

Tabloid Proteome

Version 1.2

Contents

1. How to Search Tabloid Proteome	2
1.1. Protein Search	2
1.1.1. Search with one protein.....	2
1.1.2. Search with two proteins.....	2
1.1.3. Multiple protein search.....	2
1.2. Gene Search	3
1.2.1. Search with one gene.....	3
1.2.2. Search with two genes	4
1.3. Pathway search:.....	5
1.4. Disease search:.....	6
2. Search Results	7
2.1. Data Table	7
2.1.1. Interactions	8
2.1.2. Common Projects.....	8
2.1.3. Pathways	9
2.1.4. Complexes.....	9
2.1.5. GO	9
2.1.6. Disease	10
2.2. Graph View.....	10
2.3. Special case	14
3. Downloads	14
4. Citation.....	15

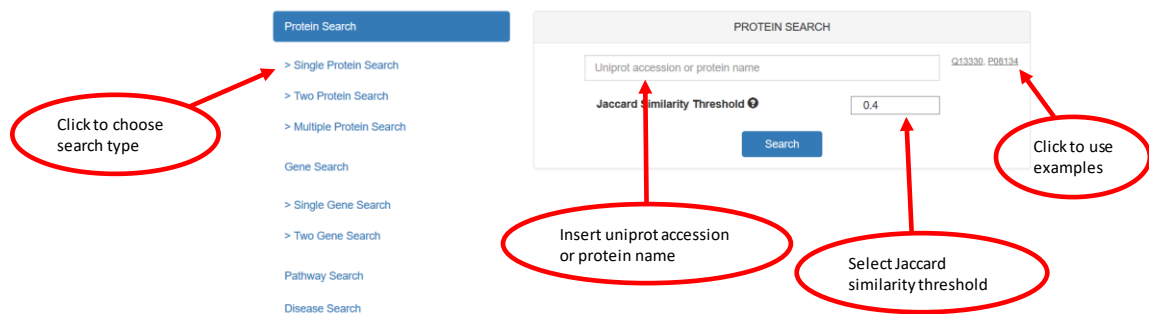
1. How to Search Tabloid Proteome

In order to perform the search, you can use the search box, provided in the search page. With the updated version of the Tabloid Proteome you can search for protein association in four ways; Protein Search, Gene Search, Pathway Search or Disease Search.

1.1. Protein Search

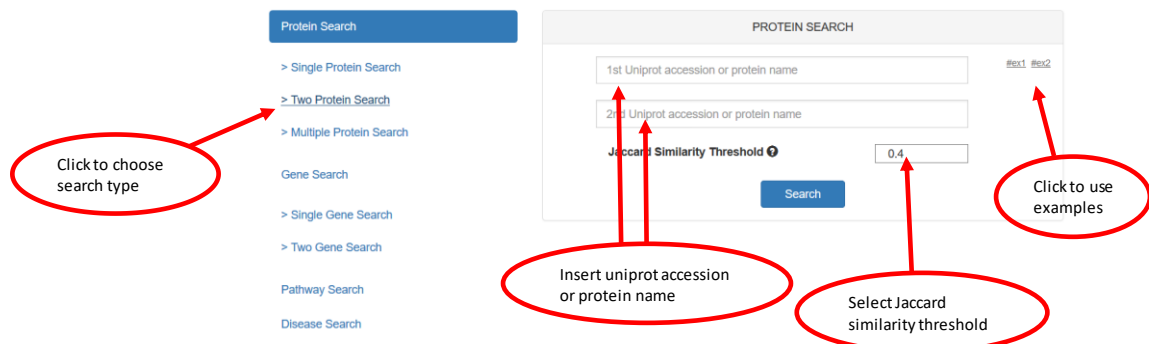
1.1.1. Search with one protein

You can search Tabloid Proteome with single protein, using Uniprot Accession or protein name. This will provide you with all possible partner of the protein that are found to be associated according to our analysis. Instead of using your own proteins, you can also use examples, provided in the upper right corner of the search box, to initiate the searches.



1.1.2. Search with two proteins

You can also start your search with two proteins, by using the 'Two Protein Search' tab in the search page. You need to define the Uniprot Accession or protein name of both proteins, as shown in image below. Similar to one protein search, you can also use examples.



1.1.3. Multiple protein search

If you have more than one protein pair to search, you can use 'Multiple Protein Search' tab under Protein Search. This will allow you to upload your own file with protein pairs as first and second

column. At the moment, multiple search only takes uniprot accession table (without header), with maximum three column.

Protein Search

- > Single Protein Search
- > Two Protein Search
- > Multiple Protein Search

Gene Search

- > Single Gene Search
- > Two Gene Search

Pathway Search

Disease Search

ONE TO ONE MULTIPLE PROTEIN SEARCH

Please upload a file that contains **protein pairs** and edge annotation (if available) in the following file formats: xls, xlsx, csv. The first two columns should be Uniprot accessions of the proteins. If you have edge annotation, you can add it to the 3rd column. The file should not include any title.

Select file to upload:

Choose file No file chosen

Include edge annotation

Search Clean

Moreover, if you would like to add your own annotation score in the resulting graph, you can include the score in the third column of the file, and click on button 'edge annotation'. You can upload file in any of the three format xls, xlsx, csv. If you have selected to add your own score, it will be shown in the third column upon upload, as shown below.

