

## Edges Software Protocol

### Description:

Edges Software can be used to find the combinations of objects within different categories. A combination is an arrangement of objects where order does not matter. The possible number of combinations can be calculated using the following equation:

$$C(n, k) = \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

$n = \text{total number of items}$   
 $k = \text{number of items being chosen at a time}$

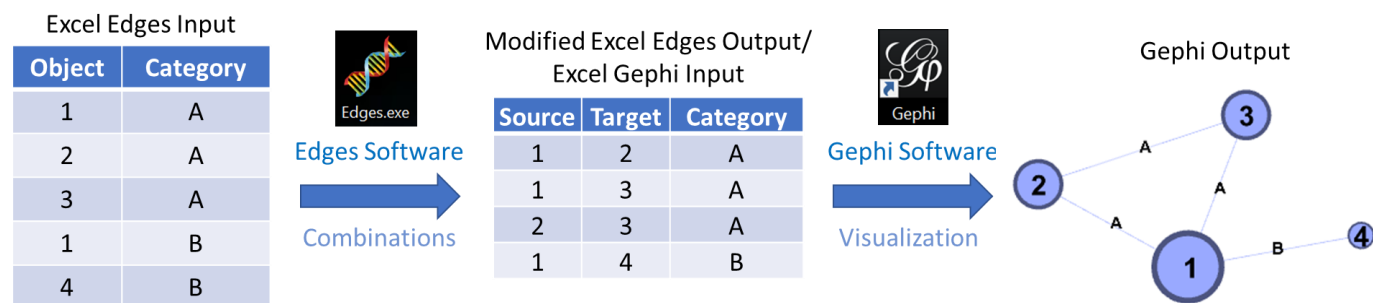
**Example 1: Determine relationship of 4 given objects that interact through 2 categories.**

$$C(4, 2) = \binom{4}{2} = \frac{4!}{2!(4-2)!} = \frac{4 \times 3 \times 2 \times 1}{2 \times 1 \times 2 \times 1} = \frac{12}{2} = 6$$

$n = 4 \text{ objects}$   
 $k = 2 \text{ categories}$

An excel file containing objects and their relationships through categories can be inputted into edges software. Edges creates an output excel file that contains the combinations of objects within the categories. After a few minor modifications, this edges output file can be inputted into the open-source graphing software, Gephi, to build a visual network of the excel file data.

**Example 2: Determine the relationship of 4 given objects (labeled 1-4) that interact through 2 categories (labeled A or B).**

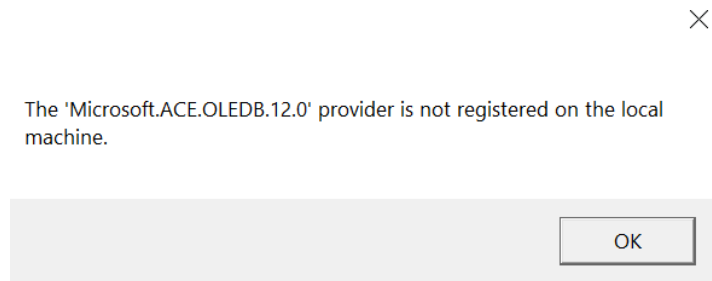


Objects 1,2, and 3 interact with each other through category A. Objects 1 and 4 interact with each other through category B. Objects 2 and 3 do not interact through category B and are therefore not connected to object 4.

This software was originally designed by Ehab Rizek to create a target-(pathway)-target network to analyze the pathway connections between protein targets, but this software can be used to find the combinations of any objects within different categories.

