DynOVis: Dynamic Omics Visualization manual

This user manual will guide you through the basic steps that will translate your interaction file into a network graph.

R libraries for DynOVis

DynOVis requires a couple of R packages to be installed:

- Shiny
- Xlsx
- rJava
- igraph
- BioNet

These libraries can be manually be installed by the user in his/her preferred local library, or the different packages can be installed by running 'DynOVisInstallPackages.R'

Run the tool on a local computer

The tool can be downloaded from the Bitbucket repository (https://bitbucket.org/mutgx/dynovis/src) and launched in either R or RStudio. Here, we advise to use Rstudio, since Rstudio makes it very easy to launch any R shiny tool.

In Rstudio, the user should navigate to his/her local directory where DynOVis is placed and open the script 'server.R'. This is the main file that will load and launch all the functionalities of the tool. In Rstudio, the runApp button will automatically launch the tool in the user's Rstudio-defined browser. Note that Rstudio will open Internet Explorer in most cases, we strongly advice each user to either copy and paste the link into their Mozilla Firefox or Google chrome browser.

Step 1: upload a network file and Dose/Time series data

After navigating from the home screen to the 'User input data' panel, the user can upload the following network file types (figure 1, panel A):

- Adjacency matrix
- Edge list
- Cytoscape sif file

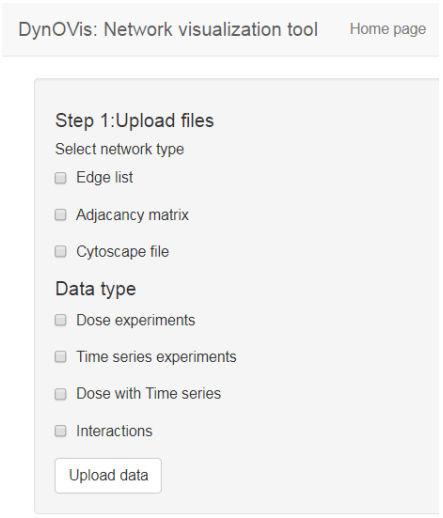
The adjacency matrix (of dimension MxM) should be a matrix containing zeros if there is no connection and ones if there is a connection between nodes (figure 1, panel B). The edge list should be a two column file containing the source node in column 1 and the target node in column 2. (Figure 1, panel C). If you have created a network earlier in Cytoscape, you can export your network as a Cytoscape sif file and import it into DynOVis. This way, you don't have to build a new edge list or adjacency matrix.

After you have made your selection, you have to click on 'Browse' and navigate to your directory where your network file is stored.

After the network is uploaded, you can choose to add experimental data to the network. This can be either time, dose or dose-over time experiments. The values from the experimental data are used as values to color the nodes in the network (example: log fold change of genes over time). Each of the different files, whether it is a dose series or time series has a comparable data file structure. The rows in the file correspond to the different nodes in the network and therefore should have the same name. The columns in the file correspond to the dose, time point or Dose X at time point Y (figure 1, panel D).

As a last option, the user can upload an interaction file, which is a file that will specify which interactions are present at different time points.

A: screen to select data



B: Adjacency matrix

name	Node 1	Node 2	Node 3	Node 4
Node 1	1	1	1	1
Node 2	1	1	0	0
Node 3	1	0	1	0
Node 4	1	0	0	1

C: edge list

Source	Target
7535	4067
7535	115650
7535	4790
7535	5970
7535	7185
7535	597

D: time series data structure

id	T2	T8	T24
2729	-0.01414	0.395255	0.690599
6540	0.049212	-0.13755	-0.1259
6534	-0.05911	-0.06443	-0.08445
3082	-0.06205	-0.09891	0.096693

E: Dose series data structure

id	D50	D100	D500
1244	-0.0114	0.130063	0.117081
2730	-0.00572	0.136712	-0.11441
55334	0.028874	0.074045	-0.0131

F: Dose-over-time series data structure

id	T2+D200	T8+D200	T24+D200	T2+D1000	T8+D1000	T24+D100	T2+D5000	T8+D5000	T24+D5000
7535	0.025349	0.077682	-0.17961	0.043212	-0.07229	-0.12132	0.071116	0.379976	0.054713
27040	0.065909	0.081641	0.117148	0.040484	-0.06663	0.049823	0.021885	0.434626	0.25041
5335	0.040706	0.014213	-0.03479	0.069422	0.010691	0.080971	0.095403	0.168321	0.217452