ONE TO ONE MULTIPLE PROTEIN SEARCH

1st Uniprot Accession	2nd Uniprot Accession	Edge Annotation
O15145	Q13404	0.45
O15230	P27658	0.38
O15379	A6NED2	0.56
O15511	P58546	0.42

Jaccard Similarity Threshold ⓘ 0.3

Select file to upload:

Choose file

Include edge annotation

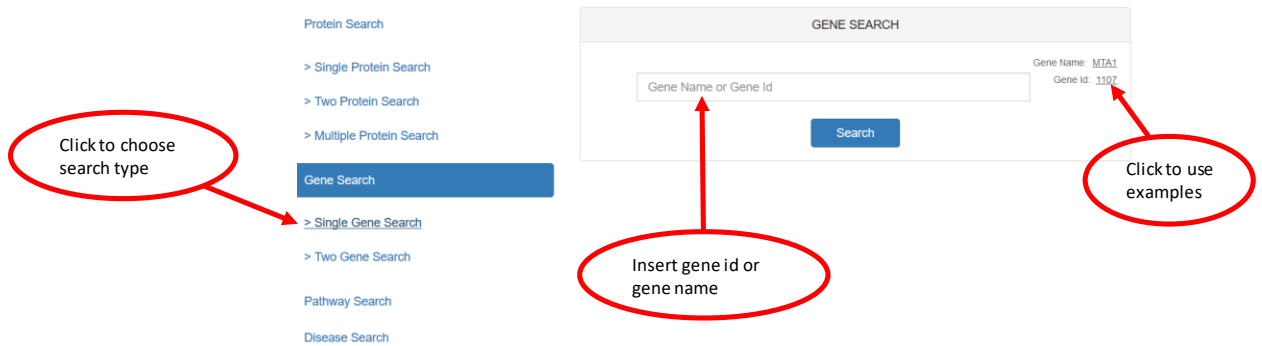
Search Clean

1.2. Gene Search

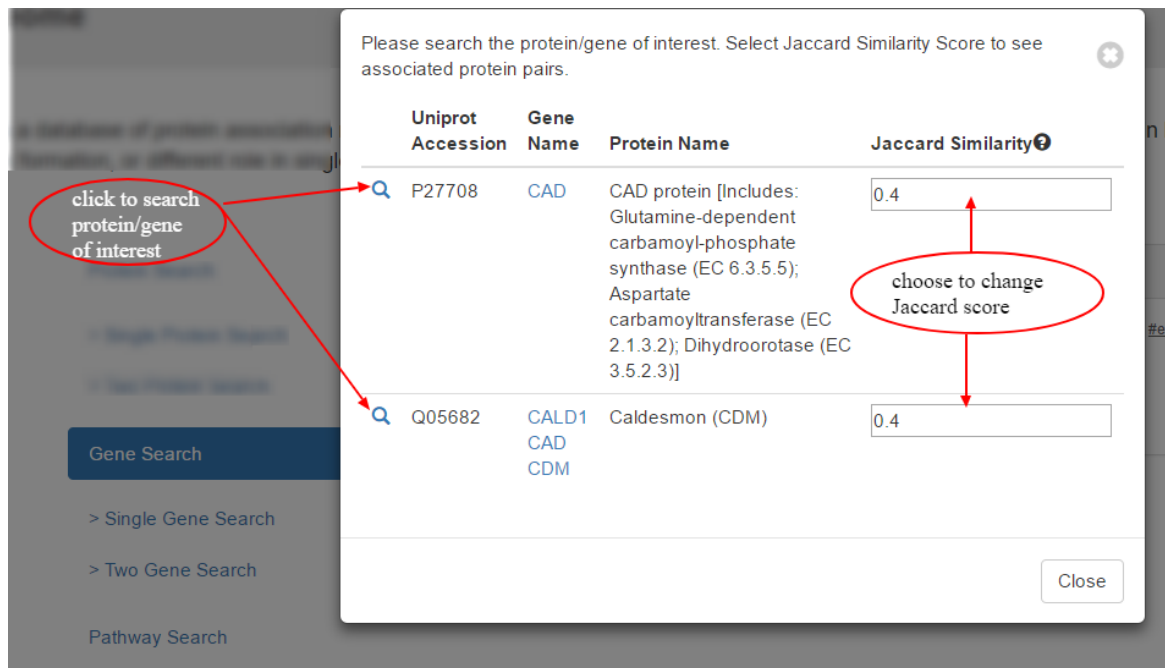
You can also search Tabloid Proteome either using gene name or NCBI RefSeq GeneID. Gene Search also allow you to search with up to two genes using with single or double gene search.

1.2.1. Search with one gene

To search with one gene, you can click on 'Single Gene Search' tab given under Gene Search. A search window for gene search with appear, as shown below. Here you can search for all possible proteins and their association with your gene of interest.



After clicking on search button, another window will appear, showing all possible proteins found for the entered gene. You can initiate your search by clicking on the search button next to the protein accession. In this window you can also change for Jaccard similarity threshold.



1.2.2. Search with two genes

Similar to the one gene search, you can also search with two genes, using 'Two Gene Search' tab.

Protein Search

- > Single Protein Search
- > Two Protein Search
- > Multiple Protein Search
- Gene Search**
 - > Single Gene Search
 - > Two Gene Search
- Pathway Search
- Disease Search

GENE SEARCH

1st Gene Name or Gene Id #ex1 #ex2

2nd Gene Name or Gene Id

Search

Click to choose search type

Insert gene id or gene name

Click to use examples

The next window will allow you to search with your interesting protein pairs and with the search button next to the protein pair you can initiate the search (as shown below).

Total number of protein pairs found: 1

Please select the protein/gene pair of interest

	1st Uniprot Accession	1st Gene Name	1st Protein Name	2nd Uniprot Accession	2nd Gene Name	2nd Protein Name	
	Q13330	MTA1	Metastasis-associated protein MTA1	Q12873	CHD3	Chromodomain-helicase-DNA-binding protein 3 (CHD-3) (EC 3.6.4.12) (ATP-dependent helicase CHD3) (MI-2 autoantigen 240 kDa protein) (MI2-alpha) (Zinc finger helicase) (hZFH)	0.43

Click here to search

1.3. Pathway search:

You can also search for all associated protein pairs available in Tabloid Proteome, for the particular pathway. For the current version of Tabloid Proteome, only include lowest level of pathways from Reactome Database. The search can be initiated either by unique Reactome ID or pathway name.

