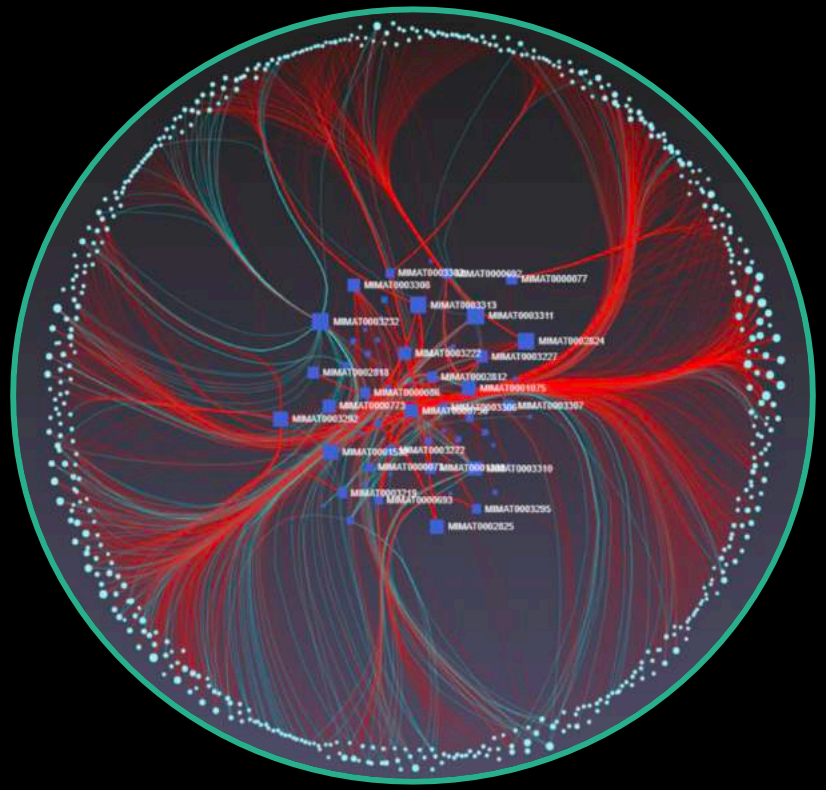
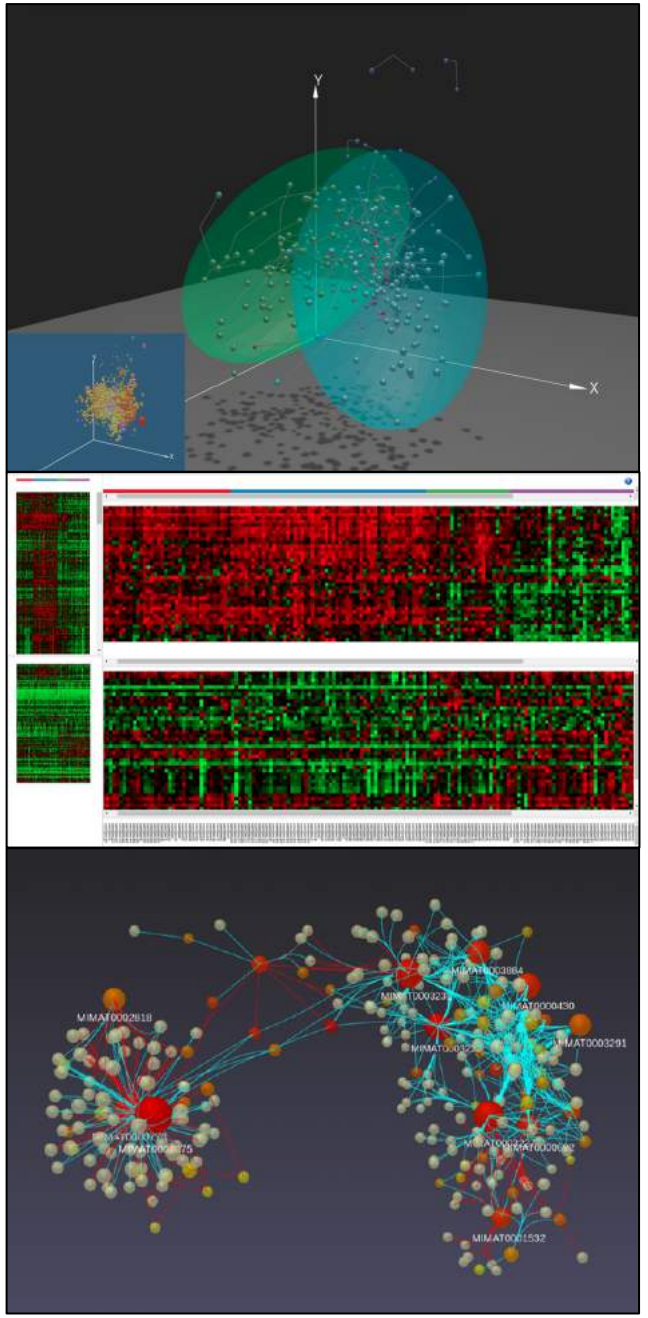
TUTORIAL 4:
Correlation + network analysis



Intro to OmicsAnalyst

- Web-based platform designed for data-driven multi-omics integration and visualization
- Designed to be accessible to bench scientists rather than bioinformatician
- Integrates well-established multivariate and univariate statistics with innovative visual analytics to support:
  - Integrative multi-omics analysis
  - Clustering and pattern discovery
  - Correlation analysis



