CrosstalkNet Documentation

Getting Started

CrosstalkNet is a web-based network visualization tool to retrieve and mine interactions in large scale co-expression networks, which are bipartite in nature. An important application of co-expression networks in biology is to uncover the effects of tumor microenvironment in the context of tumor epithelial-stromal interactions. Below is a tutorial of the web application. The manual is written according to the functionality of each tab in the application.

For guest users, the application has a case study from Oh et al. work (Oh et al. 2015). In order to upload customized network files, users have to register and the credentials can be obtained from bhaibeka@uhnresearch.ca **or** benjamin.haibe.kains@utoronto.ca

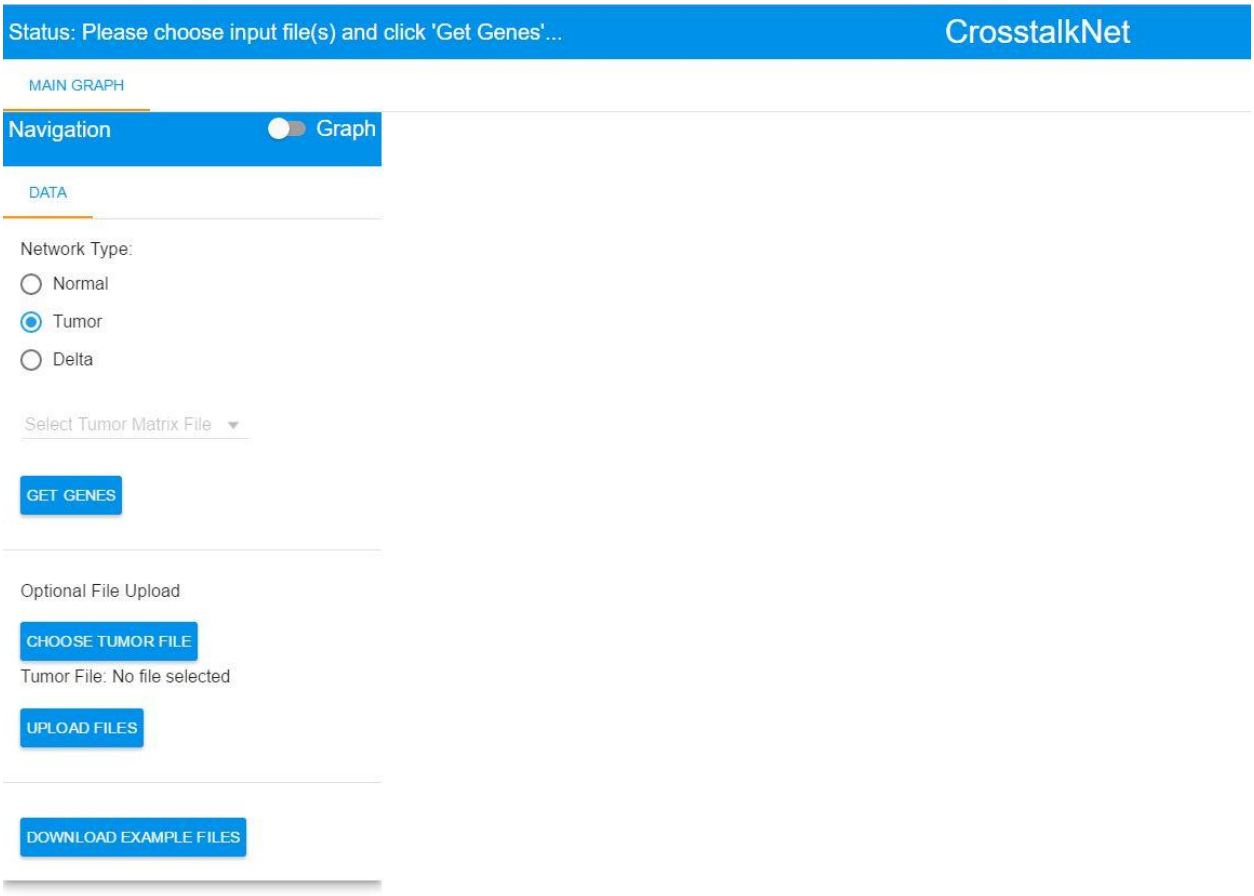


Figure 1: Application view when logged in.

Main Graph

Data

Selecting a Dataset

Before any analysis of the network can be done, a network type and associated file(s) must be chosen from Main Graph ---> Data. Once a the network type and file(s) are selected, the “Get Genes” button needs to be clicked in order to enable the other tabs and load the list of genes associated with the specified file(s).

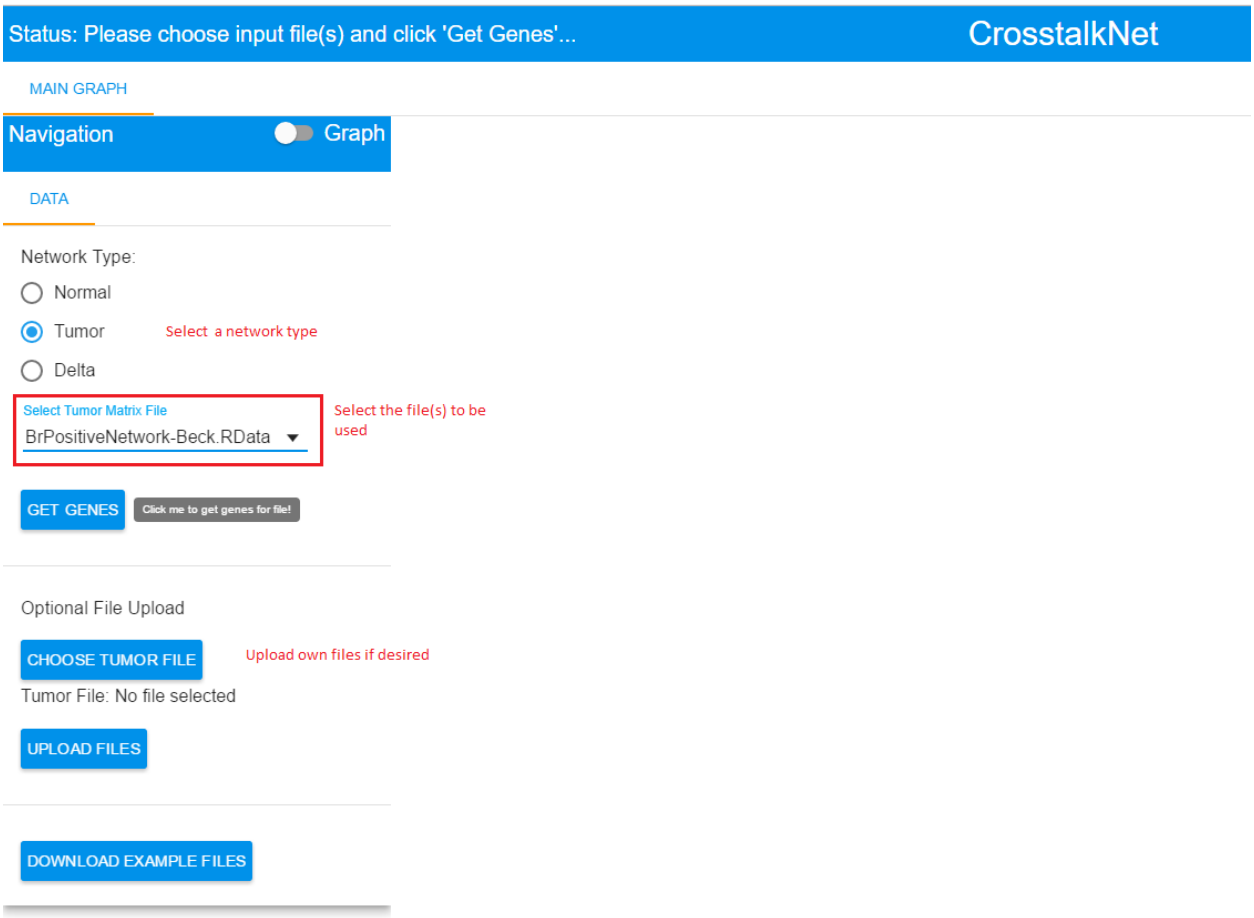


Figure 2: Data subtab in the Main Graph tab. Data files are selected here.

