

1)Data aquisition

1.1.) If we dont have our gene list, we can generate with the metasearch tool STRING 10 [<https://string-db.org/>]. We can choose **one gene or multiple** gene lists (or protein names).After that, we need to choose **organism** (*Mus musculus*, *Homo sapiens*,...). O

The image shows the 'SEARCH' section of the STRING database interface. On the left, a sidebar lists search methods: 'Protein by name', 'Protein by sequence', 'Multiple proteins', 'Multiple sequences', 'Organisms', 'Protein families ("CODs")', 'Examples', and 'Random entry'. The first four methods are marked with green arrows pointing to the right, while the last four are crossed out with a large orange 'X'. The main area is titled 'Single Protein by Name / Identifier' and contains two input fields: 'Protein Name:' and 'Organism:'. The 'Protein Name' field has a red arrow pointing to it and a hint '(examples: #1 #2 #3)'. The 'Organism' field has a red arrow pointing to it and a dropdown menu currently showing 'auto-detect'. Below these fields is a purple 'SEARCH' button with a blue arrow pointing to it.

Given the gene list

IDUA
IDS
NAGLU
ARSB
ARSA
HGSNAT
GNPTAB

Select the organism - Homo sapiens and press enter.
The followed output was generate:

The following proteins in *Homo sapiens* appear to match your input. Please review the list, then click 'Continue' to proceed.

[<- BACK](#)
[↓ MAPPING](#)
[CONTINUE ->](#)

1) 'IDUA':

☒ [IDUA](#) - Alpha-L-iduronidase; Iduronidase, alpha-L-

2) 'IDS':

☒ [IDS](#) - Iduronate 2-sulfatase; Required for the lysosomal degradation of heparan sulfate and dermatan sulfate; Sulfatases

☐ [ID2](#) - DNA-binding protein inhibitor [ID-2](#); Transcriptional regulator (lacking a basic DNA binding domain) which negatively regulates the basic helix-loop-helix (bHLH) transcription factors by forming heterodimers and inhibiting their DNA binding and transcriptional activity. Implicated in regulating a variety of cellular processes, including cellular growth, senescence, differentiation, apoptosis, angiogenesis, and neoplastic transformation. Inhibits skeletal muscle and cardiac myocyte differentiation. Regulates the circadian clock by repressing the transcriptional activator activity of the C [...] [a.k.a. [BHLHB26](#), [DNA-binding protein inhibitor ID2](#), [EAX01019.1](#)]

☐ [TCF23](#) - Transcription factor 23; Inhibits E-box-mediated binding and transactivation of bHLH factors. Inhibitory effect is similar to that of [ID](#) proteins. Inhibits the formation of TCF3 and MYOD1 homodimers and heterodimers. Lacks DNA binding activity. Seems to play a role in the inhibition of myogenesis (By similarity); Basic helix-loop-helix proteins [a.k.a. [BHLHA24](#), [ENST00000407815](#), [TCF23_HUMAN](#)]

truncated ... [show all 9 options](#)

3) 'NAGLU':

☒ [NAGLU](#) - Alpha-N-acetylglucosaminidase; Involved in the degradation of heparan sulfate

4) 'ARSB':

☒ [ARSB](#) - Arylsulfatase B; Removes sulfate groups from chondroitin-4-sulfate (C4S) and regulates its degradation. Involved in the regulation of cell adhesion, cell migration and invasion in colonic epithelium. In the central nervous system, is a regulator of neurite outgrowth and neuronal plasticity, acting through the control of sulfate glycosaminoglycans and neurocan levels (By similarity); Belongs to the sulfatase family