Figure 1 Upload data interface. Panel A: here the user can select the network type; this can either be an adjacency matrix (panel B), edge list (panel C) or a Cytoscape sif file. Furthermore, experimental data can be uploaded: time series data (panel D), dose series data (panel E) or dose-over-time series data (panel F)

Step 2: set your network characteristics

After you have uploaded the data successfully a new step will popup: you should select if your network is undirected or directed (Figure 2). In case of a directed network, arrow heads will be placed on the edges in the direction from source node to target node.

Step 3: Connect the database to your network

Finally, you should select which organism your network data originates from: Homo sapiens, Mus Musculus and Ratus Norvegicus. The second step will be to select if your node IDs are either Gene Symbol or Entrez ID. After these two selections you can press ‘build network’ to construct your final network (figure 3).

If you don’t want to connect the database to your network, you should leave the field ‘map nodes to database’ unchecked.

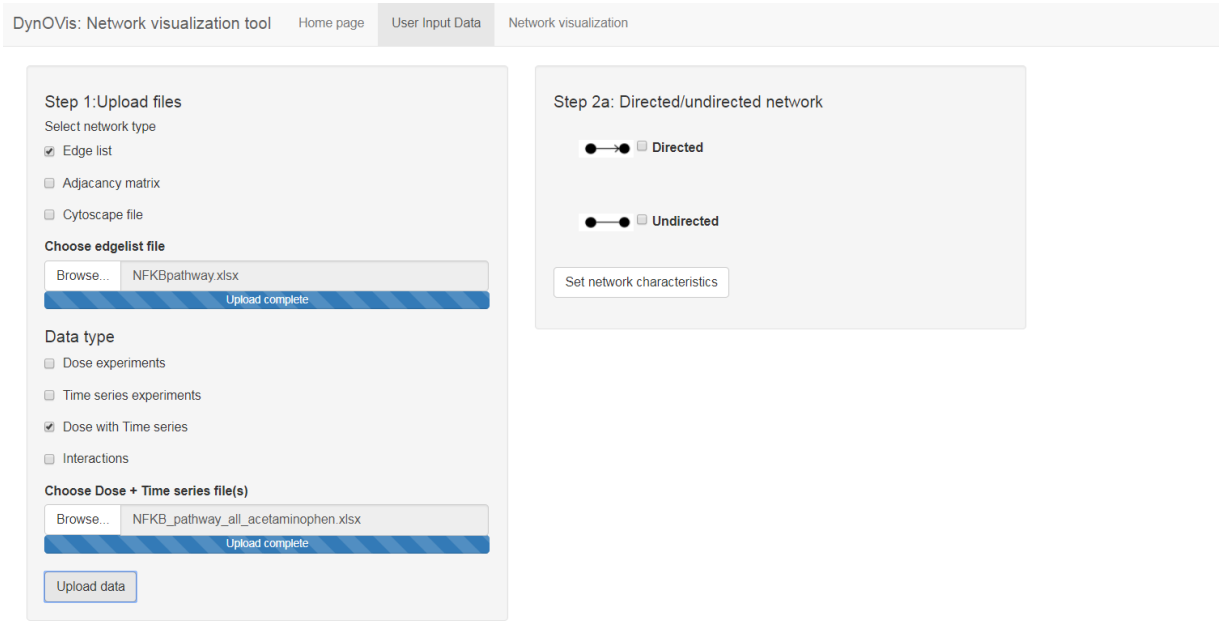


Figure 2 after the data has been uploaded, the user needs to specify if the network has directed or undirected edges. In case of directed edges, an arrow marker will be placed on the edge between two nodes