Upload

Users can upload their own adjacency matrices to be used for visualization. The adjacency matrix should be converted to a sparse matrix using Matrix package in *R*, and the file should be saved as an RData file using saveRDS function in *R* before uploading the files. Here are the requirements that must be satisfied for a successful file upload:

- Files must contain a single dgCMatrix from the Sparse Matrix package
- No matrix is allowed to have NA values
- The rownames of a matrix must match its colnames
- The row names and column names of a matrix must not contain NA values
- Files must be saved using the saveRDS function
- Files must have an Rdata extension
- Files must be smaller than 20MB in size

It is important to note that there are 3 files required in order to upload a delta network: the normal, tumor, and differential file.

Types of Networks

The user can upload normal and tumor co-expression networks which are sparse. In addition, the sparse delta networks (the difference of tumor and normal) can also be uploaded, which are obtained from matched tumor and normal epi-stroma pairs.

Query

Selection of genes

From the Query sub-tab, the user can find out information about the first and second neighbours of a group of selected genes. To select a gene, click on the "Choose Genes" box and start typing in the name of a gene. The box will filter down results based on what you've typed so far. You will notice that all gene names have a -E or -S appended to them followed by a number. The -E indicates epi, the -S indicates stroma, and the number indicates the number of first neighbours for that gene.

As an example, let us select CD8A in the stroma, GZMB in the stroma, and CXCL13 expressed in the epithelial. As you begin typing CD8A, you will notice that the drop down list has fewer and fewer results. To choose CD8A-S, simply click on it.

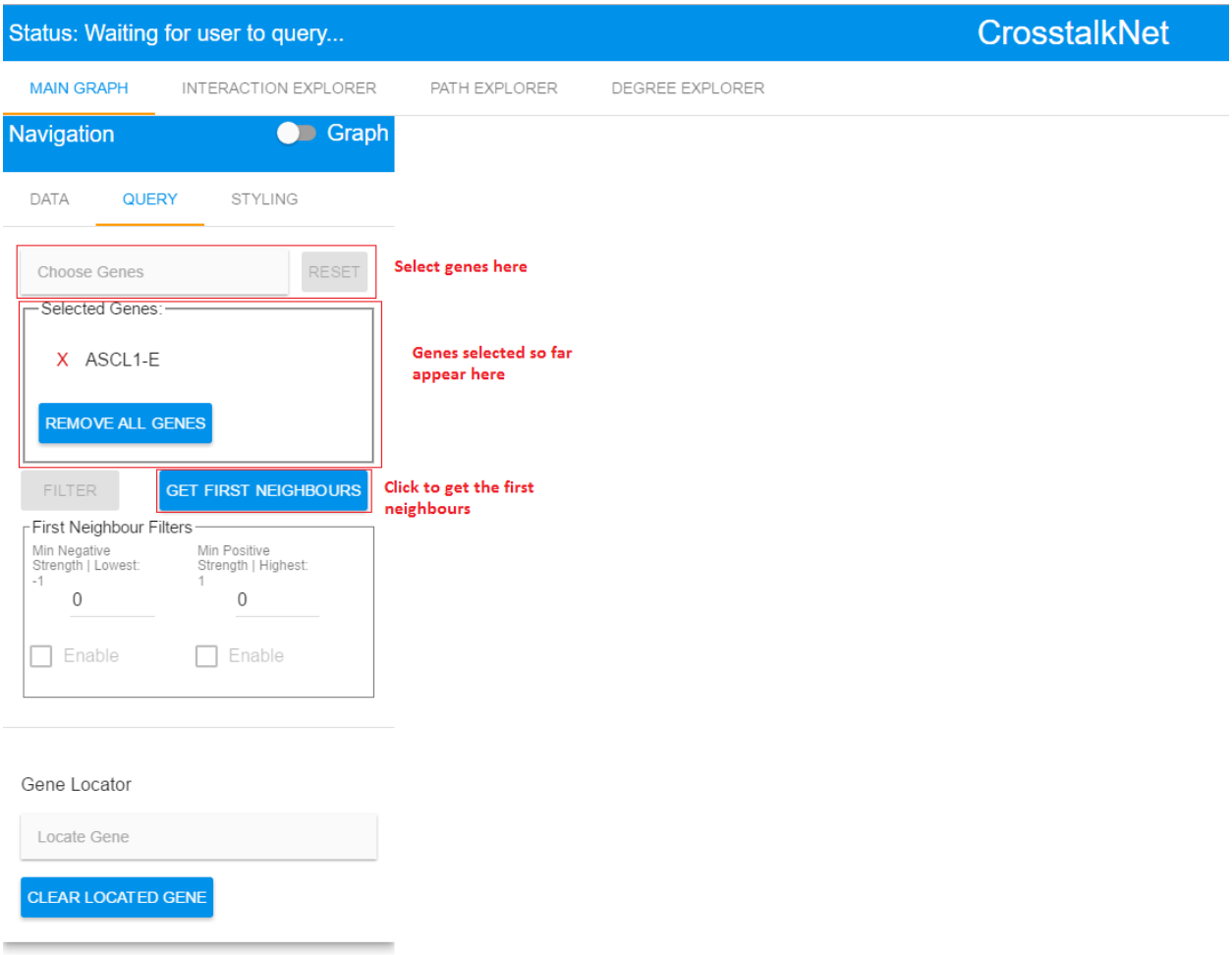


Figure 3: The Main Graph tab with a single gene selected, but graph not yet displayed.

First and Second neighbors

Once the user selects the genes of interest and clicks on the “Get First neighbors” button, all the interactions pertained to those selected genes will appear as a graph along with a tabular view of results. Once the “Get First Neighbors” button is clicked, the filters will become usable. There are values indicating the lowest negative interaction strength and highest positive interaction strength. A description of how to use the filter feature is mentioned in the subsection Filter.

Furthermore, the “Get Second Neighbors” button will be automatically enabled once “Get First Neighbors” button is clicked.. When the user clicks on the “Get Second Neighbors” button, the second neighbors of the selected genes will be displayed in addition to the already displayed first neighbors.

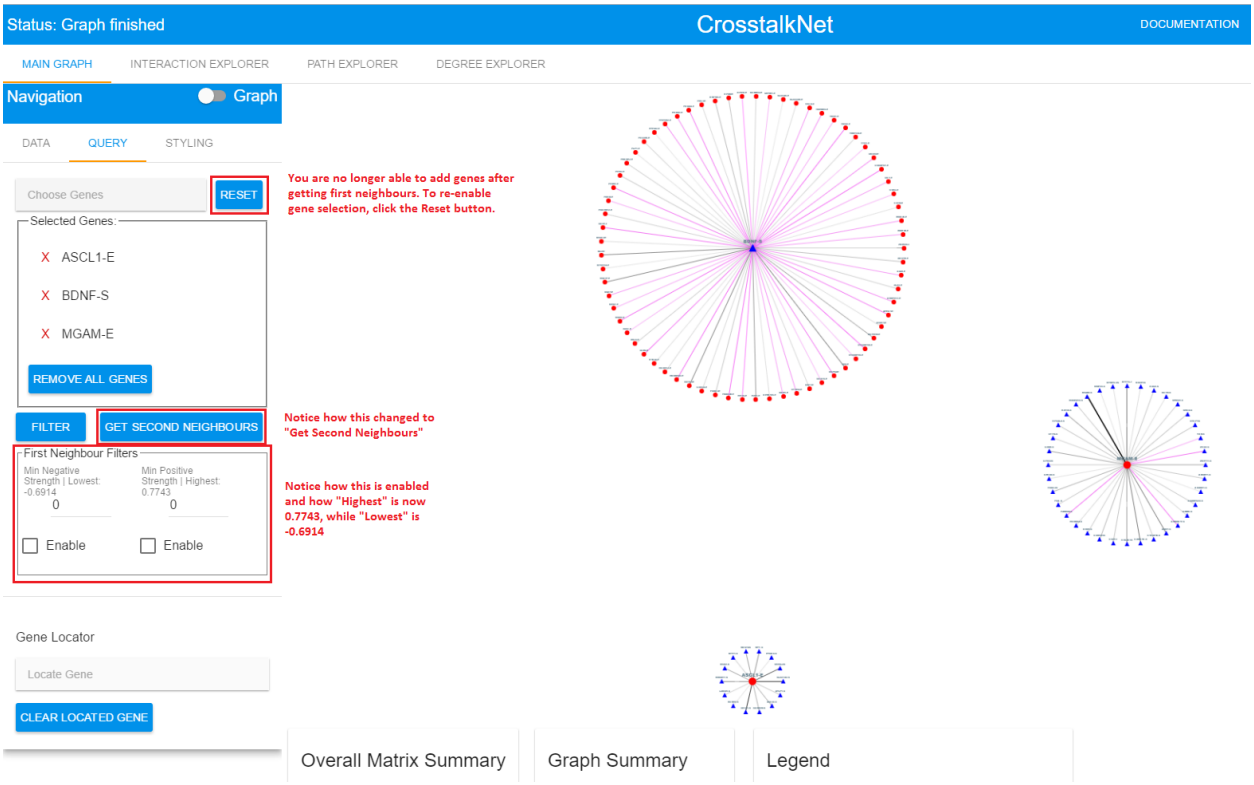


Figure 4: First neighbours of ASCL1-E, BDNF-S, MGAM-E.

