Protein-Protein Interactions

Outline

1. What is a protein-protein interaction (PPI)?
2. Why are PPI data important?
3. How can we describe a PPI?
4. Where can we find PPI data?

1.
2.
3.
4.

What is a protein-protein interaction (PPI)?

2 proteins in a binary physical contact or more participants in a homo- or hetero-oligomeric complex.

Protein–protein interactions (PPIs) are the physical contacts of high specificity established between two or more protein molecules as a result of biochemical events steered by electrostatic forces including the hydrophobic effect. (Wikipedia)

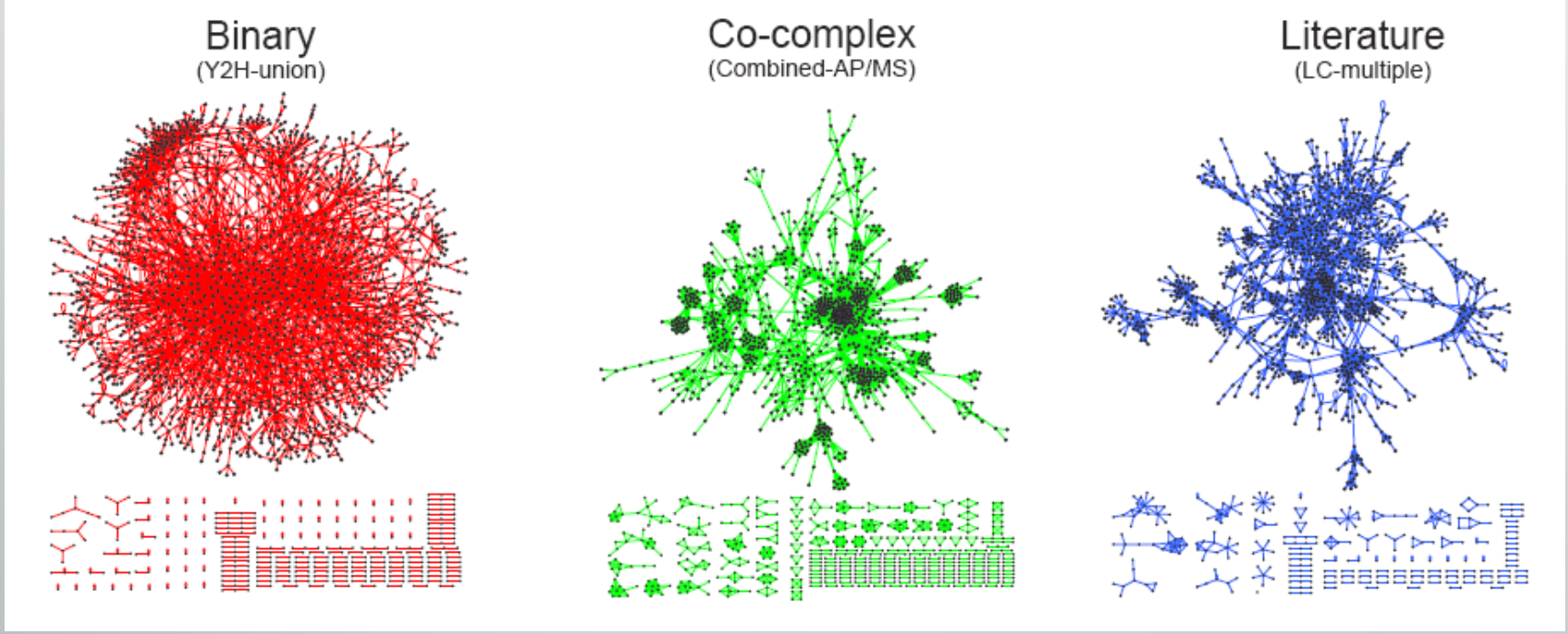
1.
2.
3.
4.

Why are PPIs important?

- The interaction is part of the functions of the proteins
  - Signalling cascades
  - Large complexes
- Interactom:
  - the whole set (network) of interacting proteins
  - also the indirect interactions among genes
  - the key to understand the biological system
- Aberrant interactions are well-known causes of diseases like Alzheimer's, cancer, etc.

1.
2.
3.
4.