Requirements

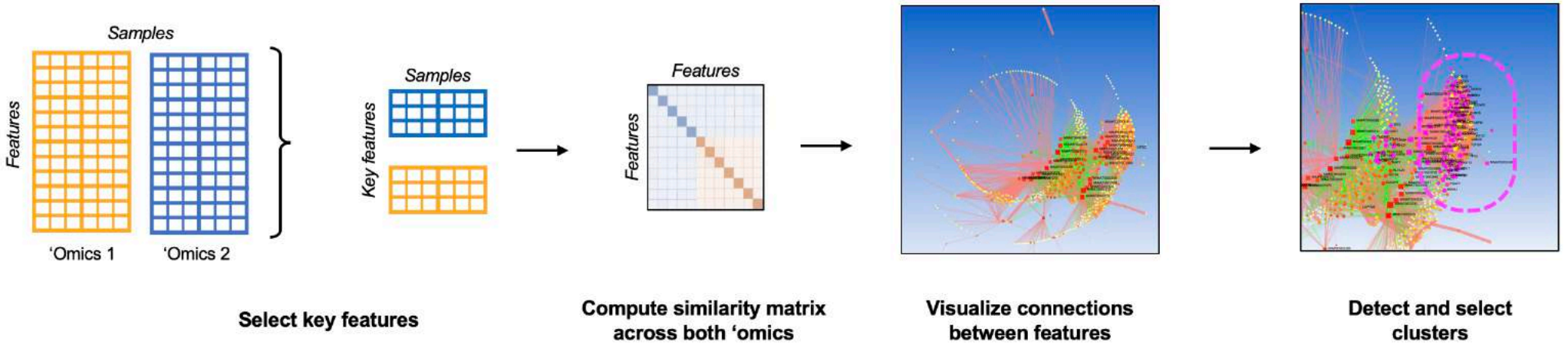
- Modern browser supporting WebGL.
- Ensure that WebGL is enabled in your browser!
  - Go to this page https://get.webgl.org/ to verify your browser supports WebGL.
  - Refer to the FAQs for instructions on how to enable WebGL.
- For the best performance and visualization, use the latest version of Google Chrome.

Goal for this tutorial

- A challenge for multi-omics analysis is to interpret highly complex and heterogeneous data without being overwhelmed by it
- Visualizing global patterns and clusters facilitate data interpretation
- The network analysis track focuses on presenting relationships within and between ‘omics layers, to understand the correlation structures between features
- The goal of this tutorial is to:
  - Introduce OmicsAnalyst’s tool for network analytics
  - Targeted analysis by performing functional enrichment analysis on selected regions from the example network

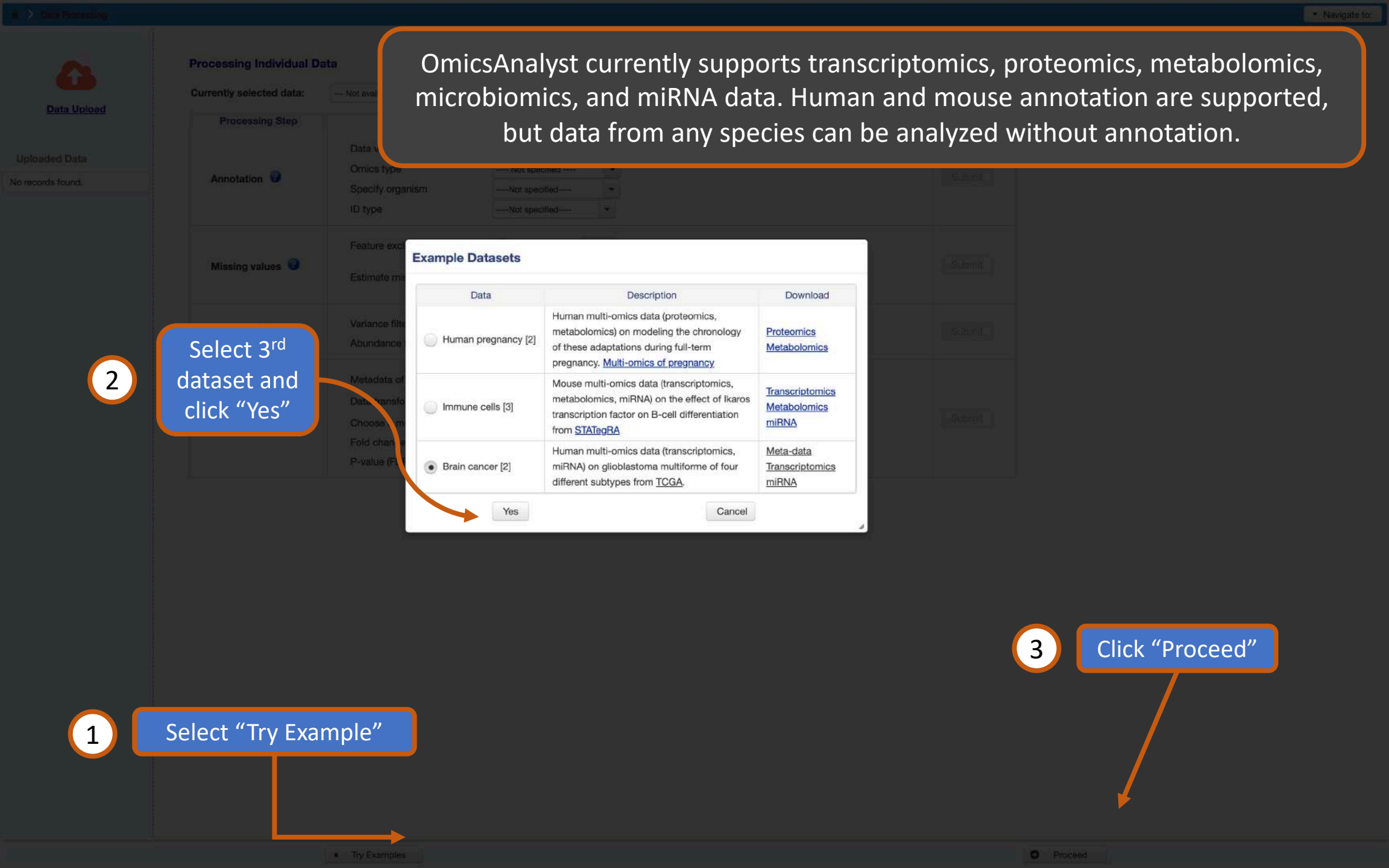
Correlation analysis + network visualization