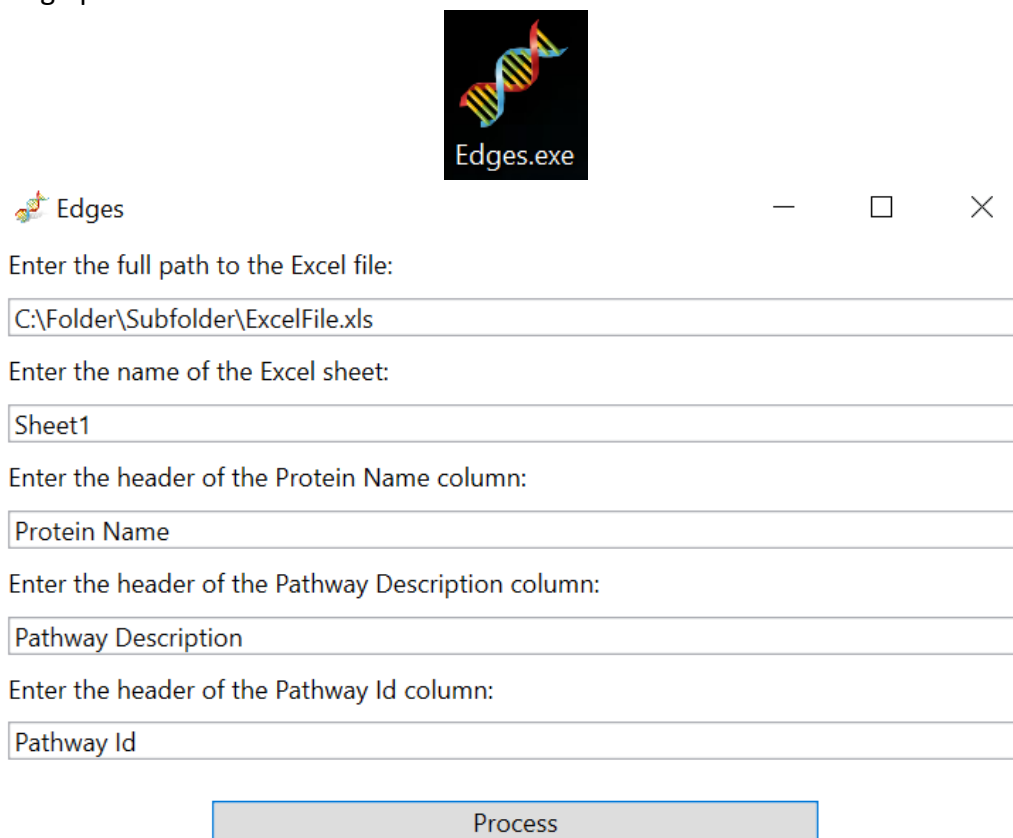
## Requirements:

Requires Access Database to run. If you do not have this the following error message will appear.



## Instructions:

1. Download Edges Software and Access Database.
2. Place the Edges Software on the desktop or in an easily accessible folder. The Edges output file will be placed in the same location as the Edges Software.
3. Open Edges Software by double clicking the icon, or right clicking on the icon and then selecting open.