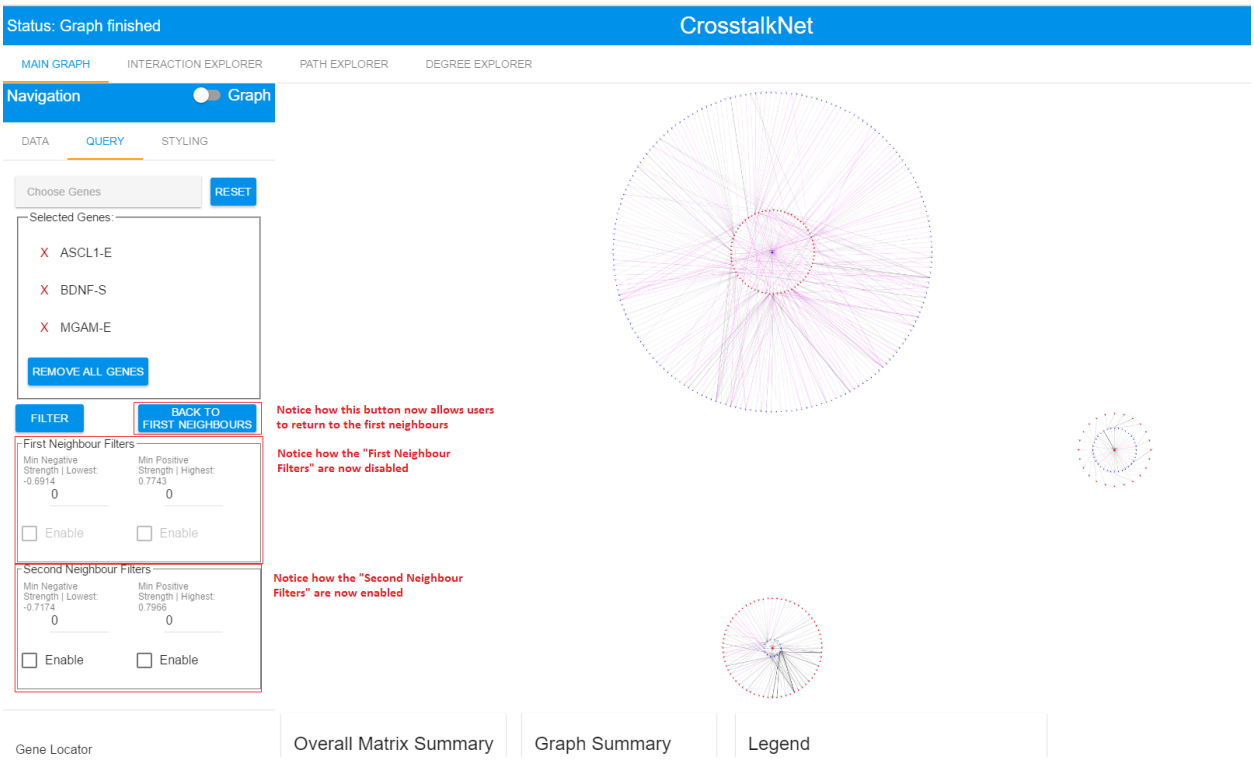


Figure 5: First and second neighbours of ASCL1-E, BDNF-S, MGAM-E.

Filter

After obtaining the first neighbours, they can be filtered so that only interactions of a specified strength are seen. There are 2 input boxes available for filtering first neighbours interactions.

1. The box labeled "Min Negative Correlation | -1" is used to control only the negative interactions. If one types -0.5 in the box, and selects enable, then only negative interactions of <= -0.5 will be shown. The number after the "|" indicates the lowest interaction in the entire graph. Entering a number lower than this or entering a positive number are both prohibited.

2. The box labeled "Min Positive Correlation | 0.8455" is used to control only the positive interactions. If one types 0.7 in the box, and selects enable, then only positive interactions of >= 0.7 will be shown. The number after the "|" indicates the highest interaction in the entire graph. Entering a number higher than this or entering a negative number are both prohibited.

If one enables both filters, then all interactions <= -0.5 and >= 0.7 will be returned. To get the filtered results, simply click the "Filter" button. If the user doesn't want to filter, then the checkboxes should remain unchecked.

The filter option works for the second neighbors as well. The following figures show the effect of filtering first neighbors before (Figure 5) and after (Figure 6) filtering.

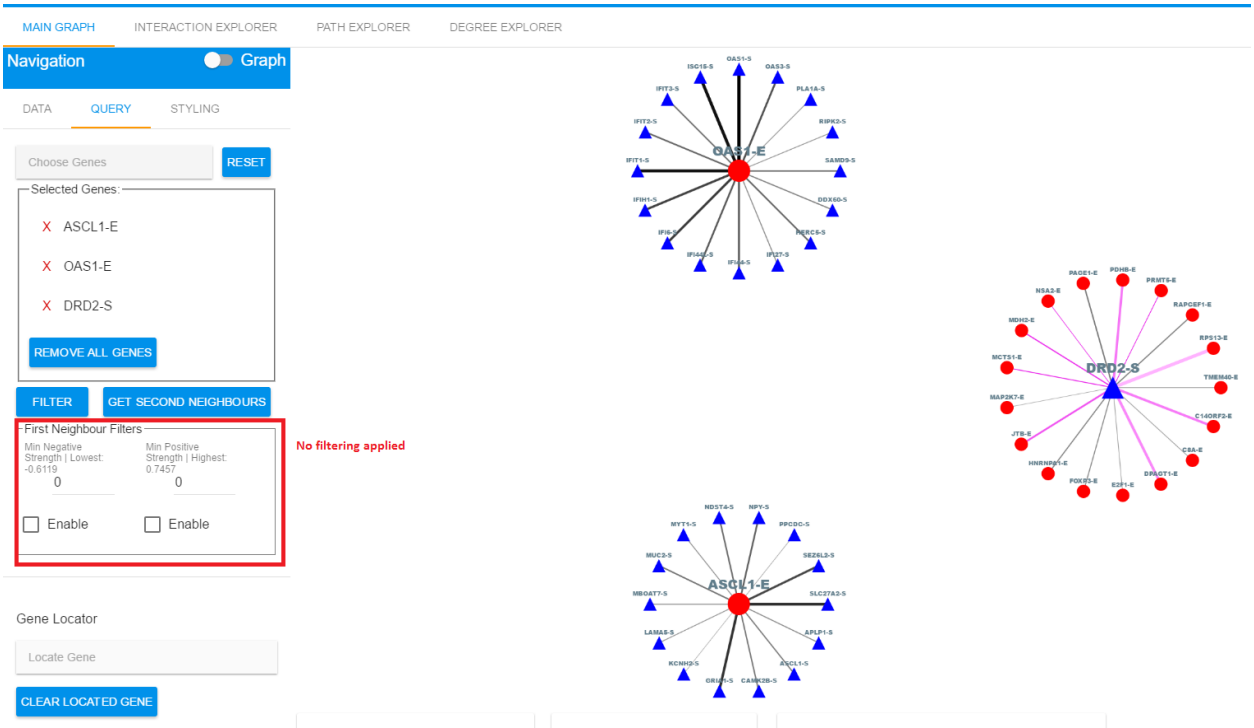


Figure 6: The first neighbors of ASCL1-E, OAS1-E, and DRD2-S before filtering.

