Tabloid Proteome

Version 1.4

Contents

1. How to Search Tabloid Proteome	3
1.1. Protein Search	3
1.1.1. Search with one protein	3
1.1.2. Search with two proteins	3
1.1.3. Multiple protein search	4
1.2. Gene Search	5
1.2.1. Search with one gene	5
1.2.2. Search with two genes	6
1.3. Pathway search:	6
1.4. Disease search:	7
1.5. Tissue Search	8
2. Search Results	9
2.1. Data Table	9
2.1.1. Interactions	10
2.1.2. Common Projects	11
2.1.3. Pathways	11
2.1.4. Complexes	11
2.1.5. GO	12
2.1.6. Disease	12
2.2. Graph View	12
2.3. Special case	17
3. Downloads	18
4. Citation	19





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.



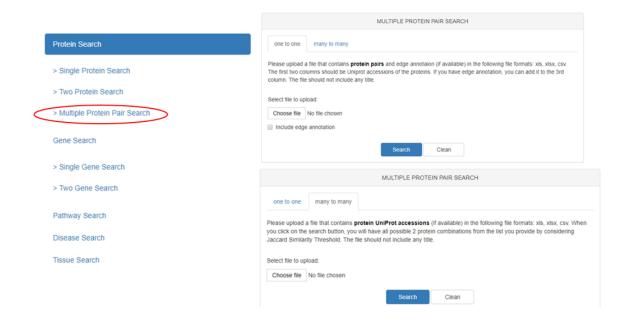


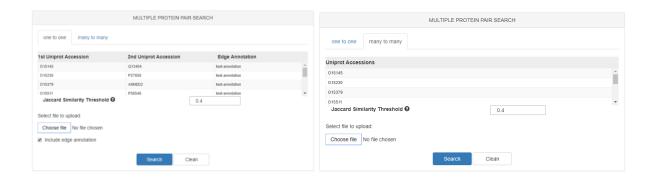
1.1.3. Multiple protein search

Multiple protein search provides two different ways to upload your own file. In the one-to-one section, you can upload file with protein pairs as the first and the second column. The pairs that are found in database considering the Jaccard threshold value, will be shown. 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 the button 'edge annotation'.

In the many-to-many section, you can upload a file that contains proteins and you will have all possible 2 protein combinations from the list you provide by considering Jaccard Similarity Threshold.

At the moment, multiple search only takes uniprot accession table (without header), with maximum three column. You can upload file in any of the three format xls, xlsx, csv.







1.2. Gene Search

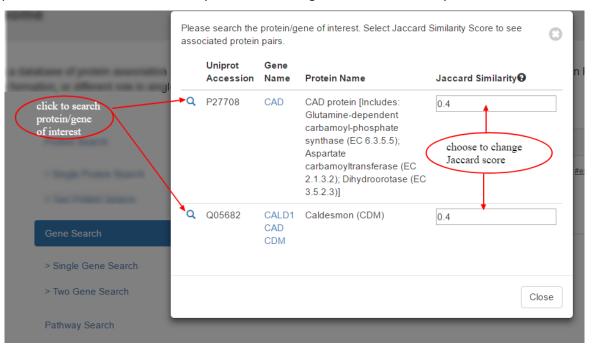
You can also search Tabloid Proteome either using gene name or NCBI RefSeq GeneID. Gene Search also allow you to search with up to two genes using with single or double gene search.

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.



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).



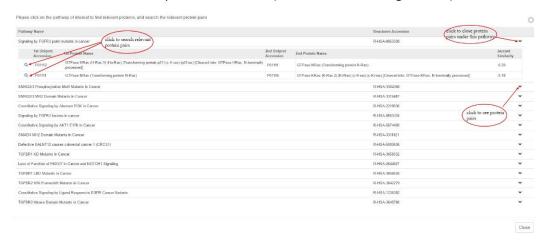
1.3. Pathway search:

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.





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).



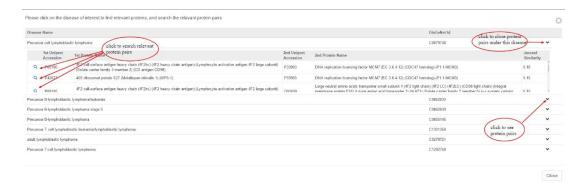
1.4. Disease search:

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.