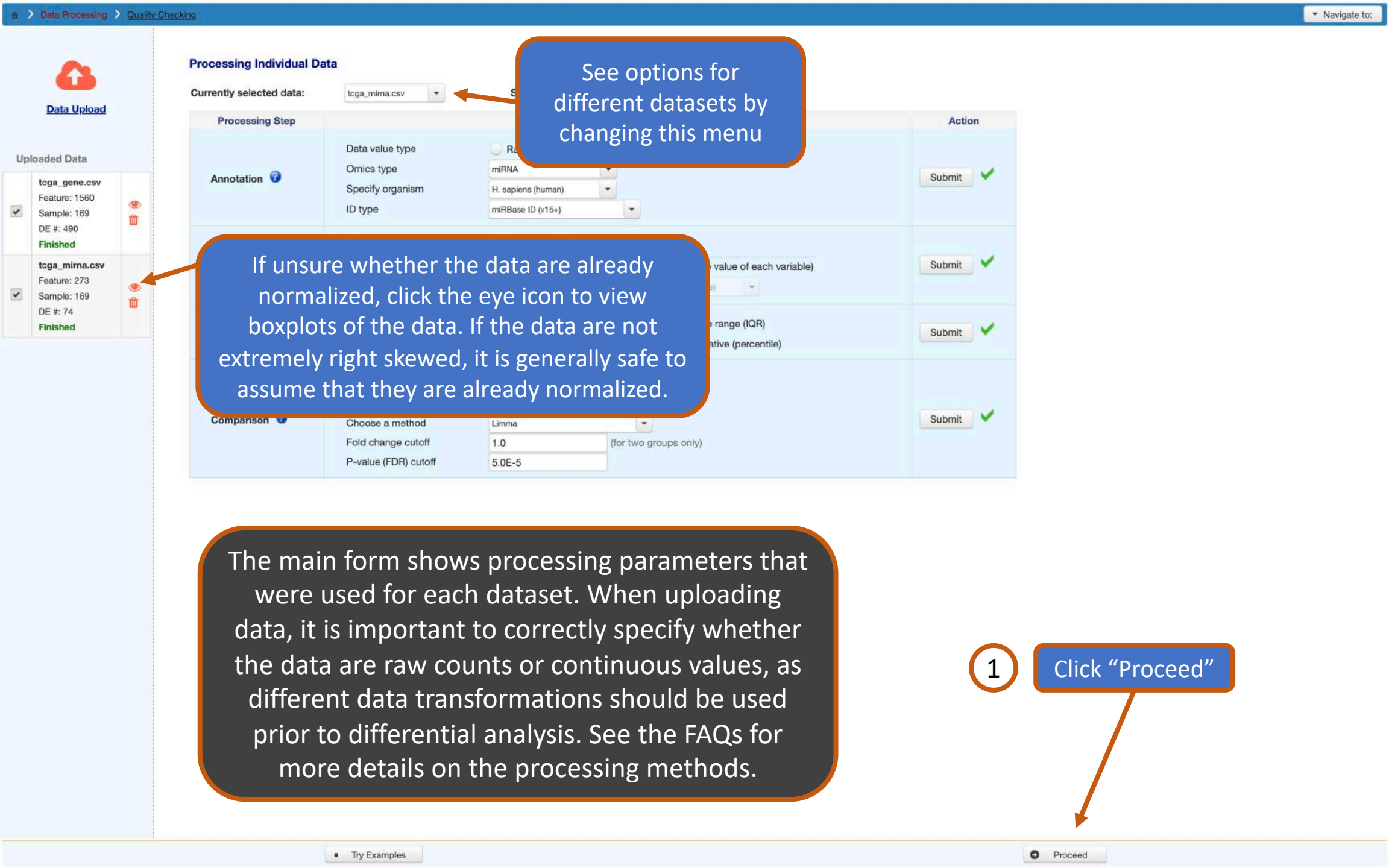
- An overview of this track is:



- To start the tutorial, click **GET STARTED** on the homepage

OmicsAnalyst currently supports transcriptomics, proteomics, metabolomics, microbiomics, and miRNA data. Human and mouse annotation are supported, but data from any species can be analyzed without annotation.




Data Processing
Quality Checking
Navigate to:
Data Upload
Uploaded Data
tcga_gene.csv
Feature: 1560
Sample: 169
DE #: 490
Finished
tcga_mirna.csv
Feature: 273
Sample: 169
DE #: 74
Finished
Processing Individual Data
Currently selected data:
tcga_mirna.csv
See options for different datasets by changing this menu
Processing Step
Action
Annotation
Data value type
Omics type
miRNA
Specify organism
H. sapiens (human)
ID type
miRBase ID (v15+)
Submit
If unsure whether the data are already normalized, click the eye icon to view boxplots of the data. If the data are not extremely right skewed, it is generally safe to assume that they are already normalized.
Comparison
Choose a method
Limma
Fold change cutoff
1.0
(for two groups only)
P-value (FDR) cutoff
5.0E-5
The main form shows processing parameters that were used for each dataset. When uploading data, it is important to correctly specify whether the data are raw counts or continuous values, as different data transformations should be used prior to differential analysis. See the FAQs for more details on the processing methods.
1
Click “Proceed”
Try Examples
Proceed

The page provides graphics to ensure that the data has been properly normalized. Since different transformation may be preferred compared to differential analysis, you can update the method here. Scaling the datasets so that have comparable distributions across 'omics types is recommended.

Uploaded Data

tcga_gene.csv
Feature: 1560
Sample: 169
DE #: 490

tcga_mirna.csv
Feature: 273
Sample: 169
DE #: 74

See the number of differential features from each dataset. Click the icon to update thresholds.

Data Quality Check

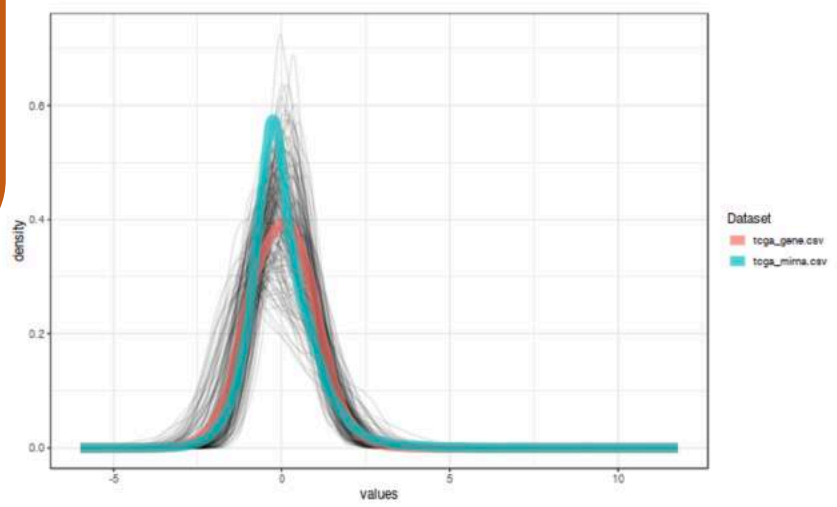
The uploaded omics datasets are summarized below:

Total number of samples:	169
Group names:	Classical; Mesenchymal; Neural; Proneural
Individual datasets:	tcga_gene.csv; tcga_mirna.csv
Corresponding feature number:	1560; 273

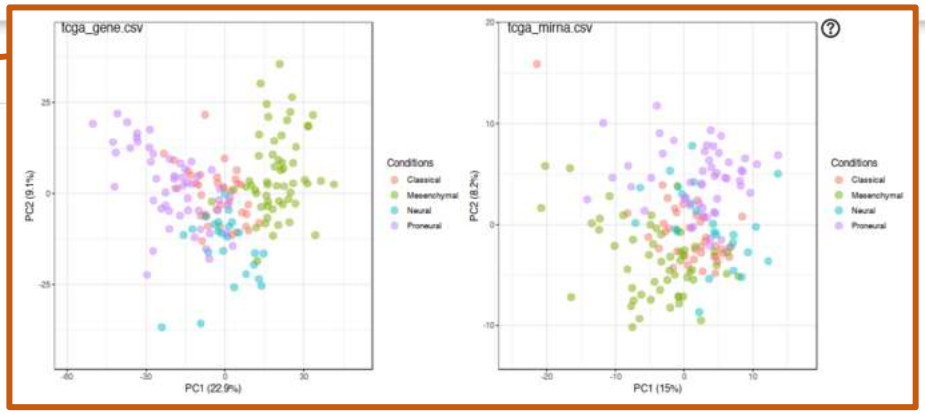
The density plot and PCA plots are generated to provide an overview of the omics datasets for qual
the distributions across different omics layers. If the overall distribution seems to be in very different ran
accounting for batch effect prior to uploading your data to OmicsAnalyst, for example with the R package COMBAT.

Current omics data: tcga_mirna.csv
Sample normalization: None
Data transformation: None
Data scaling: Auto scaling
Update

Density plot | PCA plot | t-SNE plot



1 Look at the "PCA plot"



2 Click "Proceed"

Previous | Proceed

The network analysis requires two selections: the first determines how key 'omics features will be selected, and the second determines how the network will be constructed from the key features.

Uploaded Data

tcga_gene.csv
Feature: 1560
Sample: 169
DE #: 490

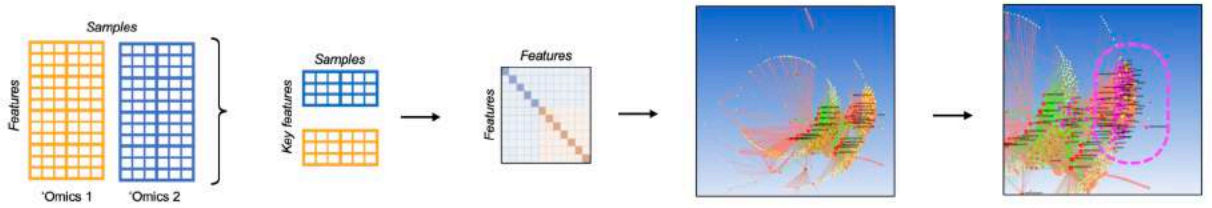
Please choose a method to proceed

Correlation Network Analysis | Cluster Heatmap Analysis | Dimension Reduction

Stay on "Correlation Network Analysis" tab

1

The objective of this analysis is to understand relationships between key features between tw... ways to do this: either we select features that vary significantly across experimental groups using the differential analysis at the data upload step, or we ...duct... sets of features that are both highly connected within individual 'omics datasets and highly associated across 'omics datasets. For the ...top ...es from each component and 'omics data are selected based on the loading scores. Next, we compute the pairwise similarity of selected ...es. Various downstream edge and node filters are also provided to help further refine the network.



Select key features | Compute similarity matrix across both 'omics | Visualize connections between features | Detect and select clusters

Feature selection method: Statistically significant features
Similarity matrix method: Spearman