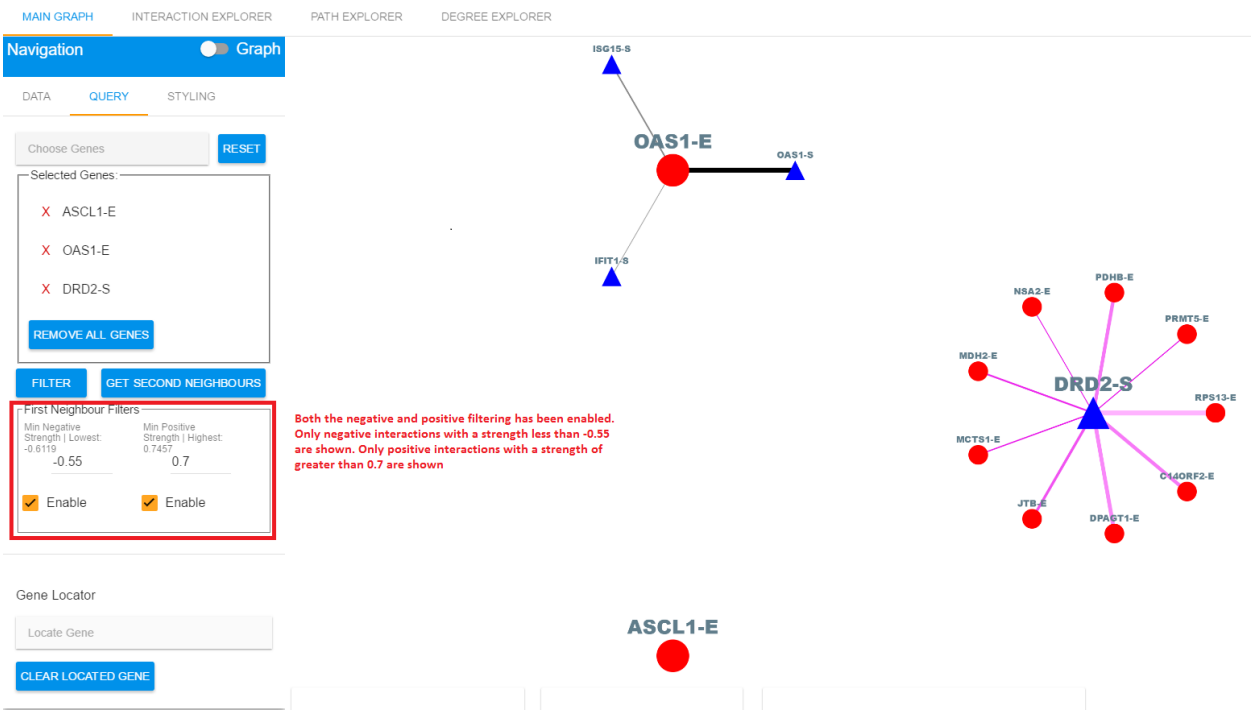


Figure 7: The first neighbors of ASCL1-E, OAS1-E, and DRD2-S after filtering.

Gene locator

The user has the ability to search for a gene of interest in the graph by typing it into the “Gene Locator” text box. If the gene is present in the graph, it will appear in the list of potential matches as the user is typing, and clicking on the gene will cause it to be highlighted green and zoomed in on.

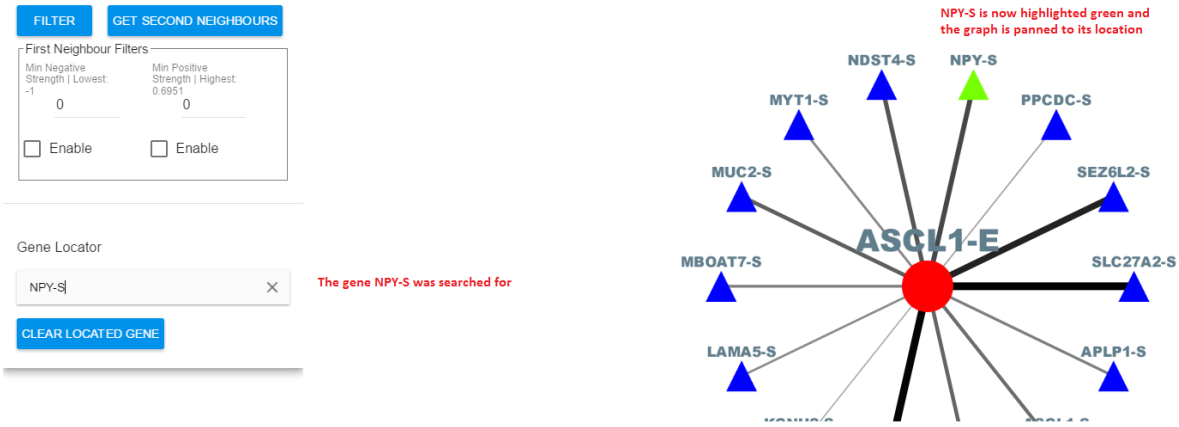


Figure 8: The gene locator used to locate the gene NPY-S.

Edge inspector

This functionality displays information about the edge weight, which is the strength of interaction between any two nodes in a given network. In the case of delta networks, the edge weight

between two nodes provides the difference between the tumor and normal interaction along with the edge weight of normal and tumor networks. To use the edge inspector, simply click on any edge, and the “Edge Inspector” tile will appear at the bottom of the web page.

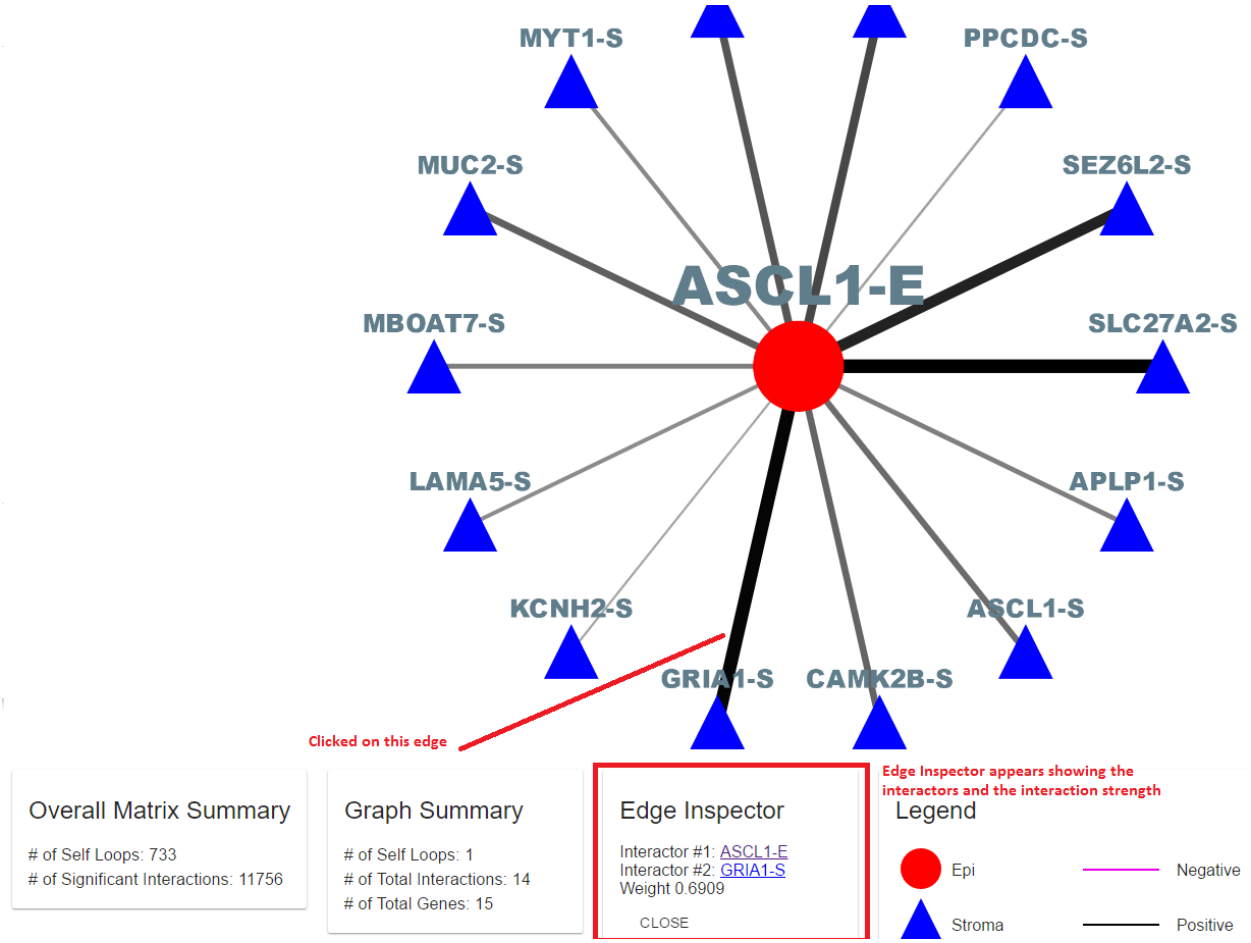


Figure 9: An example of the edge inspector displayed for the edge between ASCL1-E and GRIA1-S.