Protein Search

- > Single Protein Search
- > Two Protein Search
- Gene Search
 - > Single Gene Search
 - > Two Gene Search
- Pathway Search**
- Disease Search

PATHWAY SEARCH

Reactome accession or pathway name #ex1 #ex2

Search

click to choose search type

insert reactome accession or pathway name

click to search with example

After the selection of the pathway, in the next window you can see multiple pathway entries. Select for your pathway of interest and click on the down arrow on the right side of the pathway name. The click on the down arrow will show you all possible protein pairs found in the Tabloid Proteome with

their Jaccard Similarity score. Select the protein pair of interest and initiate the search, by clicking on search button next to the protein accession (as shown in the image below).

Please click on the pathway of interest to find relevant proteins, and search the relevant protein pairs

Pathway Name	Reactome Accession	Jaccard Similarity		
Signaling by FGFR3 point mutants in cancer	R-HSA-8853338			
1st UniProt Accession	1st Protein Name	2nd UniProt Accession	2nd Protein Name	Jaccard Similarity
P01112	GTPase HRas (H-Ras-1) (Ha-Ras) (Transforming protein p21) (c-H-ras) (p21ras) [Cleared into: GTPase HRas, N terminally processed]	P01111	GTPase NRas (Transforming protein N-Ras)	0.29
P01111	GTPase NRas (Transforming protein N-Ras)	P01115	GTPase KRas (K-Ras-2) (K-Ras) (c-K-ras) [Cleared into: GTPase KRas, N terminally processed]	0.18
SMAD2/3 Phosphorylation Motif Mutants In Cancer	R-HSA-3304366			
SMAD2/3 MH2 Domain Mutants In Cancer	R-HSA-3315487			
Constitutive Signaling by Aberrant PI3K In Cancer	R-HSA-2219538			
Signaling by FGFR3 fusions in cancer	R-HSA-8853334			
Constitutive Signaling by AKT1 E17K In Cancer	R-HSA-5674400			
SMAD4 MH2 Domain Mutants In Cancer	R-HSA-3311021			
Defective GALNT12 causes colorectal cancer 1 (CRC51)	R-HSA-5883636			
TGFBR1 KD Mutants In Cancer	R-HSA-3656532			
Loss of Function of FBXW7 In Cancer and NOTCH1 Signaling	R-HSA-2844607			
TGFBR1 LBD Mutants In Cancer	R-HSA-3656535			
TGFBR2 M51 Frameshift Mutants In Cancer	R-HSA-3642279			
Constitutive Signaling by Ligand Responsive EGFR Cancer Variants	R-HSA-1336382			
TGFBR2 Kinase Domain Mutants In Cancer	R-HSA-3646796			

Close

1.4. Disease search:

Similar to the pathway search, you can also search for the associated protein pairs available in the Tabloid Proteome for particular disease. Tabloid Proteome include all human diseases from DisGenNET database. Here you can search, by clicking on 'Disease Search', either by using the disease name or unique DisGeNET ID.

Protein Search

- > Single Protein Search
- > Two Protein Search

Gene Search

- > Single Gene Search
- > Two Gene Search

Pathway Search

Disease Search

DISEASE SEARCH

#ex1 #ex2

Search

Insert DisGeNET id or disease name

click to search with example

After the selection of the disease, in the next window it will show you, all possible disease entries found for your search (as shown below). With the down arrow you can check for all possible associated protein pairs with minimum 0.4 Jaccard similarity score. With the protein pair of interest you can initiate the search, by clicking on the search button next to the Unique DisGeNET Id.

Please click on the disease of interest to find relevant proteins, and search the relevant protein pairs

Disease Name		DisGeNet Id	
Precursor cell lymphoblastic lymphoma		C1079748	
1st Uniprot Accession	1st Protein Name	2nd Uniprot Accession	2nd Protein Name
P08196	CD4 cell-surface antigen heavy chain (4F2c) (4F2 heavy chain antigen) (Lymphocyte activation antigen 4F2 large subunit) (Solite carrier family 3 member 2) (CD antigen CD98)	P33993	DNA replication licensing factor MCM7 (EC 3.6.4.12) (CDC47 homolog) (P1.1-MCM3)
P42517	48S ribosomal protein S27 (Metalloprotein-stimulin 1) (MPS-1)	P33993	DNA replication licensing factor MCM7 (EC 3.6.4.12) (CDC47 homolog) (P1.1-MCM3)
P58194	4F2 cell-surface antigen heavy chain (4F2c) (4F2 heavy chain antigen) (Lymphocyte activation antigen 4F2 large subunit)	P33993	Large neutral amino acids transporter small subunit 1 (4F2 light chain) (4F2 LC) (CD98 light chain) (Integral membrane protein F 50) (Lysine amino acid transporter 1) (LNT1) (Solite carrier family 7 member 6) (Lysine carrier)
Precursor B-lymphoblastic lymphoma/leukemia		C0862630	
Precursor B-lymphoblastic lymphoma stage I		C0862639	
Precursor B-lymphoblastic lymphoma		C0855146	
Precursor T-cell lymphoblastic leukemia/lymphoblastic lymphoma		C1301359	
adult lymphoblastic lymphoma		C0278721	
Precursor T-cell lymphoblastic lymphoma		C1292758	

Close

2. Search Results

If the protein(s) searched either through protein, gene, pathway or search, pathway search is present in the database and have association with any another protein, Tabloid Proteome will provide result in two forms; Data Table and Graph View.

2.1. Data Table

Data table view gives association information between two or more proteins. If searched for one protein, the resulting table will show you one or more rows, depending on the association found for the searched protein. In the data table, each row represents the possible biological connection for each partner proteins (as shown below). For the double protein search, the resulting table will provide user with one row, showing possible biological connection between two proteins.