4. Obtain objects, categories, and define relationship between objects and categories.

**Objects**

Protein Name
CDK1
FGF2
JAK2
VEGFA

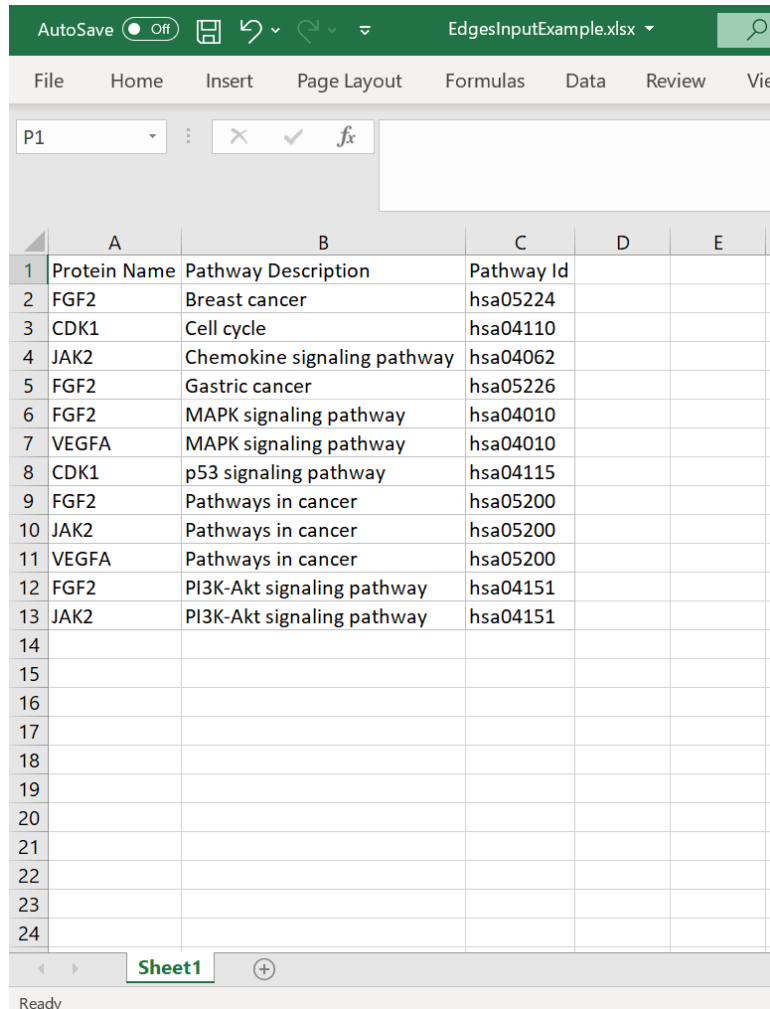
**Categories**

Pathway Description	Pathway Id
Breast cancer	hsa05224
Cell cycle	hsa04110
Chemokine signaling pathway	hsa04062
Gastric cancer	hsa05226
MAPK signaling pathway	hsa04010
p53 signaling pathway	hsa04115
Pathways in cancer	hsa05200
PI3K-Akt signaling pathway	hsa04151

**Object-Category Relationship**

Protein Name	Pathway Description	Pathway Id
FGF2	Breast cancer	hsa05224
CDK1	Cell cycle	hsa04110
JAK2	Chemokine signaling pathway	hsa04062
FGF2	Gastric cancer	hsa05226
FGF2	MAPK signaling pathway	hsa04010
VEGFA	MAPK signaling pathway	hsa04010
CDK1	p53 signaling pathway	hsa04115
FGF2	Pathways in cancer	hsa05200
JAK2	Pathways in cancer	hsa05200
VEGFA	Pathways in cancer	hsa05200
FGF2	PI3K-Akt signaling pathway	hsa04151
JAK2	PI3K-Akt signaling pathway	hsa04151

5. Format the Edges input file to look like the following excel sheet:



	A	B	C	D	E
1	Protein Name	Pathway Description	Pathway Id		
2	FGF2	Breast cancer	hsa05224		
3	CDK1	Cell cycle	hsa04110		
4	JAK2	Chemokine signaling pathway	hsa04062		
5	FGF2	Gastric cancer	hsa05226		
6	FGF2	MAPK signaling pathway	hsa04010		
7	VEGFA	MAPK signaling pathway	hsa04010		
8	CDK1	p53 signaling pathway	hsa04115		
9	FGF2	Pathways in cancer	hsa05200		
10	JAK2	Pathways in cancer	hsa05200		
11	VEGFA	Pathways in cancer	hsa05200		
12	FGF2	PI3K-Akt signaling pathway	hsa04151		
13	JAK2	PI3K-Akt signaling pathway	hsa04151		
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					
24					

