Rossella De Cegli, PhD

(BAD days)

13/07/2022

FONDAZIONE

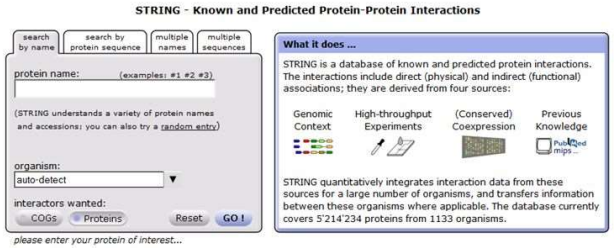
https://string-db.org

STRING: The functional protein association networks tool

- The STRING database aims to collect and integrate information about **functional interactions** between the expressed proteins, by consolidating known and predicted protein–protein association data for a large number of organisms.
- The associations in STRING include **direct (physical)** interactions, as well as **indirect (functional)** interactions

STRING Database

- First released in 2000
- Collection of data showing protein-protein interactions
- Predicts protein-protein interactions using multiple sources

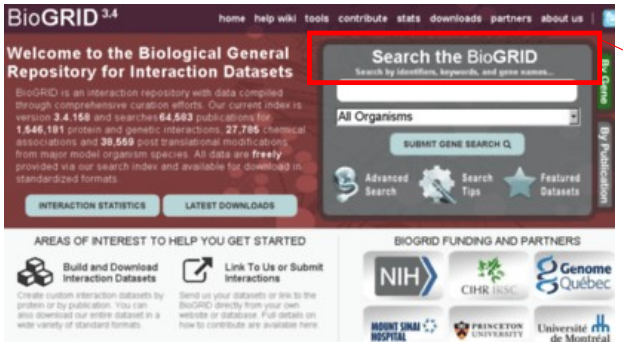


BioGRID | Biological General Repository for Interaction Datasets

- BioGRID is an **interaction repository** with data compiled through comprehensive curation efforts.
- Includes 75,988 publications for **2,005,220** protein and genetic interactions, **29,093** chemical associations and **968,210** post translational modifications from major model organism species.
- All data are freely available for download

STRING_gkw937.pdf and BIOGRID_gkw1102.pdf

BioGRID ❖ https://thebiogrid.org



PPARGC1A Result Summary | BioGRID 19/03/2018, 16:36

BioGRID 3.4 home help wiki tools contribute stats downloads partners about us

Gene / Identifier Search

PPARGC1A

Homo sapiens GO

Result Summary

PPARGC1A *Homo sapiens*

LEM6, PGC-1(alpha), PGC-1v, PGC1, PGC1A, PPARGC1

peroxisome proliferator-activated receptor gamma, coactivator 1 alpha

UBI SUMO

GO Process (32) **GO Function** (12) **GO Component** (3)

EXTERNAL DATABASE LINKOUTS

HGNC | OMIM | VEGA | Entrez Gene | RefSeq | UniprotKB | Ensembl | HPRD

! Download 137 Published Interactions For This Protein

Stats & Options

Current Statistics **Publications:** 59

High Throughput		Low Throughput
3 (2%)	134 Physical Interactions	131 (98%)
0 (0%)	3 Genetic Interactions	3 (100%)

Search Filters Customize how your results are displayed...

No Filter: Show All Associations

Your BioGRID Download File Has Been Successfully Created!

Thank you for your patience as your download file was created! It is now available for you to download in the format you requested. To retrieve your file, simply click on the **Download Your File** button below and pick the destination location on your computer to begin the download process. If you have any problems retrieving the file, please consult our **help wiki** for more information on our files, formats, and build techniques or alternatively, simply utilize one of our many pre-compiled downloads available on our **download page**.