Protein Association Data Table

Download Graph View

1st protein	2nd protein	1st protein gene name	2nd protein gene name	Jacc similarity score	Interaction	Ensembl paralog	# common complex	# common pathway	# common projects	# common diseases
Q13330	Q8TD00	MTA1	Chromodomain-helicase-DNA-binding protein 5 (CHD-5) (EC 3.6.4.12) (ATP-dependent helicase CHD5)	0.43	yes	no	0	0	5	20
Q13330	Q12873	MTA1	Chromodomain-helicase-DNA-binding protein 3 (CHD-3) (EC 3.6.4.12) (ATP-dependent helicase CHD3) (M2)	0.43	yes	no	3	3	6	2
Q13330	Q8V909	MTA1	Transcriptional repressor p66-beta (GATA zinc finger domain-containing protein 26) (p66/p68)	0.48	yes	no	0	3	7	8
Q13330	Q96573	MTA1	Paired amphipathic helix protein Sin3a (Histone deacetylase complex subunit Sin3a) (Transcriptional repressor Sin3a)	0.4	yes	no	1	0	6	9

Uniprot accession column (blue) provides direct link to Uniprot website. You can also sort each column with high to low score or with alphabetical order. With the click on each protein pair, you can find more information about the association between two proteins (as shown in image below) like Interactions, Projects, Pathways, Complexes, Disease, and GO.

Protein Association Data Table

Download Graph View

1st protein ID	2nd protein ID	1st protein name	1st protein gene name	2nd protein name	2nd protein gene name	Jacc similarity score	Interaction	Ensembl paralog	# common complex	# common pathway	# common projects	# common diseases
Q13330	Q8TD00	Metastasis-associated protein MTA1	MTA1	Chromodomain-helicase-DNA-binding protein 5 (CHD-5) (EC 3.6.4.12) (ATP-dependent helicase CHD5)	CHD5, KIAA0444	0.43	yes	no	0	0	5	20
Q13330	Q12873	Metastasis-associated protein MTA1	MTA1	Chromodomain-helicase-DNA-binding protein 3 (CHD-3) (EC 3.6.4.12) (ATP-dependent helicase CHD3) (M-2 autoantigen; 240 kDa protein) (M2-alpha) (Zinc finger helicase) (ZNFH)	CHD3	0.43	yes	no	3	3	6	2
Q13330	Q8V009	Metastasis-associated protein MTA1	MTA1	Transcriptional repressor p66-beta (GATA zinc finger domain-containing protein 2B) (p66beta)	GATAD2B, KIAA1150	0.48	yes	no	0	3	7	8

Interactions

IntAct : yes BioGRID : yes

Intact Confidence	Intact Detection	Intact Interaction Type
0.35	MI0007_antib tag immunoprecipitation	MI0914_association

Common Projects

Pathways

Complexes

GO

Diseases

Q13330	Q8V009	Metastasis-associated protein MTA1	MTA1	Paired amphiphilic helix protein Sin3a (Histone deacetylase complex subunit Sin3a) (Transcriptional repressor Sin3a)	SIN3A	0.4	yes	no	1	0	6	9
--------	--------	------------------------------------	------	--	-------	-----	-----	----	---	---	---	---

With the further click on the expandable buttons, at the end of each row, you can get more information about each type of association, in separate section. More information about each section is described below.

2.1.1. Interactions

Interaction section gives information about protein-protein interaction between two proteins. All the information about the interaction is collected from IntAct and BioGRID database. If the interaction is found in the IntAct database, then more information about the interaction is shown in this section (as shown in image below).

Interactions

IntAct : yes BioGRID : no

Intact Confidence	Intact Detection	Intact Interaction Type
0.59	MI0398_two hybrid pooling approach	MI0915_physical association
0.59	MI0018_two hybrid	MI0915_physical association
0.59	MI0663_confocal microscopy	MI0403_colocalization

2.1.2. Common Projects

Common Projects section shows all common projects between two proteins (shown in image below). By clicking on one of the project accessions (marked with blue), it will redirect to PRIDE website.

Common Projects

No	Project Accession	Keywords	Tissue	Tags	Instruments
1	PXD000314	Sequential IMAC, human lung cancer tissue, phosphoproteomics	cell culture, lung		instrument model
2	PXD001572	DNA sensing, protein interactions, IFX	cell culture	Biomedical	LTQ Orbitrap Velos
3	PRD000711	Not available	293T cells	Technical, PRIME-XS Project	instrument model, LTQ Orbitrap Velos

2.1.3. Pathways

Pathways section shows the common pathways found between two proteins, according to Reactome database. By clicking on Reactome accession (marked in blue), it will re-direct to Reactome's 'Pathway Browser' page.

No	Reactome Accession	Pathway Name	Evidence Code
1	R-HSA-1445148	Translocation of GLUT4 to the plasma membrane	TAS
2	R-HSA-75035	Chk1/Chk2(Cds1) mediated inactivation of Cyclin B:Cdk1 complex	TAS
3	R-HSA-5625740	RHO GTPases activate PKNs	TAS
4	R-HSA-5628897	TP53 Regulates Metabolic Genes	TAS

2.1.4. Complexes

Complexes section, as shown below in the image, will provide you with list of complexes where both proteins are present, according to CORUM database. Here, with the click to Pubmed Id, it will re-direct to the link to PubMed webpage and click on the CORUM id, (marked in blue) will re-direct to the CORUM database.