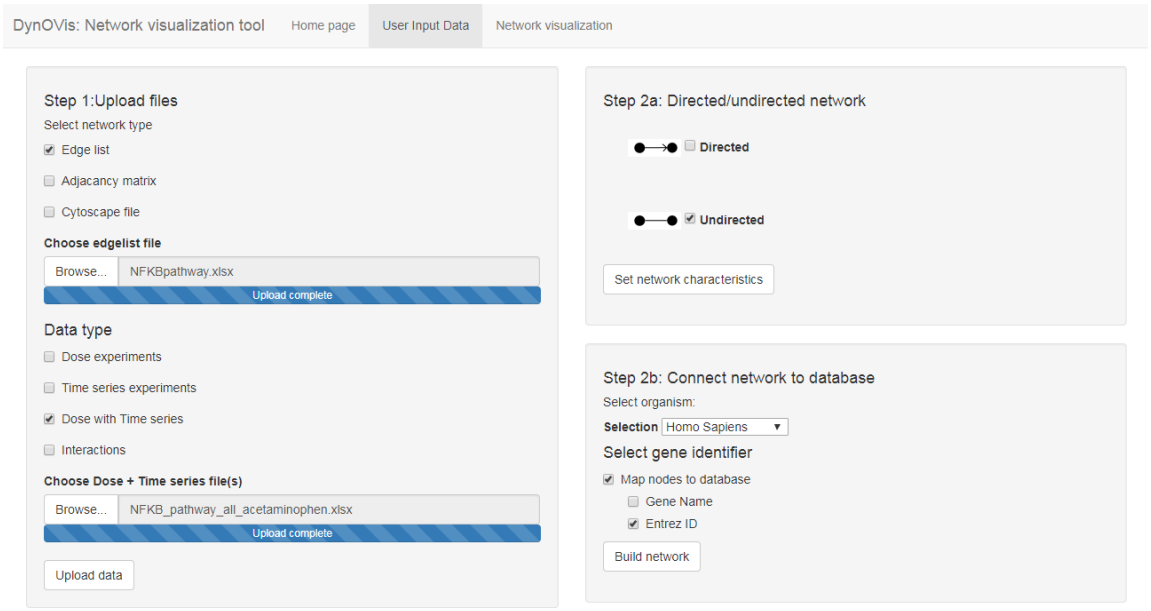


Figure 3 to connect the database to your network, the user has to select the species and gene identifier. When this selection has been made, the network can be build

The visualization of your network

After the network is build, it will be visible in the section tab 'Network visualization'. Here you can interact with your network, for instance by dragging and placing nodes at different positions. Moreover, with a right-mouse-click you can pull the modal that will show the biological information associated with the node you just clicked (figure 4).

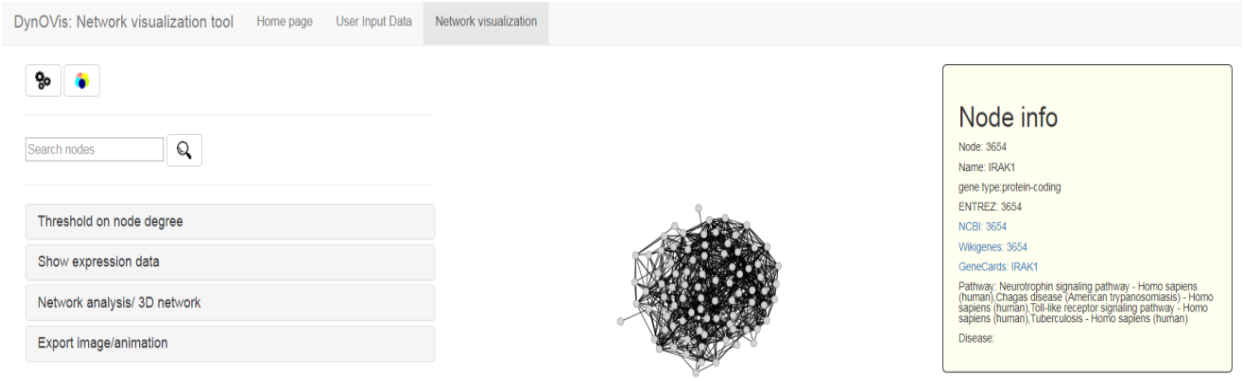


Figure 4 Network visualization panel with the different visualization options (left) and the node info hub (right)

<u>Threshold the network on node degree</u>

Since it is possible that your network has a high density in edges between nodes, it can help to set a threshold to filter out nodes that have a low degree of edges.