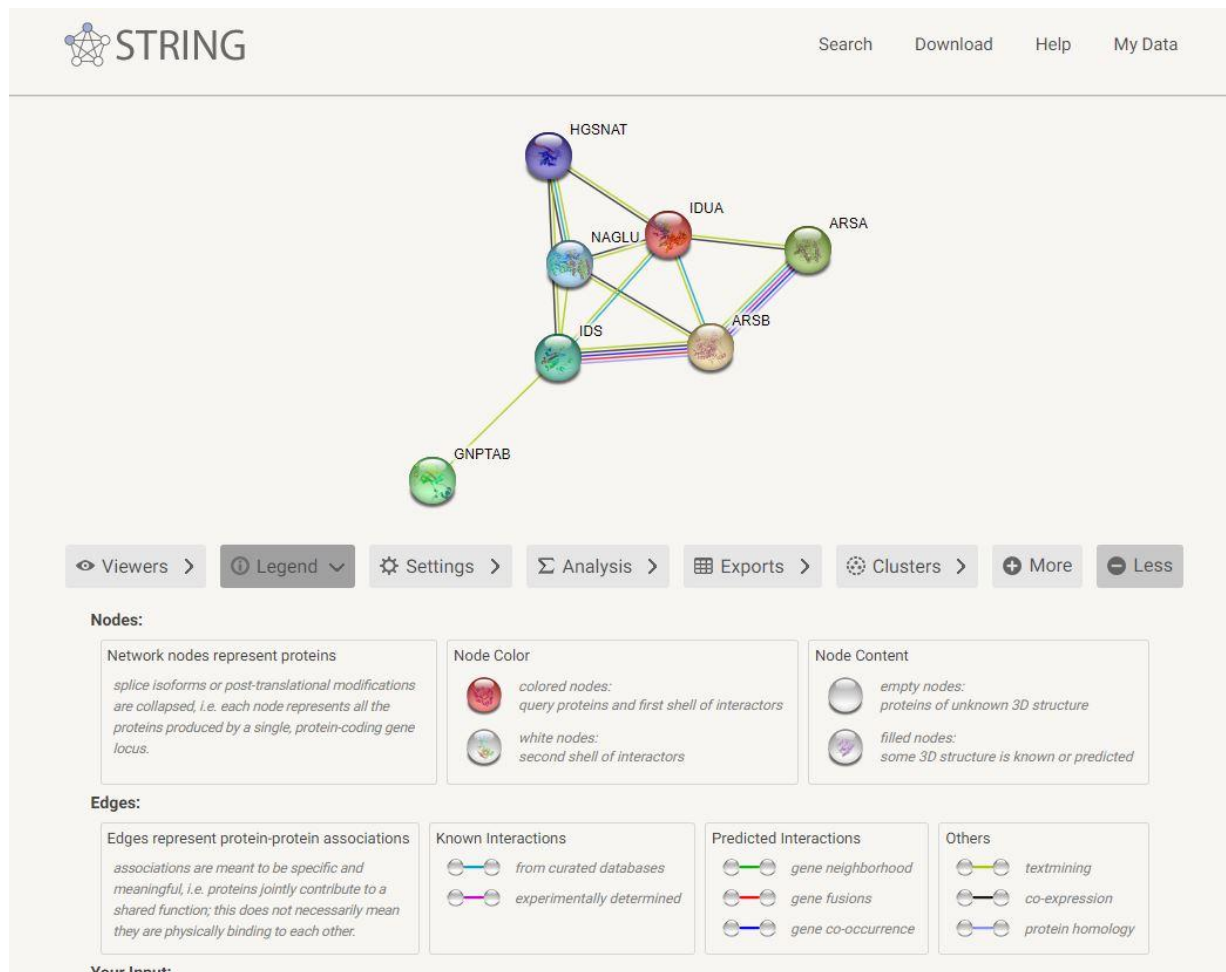
☐ [SLURP1](#) - Secreted Ly-6/uPAR-related protein 1; Has an antitumor activity. Was found to be a marker of late differentiation of the skin. Implicated in maintaining the physiological and structural integrity of the keratinocyte layers of the skin. In vitro down-regulates keratinocyte proliferation; the function may involve the proposed role as modulator of nicotinic acetylcholine receptors (nAChRs) activity. In vitro inhibits alpha-7-dependent nAChR currents in an allosteric manner. In T cells may be involved in regulation of intracellular Ca(2+) signaling. Seems to have a immunomodulatory function [...] [a.k.a. [ARS](#), [57152](#), [2MUO](#), [ArsB](#)]

5) 'ARSA':

☒ [ARSA](#) - Arylsulfatase A; Hydrolyzes cerebroside sulfate; Belongs to the sulfatase family

☐ [ASNA1](#) - Guided entry of tail-anchored proteins factor 3, atpase; ATPase ASNA1; ATPase required for the post-translational delivery of tail-anchored (TA)