Yeast interactom

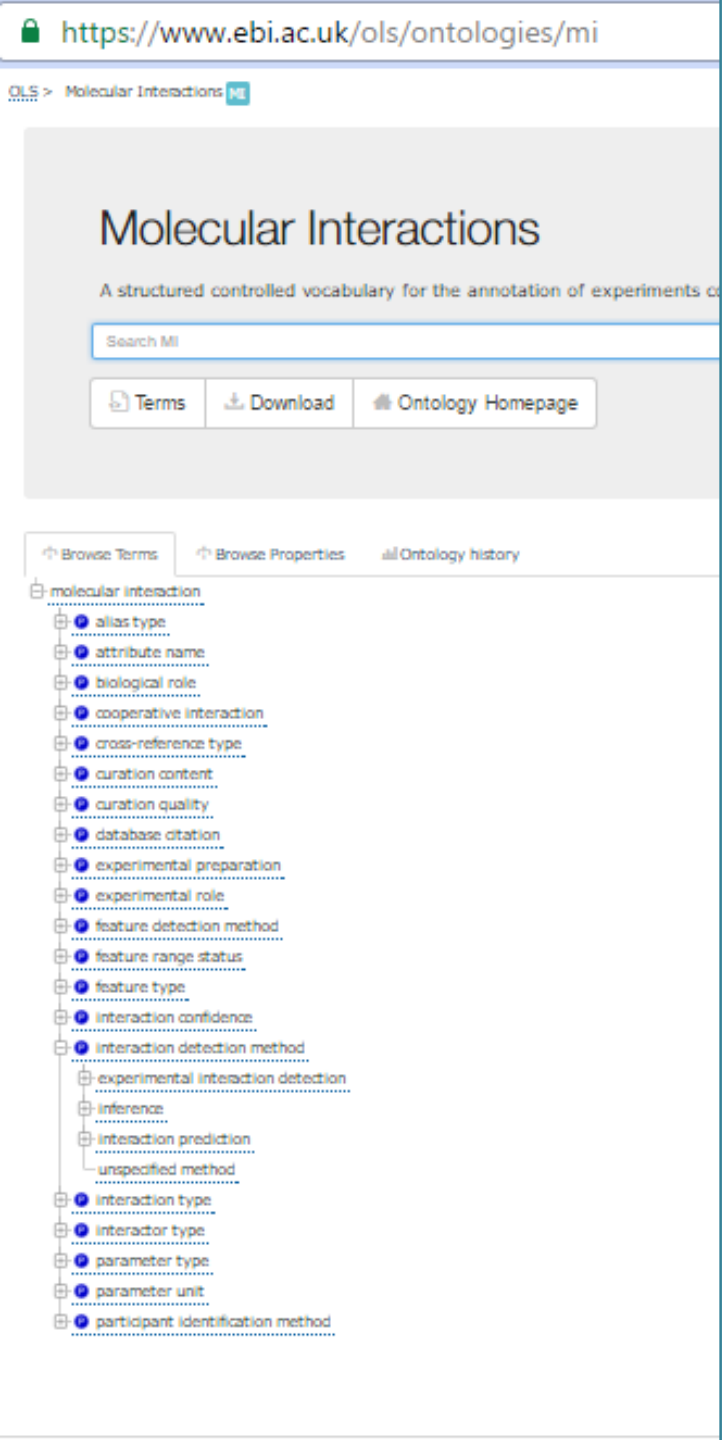


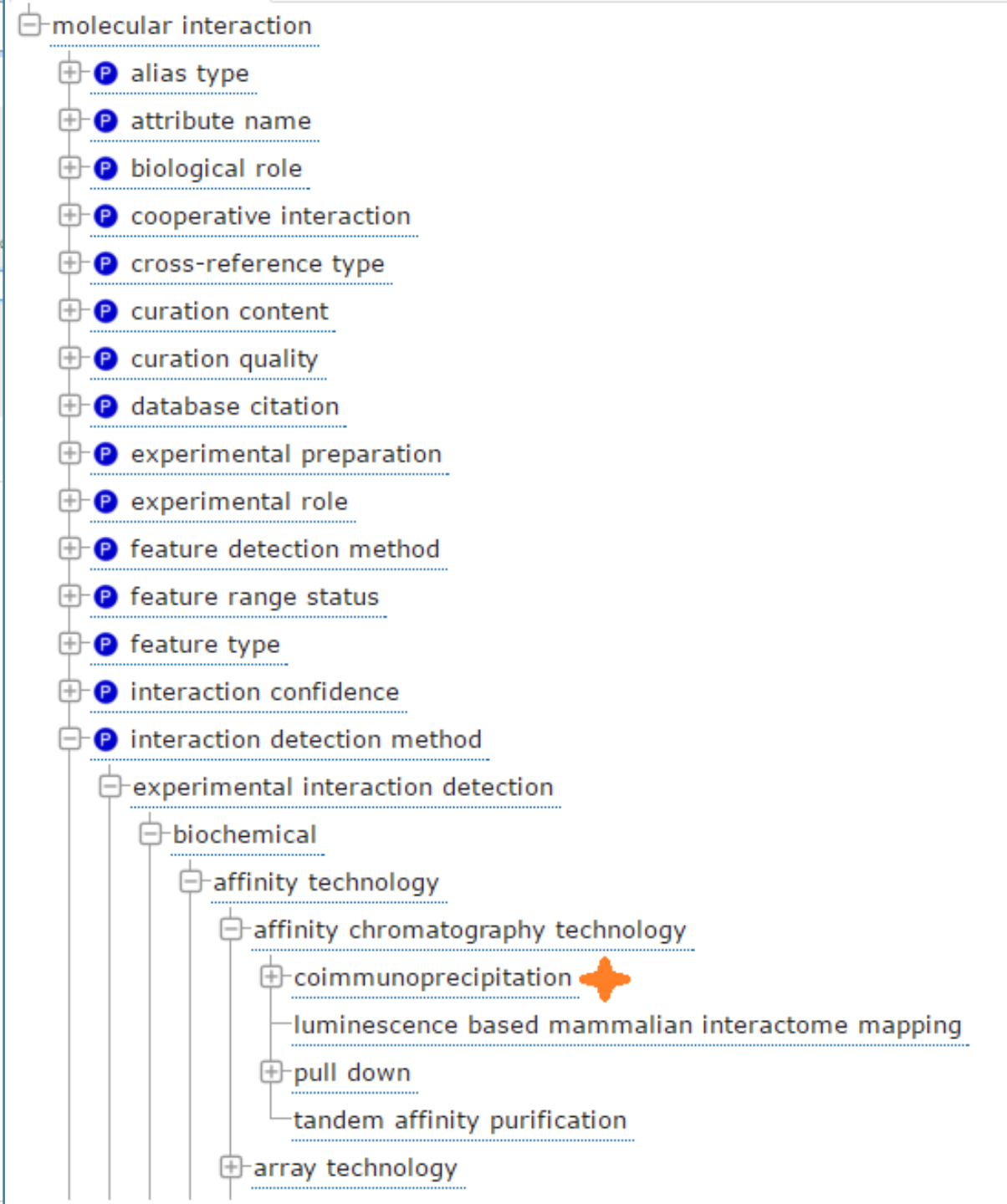
http://interactome.dfci.harvard.edu/S_cerevisiae/

1.
2.
3.
4.

How can we describe a PPI?

- Interaction partners (name, taxon, database accession/ID)
- Detection method
- Association type (direct, association, physical association)
  - PSI-MI: Proteomics Standard Initiative Molecular Interaction
- Direction and functionality (e.g X phosphorilates Y)


https://www.ebi.ac.uk/ols/ontologies/mi
OLS > Molecular Interactions MI
Molecular Interactions
A structured controlled vocabulary for the annotation of experiments c
Search MI
Terms
Download
Ontology Homepage
Browse Terms
Browse Properties
Ontology history
molecular interaction
alias type
attribute name
biological role
cooperative interaction
cross-reference type
curation content
curation quality
database citation
experimental preparation
experimental role
feature detection method
feature range status
feature type
interaction confidence
interaction detection method
experimental interaction detection
inference
interaction prediction
unspecified method
interaction type
interactor type
parameter type
parameter unit
participant identification method


molecular interaction
alias type
attribute name
biological role
cooperative interaction
cross-reference type
curation content
curation quality
database citation
experimental preparation
experimental role
feature detection method
feature range status
feature type
interaction confidence
interaction detection method
experimental interaction detection
biochemical
affinity technology
affinity chromatography technology
coimmunoprecipitation
luminescence based mammalian interactome mapping
pull down
tandem affinity purification
array technology

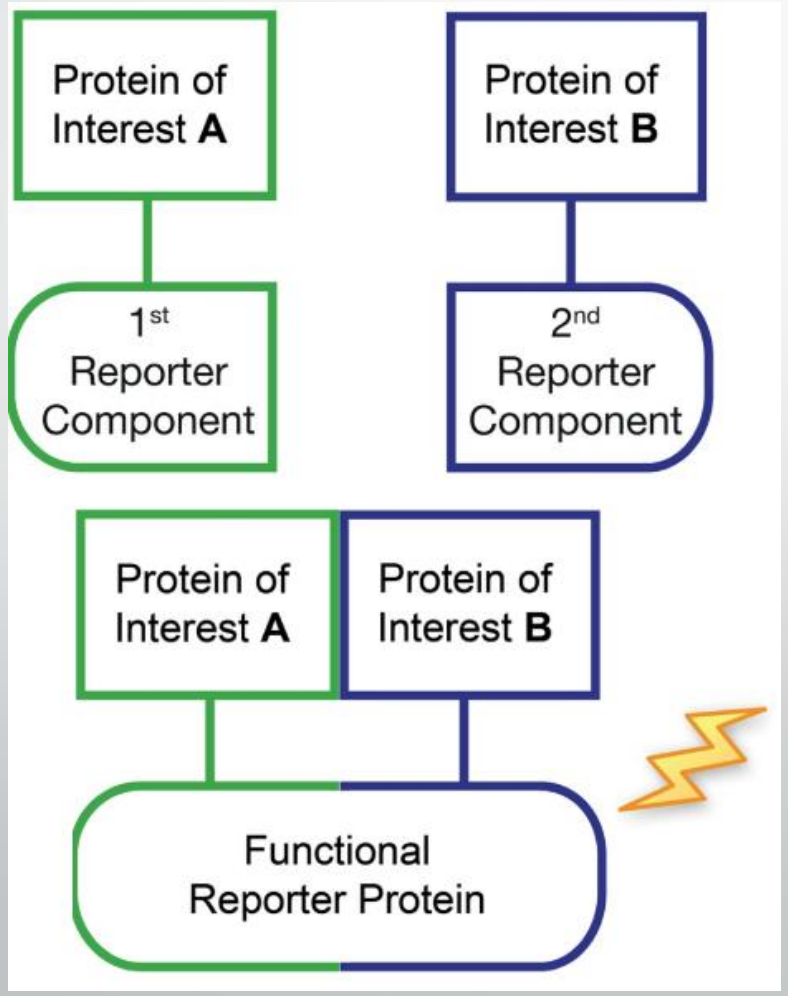
1.
2.
3.
4.