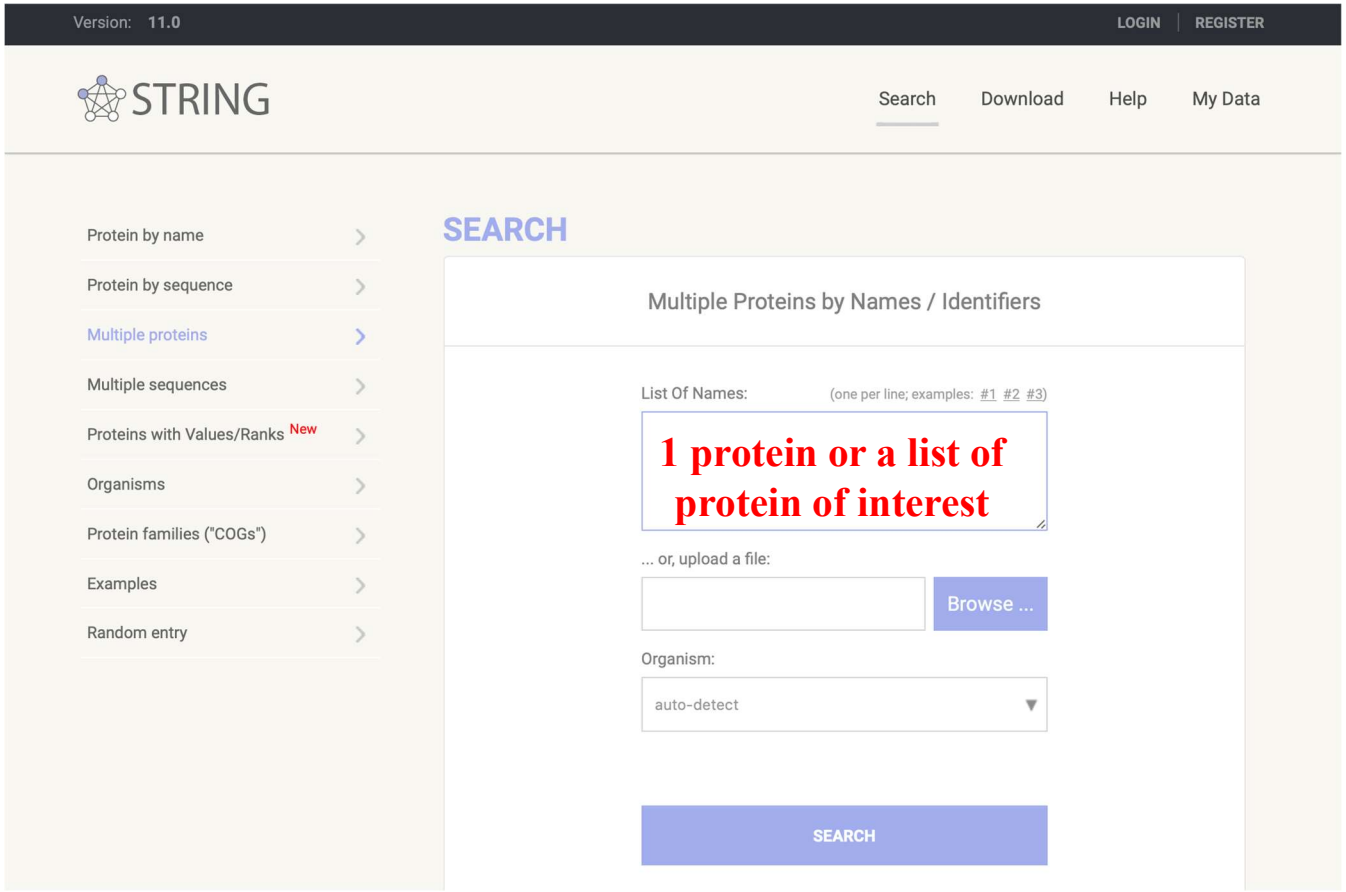
DOWNLOAD YOUR FILE #

Official Symbol Interactor A	Official Symbol Interactor B	Experimental System	Experimental System Type	Author	Publication Source	Organism Interactor A	Organism Interactor B	Throughput
CDK6	PPARGC1A	Biochemical Activity	physical	Anders L (2011)	PUBMED:22094256	9606	9606	High Throughput
CAPNS1	PPARGC1A	Two-hybrid	physical	Wang J (2011)	PUBMED:21988832	9606	9606	High Throughput
KIAA1429	PPARGC1A	Affinity Capture-RNA	physical	Yue Y (2018)	PUBMED:29507755	9606	9606	High Throughput
EGFR	PPARGC1A	Affinity Capture-MS	physical	Foerster S (2013)	PUBMED:23956138	9606	9606	High Throughput|Low Throughput
PPARGC1A	NR1H4	Reconstituted Complex	physical	Zhang Y (2004)	PUBMED:14729567	9606	9606	Low Throughput
PPARGC1A	NR1H4	Affinity Capture-Western	physical	Zhang Y (2004)	PUBMED:14729567	9606	9606	Low Throughput
PPARGC1A	MED1	Affinity Capture-Western	physical	Wallberg AE (2003)	PUBMED:14636573	9606	9606	Low Throughput