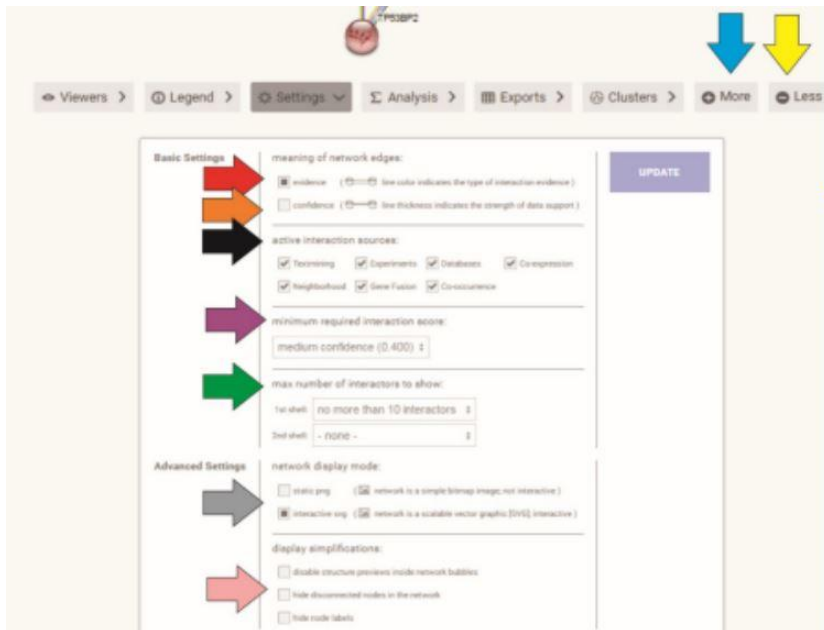
The following proteins in *Homo sapiens* appear to match your input. Please review the list, then click 'Continue' to proceed.



1.2.) Configuration to generate the network

In SETTINGS below the network we can see the available connector view types. We have two options to describe edges.

- **“Confidence”** the most reliable (trust) associations will be illustrated by thicker connectors, and thinner ones by less reliable ones;
- **“Evidence”** string will place several connectors between pairs of proteins that indicate where the information about the interaction of these proteins was taken from. This will be discussed later



We have more two options:

- “More”, we can add more than 10 nodes (genes or proteins). This is not random, its based on confidence.
- “Less”, remove the nodes add by “More” option.

In “Active Interaction Sources” it is possible to choose which of the types of connection information is desired. Information describing what each type does can be seen in Figure 4. It is recommended to always exclude the “text mining” option, as it is the least reliable. The “gene-fusion” option is normally only used when the target organism is a prokaryote, if not, it is advisable to exclude it from the search. Information about each type of interaction can be seen on the LEGEND tab.

In “Minimum required interaction score” it is possible to choose the degree of confidence of the connections, ranging from 0 to 1 (the closer to 1, the greater the confidence). To follow the protocol presented here, leave it at the default of 0.700.

In the option “Max number of interactions to shown” we can determine how many interactions per protein is desired. The more you choose, the more loaded the network becomes, so in the protocol, choose “no more than 10 interactions” and nothing in the second bottom option.

In “Network display mode”, we can use “static png”, to generate the network image or “interactive svg” to browse into the interactive network display.

After this, click on UPDATE.

Search Download Help My Data

Viewers Legend Settings Analysis Exports Clusters More Less

Basic Settings

Network type:

☒ full STRING network (the edges indicate both functional and physical protein associations)

☐ physical subnetwork (the edges indicate that the proteins are part of a physical complex)

meaning of network edges:

☒ evidence (line color indicates the type of interaction evidence)

☐ confidence (line thickness indicates the strength of data support)

active interaction sources:

☒ Textmining
 ☒ Experiments
 ☒ Databases
 ☒ Co-expression
☒ Neighborhood
☒ Gene Fusion
☒ Co-occurrence

minimum required interaction score:

medium confidence (0.400)

max number of interactors to show:

1st shell: - none / query proteins only -

2nd shell: - none -

Advanced Settings

network display mode:

☐ static png (network is a simple bitmap image; not interactive)

☒ interactive svg (network is a scalable vector graphic [SVG]; interactive)

UPDATE

1.3.) Save: Exports option

Viewers Legend Settings Analysis Exports Clusters More Less

Export your current network:

[download](#) ... as a bitmap image: file format is PNG; portable network graphic
[download](#) ... as a high-resolution bitmap: same PNG format, but at higher resolution
[download](#) ... as a vector graphic: SVG scalable vector graphic - can be opened and edited in Illustrator, CorelDraw, Inkscape etc.
[download](#) ... as short tabular text output: TSV tab separated values - can be opened in Excel and Cytoscape (data only one-way edges: A-B)
[download](#) ... as tabular text output: TSV tab separated values - can be opened in Excel (data only one-way edges: A-B)
[download](#) ... as an XML summary: structured XML, interaction data, according to the PSM data standard
[download](#) ... protein node degrees: node degree of proteins in your network (given the current score cut-off)
[download](#) ... network coordinates: a flat file format describing the coordinates and colors of nodes in the network
[download](#) ... protein sequences: MFA: multi-fasta format - containing the aminoacid sequences in the network
[download](#) ... protein annotations: a tab-delimited file describing the names, domains and descriptions of proteins in your network
[download](#) ... functional annotations: a tab-delimited file containing all known functional terms of proteins in your network

[Send network to Cytoscape](#)

Browse interactions in tabular form:

*node1	node2	node1_accession	node2_accession	node1_annotation	node2_annotation	score
ARSA	ARSB	ENSP00000216124	ENSP00000264914	Arylsulfatase A; Hydrolyzes cerebroside ...	Arylsulfatase B; Removes sulfate groups ...	0.883
ARSA	IDUA	ENSP00000216124	ENSP00000247933	Arylsulfatase A; Hydrolyzes cerebroside ...	Alpha-L-iduronidase; Iduronidase, alpha-L ...	0.606
ARSB	ARSA	ENSP00000264914	ENSP00000216124	Arylsulfatase B; Removes sulfate groups ...	Arylsulfatase A; Hydrolyzes cerebroside ...	0.883
ARSB	IDS	ENSP00000264914	ENSP00000239801	Arylsulfatase B; Removes sulfate groups ...	Iduronate 2-sulfatase; Required for the ly ...	0.471
ARSB	IDUA	ENSP00000264914	ENSP00000247933	Arylsulfatase B; Removes sulfate groups ...	Alpha-L-iduronidase; Iduronidase, alpha-L ...	0.973
ARSB	NAGLU	ENSP00000264914	ENSP00000225927	Arylsulfatase B; Removes sulfate groups ...	Alpha-N-acetylglucosaminidase; Involve ...	0.572
GNPTAB	IDS	ENSP00000299314	ENSP00000339801	N-acetylglucosamine-1-phosphotransfer ...	Iduronate 2-sulfatase; Required for the ly ...	0.444
HGSNAT	IDS	ENSP00000368965	ENSP00000339801	Heparan-alpha-glucosaminide N-acetyl ...	Iduronate 2-sulfatase; Required for the ly ...	0.444
HGSNAT	IDUA	ENSP00000368965	ENSP00000247933	Heparan-alpha-glucosaminide N-acetyl ...	Alpha-L-iduronidase; Iduronidase, alpha-L ...	0.531
HGSNAT	NAGLU	ENSP00000368965	ENSP00000225927	Heparan-alpha-glucosaminide N-acetyl ...	Alpha-N-acetylglucosaminidase; Involve ...	0.991
IDS	ARSB	ENSP00000339801	ENSP00000264914	Iduronate 2-sulfatase; Required for the ly ...	Arylsulfatase B; Removes sulfate groups ...	0.471
IDS	GNPTAB	ENSP00000247933	ENSP00000299314	Iduronate 2-sulfatase; Required for the ly ...	N-acetylglucosamine-1-phosphotransfer ...	0.444
IDS	HGSNAT	ENSP00000339801	ENSP00000368965	Iduronate 2-sulfatase; Required for the ly ...	Heparan-alpha-glucosaminide N-acetyl ...	0.444
IDS	IDUA	ENSP00000339801	ENSP00000247933	Iduronate 2-sulfatase; Required for the ly ...	Alpha-L-iduronidase; Iduronidase, alpha-L ...	0.976
IDS	NAGLU	ENSP00000339801	ENSP00000225927	Iduronate 2-sulfatase; Required for the ly ...	Alpha-N-acetylglucosaminidase; Involve ...	0.563
IDUA	ARSA	ENSP00000247933	ENSP00000216124	Alpha-L-iduronidase; Iduronidase, alpha-L ...	Arylsulfatase A; Hydrolyzes cerebroside ...	0.606
IDUA	ARSB	ENSP00000247933	ENSP00000264914	Alpha-L-iduronidase; Iduronidase, alpha-L ...	Arylsulfatase B; Removes sulfate groups ...	0.973
IDUA	HGSNAT	ENSP00000247933	ENSP00000368965	Alpha-L-iduronidase; Iduronidase, alpha-L ...	Heparan-alpha-glucosaminide N-acetyl ...	0.531
IDS	IDS	ENSP00000247933	ENSP00000339801	Alpha-L-iduronidase; Iduronidase, alpha-L ...	Iduronate 2-sulfatase; Required for the ly ...	0.976
IDUA	NAGLU	ENSP00000247933	ENSP00000225927	Alpha-L-iduronidase; Iduronidase, alpha-L ...	Alpha-N-acetylglucosaminidase; Involve ...	0.779