The degree of a node is defined by the following equations:

$$\text{Total degree:} \quad \text{Node}_{\text{i,Total degree}} = \sum \text{incoming edges} + \text{outgoing edges}$$

$$\text{Inner degree:} \quad \text{Node}_{\text{i,Inner degree}} = \sum \text{incoming edges}$$

$$\text{Outer degree} \quad \text{Node}_{\text{i,Outer degree}} = \sum \text{outgoing edges}$$

Changing one of the threshold values will change the opacity of the nodes that fail to meet this threshold and thus highlight the nodes that are above or equal to the threshold (figure 5, panel A).

Furthermore, you can set the node size based on the total degree, inner degree or outer degree (figure 5, panel B). This will allow you to quickly identify all the nodes that could play a major role in your network (based on the hypothesis that important nodes have a high number of connections).

Furthermore, under the tab “Network analysis /3D network” the user has the option to construct a table with the different degree for each gene by clicking the button ‘HUB nodes’ (figure 6).

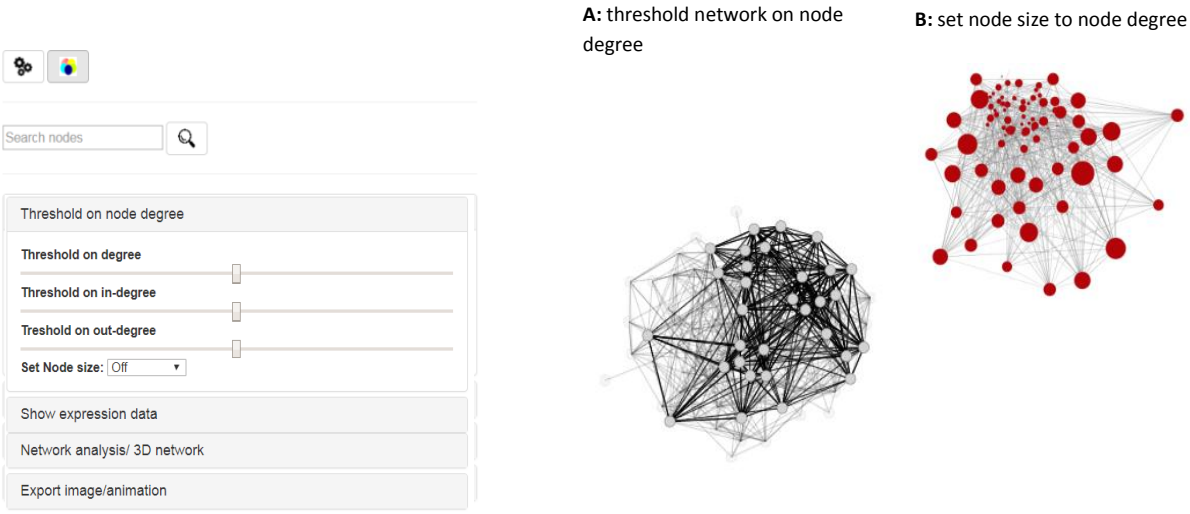


Figure 5 panel A: The menu options showing the different node degree thresholds. Panel B: when the node size is set to a certain value, the different nodes will be sized according to their degree.