Detection methods for PPIs

- NMR, X-ray christallography, circular dicroism (CD)
- FRET (Förster resonance energy transfer)
- SPR (surface plasmon resonance)
- Coimmunoprecipitation, Pull-down, TAP
- Enzymatic assays
- PCA (protein complementation assay), pl. Y2H
- Cross-linking
- …

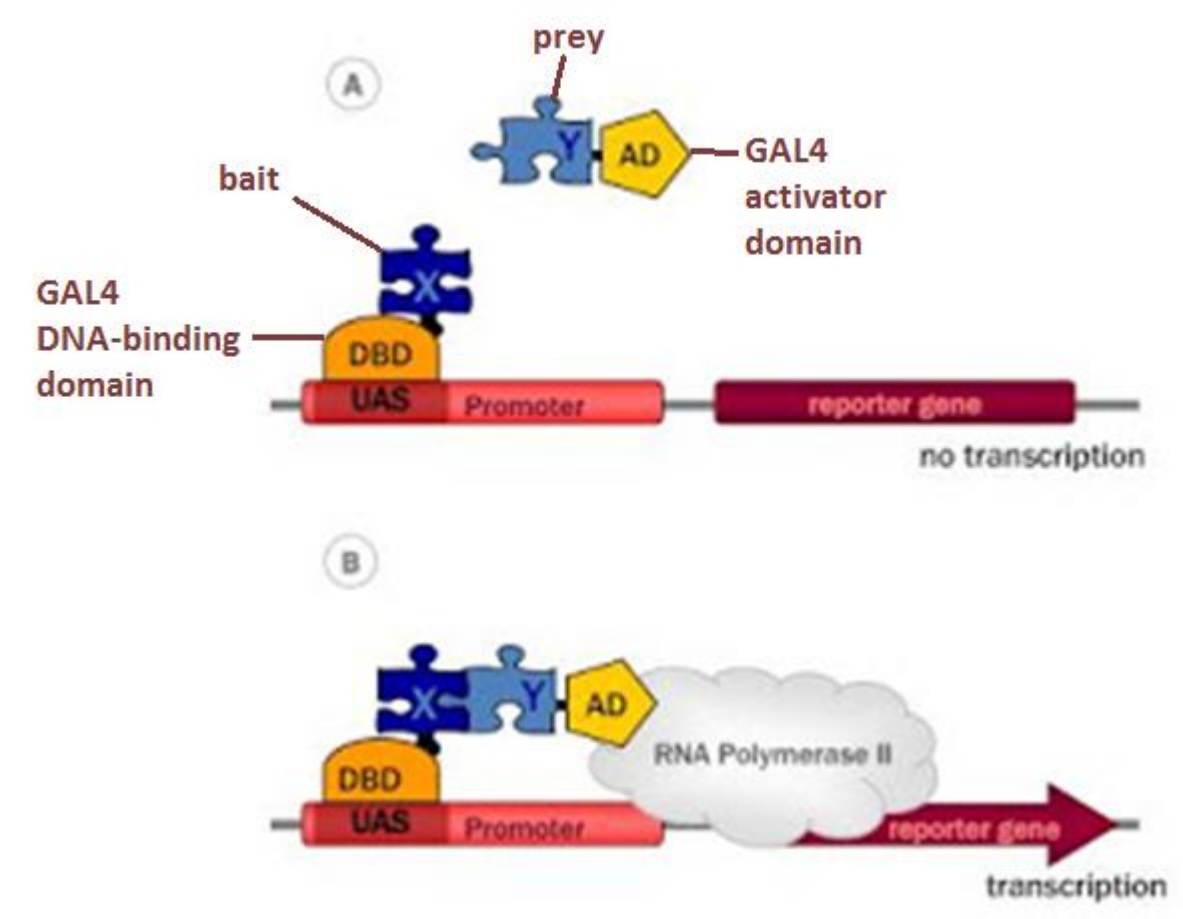
PCA
Protein-fragment complementation assay

1.
2.
3.
4.



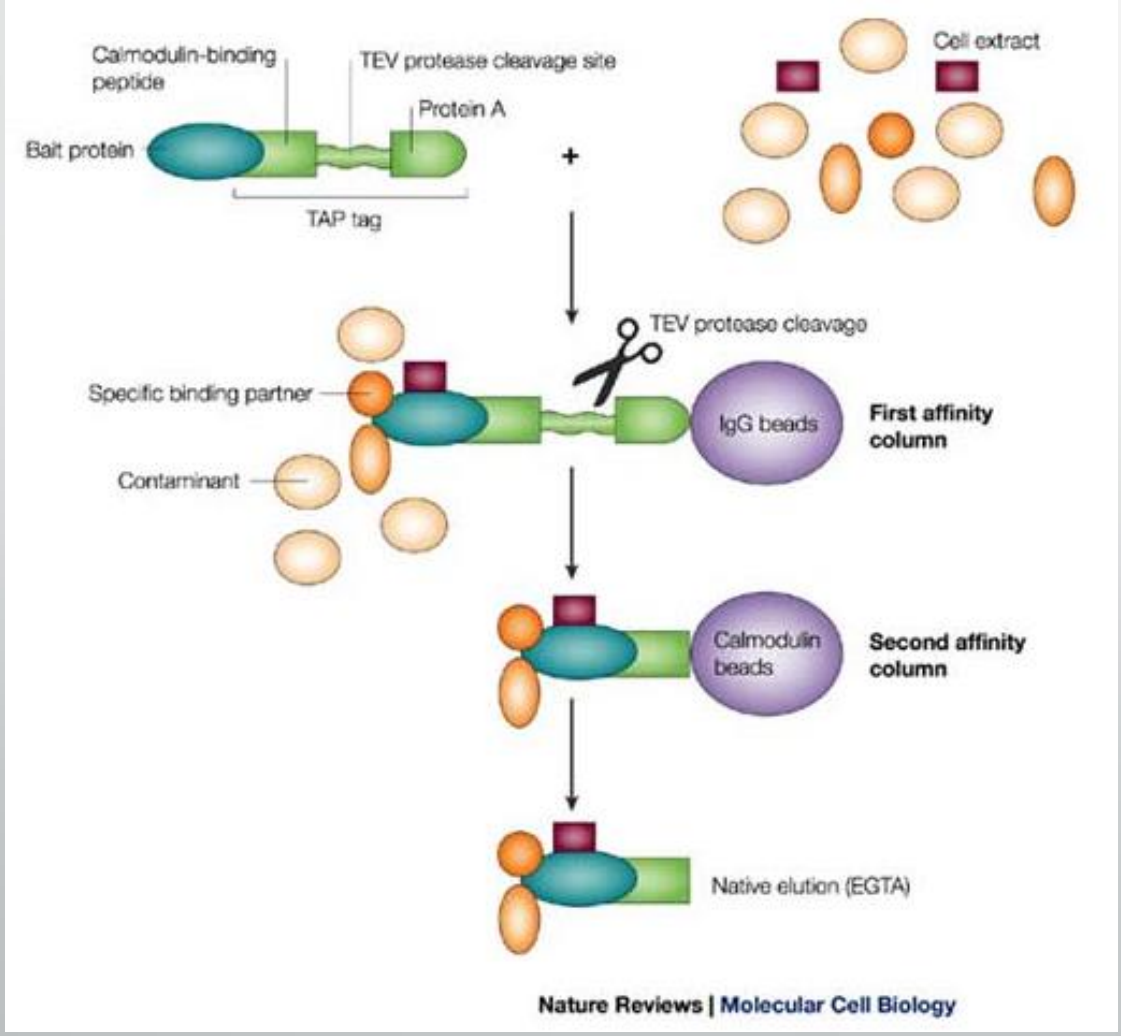
1.
2.
3.
4.

The yeast 2 hybrid method



1.
2.
3.
4.