Styling

Layouts

This option is located in the Styling subtab of the Main Graph tab. There are 3 layouts to display the graph, namely, concentric, random and bipartite. The user can select a layout of choice and the graph gets refreshed automatically. Zoom in and out features are enabled on the graph and the “Reset Zoom” button resets the zoom level (in case the user zooms in too much or too far away and has lost the graph). The “Reset Graph” button will place all the nodes in their original position in addition to resetting the zoom level.

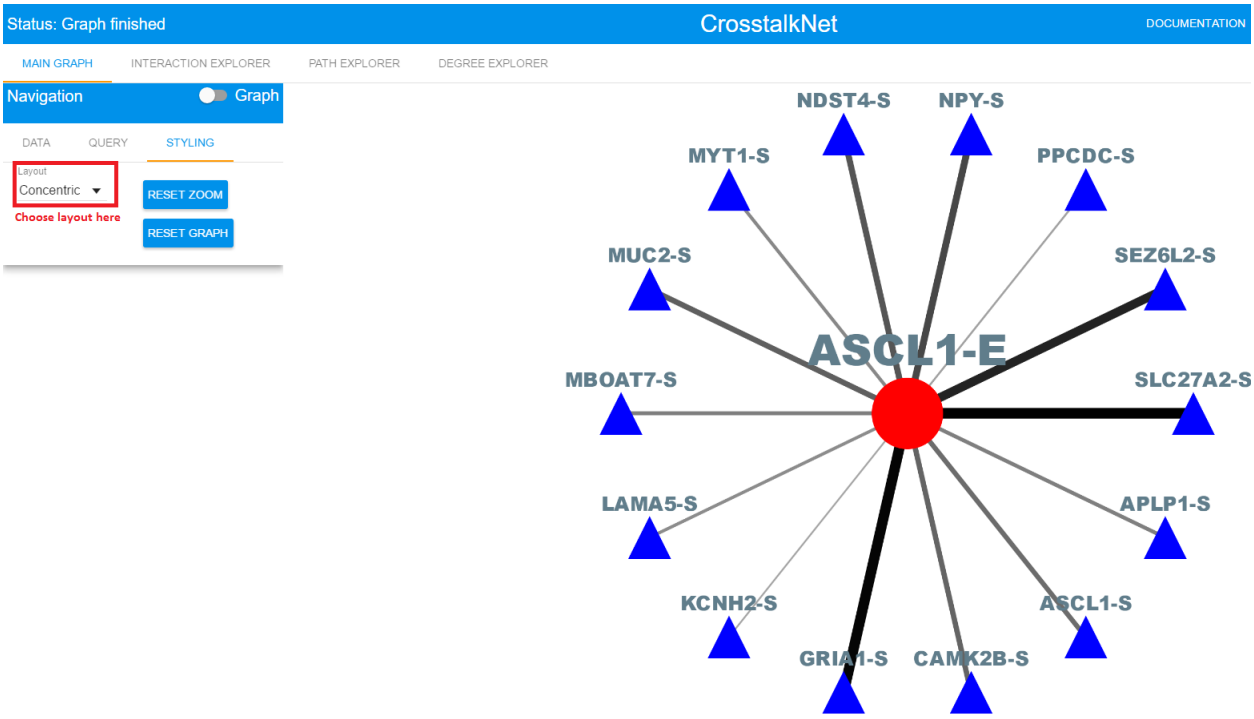


Figure 10: The layout dropdown in the Styling subtab of the Main Graph tab.

Edge Attributes

The color and thickness that an edge has, gives it semantic meaning. Grey to black edges represent positive interactions. Light magenta to magenta edges represent negative interactions. The darkness darkness and thickness of an edge represents its relative magnitude to other interactions. For example, say that a graph has only 4 edges in it with the following correlations: -0.8, -0.6, 0.5, 0.65. The -0.8 edge will be dark magenta and will be thick whereas the -0.6 interaction will be light magenta and will be thin. The 0.65 edge will be dark black and will be thick whereas the 0.5 interaction will be light grey and will be thin.

Network Statistics

There are a total of 4 cards at the bottom of the graph, 3 of which are visible by default and the fourth can be seen when an edge is clicked.

1. **Overall network Summary**: Displays the number of self loops and significant interactions.
2. **Graph Summary**: Shows statistics about the graph that is currently being displayed.
3. **Edge Inspector**: Shows information about a selected edge. Appears only when an edge is clicked.
4. **Legend:** Shows information about the styling in the graph and the meaning of it.

Figure 10 shows these cards.

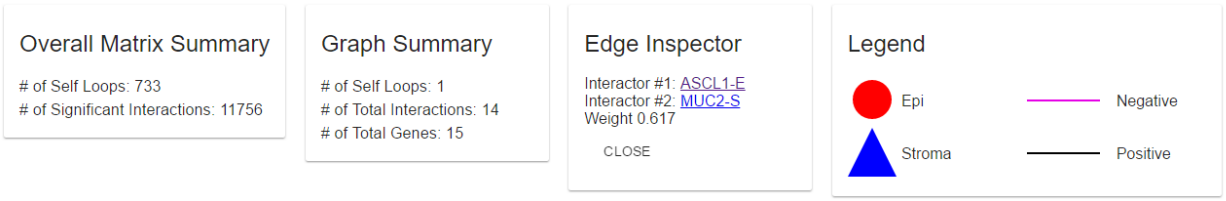
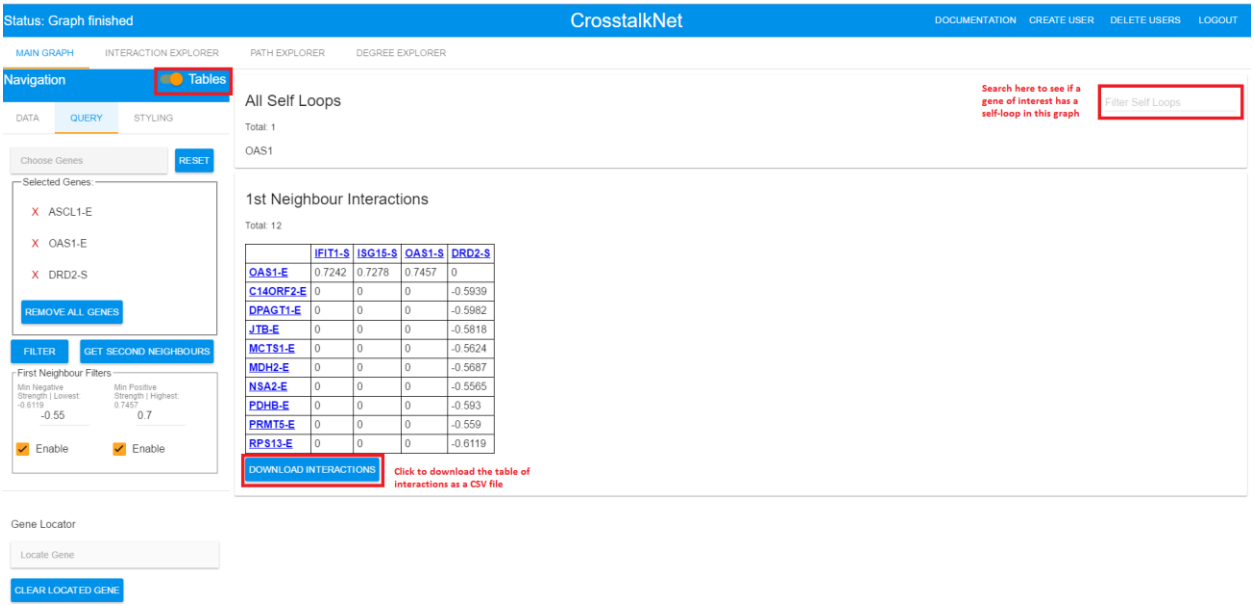


Figure 11: The 4 network statistic cards.