Proceed

3

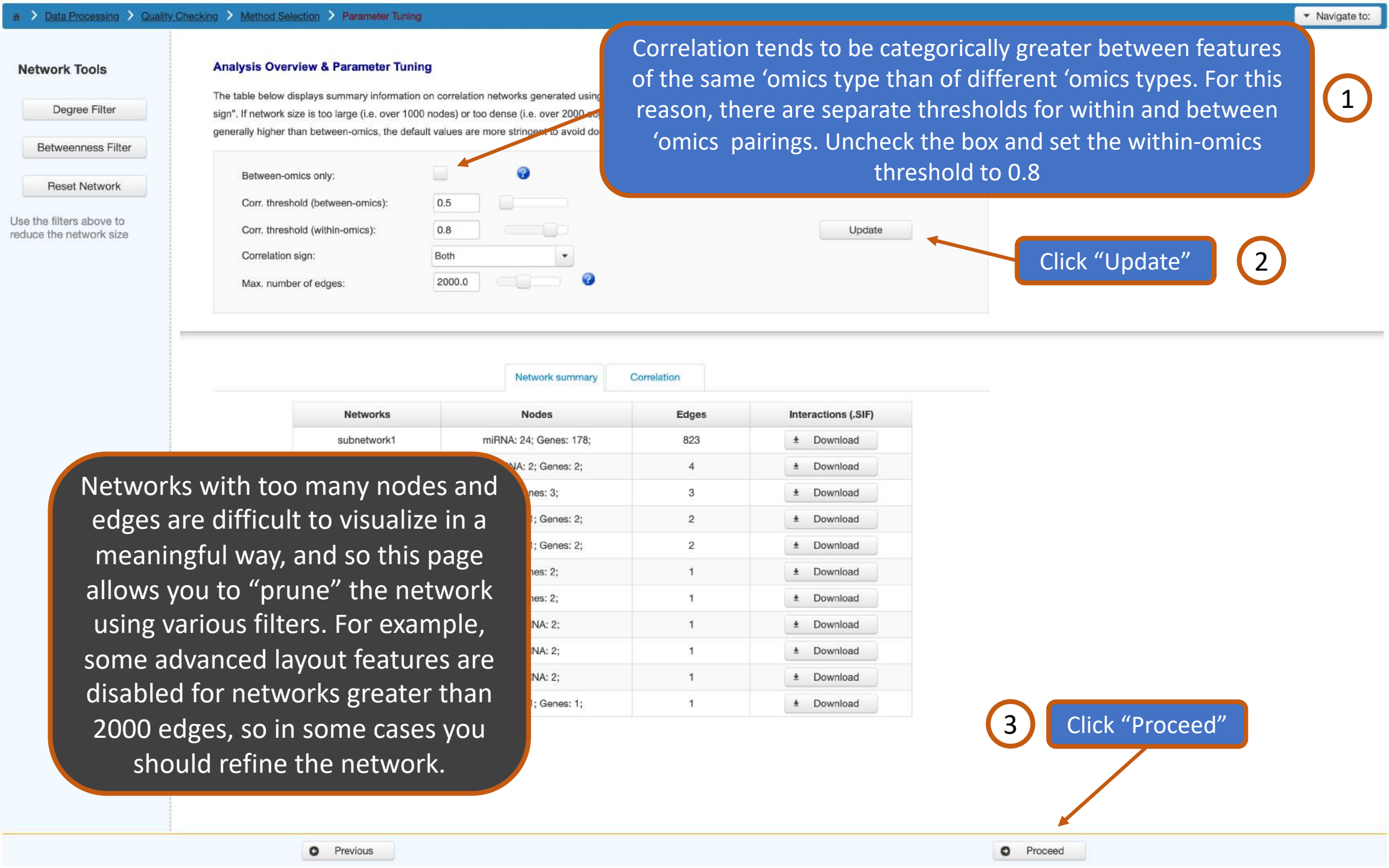
Click "Proceed"

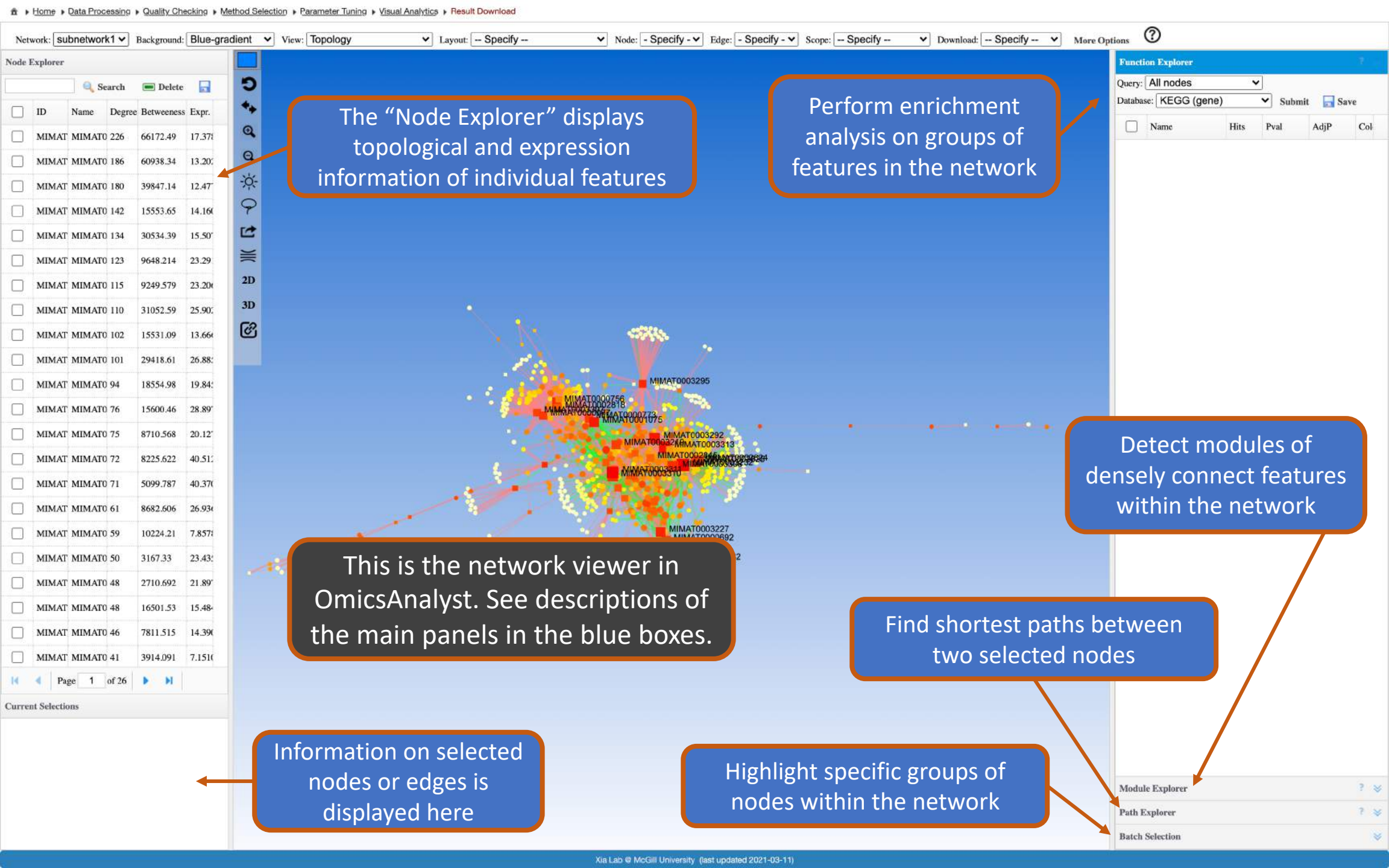
Feature selection method: This option selects key features to construct the correlation network based on the statistically signif... conducted on the previous page. To change the number of seed features, click the **Editor icon** of the corresponding data on the left panel.