PPARGC1A	MED16	Affinity Capture-Western	physical	Wallberg AE (2003)	PUBMED:14636573	9606	9606	Low Throughput

Refer to "BIOGRID-GENE-116097-3.5.184._PPARGC1A.txt"

STRING

https://string-db.org



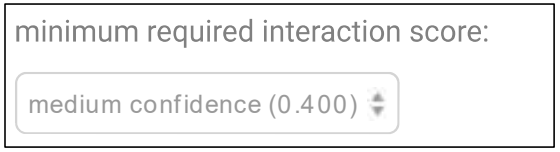
For each protein–protein association stored in STRING, a score is provided.

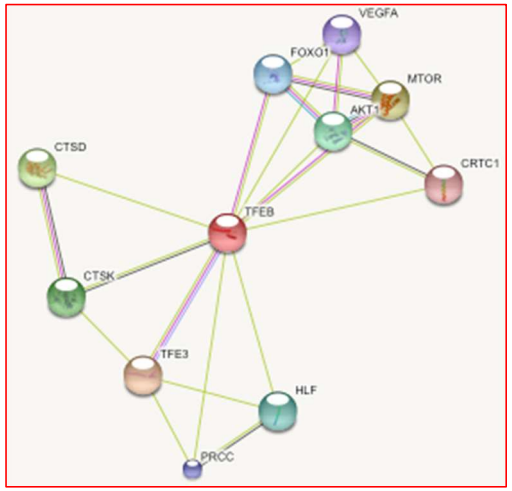
- These scores (i.e., the 'edge weights' in each network) are scaled between **zero and one**.
- They indicate the estimated likelihood that a given interaction is biologically meaningful, specific and reproducible, given the supporting evidence.

STRING does not discriminate between different splice isoforms or post-translationally modified forms.

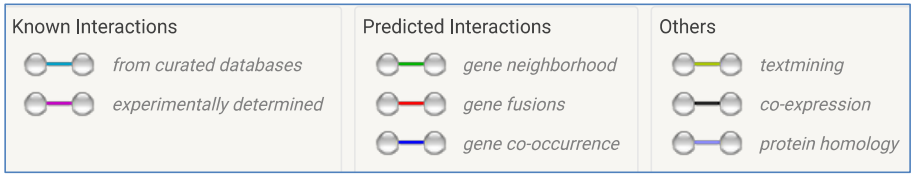
There are 4 different level of confidence and each of them have a specific score:

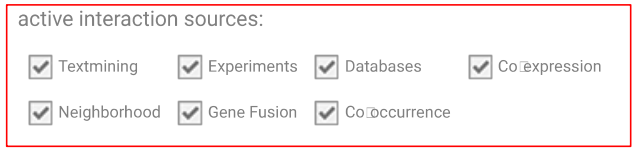
- Highest;
- High;
- Medium;
- Low





The 7 STRING SOURCES:

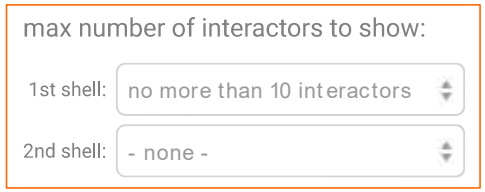




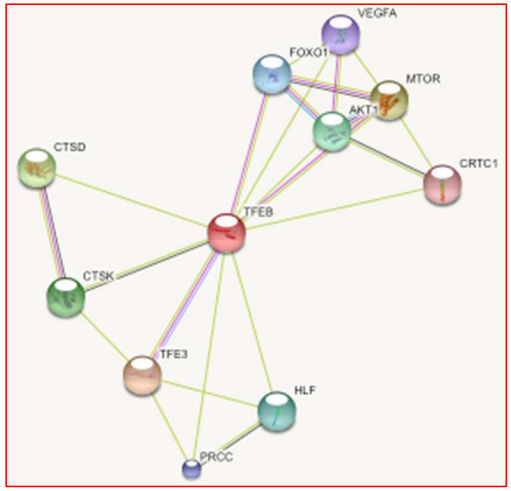
1. **Textmining**: shows a list of significant protein interaction groups, extracted from the scientific abstracts
2. **Experiments**: show a list of significant protein interaction datasets, gathered from other protein-protein interaction databases (**BioGRID**).
3. **Databases**: This view shows a list of significant protein interaction groups, gathered from curated databases.
4. **Co-expression**: shows the genes that are co-expressed in the same or in other species (by homology).
5. **Co-occurrence**: shows the presence or absence of linked proteins across species.
6. **Fusion**: Genes that are sometimes fused into single open reading frames.
7. **Neighborhood:** Groups of genes that are frequently observed in each other's genomic neighborhood

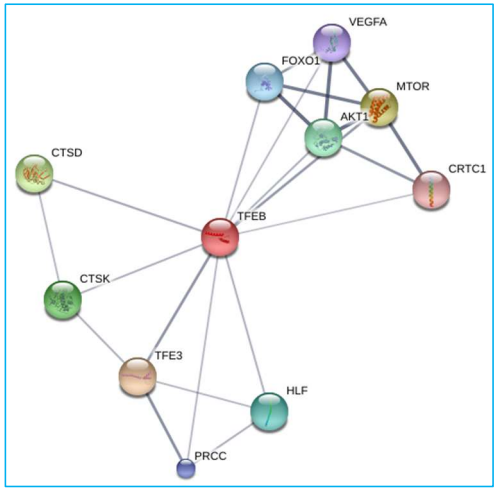
There are 5 different interaction choices:

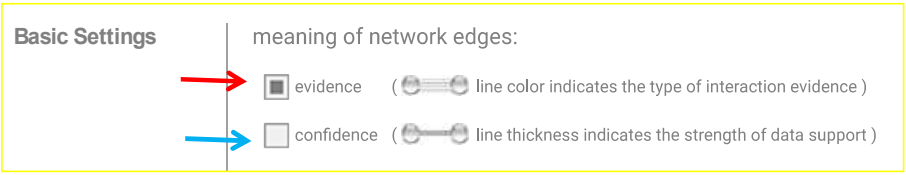
- NONE (query only);
- No more than 5 interactors;
- No more than 10 interactors;
- No more than 20 interactors;
- No more than 50 interactors
- Custom