Tabular View

In addition to the graph, the interactions and self-loops can also be seen in a table-view which is more convenient for doing analysis with. To go to the table view, click on the switch shown in the below example:



	IFIT1-S	ISG15-S	OAS1-S	DRD2-S
OAS1-E	0.7242	0.7278	0.7457	0
C14ORF2-E	0	0	0	-0.5939
DPAGT1-E	0	0	0	-0.5982
JTB-E	0	0	0	-0.5818
MCTS1-E	0	0	0	-0.5624
MDH2-E	0	0	0	-0.5687
NSA2-E	0	0	0	-0.5565
PDHB-E	0	0	0	-0.593
PRMT5-E	0	0	0	-0.559
RPS13-E	0	0	0	-0.6119

Figure 12: Tabular view of the Main Graph tab showing self loops as well as the interactions from the graph in a tabular format.

The user can download the table of interactions in a CSV file, using the “Download Interactions” button located at the bottom of a table.

Interaction Explorer

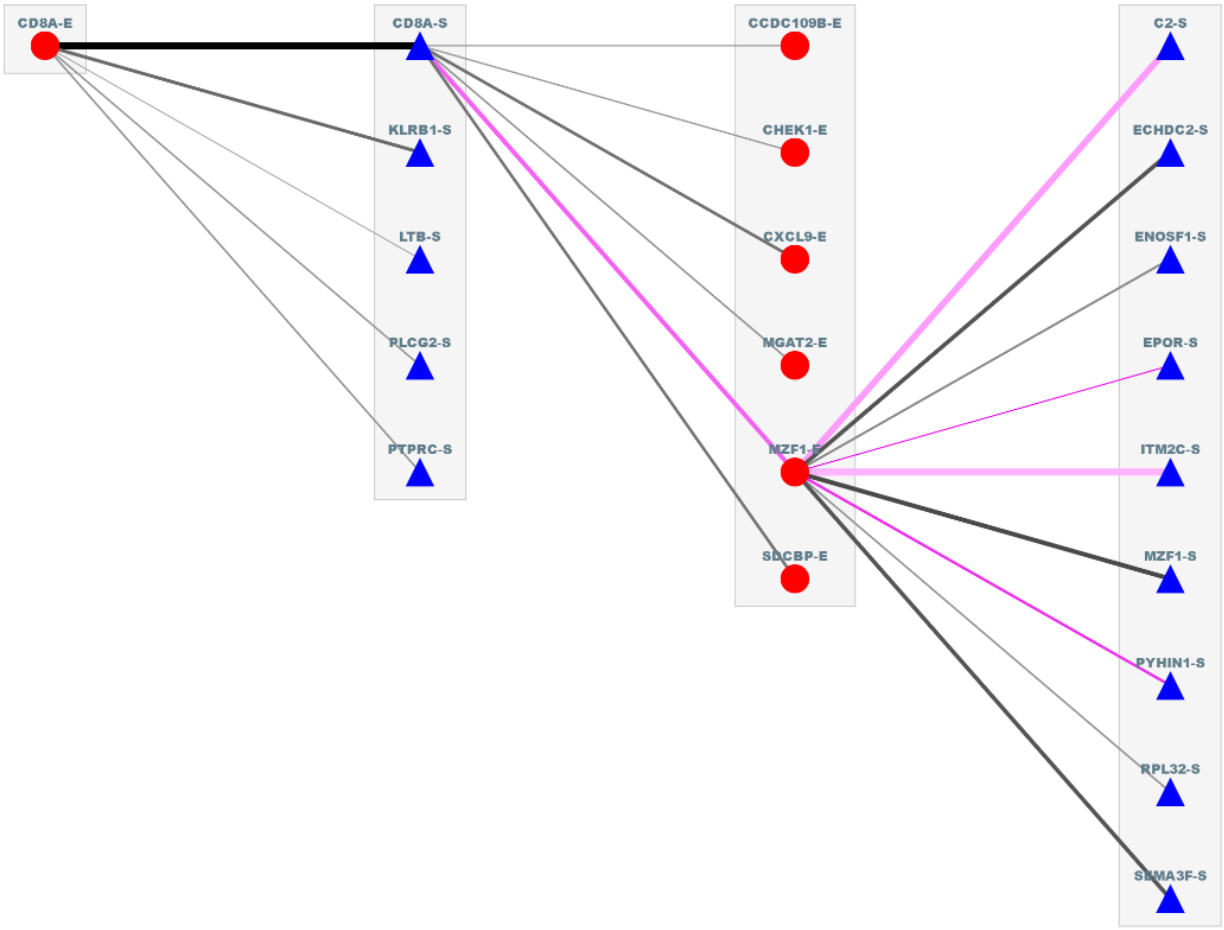


Figure 13: Example graph generated by the interaction explorer.

This tab is used for a more targeted approach to explore the network. First, a user must choose a single gene of interest and then click the "Get Data" button. Once the graph is loaded, the genes available for selection are simply the neighbours of the selected gene. This process can be repeated over and over again to go beyond the 2nd neighbour level available in the main graph. An example is presented below, in which the interactions are displayed for a depth of 3 levels.

Query

Selection of First Gene

To begin using the Interaction explorer, simply type in a gene of interest in the text box and the click the “Get Data” button. Figures 12 and 13 show the result of doing this.

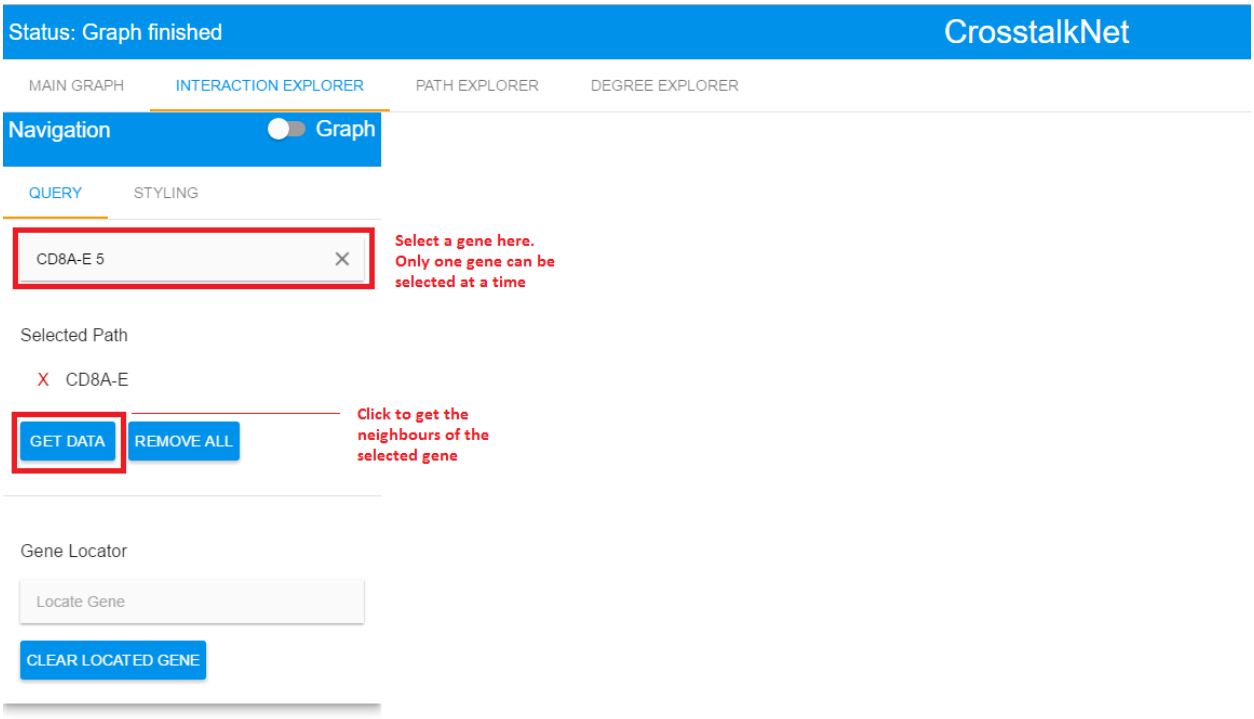


Figure 14: Gene CD8A-E selected in the Interaction explorer as the first gene of interest. Graph has not yet been obtained.