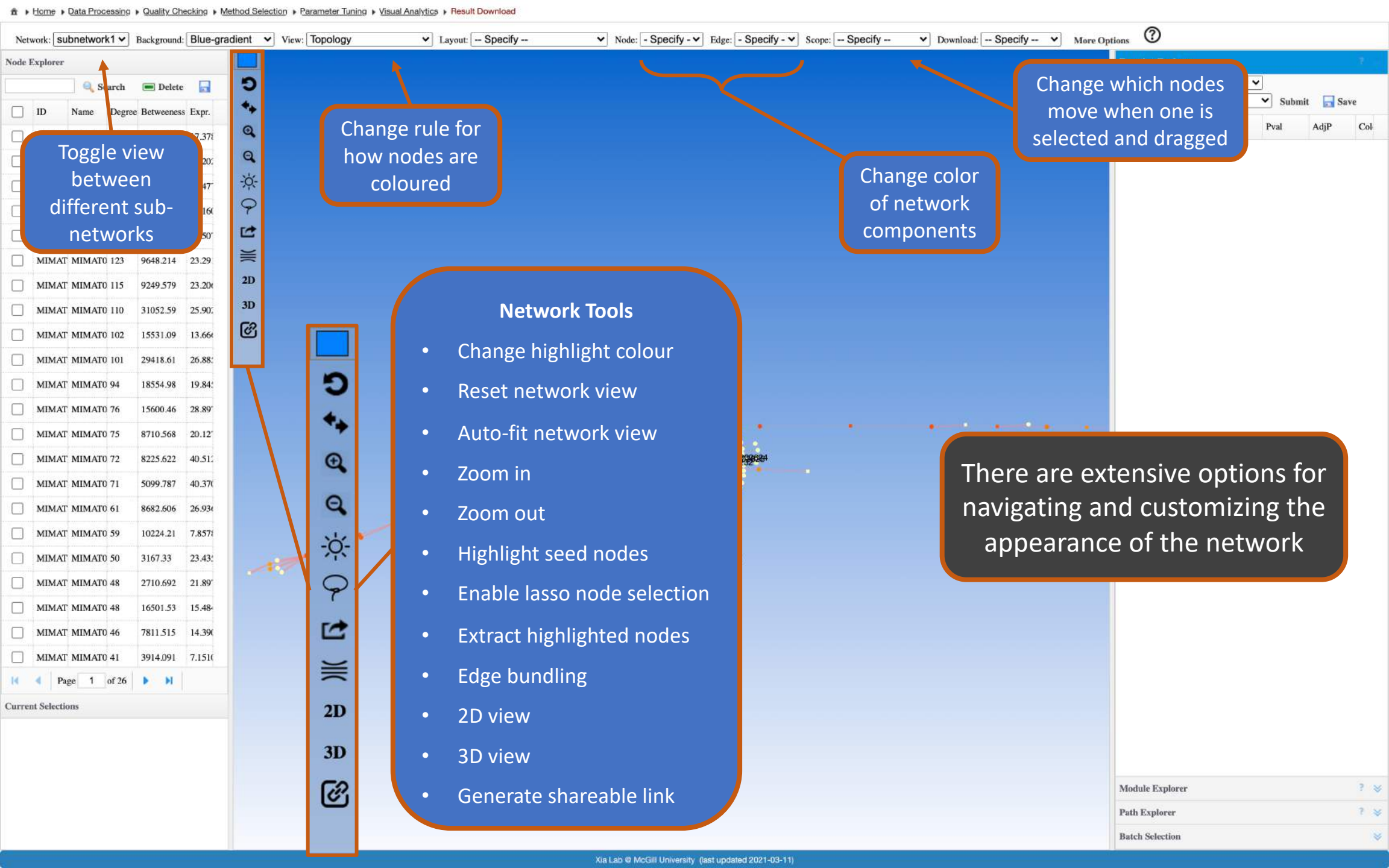
Similarity matrix method: classical Spearman correlation analysis is used to compute the pairwise similarity between selected features. Spearman correlation is a measure of the strength and direction of the monotonic relationship between two variables. It is rank-based, and therefore does not assume normality or linearity and can do a better job at representing non-linear relationships than the more commonly used Pearson correlation. The resulting pairwise Spearman correlations are used to form a network, where the nodes represent features and edges repre...