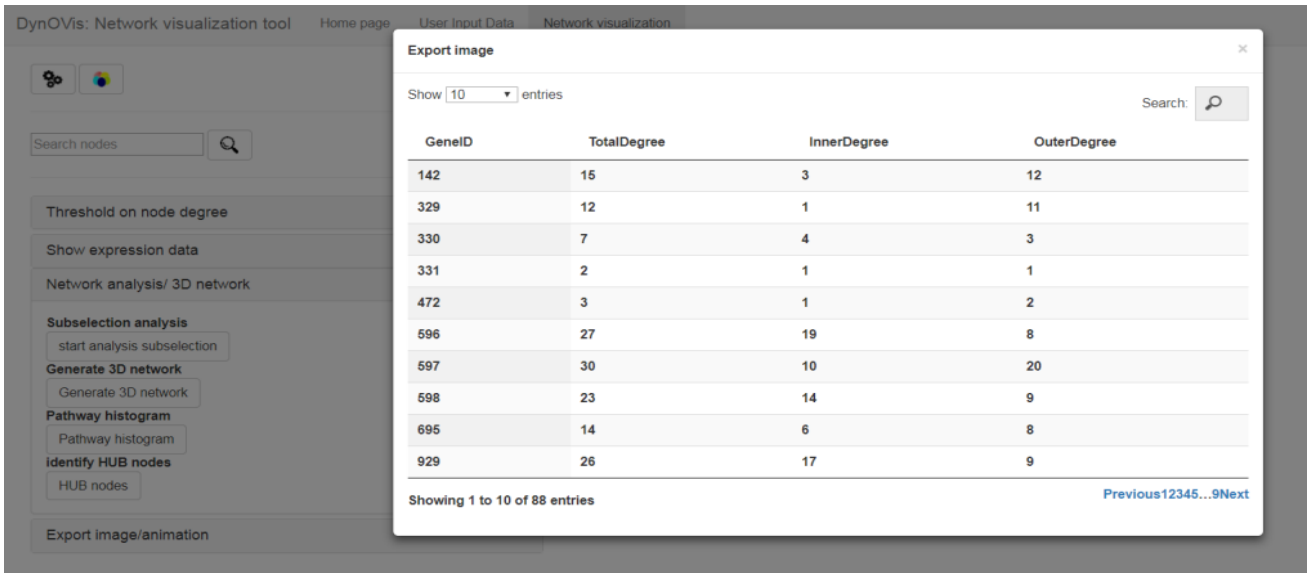


Figure 6 Network hub table: under network analysis, the user can select 'identify hub nodes' which will launch the table containing the information about node degree.

Show interaction data on network

If you have uploaded expression data during step 1, it is possible to color the nodes in the network based on the expression values from this data file. This data is mapped onto the network by using a linear color scale function which automatically detects the range of the data. Animation over time, but also frame-by-frame view can be used to study the changes in expression in the network.

There are two linear color scales that can be applied to visualize the experimental data: Red-Green scale and a Red-Blue scale (figure 7). To determine the domain of the color scale, the maximum absolute value of the experimental data is determined and used as the left and right boundary of the scale. By using the absolute maximum value, we make sure that a dark red color (meaning very low expression) is equal to a bright green color (meaning high expression).

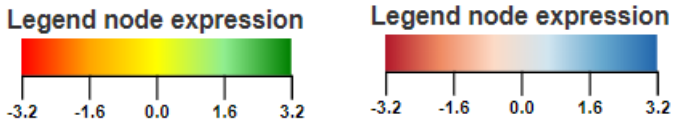


Figure 7 (Left) color scale for the Red-Green color gradient (Right) color scale for the Red-Blue color gradient.

Highlight the neighboring nodes

To highlight the neighboring nodes of the node you are interested in, you only have to do a 'double-left-mouse-click' on your node of interest. As an example, if we are interested to find out which are the neighboring nodes (or in this specific case genes) of our node of interest (encircled node in figure 8), we can highlight these neighboring nodes. Now we can follow the change of expression values over time if we run our dynamic animation, or can create a 2D line graph of the changes over expression in dose/time/dose-over-time for our selection.

Histogram of the associated biological pathways in the network

Under the Network analysis / 3D network tab you have the option to create a histogram of the top 10 pathways that are associated with the genes in your network (figure 9).

For each gene in the network, the associated pathways are searched for in the database that is attached to DynOVis. Since one gene can be associated to multiple pathways, the number of different pathways can be very high, depending on the number of genes in your network. Therefore, a ranking is made were the frequency of the biological pathway occurring in the total list of biological pathways found in the network.