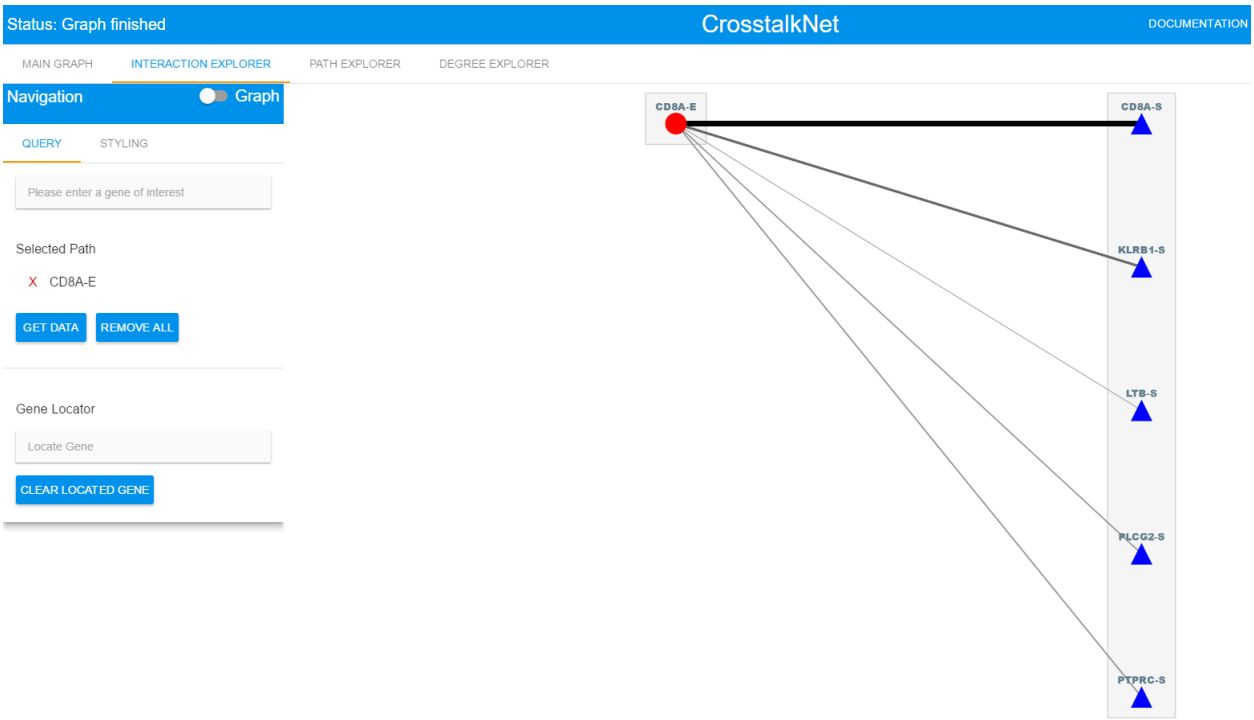


Figure 15: Neighbors obtained in the Interaction Explorer for the gene CD8A-E.

Selection of Additional Genes

Once the neighbours for the first selected gene have been obtained, a user can then obtain the neighbours of any one of those neighbours. When a user goes to enter a gene, the list of available genes is restricted to the most recent group of neighbours obtained. Figures 15 demonstrates this.

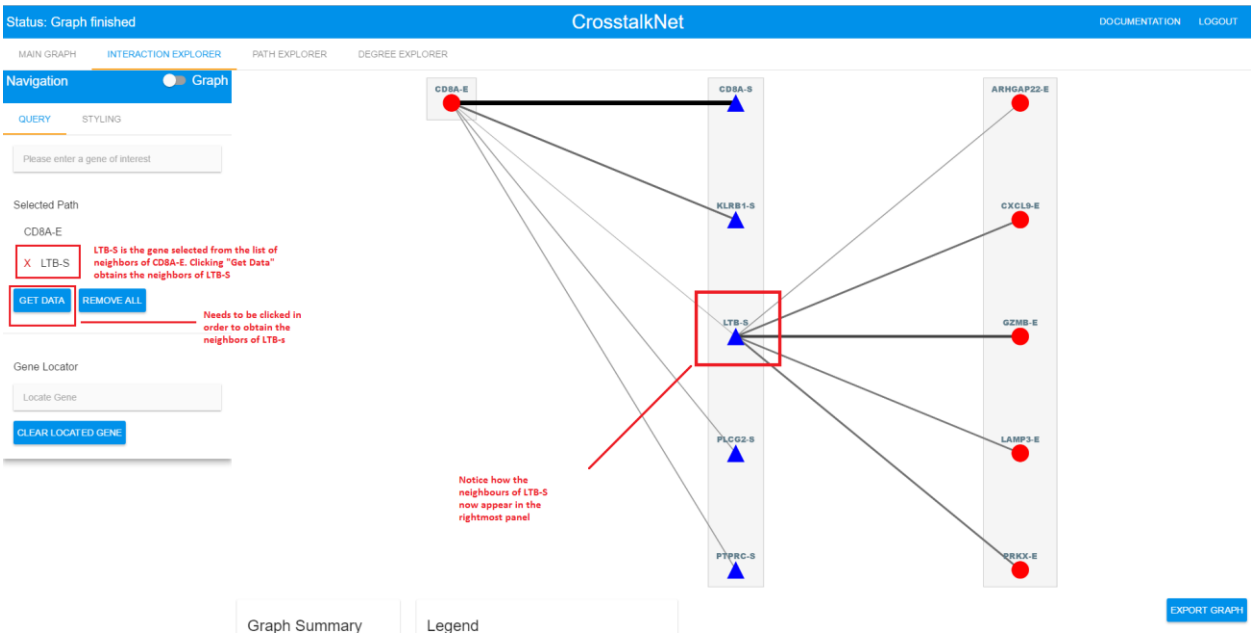


Figure 16: Neighbours obtained in the interaction explorer for the gene LTB-S, where LTB-S is a neighbour of the already selected gene CD8A-E.

Tabular View

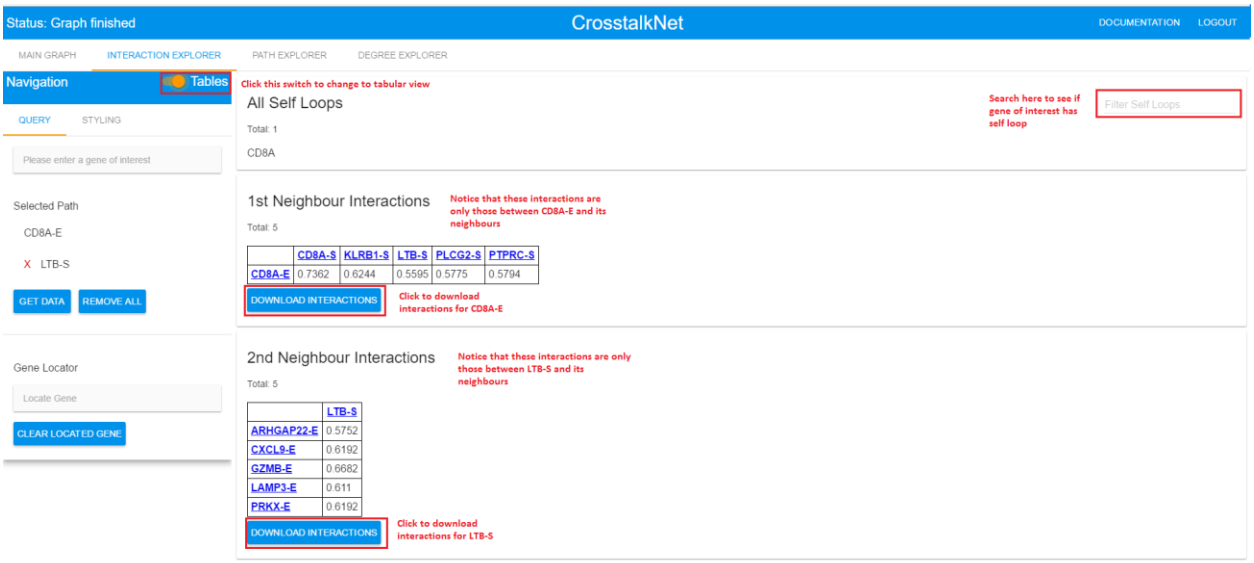


Figure 17: Tabular view of the Interaction Explorer tab showing self loops as well as interactions from the graph in a tabular format.