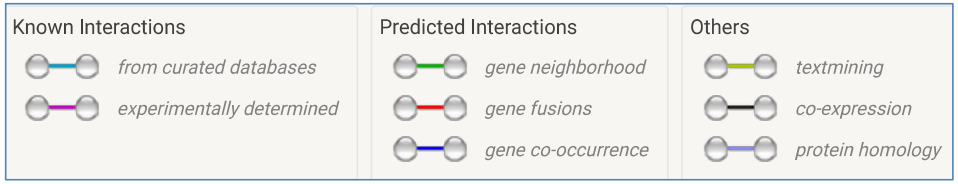


EDGEs visualization: EVIDENCE or CONFIDENCE

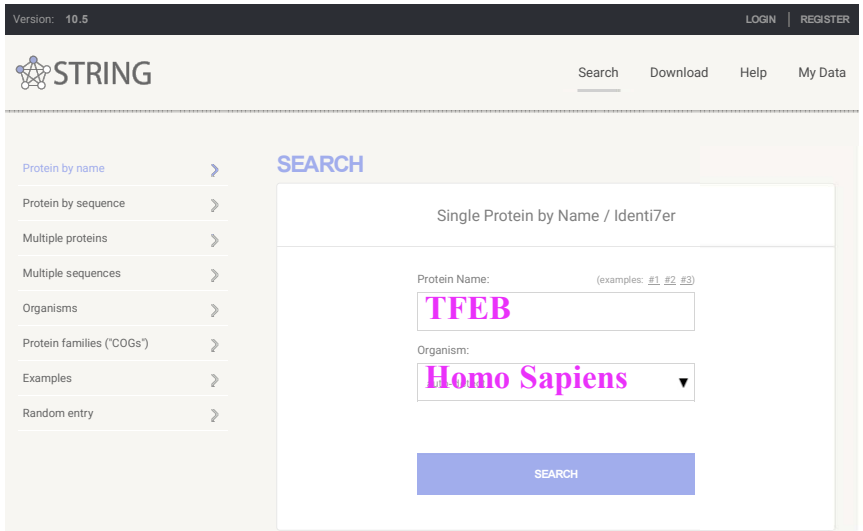


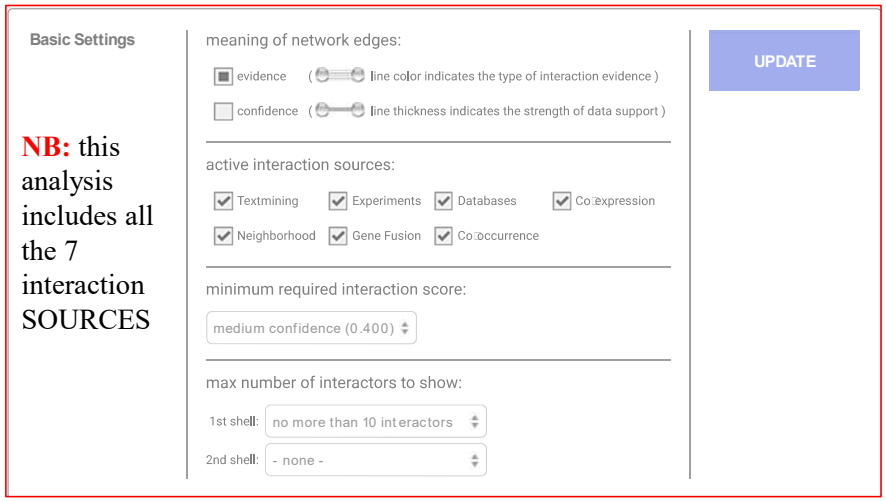