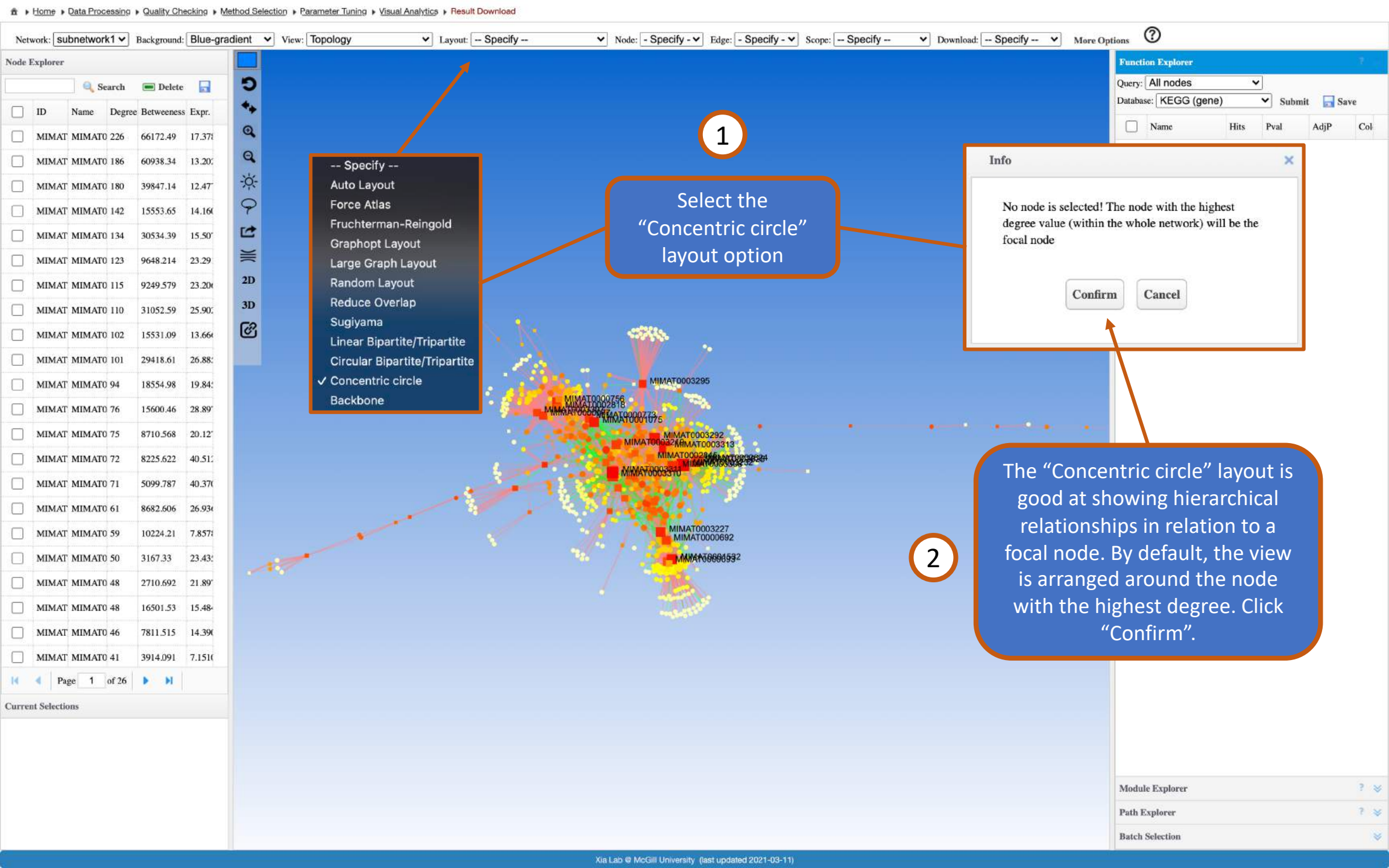
2

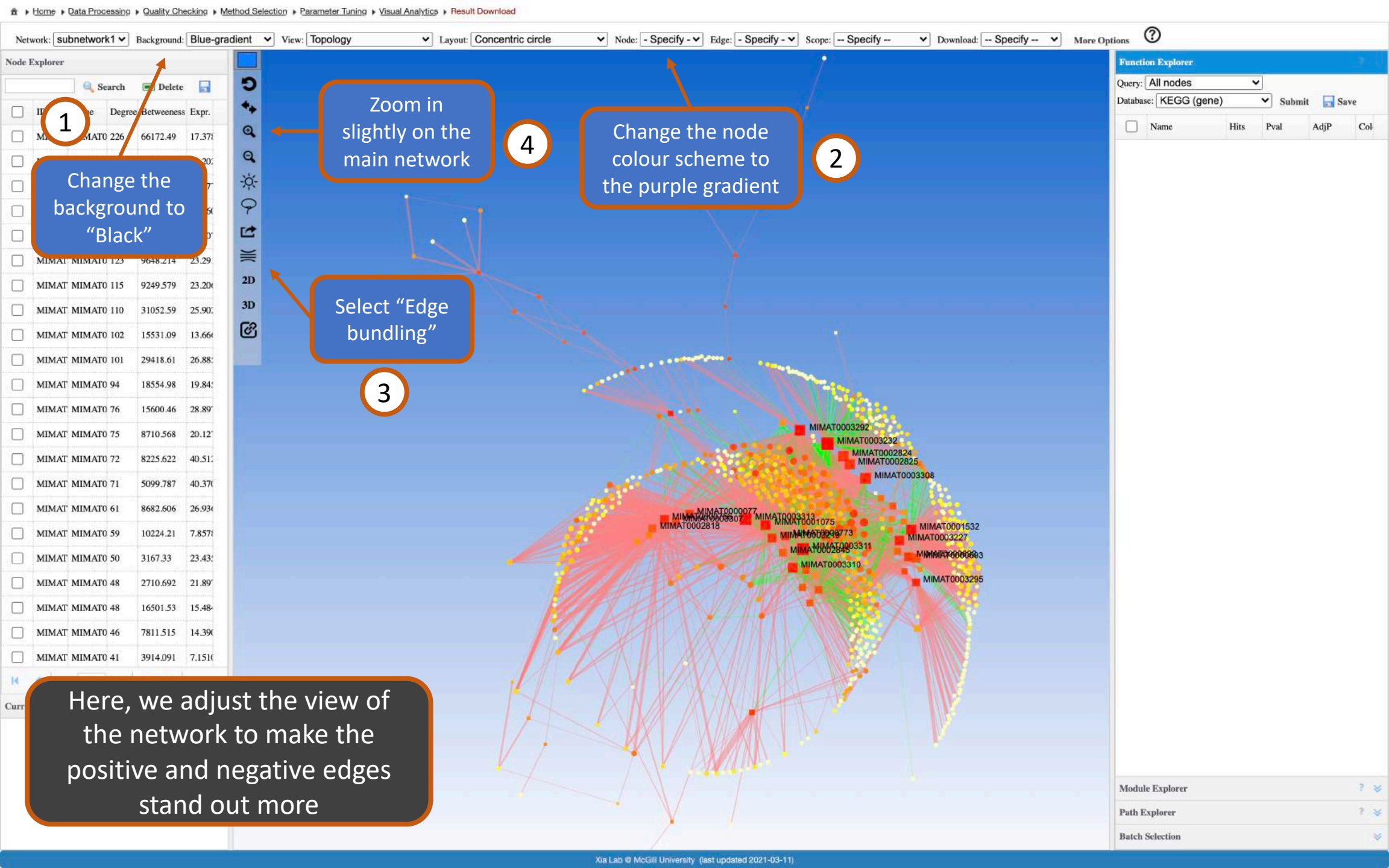
Select "Statistically significant features" and "Spearman"