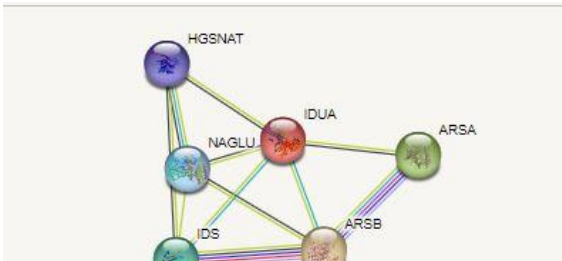
To open this generated network into Cytoscape, choose “save as simple tabular output” and Download.

The followed table be generated:

#node1	node2	node1_string_id	node2_string_id	neighborhood_on_chromosome	gene_fusion	phylogenetic_cooccurrence	homology	coexpression
experimentally_determined_interaction				database_annotated	automated_textmining	combined_score		
ARSA	IDUA	9606.ENSP00000216124	9606.ENSP00000247933	0	0	0.143	0	0.560 0.606
ARSA	ARSB	9606.ENSP00000216124	9606.ENSP00000264914	0	0.447	0.676	0	0.115 0.800 0.836 0.883
ARSB	NAGLU	9606.ENSP00000264914	9606.ENSP00000225927	0	0	0.062	0	0.563 0.572
ARSB	IDUA	9606.ENSP00000264914	9606.ENSP00000247933	0	0	0	0.900	0.743 0.973
ARSB	IDS	9606.ENSP00000264914	9606.ENSP00000339801	0	0.003	0.442	0.564	0.062 0 0.725 0.471
GNPTAB	IDS	9606.ENSP00000299314	9606.ENSP00000339801	0	0	0	0.444	0.444
HGSNAT	NAGLU	9606.ENSP00000368965	9606.ENSP00000225927	0	0	0.064	0	0.900 0.921 0.991
HGSNAT	IDUA	9606.ENSP00000368965	9606.ENSP00000247933	0	0	0.077	0	0.513 0.531
HGSNAT	IDS	9606.ENSP00000368965	9606.ENSP00000339801	0	0	0.062	0	0.431 0.444
IDS	NAGLU	9606.ENSP00000339801	9606.ENSP00000225927	0	0	0	0.563	0.563
IDS	IDUA	9606.ENSP00000339801	9606.ENSP00000247933	0	0	0	0.900	0.777 0.976
IDUA	NAGLU	9606.ENSP00000247933	9606.ENSP00000225927	0	0	0.138	0	0.710 0.739

Node 1 and Node 2 columns are the most important. Others are metadata generated by string.