**Headers should describe the objects and categories. Headers must be located in the first row of the excel sheet.**

**Objects and categories can be in any column.**

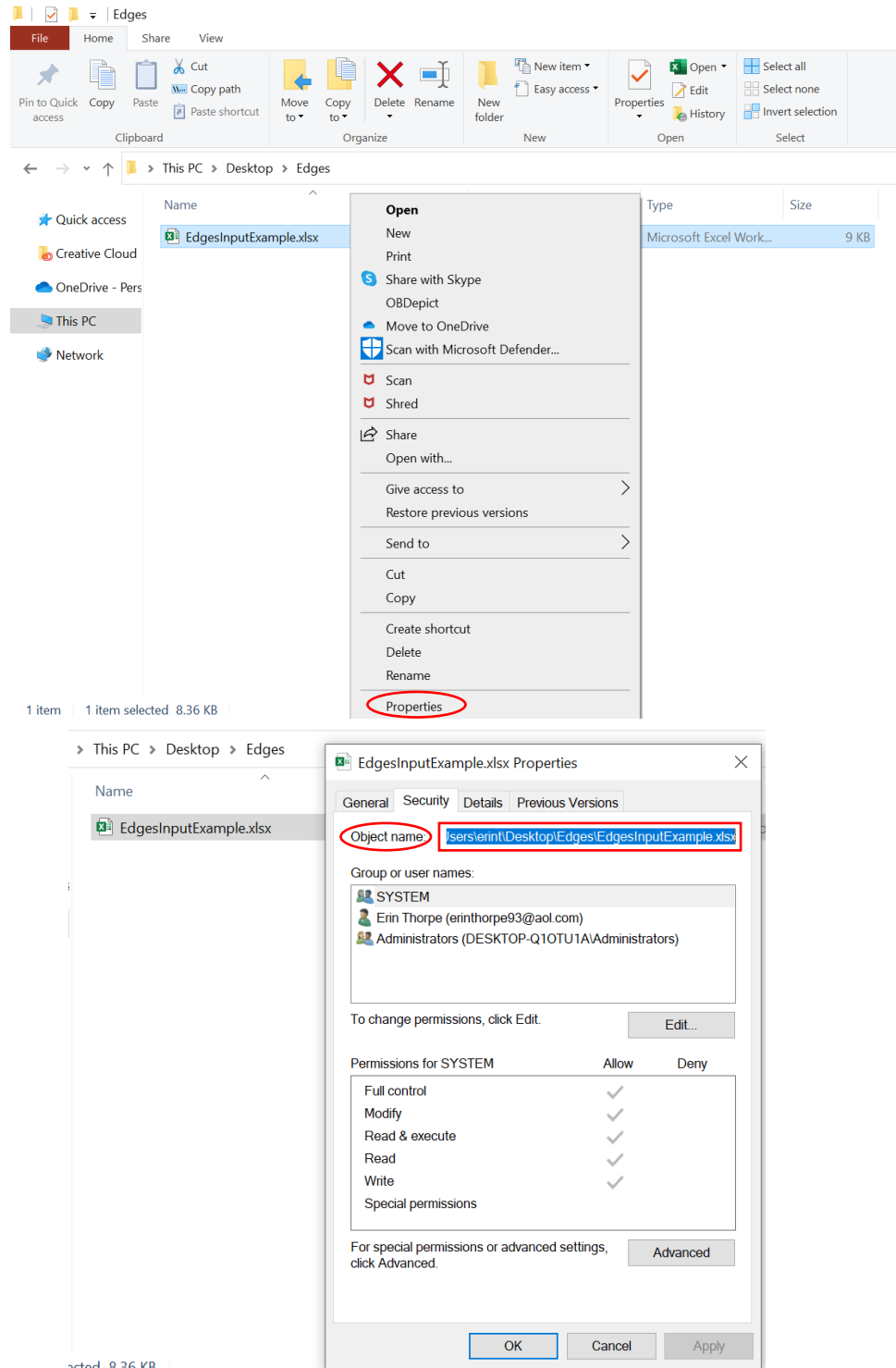
**Objects can be numerical or a combination of strings (words).**

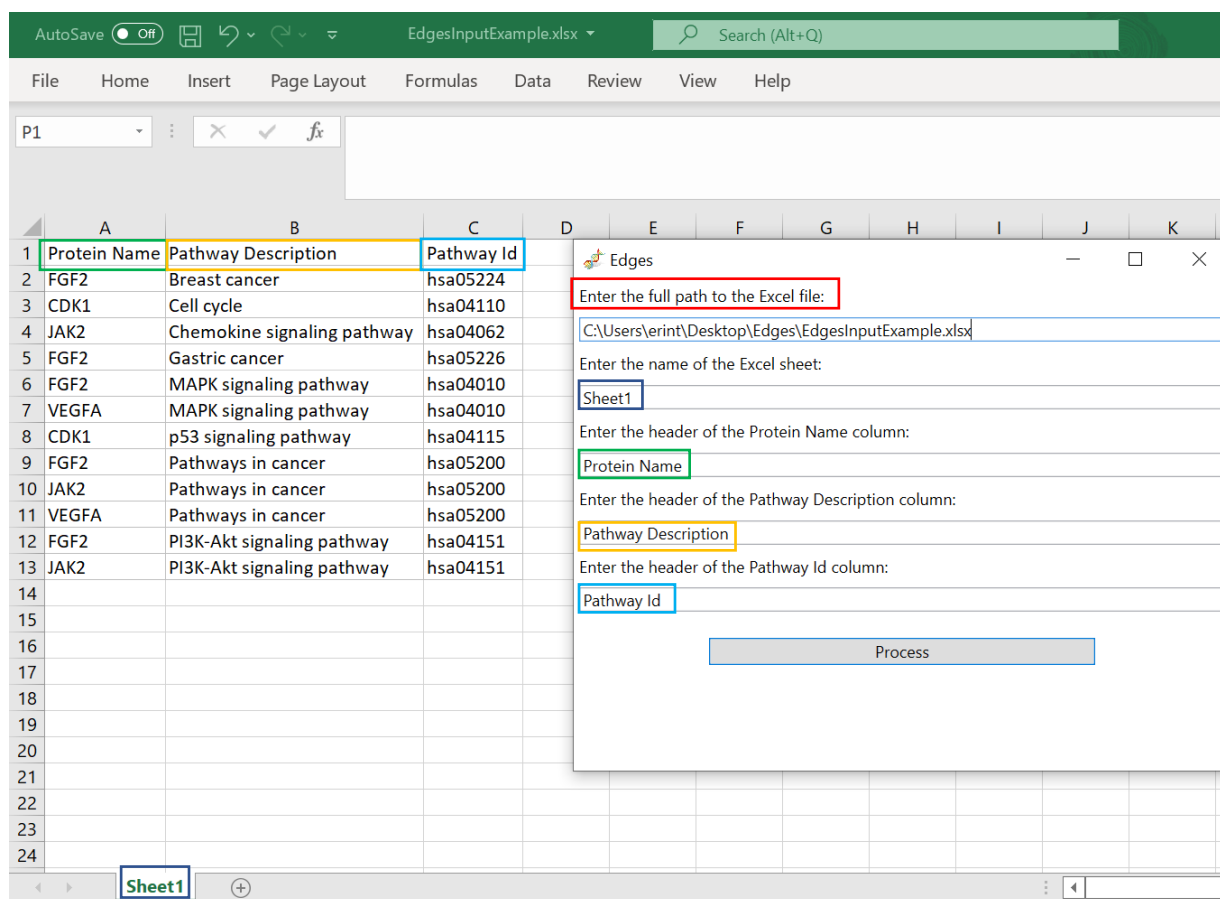
**All columns shown are necessary for the program to run successfully. If there is not a secondary identifier for your category, like how pathway Id is the secondary identifier for the pathway description, then input the same data into both the Pathway Description and Pathway Id columns.**