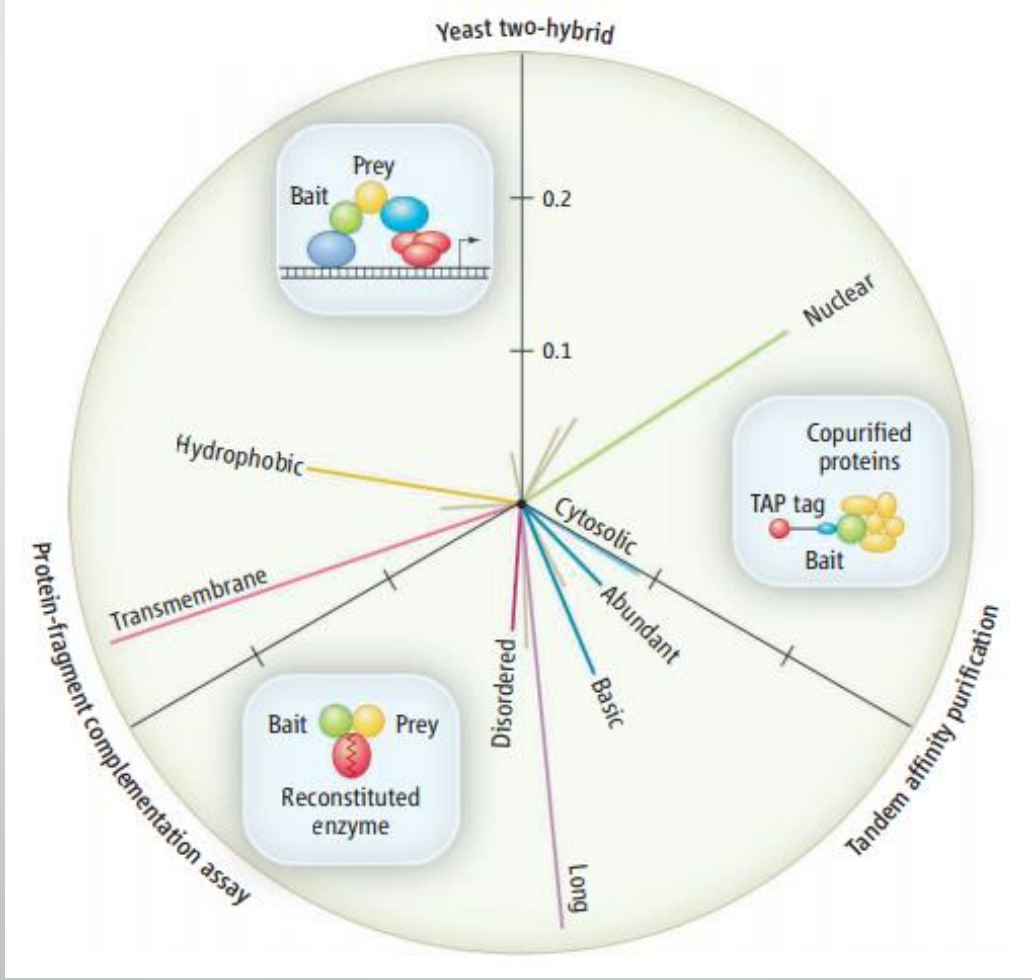
TAP
tandem affinity purification



Huber.
Nature, 2003

1.
2.
3.
4.

Hidden biases in different methods



Jensen & Borg. Science, 2008

1.
2.
3.
4.

Be cautious with false positives

WHEN? **WHERE?**

Is the encounter of the two proteins even possible?

Temporal and spacial regulation!

1.
2.
3.
4.

Association types

- Direct interaction:
  - Interaction of **2 highly purified molecules** (or a homomultimer), when there is no possibility of a third, unseen or ancillary molecule acting as a bridge between the two molecules of interest
- Physical Association:
  - the **participants form an n-ary complex** which has remained in a complex following a relatively stringent purification protocol (i.e. a more than a single pulldown or coimmunoprecipitation)
  - or a binary complex which **may involve more interactors** than those identified in the experiment
- Association:
  - interactions affinity chromatography (or one of its children e.g. coimmunoprecipitation, pulldown, TAP) with **1 bait and MORE THAN 1 prey**

1.
2.
3.
4.

What else can we find in a PPI database?

Examples

- Structure of the complex
  - Whole colmplex or at least the interface
  - Based on real measurement or docking
  - Few known
- Confidence score
- Source database
- Interaction ID

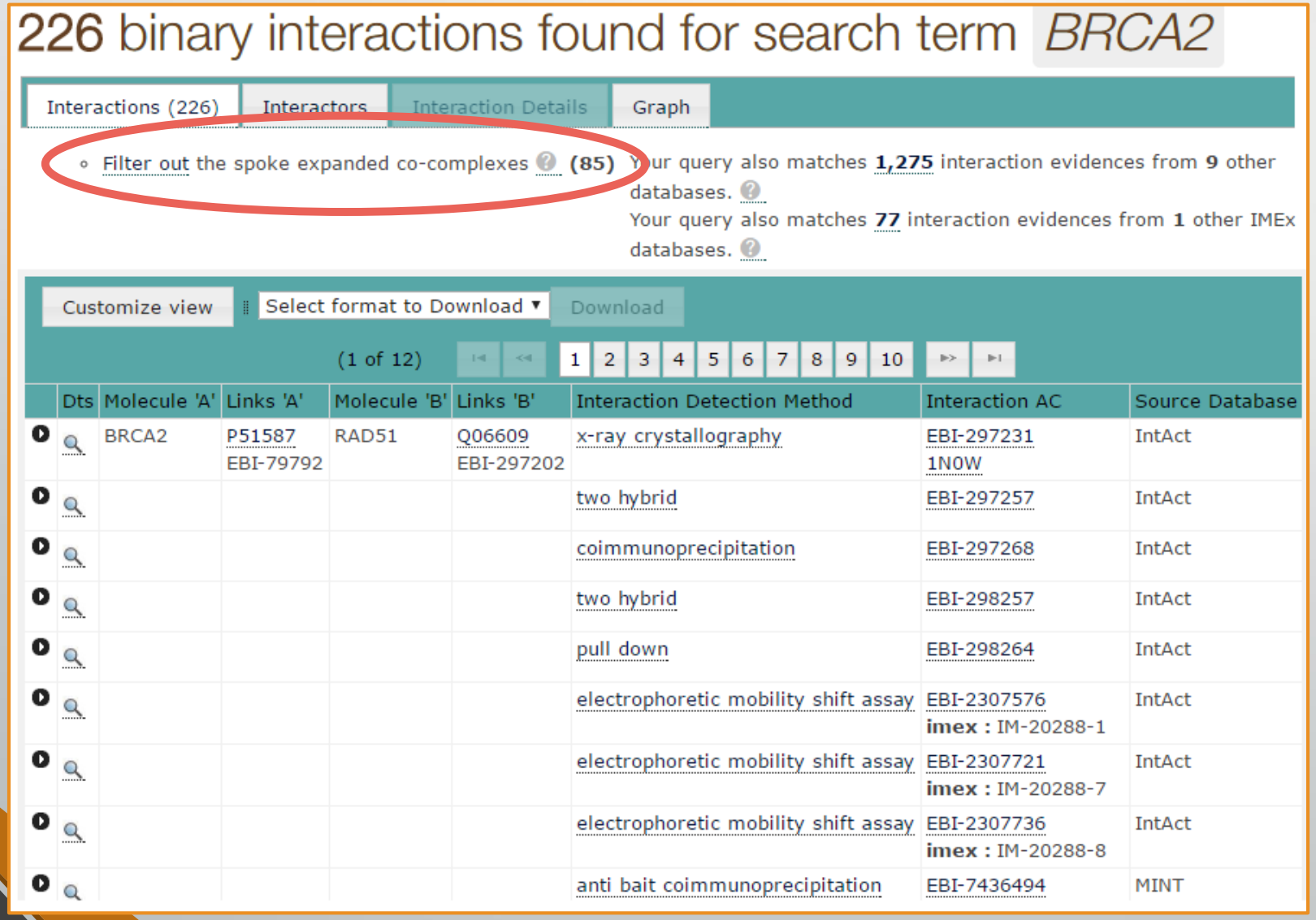
1.
2.
3.
4.

What databases are there?