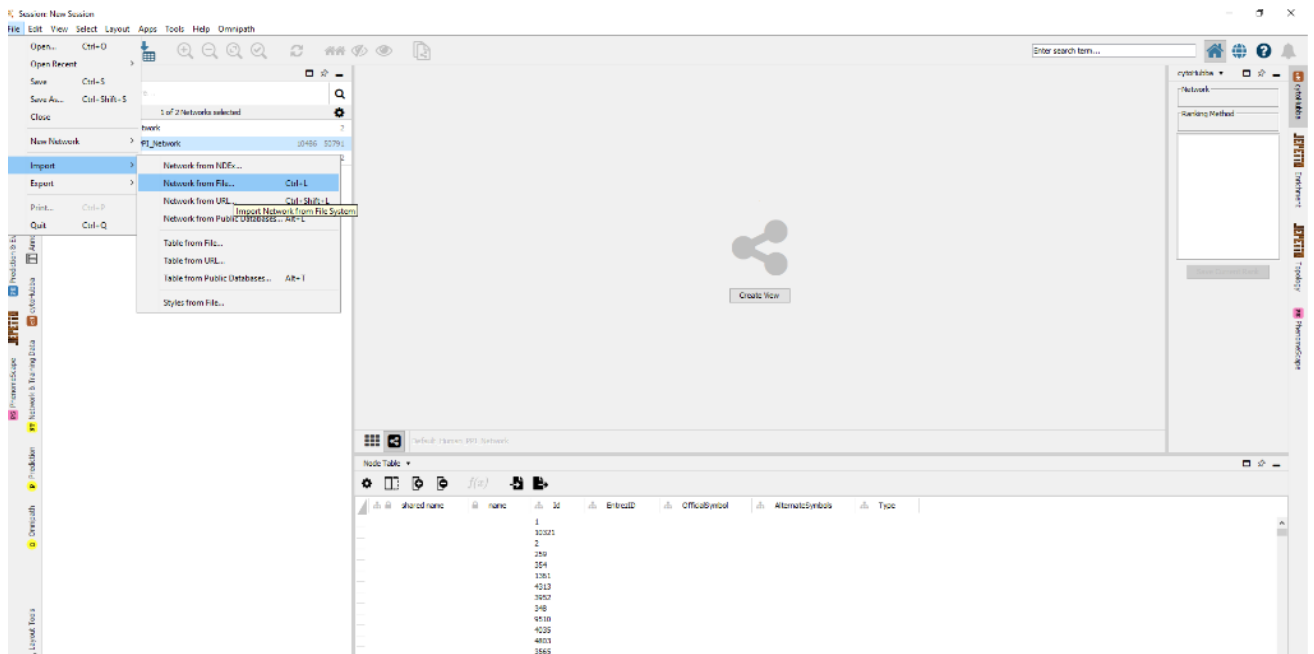
The meaning is ARSA gene found in Node1 column interacts with IDUA gene



