahn	Anja	STRING	2
Bayer	Marie Elisabeth	STRING	3
Brühl	Franziska	STRING	4
Buwen	Alena	STRING	5
Cañon Duque	Santiago Luis	STRING	6
Coelle	Mara	STRING	7
Maria	De la Cal Gonzalez	STRING	8
Duchmann	Tamara	STRING	9
Efinger	Leila	STRING	9
Ehlers	Yannick	STRING	3
Hosseini	Mersedeh	STRING	5
Jager	Loreen Maria	IntAct	1
Kaya	Can	IntAct	2
Längst	Vivian Teresa Christin	IntAct	3
Lösch	Neele	IntAct	4
Maurer	Svenja	IntAct	5
Neuroth	Sarah	IntAct	6
Rein	Ann Sophie	IntAct	7
Sarpong	Benenita Cindy	IntAct	8
Schäfer	Helena Sophie	IntAct	9
Schapelhouman	Andy	IntAct	9
Scherer	Lena	IntAct	5
Schorsch	Eva-Maria	BioGRID	1
Schulz	Dominique	BioGRID	2
Sengouni	Maria-Vasiliki	BioGRID	3
Simon	Niaef Beate	BioGRID	4
Simon Sanchez	Claudia	BioGRID	5
Solbach	Julian	BioGRID	6
Stastny	Anna	BioGRID	7
Sturm	Maximilian	BioGRID	8
Torrisi	Isabella Maria	BioGRID	9
Urschel	Jule	BioGRID	9
Wirtz Martin	Katja Montserrat	BioGRID	3
Zamor	Julie Catherine	BioGRID	7

1st part:

- Understand content and usage of database
- Fill google sheet with answers:
<https://tinyurl.com/53azz5yu>

2nd part:

Discussion of results with everyone

Exercise: Explore protein interaction databases

Student - database assignment:

Last name	First Name	DB	Question
Bächle	Natalie	STRING	1
Bahn	Anja	STRING	2
Bayer	Marie Elisabeth	STRING	3
Brühl	Franziska	STRING	4
Buwen	Alena	STRING	5
Cañon Duque	Santiago Luis	STRING	6
Coelle	Mara	STRING	7
Maria	De la Cal Gonzalez	STRING	8
Duchmann	Tamara	STRING	9
Efinger	Leila	STRING	9
Ehlers	Yannick	STRING	3
Hosseini	Mersedeh	STRING	5
Jager	Loreen Maria	IntAct	1
Kaya	Can	IntAct	2
Längst	Vivian Teresa Christin	IntAct	3
Lösch	Neele	IntAct	4
Maurer	Svenja	IntAct	5
Neuroth	Sarah	IntAct	6
Rein	Ann Sophie	IntAct	7
Sarpong	Benenita Cindy	IntAct	8
Schäfer	Helena Sophie	IntAct	9
Schapelhouman	Andy	IntAct	9
Scherer	Lena	IntAct	5
Schorsch	Eva-Maria	BioGRID	1
Schulz	Dominique	BioGRID	2
Sengouni	Maria-Vasiliki	BioGRID	3
Simon	Niaef Beate	BioGRID	4
Simon Sanchez	Claudia	BioGRID	5
Solbach	Julian	BioGRID	6
Stastny	Anna	BioGRID	7
Sturm	Maximilian	BioGRID	8
Torrisi	Isabella Maria	BioGRID	9
Urschel	Jule	BioGRID	9
Wirtz Martin	Katja Montserrat	BioGRID	3
Zamor	Julie Catherine	BioGRID	7

1st part:

- Understand content and usage of database
- Fill google sheet with answers:
<https://tinyurl.com/53azz5yu>

30 min

2nd part:

Discussion of results with everyone

Exercise: Explore protein interaction databases

Student - database assignment:

Last name	First Name	DB	Question
Bächle	Natalie	STRING	1
Bahn	Anja	STRING	2
Bayer	Marie Elisabeth	STRING	3
Brühl	Franziska	STRING	4
Buwen	Alena	STRING	5
Cañon Duque	Santiago Luis	STRING	6
Coelle	Mara	STRING	7
Maria	De la Cal Gonzalez	STRING	8
Duchmann	Tamara	STRING	9
Efinger	Leila	STRING	9
Ehlers	Yannick	STRING	3
Hosseini	Mersedeh	STRING	5
Jager	Loreen Maria	IntAct	1
Kaya	Can	IntAct	2
Längst	Vivian Teresa Christin	IntAct	3
Lösch	Neele	IntAct	4
Maurer	Svenja	IntAct	5
Neuroth	Sarah	IntAct	6
Rein	Ann Sophie	IntAct	7
Sarpong	Benenita Cindy	IntAct	8
Schäfer	Helena Sophie	IntAct	9
Schapelhouman	Andy	IntAct	9
Scherer	Lena	IntAct	5
Schorsch	Eva-Maria	BioGRID	1
Schulz	Dominique	BioGRID	2
Sengouni	Maria-Vasiliki	BioGRID	3
Simon	Niaef Beate	BioGRID	4
Simon Sanchez	Claudia	BioGRID	5
Solbach	Julian	BioGRID	6
Stastny	Anna	BioGRID	7
Sturm	Maximilian	BioGRID	8
Torrisi	Isabella Maria	BioGRID	9
Urschel	Jule	BioGRID	9
Wirtz Martin	Katja Montserrat	BioGRID	3
Zamor	Julie Catherine	BioGRID	7