No	Corum Id	Complex Name	Complex Comment	Cell Line	Disease Comment	Subunit Comment	Pubmed Id	Purification Method
1	5199	Kinase maturation complex 1	None	None	None	None	14743216	MI:0676- tandem affinity purification
2	5615	Emerin complex 52	Complexes are named on the basis of their S300 elution fraction number. Subunits 8-22 were	None	Emerin is involved in Emery-Dreifuss muscular dystrophy	None	17820012	MI:0004- affinity chromatography technologies, MI:0226- ion exchange

2.1.5. GO

Under this section, you can find information regarding common Biological Process, Molecular Function and Cellular Component of two proteins, defined by GO. These three information is provided with three different subsection with the respective names (as shown in image below). All section includes GO Id (marked with blue) which will re-direct to EMBL-EBI Gene Ontology service.

Molecular Function		
No	MF Id	MF Name
Biological Process		
No	BP Id	BP Name
1	GO:0061024	membrane organization
2	GO:1900740	positive regulation of protein insertion into mitochondrial membrane involved in apoptotic signaling pathway
3	GO:0006605	protein targeting
Cellular Component		
No	CC Id	CC Name
1	GO:0005739	mitochondrion
2	GO:0016020	membrane
3	GO:0005925	focal adhesion
4	GO:0070062	extracellular exosome

2.1.6. Disease

In the disease section, you can find the list of diseases, where both proteins are found to be involved in. All the information about the disease and protein involved in each disease are collected from DisGeNET database. With the click on the DisGeNET ids (marked with blue), it will re-direct to DisGeNET search webpage.

No	Disgenet Id	Disease Name
1	C0025202	melanoma
2	C2239176	Liver carcinoma
3	C0242379	Malignant neoplasm of lung
4	C0376358	Malignant neoplasm of prostate

2.2. Graph View

In Tabloid proteome, you can also visualize your results in form of graph, where each node represent proteins and edges represents associations. The link to the graph can be accessed through the upper right corner of the Data Table result page.

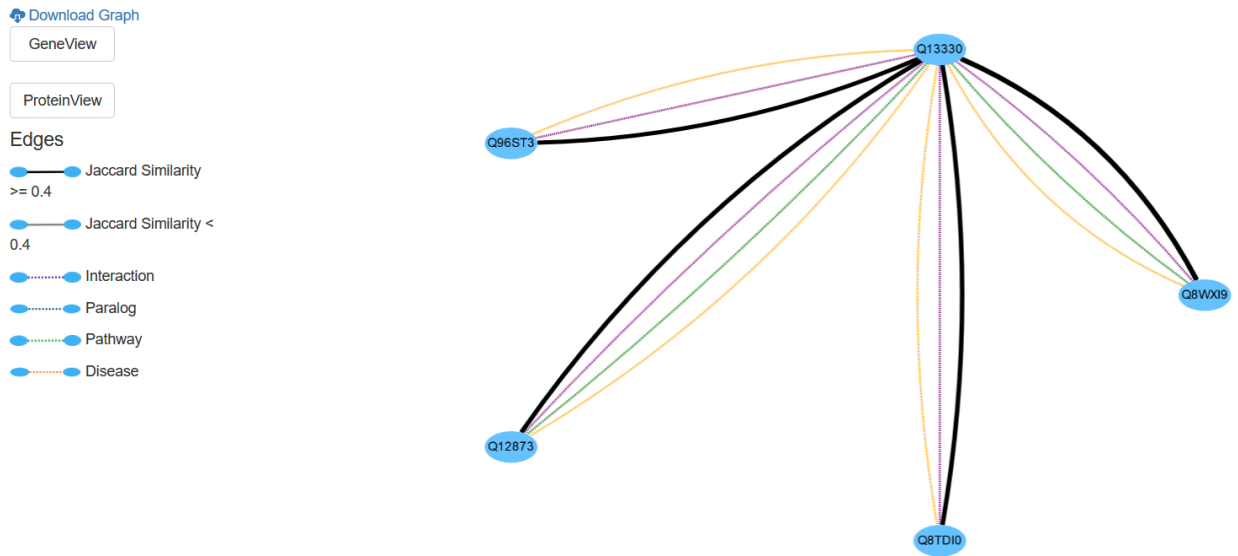
Protein Association Data Table

[Download](#) [Graph View](#)

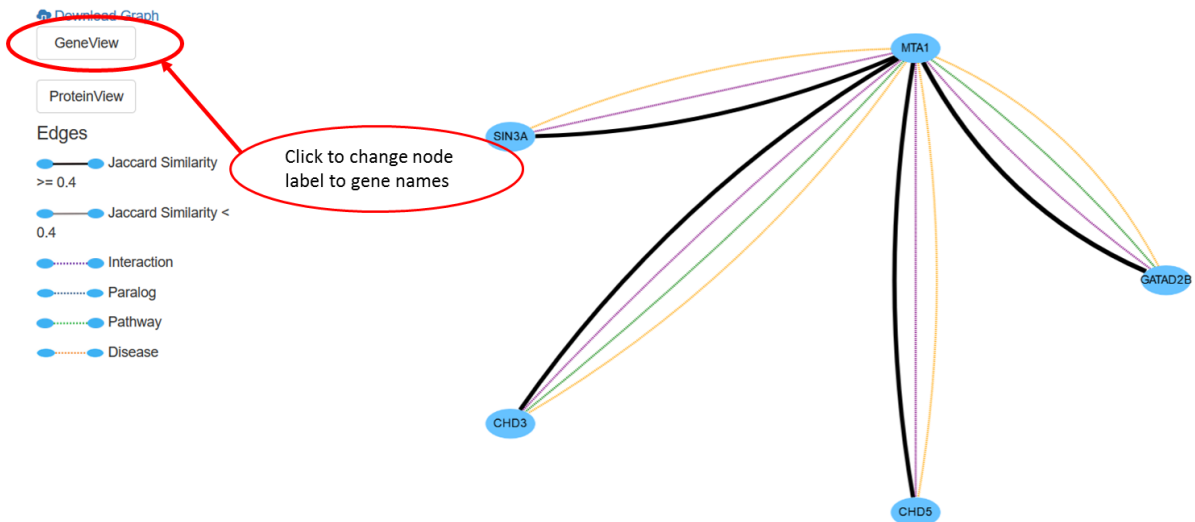
2nd protein name #	2nd protein gene name #	Jacc similarity score #	Interaction #	Ensembl paralog #	# common complex #	# common pathway #	# common projects #	# common diseases #
Chromodomain-helicase-DNA-binding protein 5 (CHD-5) (EC 3.6.4.12) (ATP-dependent helicase CHD5)	CHD5, KIAA0444	0.43	yes	no	0	0	5	20
Chromodomain-helicase-DNA-binding protein 3 (CHD-3) (EC 3.6.4.12) (ATP-dependent helicase CHD3) (M-2 autoantigen 240 kDa protein) (M2-alpha) (Zinc finger helicase) (hZFH)	CHD3	0.43	yes	no	3	3	6	2
Transcriptional repressor p66-beta (GATA zinc finger domain-containing protein 28) (p66/p68)	GATAD2B, KIAA1150	0.48	yes	no	0	3	7	8
Paired amphipathic helix protein Sin3a (Histone deacetylase complex subunit Sin3a) (Transcriptional corepressor Sin3a)	SIN3A	0.4	yes	no	1	0	6	9