**Alphabetical order is not required.**

**Trailing and leading spaces need to be removed from dataset prior to analysis for accurate results.**

5a.) Enter the full path to the excel file by going to where the excel file is located, clicking on the excel file to highlight it, then right click on it, and click on properties. Go to the security tab and copy and paste the object code into the full path input on the Edges Software interface.





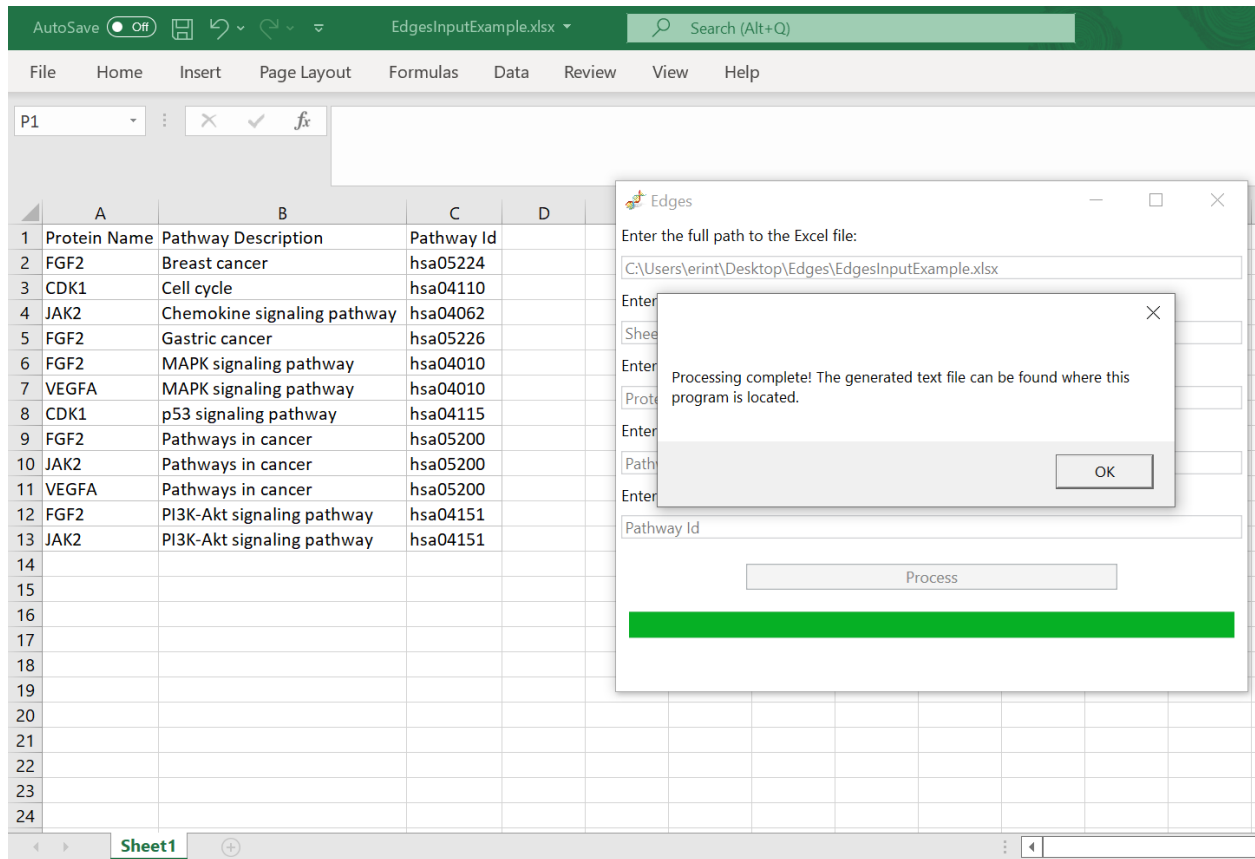
5b.) The name of excel sheet needs to match the name that you have given your excel sheet. In this example, the excel sheet name is the default “Sheet1”.