1st part:

- Understand content and usage of database
- Fill google sheet with answers:
<https://tinyurl.com/53azz5yu>

30 min

2nd part:

Discussion of results with everyone

30 min

Exercise: Explore protein interaction databases

1st part:

Understand content and usage of database

Exercise: Explore protein interaction databases

1st part:

Understand content and usage of database

1. For which/how many organisms can you obtain interaction data?

Exercise: Explore protein interaction databases

1st part:

Understand content and usage of database

1. For which/how many organisms can you obtain interaction data?
2. Where does the protein interaction data come from?

Exercise: Explore protein interaction databases

1st part:

Understand content and usage of database

1. For which/how many organisms can you obtain interaction data?
2. Where does the protein interaction data come from?
3. Which information can you gather about the interaction between human FUBP1 and U2AF2?

Exercise: Explore protein interaction databases

1st part:

Understand content and usage of database

1. For which/how many organisms can you obtain interaction data?
2. Where does the protein interaction data come from?
3. Which information can you gather about the interaction between human FUBP1 and U2AF2?
4. Can you download all interactions from the database?

Exercise: Explore protein interaction databases

1st part:

Understand content and usage of database

1. For which/how many organisms can you obtain interaction data?
2. Where does the protein interaction data come from?
3. Which information can you gather about the interaction between human FUBP1 and U2AF2?
4. Can you download all interactions from the database?
5. Is there a confidence score or weight for each interaction? If yes, what is it based on?

Exercise: Explore protein interaction databases

1st part:

Understand content and usage of database

1. For which/how many organisms can you obtain interaction data?
2. Where does the protein interaction data come from?
3. Which information can you gather about the interaction between human FUBP1 and U2AF2?
4. Can you download all interactions from the database?
5. Is there a confidence score or weight for each interaction? If yes, what is it based on?
6. What are your search options (1 protein, multiple proteins, ...)?

Exercise: Explore protein interaction databases

1st part:

Understand content and usage of database

1. For which/how many organisms can you obtain interaction data?
2. Where does the protein interaction data come from?
3. Which information can you gather about the interaction between human FUBP1 and U2AF2?
4. Can you download all interactions from the database?
5. Is there a confidence score or weight for each interaction? If yes, what is it based on?
6. What are your search options (1 protein, multiple proteins, ...)?
7. Which options do you have to filter your results?

Exercise: Explore protein interaction databases

1st part:

Understand content and usage of database

1. For which/how many organisms can you obtain interaction data?
2. Where does the protein interaction data come from?
3. Which information can you gather about the interaction between human FUBP1 and U2AF2?
4. Can you download all interactions from the database?
5. Is there a confidence score or weight for each interaction? If yes, what is it based on?
6. What are your search options (1 protein, multiple proteins, ...)?
7. Which options do you have to filter your results?
8. Which options do you have to display and download your results?

Exercise: Explore protein interaction databases

1st part:

Understand content and usage of database

1. For which/how many organisms can you obtain interaction data?
2. Where does the protein interaction data come from?
3. Which information can you gather about the interaction between human FUBP1 and U2AF2?
4. Can you download all interactions from the database?
5. Is there a confidence score or weight for each interaction? If yes, what is it based on?
6. What are your search options (1 protein, multiple proteins, ...)?
7. Which options do you have to filter your results?
8. Which options do you have to display and download your results?
9. Can you perform analyses on your search results? If yes, which ones?

Exercise: Explore protein interaction databases

1st part:

Understand content and usage of database

1. For which/how many organisms can you obtain interaction data?
2. Where does the protein interaction data come from?
3. Which information can you gather about the interaction between human FUBP1 and U2AF2?
4. Can you download all interactions from the database?
5. Is there a confidence score or weight for each interaction? If yes, what is it based on?
6. What are your search options (1 protein, multiple proteins, ...)?
7. Which options do you have to filter your results?
8. Which options do you have to display and download your results?
9. Can you perform analyses on your search results? If yes, which ones?

-> To find answers to some of these questions, you can use proteins from the Kaplanis files.

Exercise: Explore protein interaction databases

Take home messages

Exercise: Explore protein interaction databases

Take home messages

- databases differ in the type of data, search, and analysis features they provide

Exercise: Explore protein interaction databases

Take home messages

- databases differ in the type of data, search, and analysis features they provide
- make sure you understand the data before you use it

Exercise: Explore protein interaction databases

Take home messages

- databases differ in the type of data, search, and analysis features they provide
- make sure you understand the data before you use it
- take the time to pick the database that is best for your needs/goals

Exercise: Explore protein interaction databases

Take home messages

- databases differ in the type of data, search, and analysis features they provide
- make sure you understand the data before you use it
- take the time to pick the database that is best for your needs/goals
- you quickly get “too much” data -> how to prioritize?