Click to go to Graph View

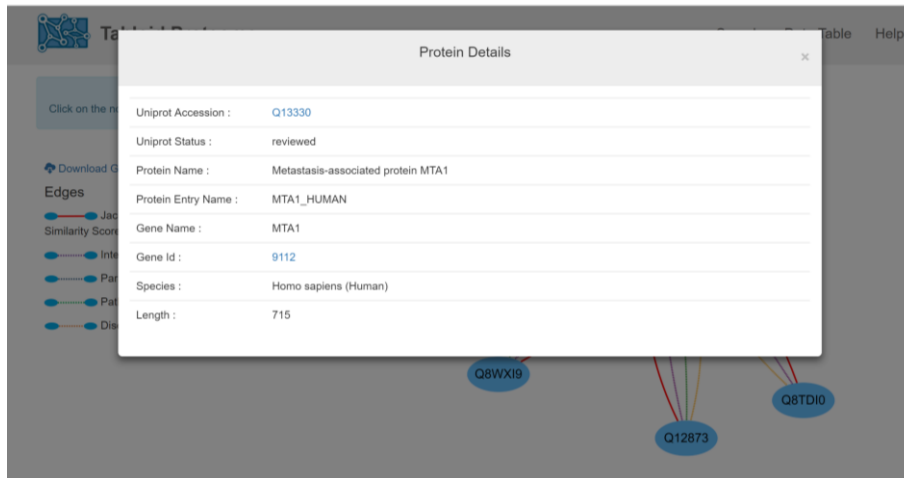
On this page, you can see dynamic graph representation of your search. All nodes and edges are repositionable. After/before changing position of the graph you can download as an image (.png) by clicking on “Download Graph” link (as shown below). For more information about the proteins and their relations, you need to click on the node or edge.



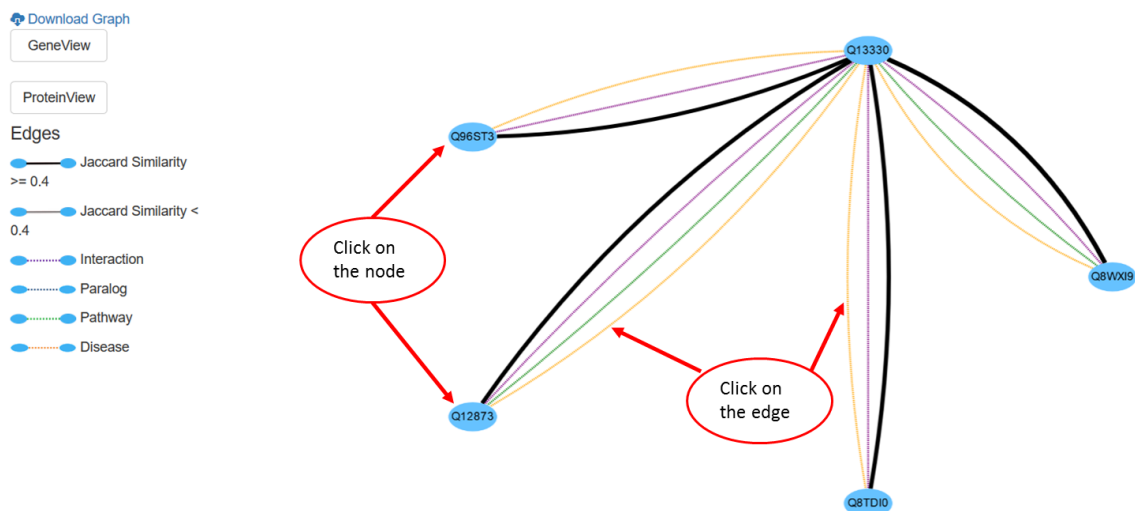
With default the nodes will be shown with the uniprot accession, however, you can change the visualization to the gene name using the “GeneView” button on the left side of the panel, as highlighted in the red in the image below.