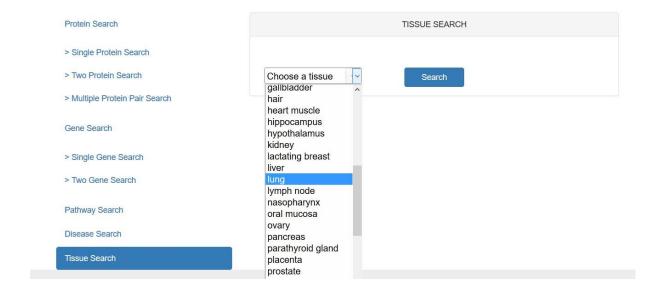
After the selection of the disease, in the next window it will show you, all possible disease entries found for your search (as shown below). With the down arrow you can check for all possible associated protein pairs with minimum 0.4 Jaccard similarity score. With the protein pair of interest you can initiate the search, by clicking on the search button next to the Unique DisGeNET Id. For protein pairs less than 0.4 can be accessed (for refer via API. more detail here http://iomics.ugent.be/tabloidproteome/tabloidApi.html)





1.5. Tissue Search

The online Tabloid Proteome also allows user to restrict their search for associated protein pair with the tissue annotations, included from The Human Protein Atlas. The tissue based searches can be initiated by selecting one of the provided 53 tissue names, from the drop down menu, for example lungs as, as shown in the figure below.



After selecting of the on the tissue from the drop down menu, click on the search button. Upon clicking on the search button, the next window will show all associated protein pairs annotated to be present in the searched tissue (as shown in figure below), according to The Human Protein Atlas.



Jaccard similarity threshold is set as 0.4 for this search. If you want to see more results with lower threshold, please contact us.

Total number of protein pairs found: 653

	1st Uniprot Accession	1st Protein Name	2nd Uniprot Accession	2nd Protein Name	Jaccard Similari y	Tissue Name
Q	P40227	T-complex protein 1 subunit zeta (TCP-1-zeta) (Acute morphine dependence-related protein 2) (CCT-zeta-1) (HTR3) (Tcp20)	P48643	T-complex protein 1 subunit epsilon (TCP-1-epsilon) (CCT-epsilon)	0.49	lung
Q	P40227	T-complex protein 1 subunit zeta (TCP-1-zeta) (Acute morphine dependence-related protein 2) (CCT-zeta-1) (HTR3) (Tcp20)	P17987	T-complex protein 1 subunit alpha (TCP-1-alpha) (CCT-alpha)	0.47	lung
Q	P40227	T-complex protein 1 subunit zeta (TCP-1-zeta) (Acute morphine dependence-related protein 2) (CCT-zeta-1) (HTR3) (Tcp20)	Q99832	T-complex protein 1 subunit eta (TCP-1-eta) (CCT-eta) (HIV-1 Nef-interacting protein) [Cleaved into: T-complex protein 1 subunit eta, N-terminally processed]	0.51	lung
Q	P40227	T-complex protein 1 subunit zeta (TCP-1-zeta) (Acute morphine dependence-related protein 2) (CCT-zeta-1) (HTR3) (Tcp20)	P50990	T-complex protein 1 subunit theta (TCP-1-theta) (CCT-theta) (Renal carcinoma antigen NY-REN-15)	0.44	lung
Q.	P40227	T-complex protein 1 subunit zeta (TCP-1-zeta) (Acute morphine dependence-related	P78371	T-complex protein 1 subunit beta (TCP-1-beta) (CCT-beta)	0.47	lung

It should be noted that, similar to disease search, here the results are also restricted to only protein pairs with at least Jaccard similarity coefficient of 0.4. However, protein pairs less than 0.4 can be accessed via API, (for more detail refer here help section or http://iomics.ugent.be/tabloidproteome/tabloidApi.html)

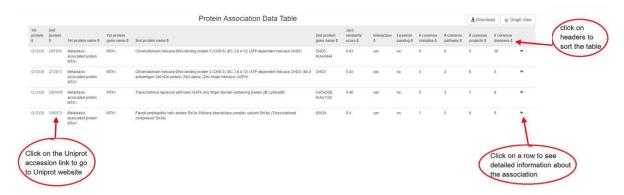
2. Search Results

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.

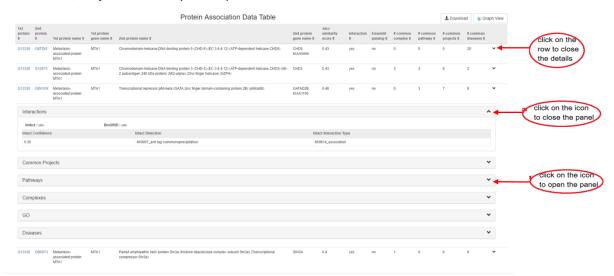
2.1. Data Table

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.





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.



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).





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.



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.