- IntAct http://www.ebi.ac.uk/intact/
- DIP (Database of Interacting Proteins) http://dip.doe-mbi.ucla.edu/
- MINT (Molecular INTeraction database) http://mint.bio.uniroma2.it/
- BioGRID http://thebiogrid.org/
- MatrixDB (extracellular matrix interaction database) http://matrixdb.univ-lyon1.fr/
- STRING http://string-db.org/
- PDB (Protein Data Bank) http://www.rcsb.org/pdb/home/home.do

IntAct

1.
2.
3.
4.



226 binary interactions found for search term *BRCA2*

Interactions (226) | Interactors | Interaction Details | Graph

- Filter out the spoke expanded co-complexes (85)

Your query also matches **1,275** interaction evidences from **9** other databases.

Your query also matches **77** interaction evidences from **1** other IMEx databases.

Customize view | Select format to Download | Download

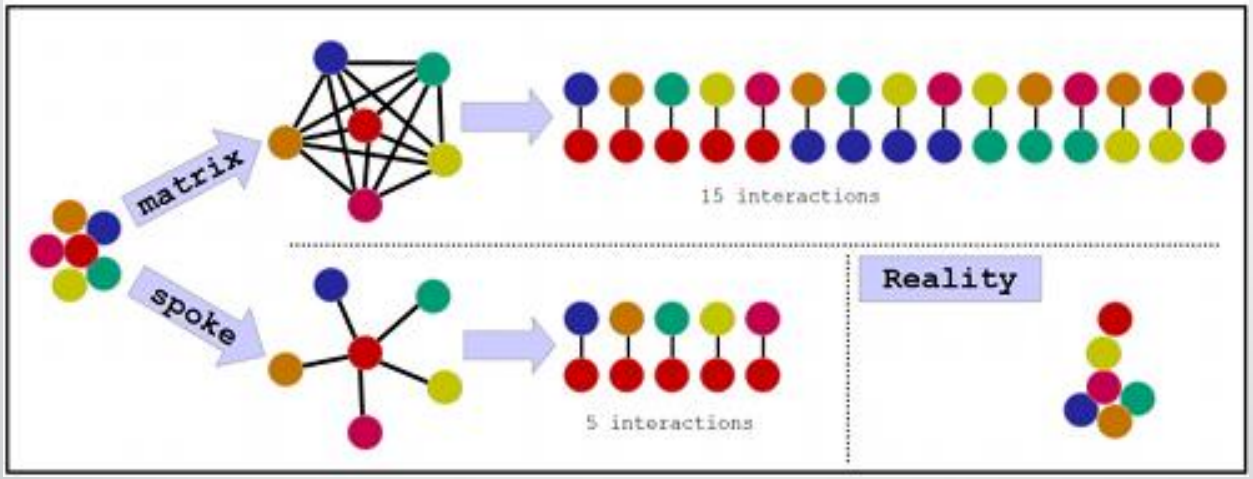
(1 of 12) 1 2 3 4 5 6 7 8 9 10

Dts	Molecule 'A'	Links 'A'	Molecule 'B'	Links 'B'	Interaction Detection Method	Interaction AC	Source Database
	BRCA2	P51587 EBI-79792	RAD51	Q06609 EBI-297202	x-ray crystallography	EBI-297231 1N0W	IntAct
					two hybrid	EBI-297257	IntAct
					coimmunoprecipitation	EBI-297268	IntAct
					two hybrid	EBI-298257	IntAct
					pull down	EBI-298264	IntAct
					electrophoretic mobility shift assay	EBI-2307576 **imex :** IM-20288-1	IntAct
					electrophoretic mobility shift assay	EBI-2307721 **imex :** IM-20288-7	IntAct
					electrophoretic mobility shift assay	EBI-2307736 **imex :** IM-20288-8	IntAct
					anti bait coimmunoprecipitation	EBI-7436494	MINT

1.
2.
3.
4.

IntAct interactions are binary

The co-complexes (n-ary interactions) are expanded into a set of binary interactions.



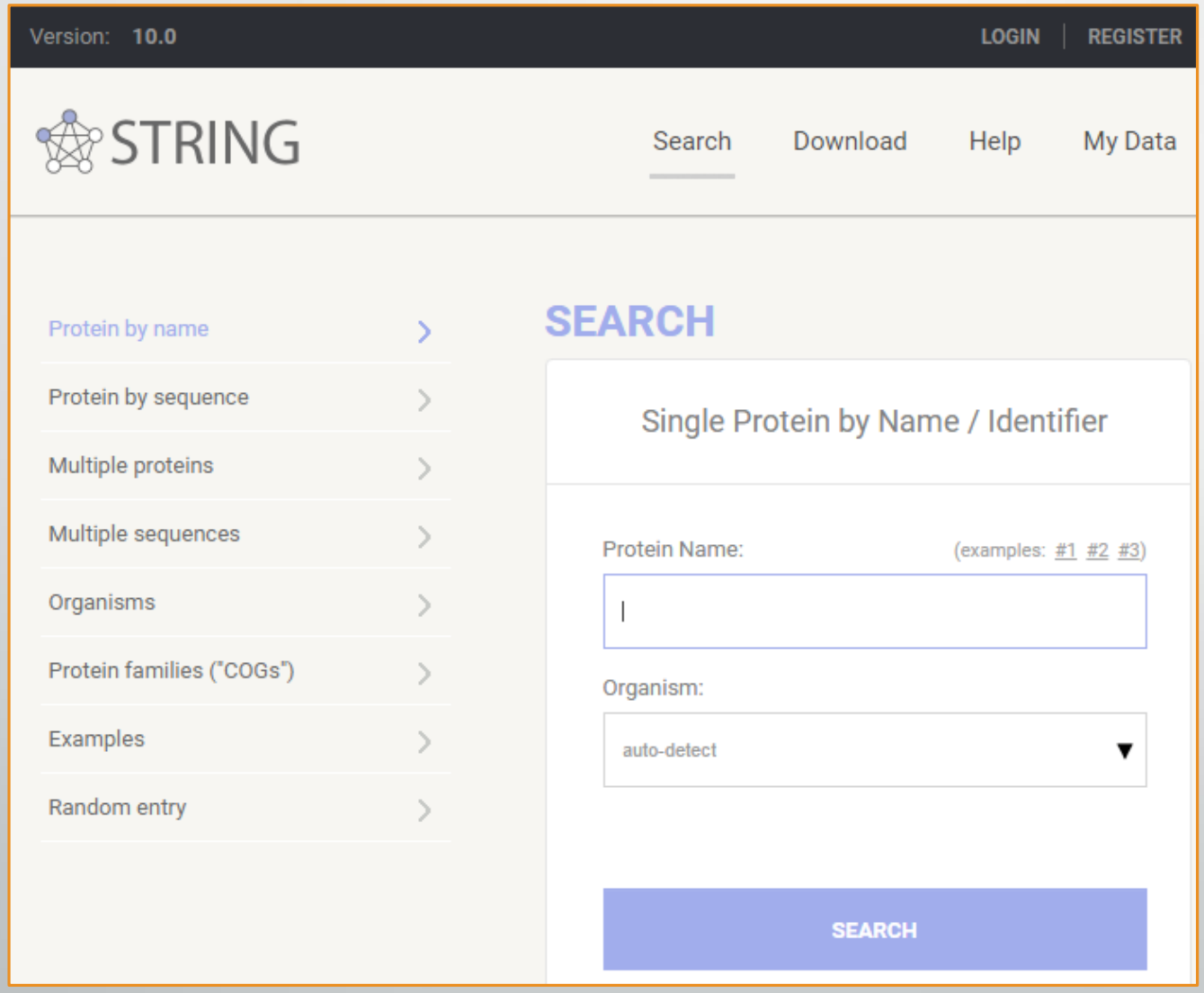
Matrix expansion: Links **all molecule to all other molecule** present in the complex. From N molecules it generates (N*(N-1))/2 binary interactions.
Spoke expansion: Links the **bait molecule to all** prey molecules. If N is the count of molecule in the complex, it generates N-1 binary interactions.