Path Explorer

This tab displays all the existing interactions between any two selected genes in the network. The results are presented in a tabular view along with the strength of interactions. The genes selected can be from the same compartment (i.e. epi-epi or stroma-stroma) or from different compartments (i.e. epi stroma). If the genes selected are from the same compartment, that means a path between the genes must go through a gene in the opposite compartment. Thus, there will be two edges (interactions) along this path. If the genes selected are from opposite compartments, there can be at most one path between them (i.e. the edge between the genes if there is one).

The paths can also be sorted according to the various columns of the table by clicking on a column header. The column header with a vertical arrow indicates which column the paths are currently sorted by and in which direction. Figure 18 shows the Path Explorer used to find paths between genes in the same compartment, and Figure 19 shows the Path Explorer used to find path between genes in opposite compartments.

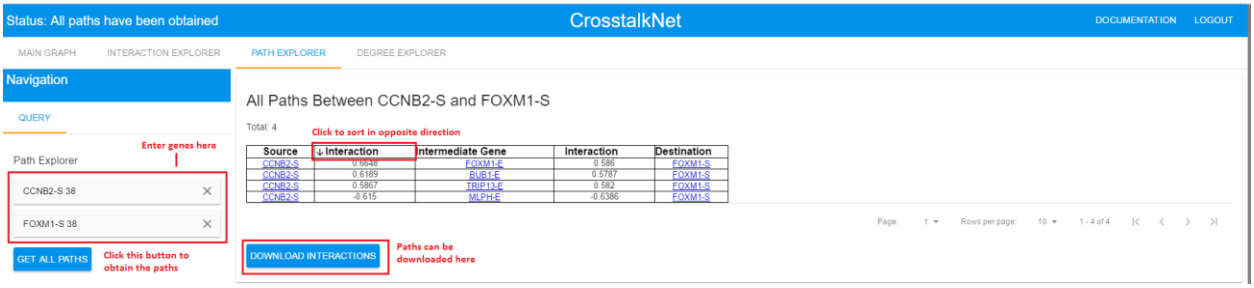


Figure 18: All paths between CCNB2-S and FOXM1-S (same compartment).

Degree Explorer

This functionality allows the user to identify genes expressed in epithelium and stroma according to the degree. The lists of genes can be downloaded in a simple csv format. There are three ways of using the Degree Explorer. The first method is to select the top x genes in the network. This will retrieve the top x epi genes as well as the top x stroma genes. The second method involves specifying a minimum degree for the genes to be returned, and both the epi genes and stroma genes satisfying the specified lower bound will be obtained. The third method is a combination of the previous two and requires a user to specify both top x genes and a minimum degree.

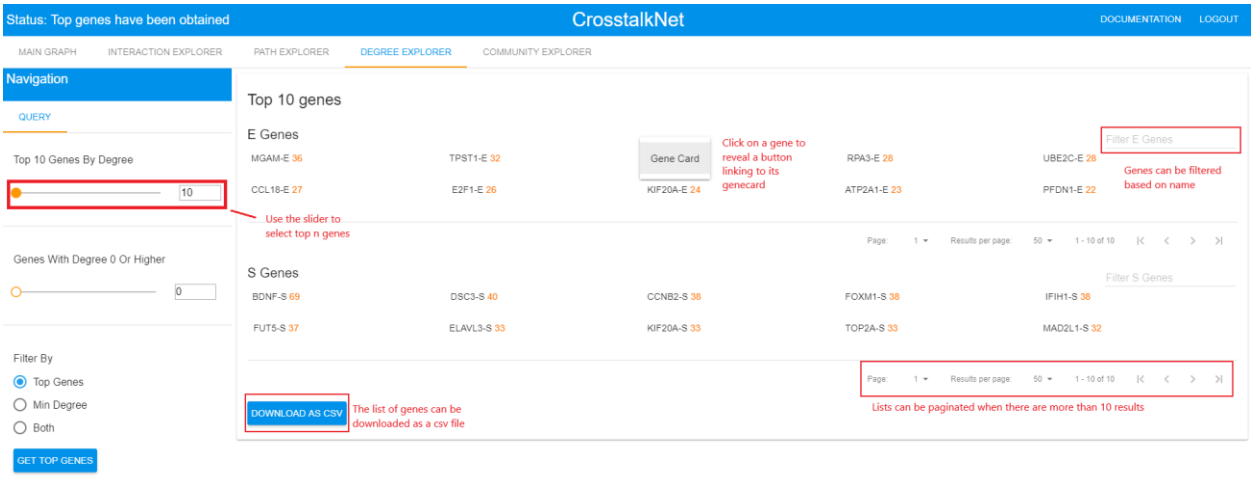


Figure 19: The top 10 genes ordered by degree.

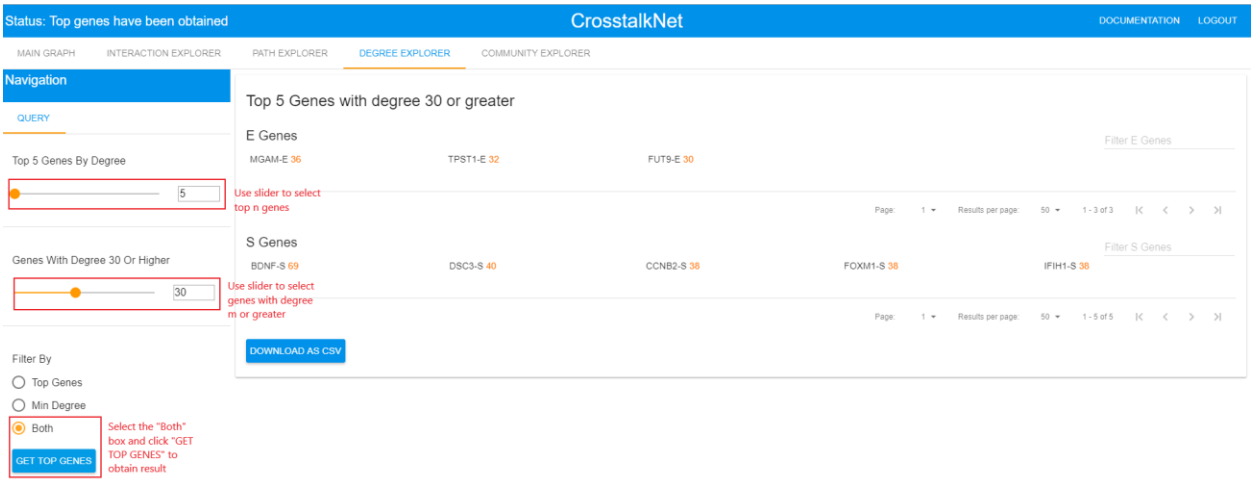


Figure 20: Top 5 genes with degree 30 or greater. Notice that epi only has 3 genes satisfying both conditions.

Browser Limitations

Edge

Using Edge results in a few styling discrepancies on the documentation page, as well as significant performance decreases throughout the app. The performance decrease seems to be due to Angular Material. Furthermore, downloading tables as CSV files does not work.

Firefox

Firefox also suffers from styling discrepancies, however the app runs much faster than when using Edge. Uploading files is the only known functionality that doesn't work with Firefox.

References

Oh, Eun-Yeong, Stephen M. Christensen, Sindhu Ghanta, Jong Cheol Jeong, Octavian Bucur, Benjamin Glass, Laleh Montaser-Kouhsari, et al. 2015. “Extensive Rewiring of Epithelial-Stromal Co-Expression Networks in Breast Cancer.” *Genome Biology* 16 (June): 128.