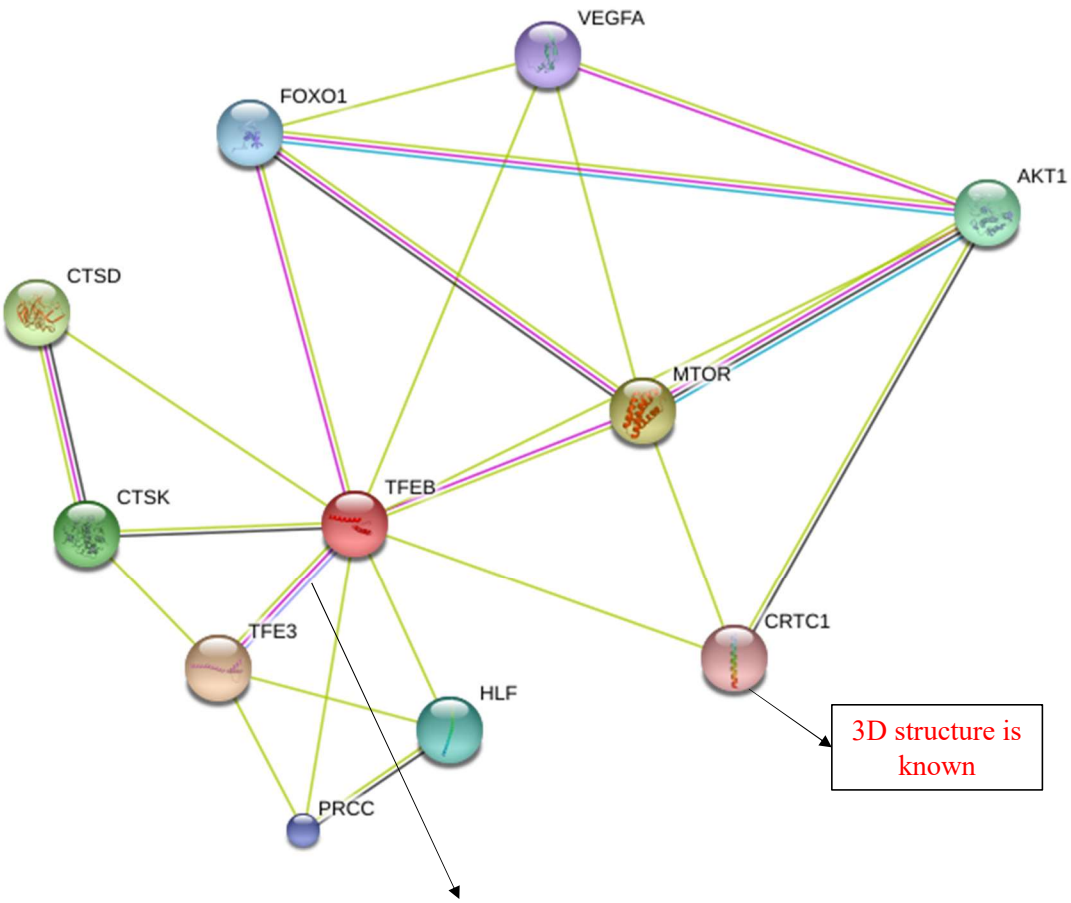
EXAMPLE: 1 gene of interest.





NB: this analysis includes all the 7 interaction SOURCES

NETWORK RESULTS:



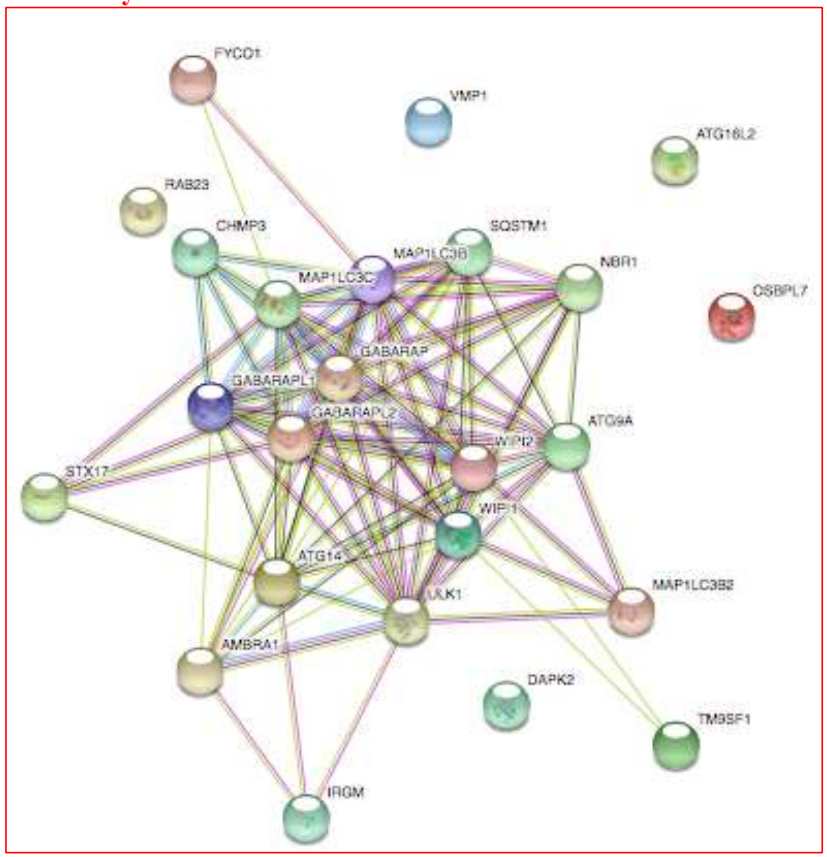
3D structure is known

It is possible to have info about:

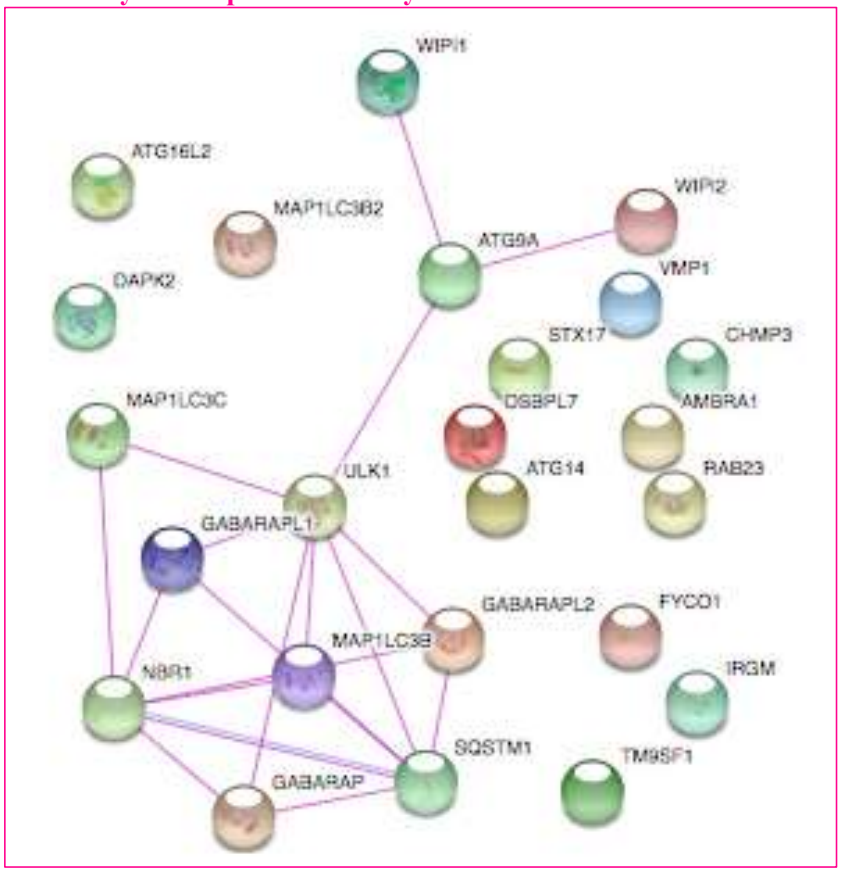
- Experimental validation of the PPI of interest
- Abstracts
- Alignment of the two sequences

Example_25genes_Autophagy_STRING

this analysis includes all the 7 interaction SOURCES



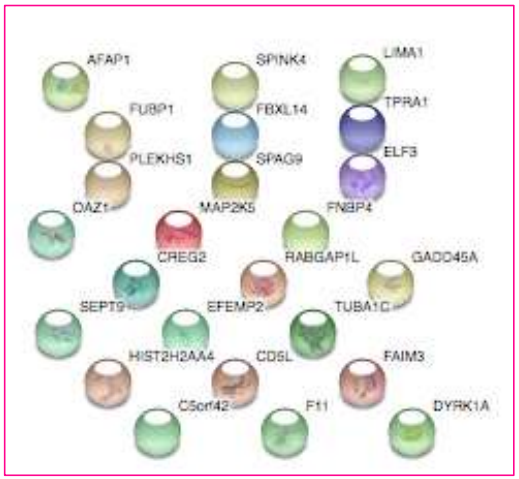
this analysis was performed only with the EXPERIMENTAL SOURCE