2.1.4. Complexes

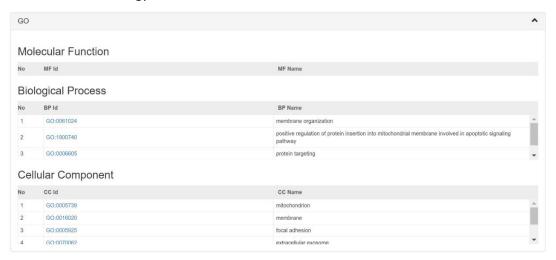
Complexes section, as shown below in the image, will provide you with list of complexes where both proteins are present, according to CORUM database. Here, with the click to Pubmed Id, it will re-direct to the link to PubMed webpage and click on the CORUM id, (marked in blue) will re-direct to the CORUM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





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 EMBLEBI Gene Ontology service.



2.1.6. Disease

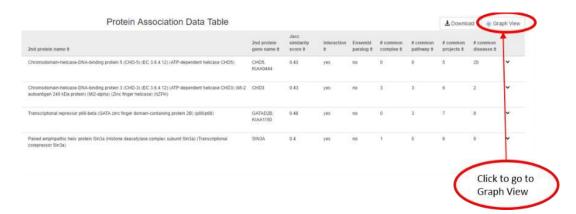
In the disease section, you can find the list of diseases, where both proteins are found to be involved in. All the information about the disease and protein involved in each disease are collected from DisGeNET database. With the click on the DisGeNET ids (marked with blue), it will re-direct to DisGeNET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



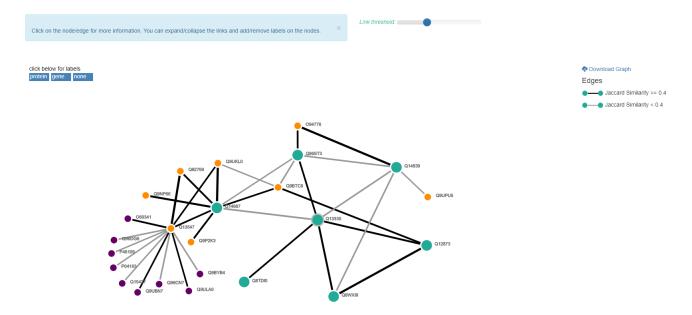
2.2. Graph View

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.



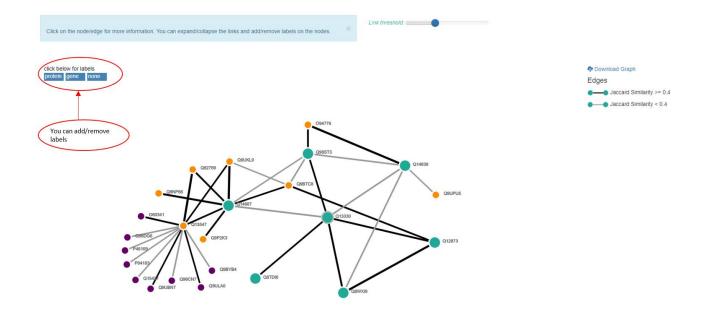


On this page, you can see dynamic graph representation of your search. All nodes and edges are repositionable. After/before changing position of the graph you can download as an image (.png) by clicking on "Download Graph" link (as shown below). For more information about the proteins and their relations, you need to click on the node or edge. You can also expand the node to see the relations with other nodes. If you want to go back to original graph you can click on the collapse button.

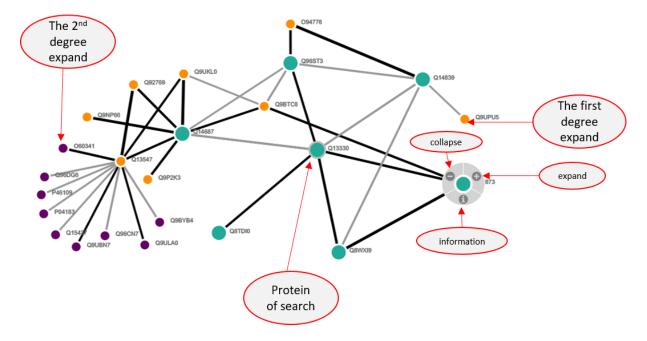


With default the nodes will be shown without any label. You can add uniprot accession with "protein" button and add gene name with "gene" button on the left hand side of the panel, as highlighted in the red in the image below. By clicking on the "none" button you can remove all labels.