5c.) The header of the Protein Name column should match the name entered to describe the column that contains the objects. In this example, the object header is “Protein Name”.

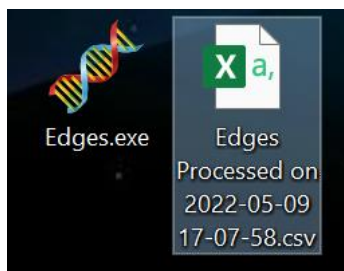
5d.) The header of the Pathway Description column should match the name entered to describe the column that contains the categories. In this example, the category header is “Pathway Description”.

5e.) The header of the Pathway Id column should match the name entered to describe the column that contains a secondary identifier for the categories (if applicable). This feature is useful if there is a secondary identifier for the categories that is unique to each category. If there is no secondary identifier, enter the same data that was in the Pathway Description column into the Pathway Id column. In this example, there is a secondary identifier for each category and the header of the secondary identifier category header is “Pathway Id”.

- Double check that all inputs are correct and then click process on the user interface of the Edges Software. This usually takes less than 2 minutes. Once the processing is complete, you will see the following message:



- The output file from the Edges software will be placed in the same location where the software is located. The file will be time stamped with the date and time according to a 24-hour clock. The output file contains the combinations of objects within categories.



AutoSave Edges Processed on 2022-05-09 17:07:58.csv Search (Alt+Q) Erin Thorpe ET

File Home Insert Page Layout Formulas Data Review View Help

E8

	A	B	C	D
1	FGF2	VEGFA	MAPK signaling pathway	hsa04010
2	FGF2	JAK2	Pathways in cancer	hsa05200
3	FGF2	VEGFA	Pathways in cancer	hsa05200
4	JAK2	VEGFA	Pathways in cancer	hsa05200
5	FGF2	JAK2	PI3K-Akt signaling pathway	hsa04151
6				
7				
8				

Edges Processed on 2022-05-09 17:07:58.csv

Note: If no combinations are possible, this data will be excluded from Edges output excel file. Two objects must interact through a common category to be included. In this example, protein CDK1 was the only protein involved in the p53 signaling pathway (hsa04115) so therefore it was excluded from the Edges output file.

8. Visualize combinations using Gephi. Add headers required by Gephi to the Edges output file and save. The first column containing the objects should be labeled “Source” and the second column containing the objects should be labeled as “Target”. The columns containing the categories and category Ids can be named the same as they were in the Edges input excel sheet. In this example, the nodes are the proteins, and the edges are the common pathways between them. Two or more common pathways between proteins are visualized as a single edge. Nodes were ranked based on degree of interaction so the more pathways a protein is involved in, the larger the node size.

	A	B	C	D
1	Source	Target	Pathway Description	Pathway Id
2	FGF2	VEGFA	MAPK signaling pathway	hsa04010
3	FGF2	JAK2	Pathways in cancer	hsa05200
4	FGF2	VEGFA	Pathways in cancer	hsa05200
5	JAK2	VEGFA	Pathways in cancer	hsa05200
6	FGF2	JAK2	PI3K-Akt signaling pathway	hsa04151
7				
8				