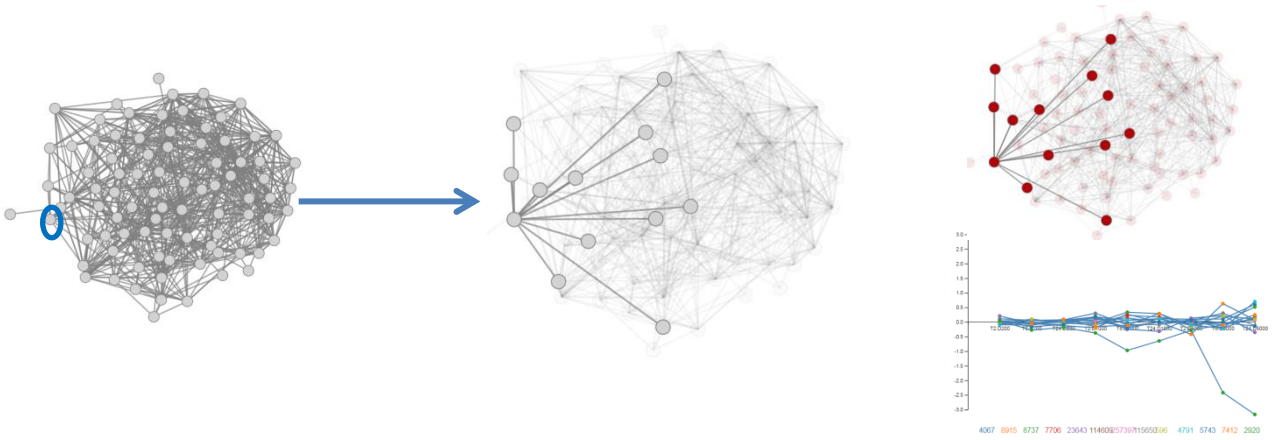


Figure 8 Highlight node with neighbors: with a left double-mouse click, the node and all the neighboring nodes will be highlighted. For this selection, the animation can be played or a frame-by-frame view can be used to visualize the experimental data on this selection. Under the network analysis menu, the user can select analyze sub-selection to get a 2D line graph of the expression over time for the selected nodes.

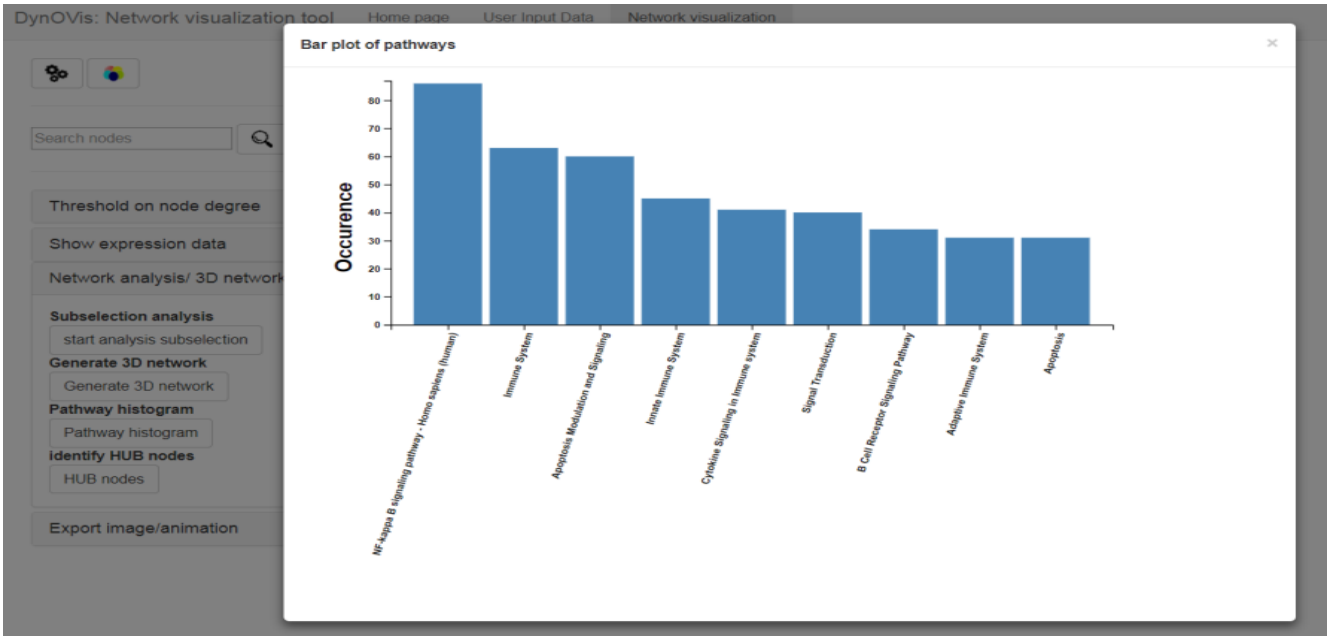


Figure 9 Pathway histogram: by selection Pathway Histogram under the network analysis menu, the user can see the top 10 pathways in his network.