Known Interactions	Predicted Interactions	Others
from curated databases	*gene neighborhood*	*textmining*
experimentally determined	*gene fusions*	*co-expression*
	gene co-occurrence	*protein homology*

Example_25genes_Autophagy_STRING

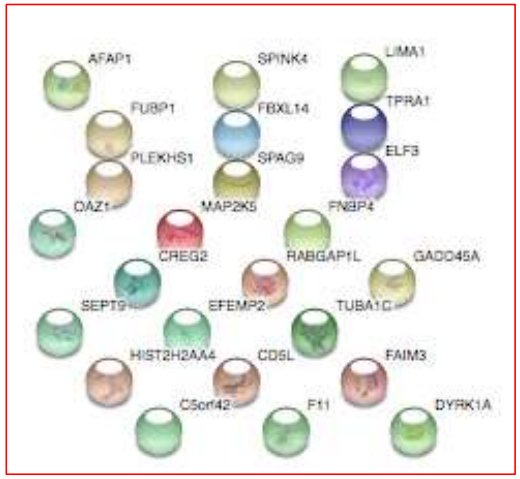
Example_25genes_RANDOM_STRING



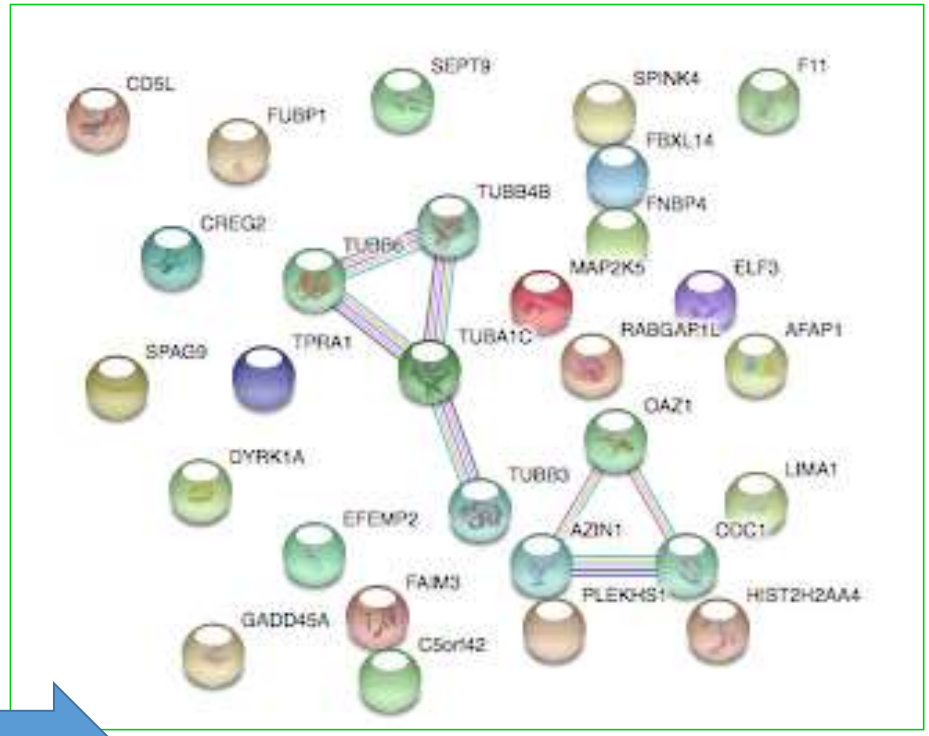
this analysis was performed only with the EXPERIMENTAL SOURCE

NO LINK found

Increase to 7 interaction SOURCES



NO LINK found



Increase the STRING available interactors to show:

max number of interactors to show:

1st shell: no more than 10 interactors

2nd shell: - none -

Example_25genes_RANDOM_STRING