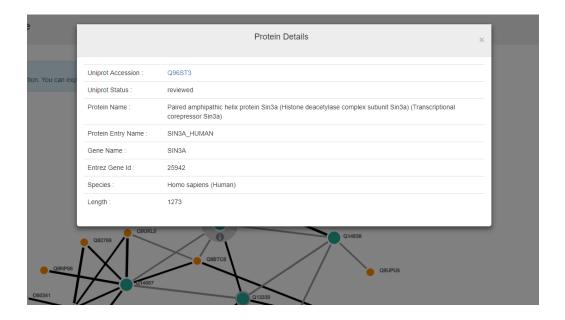
With the click on the node, there will be 3 options: Expand or collapse the relations and protein information. To remove the buttons around the node, click on the node again.



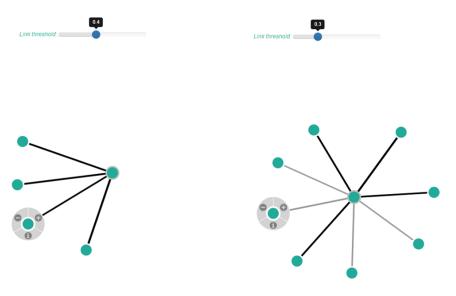
Protein of the search is highlighted with gray circle. Orange nodes are the first degree new nodes after clicking expand button of the one of the original nodes of the graph. Purple nodes are 2nd or higher degree new nodes. The graph is expandable as much as possible. If you want to remove the new nodes



and new links, collapse button will remove all related links and nodes except the original search. When the click on the information button, 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-direct to the Swissprot/Uniprot webpage. Similarly, click on Gene Id (marked with blue), it will re-direct to detailed view of the gene on NCBI Gene & Expression website.



Another new feature of this version is to be able to change Jaccard Similarity threshold and see the change on the graph immediately. User can change the threshold by using slider on top of the graph.





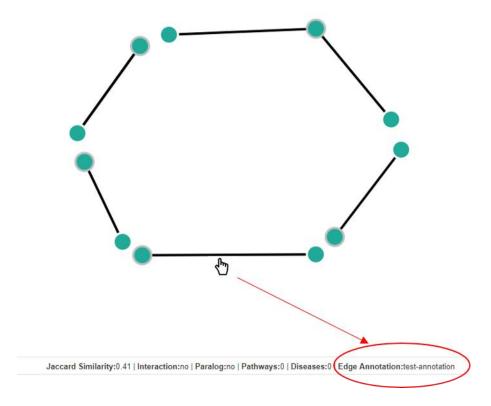
The edges between the proteins represents biological association. When user hovers on the edge, different biological association between proteins are visible below the graph. Edge with black color represent the Jaccard similarity association found with our method which is 0.4 and above, and gray color shows association having Jaccard similarity score smaller than 0.4. This information is also shown in a panel, on the right 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 below the graph when user hover on the edge. 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).

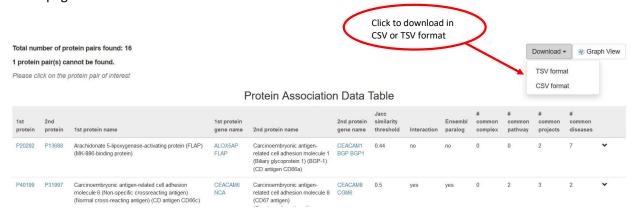


For further click on the "Data Table" option, you can view the result about the search in the Data Table view, with all information explained above. With the Data Table view, you can go to Graph View, to visualize you result with the same information and changes, as explained above.

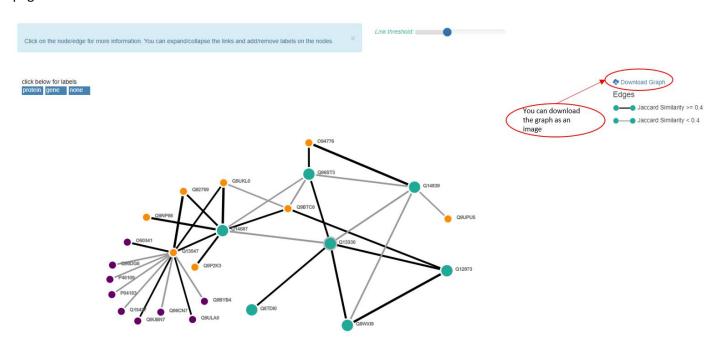


3. Downloads

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.



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: <u>10.1021/acs.jproteome.6b01066</u> PMID: 28480704

Gupta*,S., Turan*,D., Tavernier,J. and Martens,L.

The online Tabloid Proteome: an annotated database of protein associations

Nucleic Acids Res., 2018, Vol. 46, Database issue D581–D585

DOI: https://doi.org/10.1093/nar/gkx930 PMID: 29040688