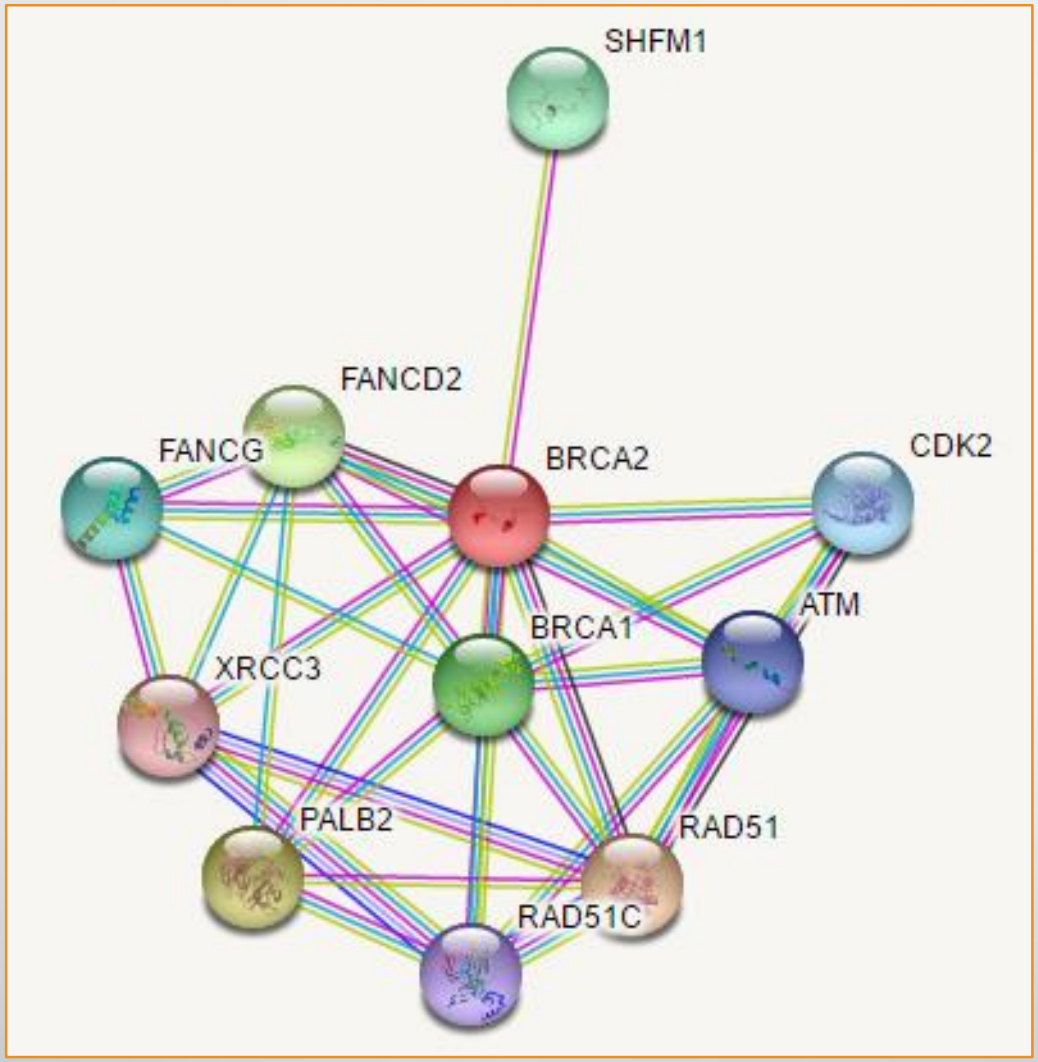
Drawback: in reality these binary interactions can be false positive. → filter out

1.
2.
3.
4.

STRING
the text mining database

- Compuational searching algorithm for finding co-occurring protein names in articles
- High throughput
- 9.643.763 proteins
  2031 organisms
  932.553.897 interactions
- Black list for obscure and ambiguous words (e.g. SDS)
- Lists of synonyms
- Identifying proteins with same name from different taxons?
- Negating statements? Contradictory publications?


Version: 10.0
LOGIN
REGISTER
STRING
Search
Download
Help
My Data
Protein by name
Protein by sequence
Multiple proteins
Multiple sequences
Organisms
Protein families ("COGs")
Examples
Random entry
SEARCH
Single Protein by Name / Identifier
Protein Name:
(examples: #1 #2 #3)
Organism:
auto-detect
SEARCH



Known Interactions

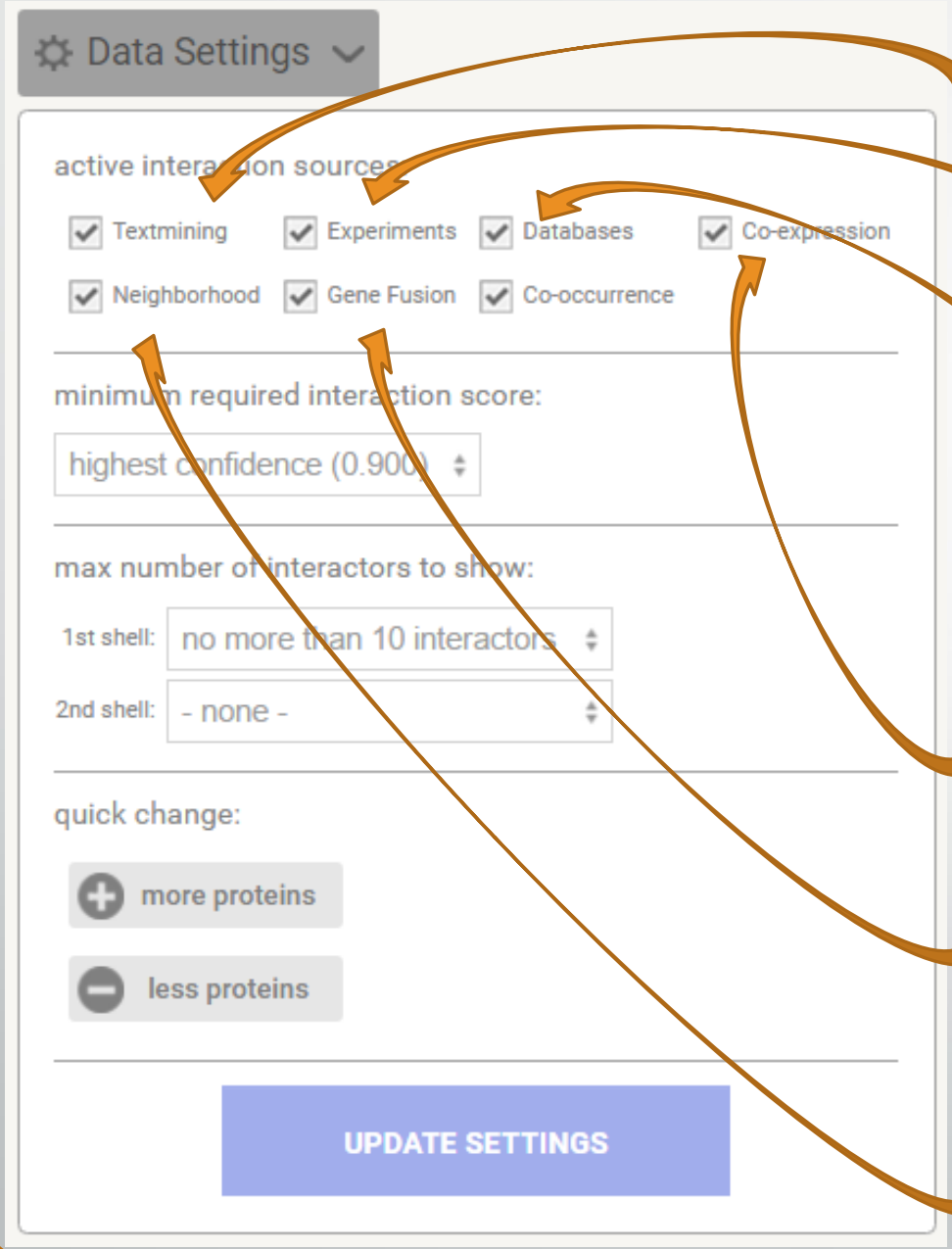
- *from curated databases*
- *experimentally determined*

Predicted Interactions

- *gene neighborhood*
- *gene fusions*
- *gene co-occurrence*

Others

- *textmining*
- *co-expression*
- *protein homology*



Automated, unsupervised textmining - proteins that are frequently mentioned together.