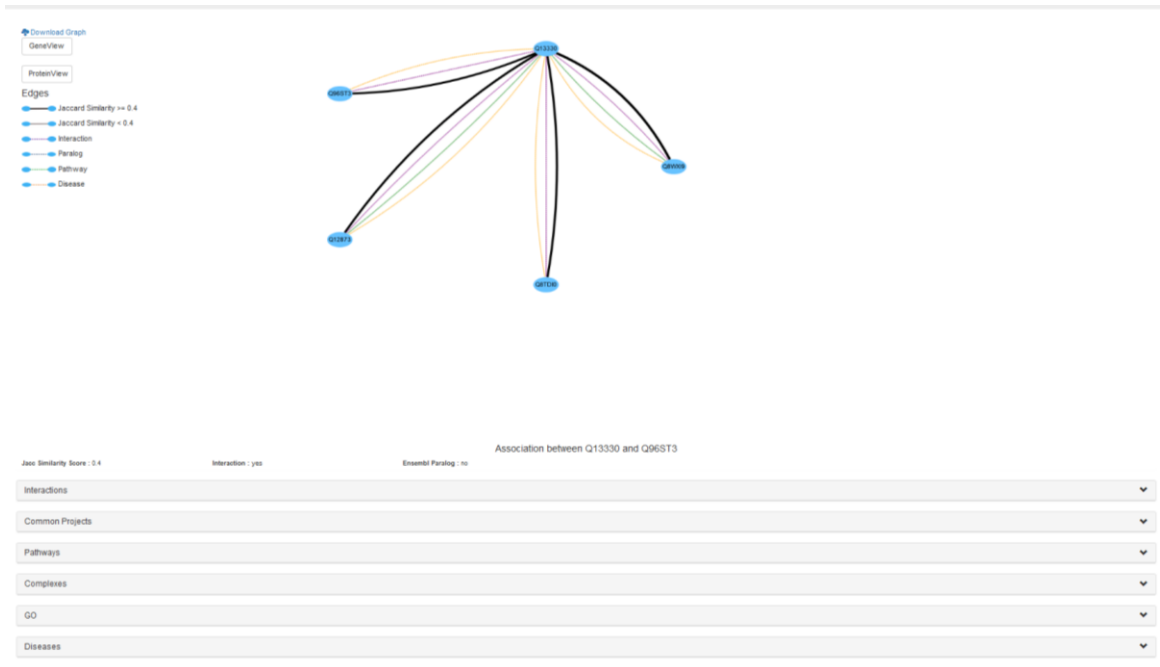
With the click on the node, a pop up window will appear which will show more information about the protein, derived from Uniprot database (as shown in the image below). With the click on the uniprot accession (marked with blue), it will re-redirect to the Swissprot/Uniprot webpage. Similarly, click on Gene Id (marked with blue), it will re-redirect to detailed view of the gene on NCBI Gene & Expression website.



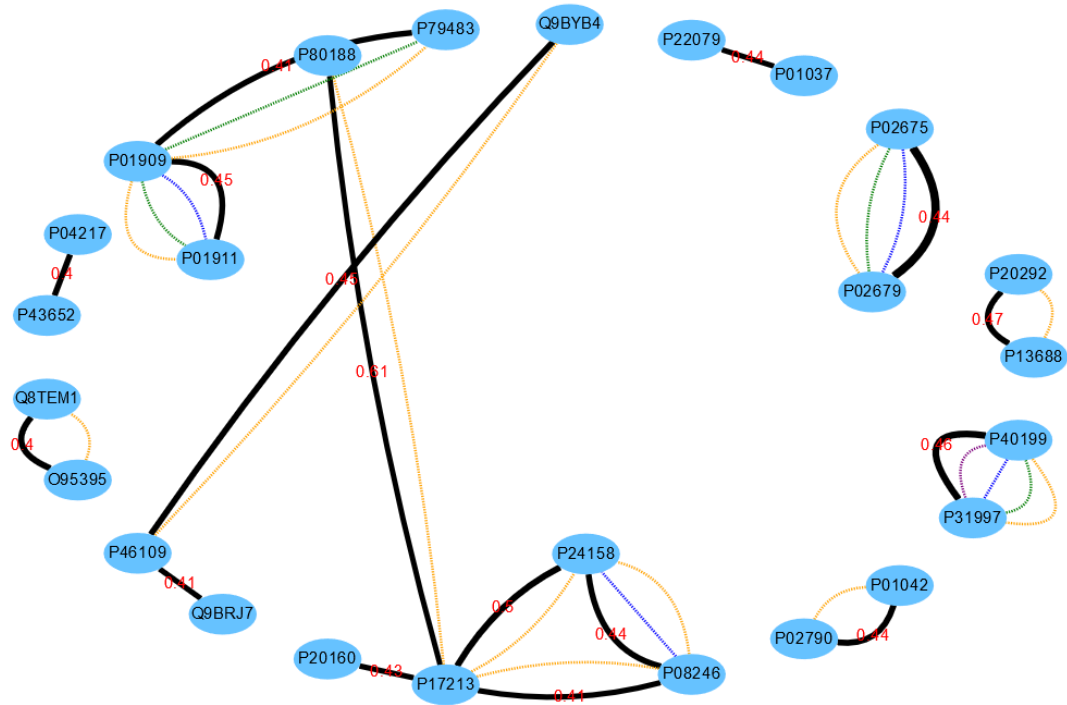
The edges between the proteins represents biological association, where the number of edges between two proteins can vary from 1 to 5. Different color of edge represents different biological association between proteins. Edge with red color represent the Jaccard similarity association found with our method, purple shows interaction, blue represent paralog, green represent that these proteins are present in same pathway, and finally orange color represent presence of both protein in same disease. This information is also shown in a panel, on the left side of the graph (as shown in the image below).



For more information about association, you can click on the edge, it will show brief information about the association, with six sub-sections; Interactions, Projects, Pathways, Complexes, Diseases, and GO (at the bottom of the graph). Each sub-section, contains same information as provided in the Data Table page (explained in more detail under section 2.1.1 to 2.1.6).