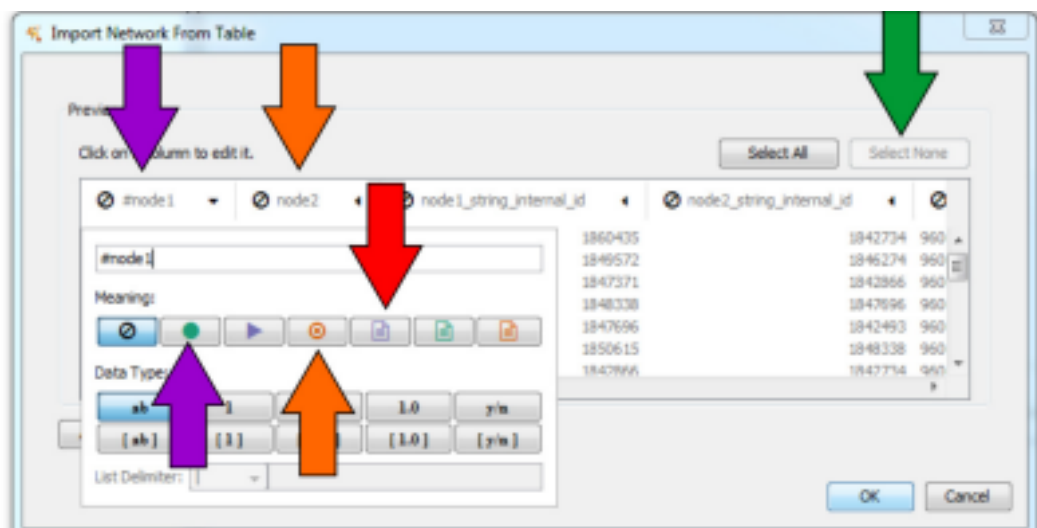
#####

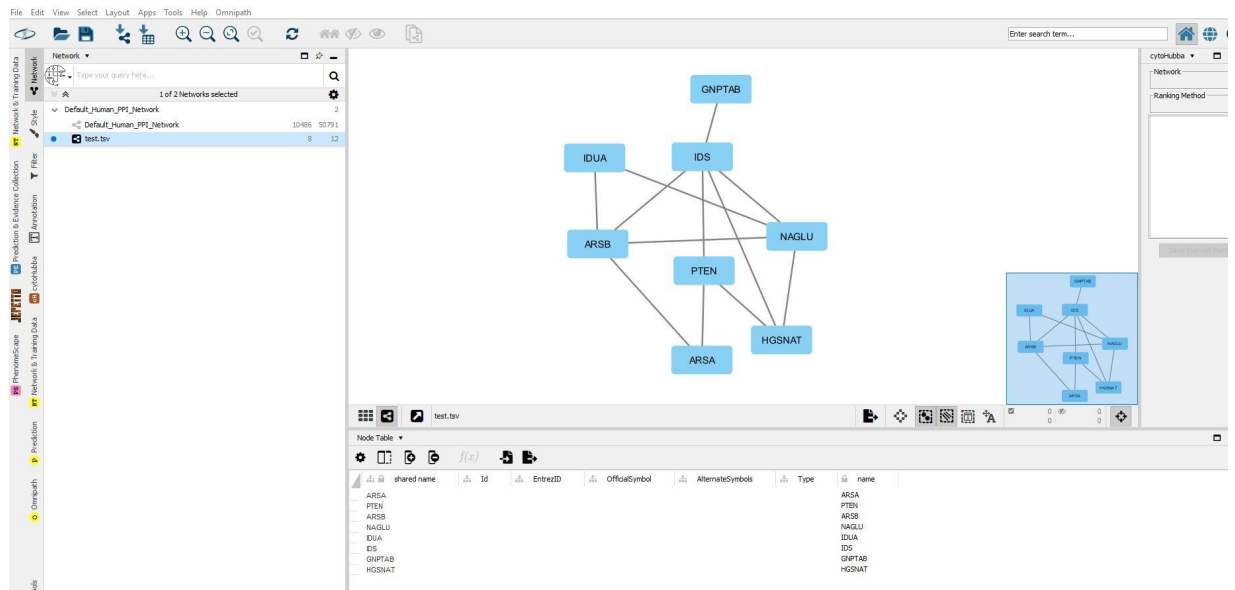
2) Importing the network file into Cytoscape

2.1.) Open cytoscape. Go to “File”, “Import”, “Network”. “File” or press Ctrl L



This window will be appear. To import the string network into Cytoscape, we need to click first in “Select None”. After “node 1”, choose “Source Node”; “Node 2”, choose “Target Node”. As mentioned before, Node 1 and Node means column 1 genes interact with column 2 genes. Next, “combined score” set to “Edge Attribute”.