Experimentally shown to bind, or to be in the same protein complex.

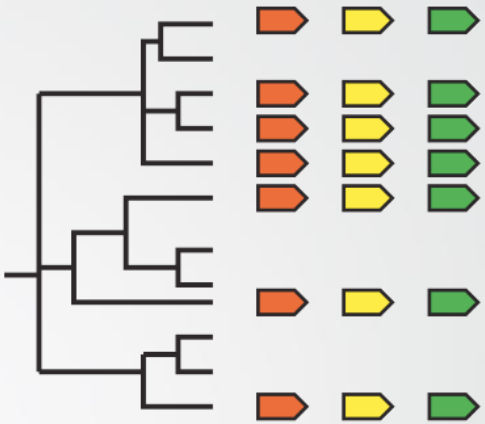
Annotated in other databases as involved together in a pathway, protein complex of biological process.

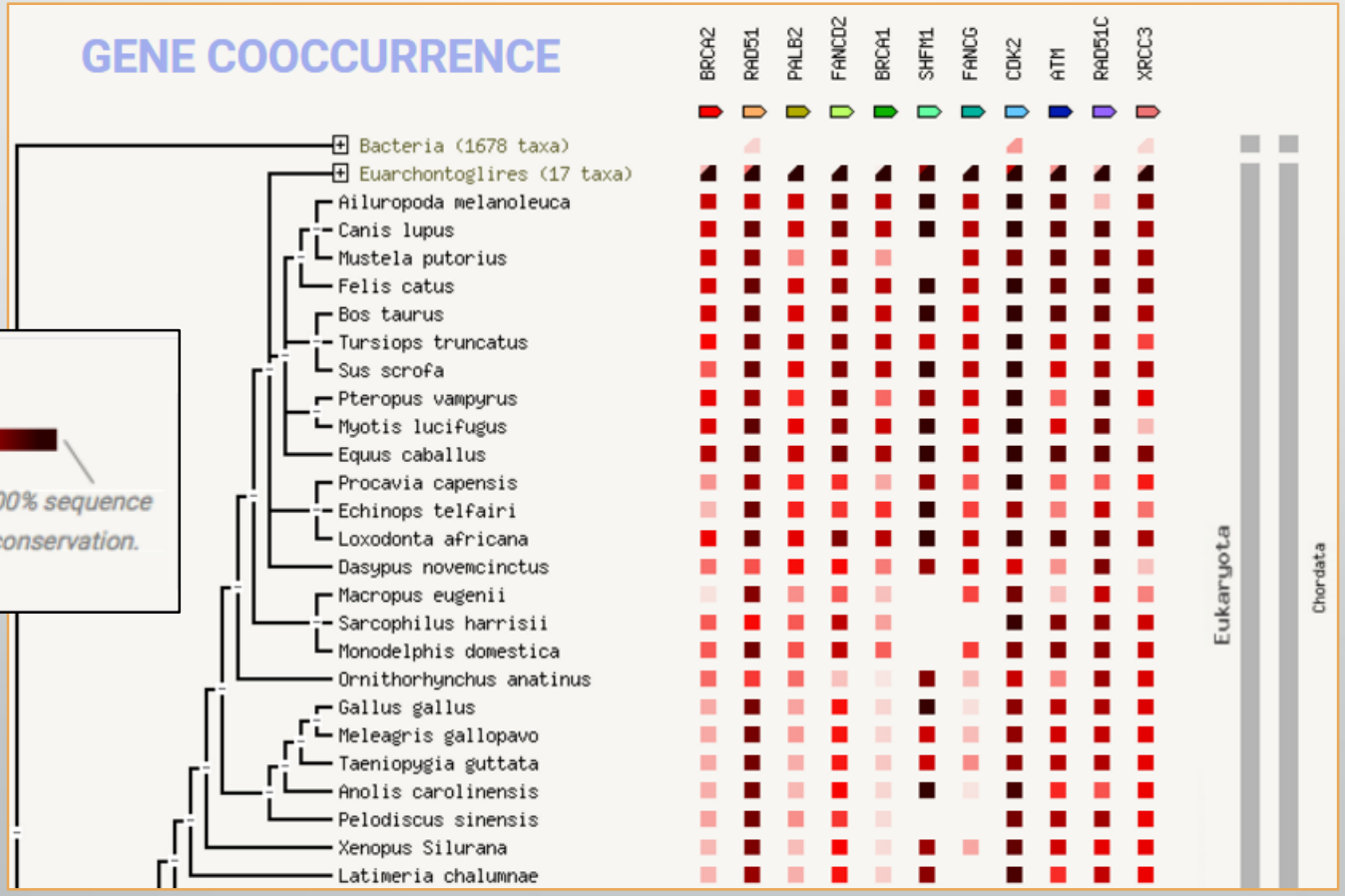
Functional association is predicted based on the expression level data in the organism.

Genes that are sometimes fused into single open reading frames.

Groups of genes that are frequently observed in each other's genomic neighborhood.