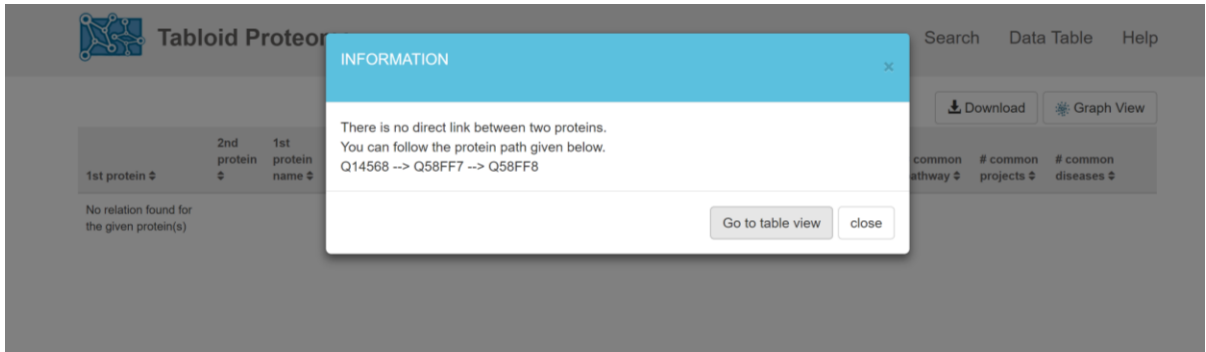


For multiple protein search, if edge annotation are added, then it will be shown in the graph, on the edges (red color). Graph shown below shows an example of the multiple protein search with added annotation.



2.3. Special case

In the double protein search, if this pair does not contain any direct biological connection in the database, Tabloid proteome will try to check for indirect association between the proteins. If the searched pair contains any indirect association, you will show a pop up window for the possible connection (as shown in the image below).



For further click on the “Data Table” option, you can view the result about the search in the Data Table view, with all information explained above. With the Data Table view, you can go to Graph View, to visualize you result with the same information and changes, as explained above.

3. Downloads

From Tabloid proteome, you can download the results, in two form; tab delimited format for the table or as an image. Tab delimited or CSV format can be downloaded from the upper right corner of Data table result page.

Total number of protein pairs found: 16
 1 protein pair(s) cannot be found.
 Please click on the protein pair of interest

Click to download in CSV or TSV format

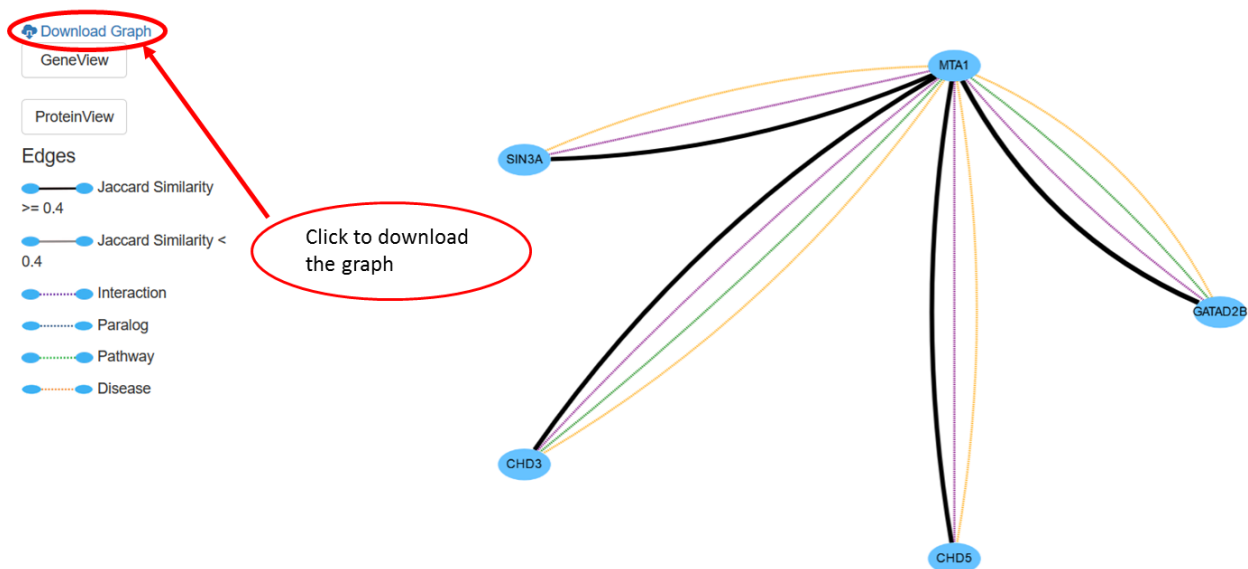
Download ▾ Graph View

TSV format
CSV format

Protein Association Data Table

1st protein	2nd protein	1st protein name	1st protein gene name	2nd protein name	2nd protein gene name	Jacc similarity threshold	Interaction	Ensembl paralog	# common complex	# common pathway	# common projects	# common diseases
P20292	P13688	Arachidonate 5-lipoxygenase-activating protein (FLAP) (MK-886-binding protein)	ALOX5AP FLAP	Carcinoembryonic antigen-related cell adhesion molecule 1 (Biliary glycoprotein 1) (BGP-1) (CD antigen CD66a)	CEACAM1 BGP BGP1	0.44	no	no	0	0	2	7
P40199	P31997	Carcinoembryonic antigen-related cell adhesion molecule 6 (Non-specific crossreacting antigen) (Normal cross-reacting antigen) (CD antigen CD68c)	CEACAM6 NCA	Carcinoembryonic antigen-related cell adhesion molecule 8 (CD67 antigen)	CEACAM8 CGM8	0.5	yes	yes	0	2	3	2

As mention above, you can also download the graph view from the upper left corner of the graph view page.



4. Citation

If you are using Tabloid Proteome, do not forget to cite us.

Surya Gupta, Kenneth Verheggen, Jan Tavernier, and Lennart Martens

Unbiased Protein Association Study on the Public Human Proteome Reveals Biological Connections between Co-Occurring Protein Pairs

Journal of Proteome Research **2017** 16 (6), 2204-2212

DOI: [10.1021/acs.jproteome.6b01066](https://doi.org/10.1021/acs.jproteome.6b01066)

PMID: 28480704