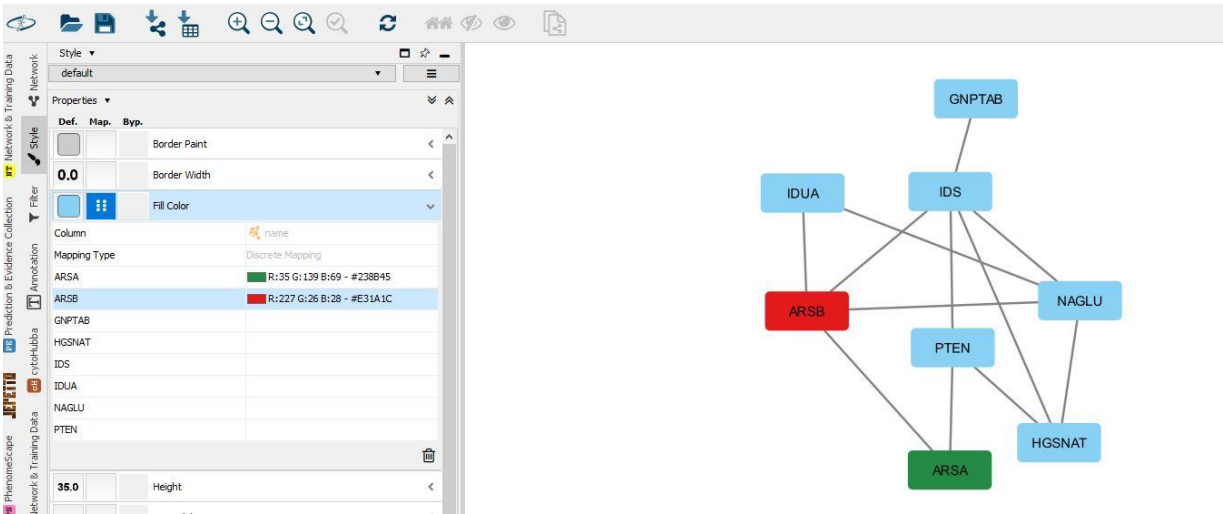




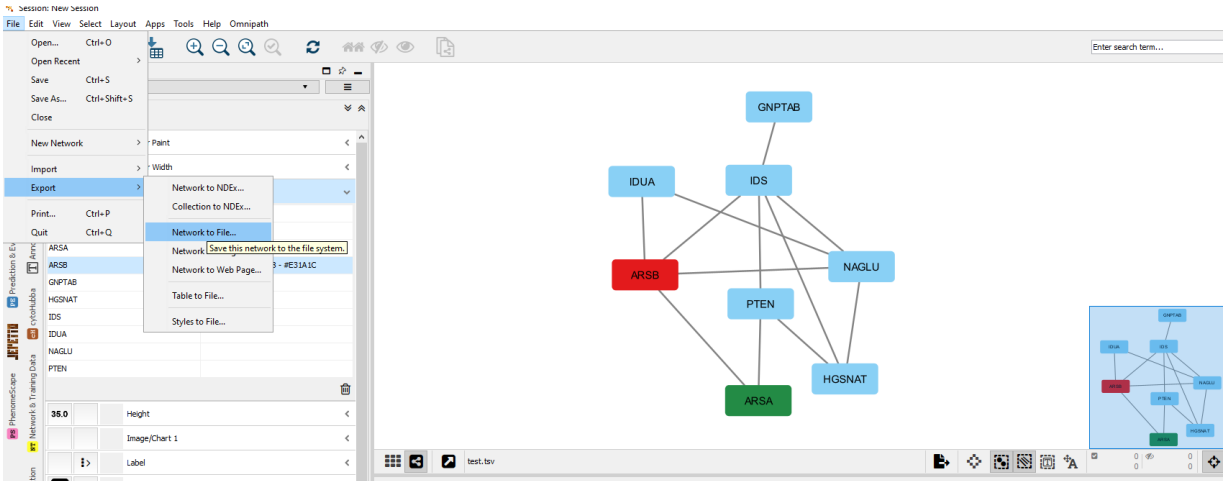
2.2.) Now we can customize the network, as follows:

The screenshot shows the 'Style' panel in Cytoscape. The panel is titled 'Style' and has a 'default' dropdown menu. Below the title is a 'Properties' section with three tabs: 'Def.', 'Map.', and 'By.'. The 'Def.' tab is selected. The properties listed are: Border Paint, Border Width (0.0), Fill Color (blue), Column (name), Mapping Type (Passthrough Mapping), Height (35.0), Image/Chart 1, Label, Label Color (black), Label Font Size (12), Shape (square), Column (name), and Mapping Type (Passthrough Mapping). The 'Style' panel is highlighted with a pink circle. The 'Fill Color' property is highlighted with a pink arrow. The 'Shape' property is highlighted with a pink arrow. The 'Label' property is highlighted with a pink arrow. The 'Label Color' property is highlighted with a pink arrow. The 'Label Font Size' property is highlighted with a pink arrow. The 'Height' property is highlighted with a pink arrow. The 'Image/Chart 1' property is highlighted with a pink arrow. The 'Mapping Type' property is highlighted with a pink arrow.

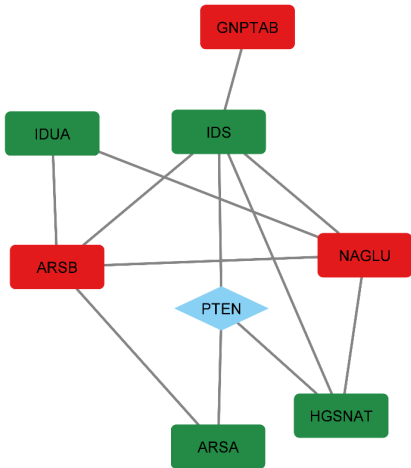
Style / Fill color / Mapping type / name / Discrete mapping / choose the colors
We adopted red to upregulate genes, green to downregulated genes, and blue to transcription factors.



After finished the customization, we can save the network in a figure (PNG or JPEG)




Congrats!



#####

Using custom tables to generate networks

If we have only a gene list with Transcription factors, miRNAs, or other elements, we need to follow this structure:



#node1	node2	n
experimentally_de		
ARSA	IDUA	9
ARSA	ARSB	9
ARSB	NAGLU	9
ARSB	IDUA	9
ARSB	IDS	9606.
GNPTAB	IDS	9606.
HGSNAT	NAGLU	9
HGSNAT	IDUA	9
HGSNAT	IDS	9606.
IDS	NAGLU	9606.
IDS	IDUA	9606.
IDUA	NAGLU	9

And save as .tsv