Phylogenetic profiles





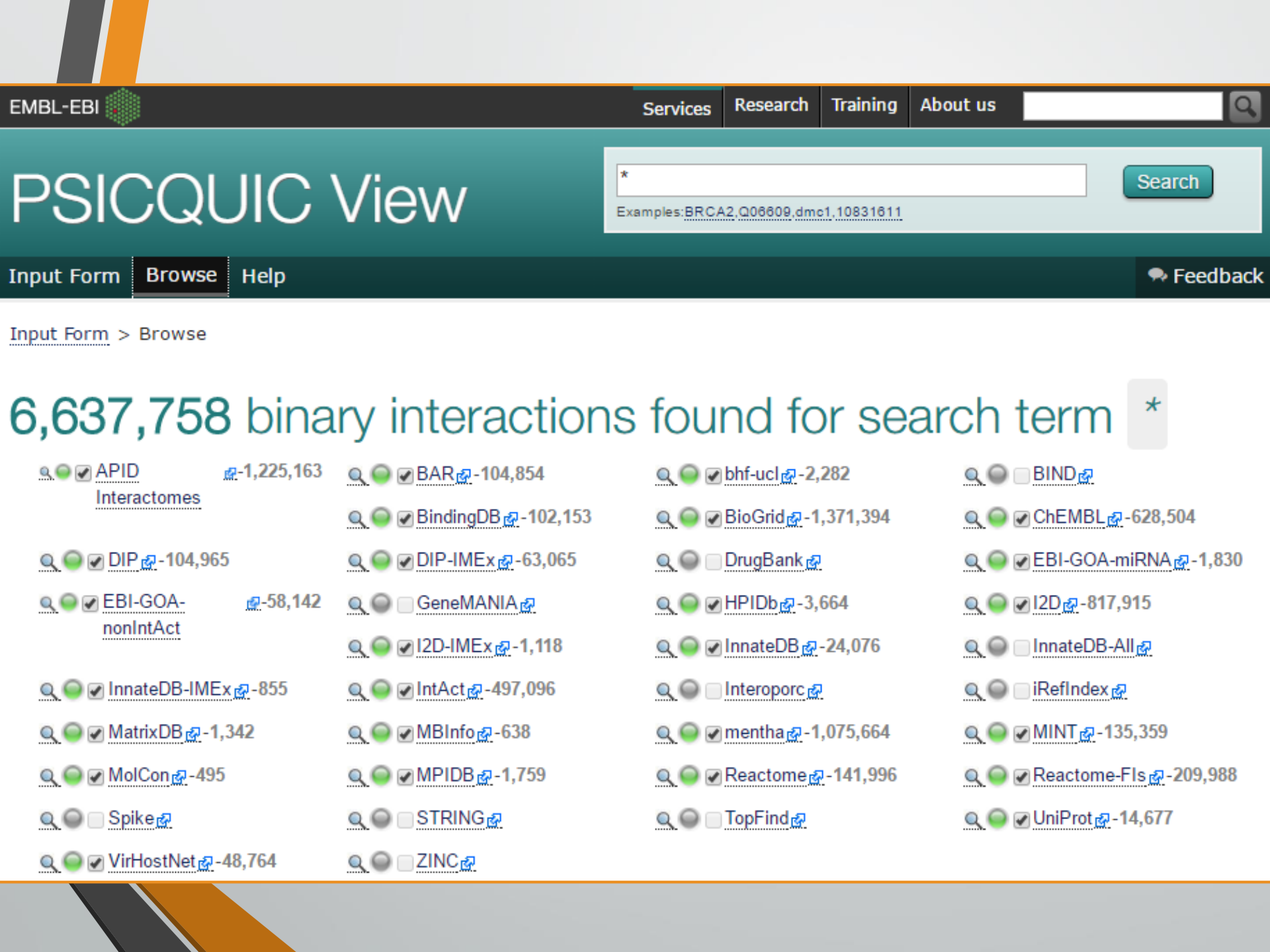
Enrichment analysis

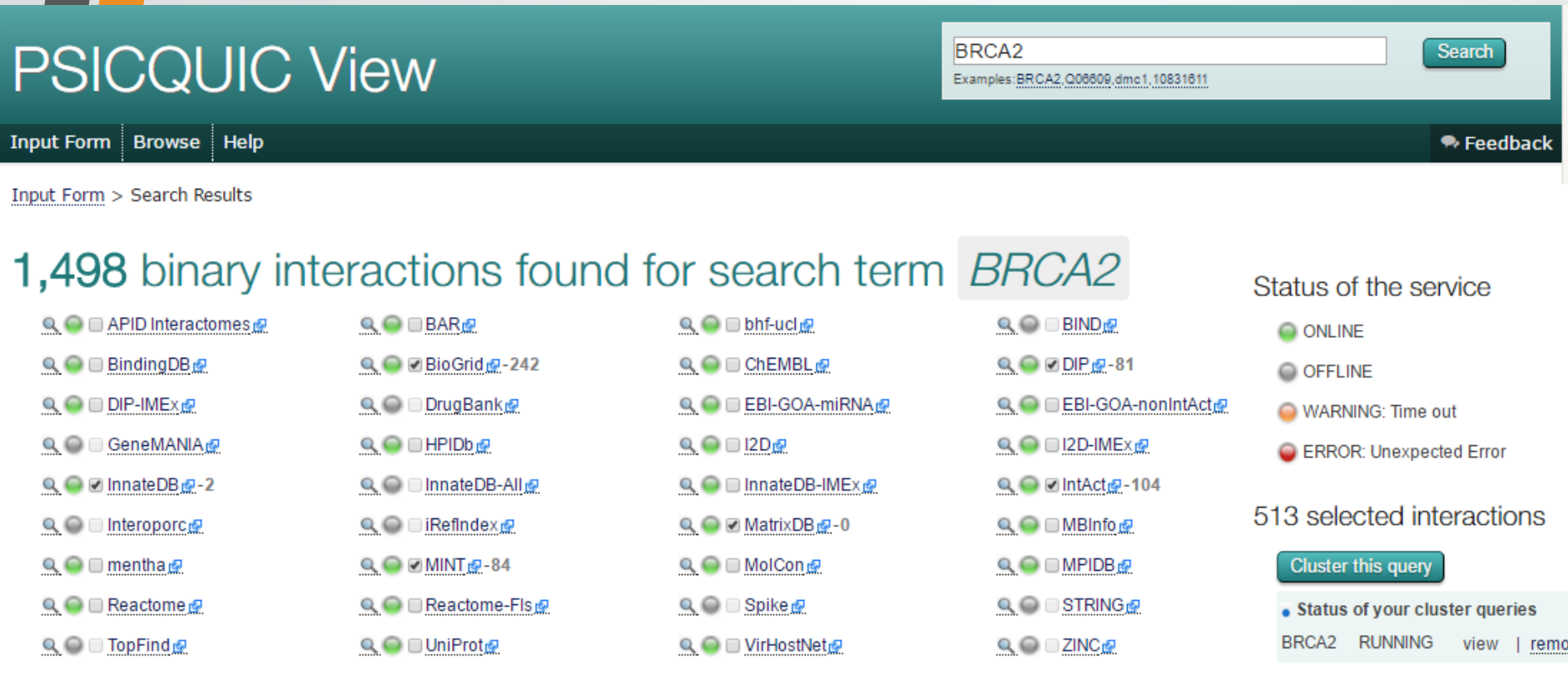
- Gene Ontology (GO) annotations
  - Molecular function
  - Biological process
  - Cell component
- KEGG pathways

1.
2.
3.
4.

Integrating databases: PSICQUIC

- Several PPI database
- Different formats
- EBI platform
- Clustering: clusters the redundant information: experimental methods, publications, source database, scoring will be aggregated in the resulting interaction.


EMBL-EBI
Services
Research
Training
About us
PSICQUIC View
*
Search
Examples:BRCA2,Q06609,dmc1,10831611
Input Form
Browse
Help
Feedback
Input Form > Browse
6,637,758 binary interactions found for search term *
APID Interactomes -1,225,163
BAR -104,854
bhf-ucl -2,282
BIND
BindingDB -102,153
BioGrid -1,371,394
ChEMBL -628,504
DIP -104,965
DIP-IMEx -63,065
DrugBank
EBI-GOA-miRNA -1,830
EBI-GOA-nonIntAct -58,142
GeneMANIA
HPIDb -3,664
I2D -817,915
I2D-IMEx -1,118
InnateDB -24,076
InnateDB-All
InnateDB-IMEx -855
IntAct -497,096
Interoporc
iRefIndex
MatrixDB -1,342
MBInfo -638
mentha -1,075,664
MINT -135,359
MolCon -495
MPIDB -1,759
Reactome -141,996
Reactome-FIs -209,988
Spike
STRING
TopFind
UniProt -14,677
VirHostNet -48,764
ZINC


PSICQUIC View
BRCA2
Search
Examples:BRCA2,Q06609,dmc1,10831611
Input Form
Browse
Help
Feedback
Input Form > Search Results
1,498 binary interactions found for search term BRCA2
APID Interactomes
BAR
bhf-ucl
BIND
BindingDB
BioGrid-242
ChEMBL
DIP-81
DIP-IMEx
DrugBank
EBI-GOA-miRNA
EBI-GOA-nonIntAct
GeneMANIA
HPIDb
I2D
I2D-IMEx
InnateDB-2
InnateDB-All
InnateDB-IMEx
IntAct-104
Interoporc
iRefIndex
MatrixDB-0
MBInfo
mentha
MINT-84
MolCon
MPIDB
Reactome
Reactome-FIs
Spike
STRING
TopFind
UniProt
VirHostNet
ZINC
Status of the service
ONLINE
OFFLINE
WARNING: Time out
ERROR: Unexpected Error
513 selected interactions
Cluster this query
Status of your cluster queries
BRCA2 RUNNING view | remove

Use of PPI databases

- Databases provide curated repositories of biological information for different organisms
- Can search for specific proteins to see what they interact with
- Can view interactions and associated data in different ways
- Can find out how many interactions exist in a specific organism
- Can overlay experimental data on PPI networks
- PPI databases use controlled vocabularies and standards for data exchange