Data Processing
Quality Checking
Method Selection
Parameter Tuning
Navigate to:
Network Tools
Degree Filter
Betweenness Filter
Reset Network
Use the filters above to reduce the network size
Analysis Overview & Parameter Tuning
The table below displays summary information on correlation networks generated using
sign". If network size is too large (i.e. over 1000 nodes) or too dense (i.e. over 2000
generally higher than between-omics, the default values are more stringent to avoid do
Between-omics only:
Corr. threshold (between-omics):
0.5
Corr. threshold (within-omics):
0.8
Correlation sign:
Both
Max. number of edges:
2000.0
Update
Correlation tends to be categorically greater between features of the same 'omics type than of different 'omics types. For this reason, there are separate thresholds for within and between 'omics pairings. Uncheck the box and set the within-omics threshold to 0.8
1
Click "Update"
2
Network summary
Correlation
Networks
Nodes
Edges
Interactions (.SIF)
subnetwork1
miRNA: 24; Genes: 178;
823
Download
4
3
2
2
1
1
1
1
1
1
Networks with too many nodes and edges are difficult to visualize in a meaningful way, and so this page allows you to "prune" the network using various filters. For example, some advanced layout features are disabled for networks greater than 2000 edges, so in some cases you should refine the network.
3
Click "Proceed"
Previous
Proceed


Home ▸ Data Processing ▸ Quality Checking ▸ Method Selection ▸ Parameter Tuning ▸ Visual Analytics ▸ Result Download
Network: subnetwork1 Background: Blue-gradient View: Topology Layout: -- Specify -- Node: - Specify - Edge: - Specify - Scope: -- Specify -- Download: -- Specify -- More Options
Node Explorer
Search Delete
ID Name Degree Betweeness Expr.
Page 1 of 26
Current Selections
2D
3D
Function Explorer
Query: All nodes
Database: KEGG (gene) Submit Save
Name Hits Pval AdjP Col
Module Explorer
Path Explorer
Batch Selection
The “Node Explorer” displays topological and expression information of individual features
Perform enrichment analysis on groups of features in the network
Detect modules of densely connect features within the network
This is the network viewer in OmicsAnalyst. See descriptions of the main panels in the blue boxes.
Find shortest paths between two selected nodes
Information on selected nodes or edges is displayed here
Highlight specific groups of nodes within the network
Xia Lab @ McGill University (last updated 2021-03-11)


Home ▸ Data Processing ▸ Quality Checking ▸ Method Selection ▸ Parameter Tuning ▸ Visual Analytics ▸ Result Download
Network: subnetwork1
Background: Blue-gradient
View: Topology
Layout: -- Specify --
Node: - Specify -
Edge: - Specify -
Scope: -- Specify --
Download: -- Specify --
More Options
Node Explorer
Search
Delete
ID
Name
Degree
Betweeness
Expr.
Page 1 of 26
Current Selections
Toggle view between different sub-networks
Change rule for how nodes are coloured
Change color of network components
Change which nodes move when one is selected and dragged
2D
3D
Network Tools
Change highlight colour
Reset network view
Auto-fit network view
Zoom in
Zoom out
Highlight seed nodes
Enable lasso node selection
Extract highlighted nodes
Edge bundling
2D view
3D view
Generate shareable link
There are extensive options for navigating and customizing the appearance of the network
Submit
Save
Pval
AdjP
Col
Module Explorer
Path Explorer
Batch Selection
Xia Lab @ McGill University (last updated 2021-03-11)


Home ▸ Data Processing ▸ Quality Checking ▸ Method Selection ▸ Parameter Tuning ▸ Visual Analytics ▸ Result Download
Network: subnetwork1 Background: Blue-gradient View: Topology Layout: -- Specify -- Node: - Specify - Edge: - Specify - Scope: -- Specify -- Download: -- Specify -- More Options
Node Explorer
Search Delete
ID Name Degree Betweeness Expr.
Page 1 of 26
Current Selections
-- Specify --
Auto Layout
Force Atlas
Fruchterman-Reingold
Graphopt Layout
Large Graph Layout
Random Layout
Reduce Overlap
Sugiyama
Linear Bipartite/Tripartite
Circular Bipartite/Tripartite
Concentric circle
Backbone
1
Select the "Concentric circle" layout option
Info
No node is selected! The node with the highest degree value (within the whole network) will be the focal node
Confirm
Cancel
2
The "Concentric circle" layout is good at showing hierarchical relationships in relation to a focal node. By default, the view is arranged around the node with the highest degree. Click "Confirm".
Function Explorer
Query: All nodes
Database: KEGG (gene)
Submit
Save
Name Hits Pval AdjP Col
Module Explorer
Path Explorer
Batch Selection
Xia Lab @ McGill University (last updated 2021-03-11)


Home ▸ Data Processing ▸ Quality Checking ▸ Method Selection ▸ Parameter Tuning ▸ Visual Analytics ▸ Result Download
Network: subnetwork1
Background: Blue-gradient
View: Topology
Layout: Concentric circle
Node: - Specify -
Edge: - Specify -
Scope: -- Specify --
Download: -- Specify --
More Options
Node Explorer
Search
Delete
Degree
Betweeness
Expr.
1
Change the background to "Black"
Zoom in slightly on the main network
4
Change the node colour scheme to the purple gradient
2
Select "Edge bundling"
3
2D
3D
Here, we adjust the view of the network to make the positive and negative edges stand out more
Function Explorer
Query: All nodes
Database: KEGG (gene)
Submit
Save
Name
Hits
Pval
AdjP
Col
Module Explorer
Path Explorer
Batch Selection
Xia Lab @ McGill University (last updated 2021-03-11)

Network: subnetwork1 | Background: Black | View: Topology | Layout: Concentric circle | Node: Color | Edge: - Specify - | Scope: -- Specify -- | Download: -- Specify -- | More Options

Notice the groups of genes that are positively and negatively correlated with key miRNAs

1

2

Click “Module Explorer”

Node Explorer

Search | Delete

ID	Name	Degree	Betweeness	Expr.
MIMAT	MIMAT0	226	66172.49	17.37
MIMAT	MIMAT0	186	60938.34	13.20
MIMAT	MIMAT0	180	39847.14	12.47
MIMAT	MIMAT0	142	15553.65	14.16
MIMAT	MIMAT0	134	30534.39	15.50
MIMAT	MIMAT0	123	9648.214	23.29
MIMAT	MIMAT0	115	9249.579	23.20
MIMAT	MIMAT0	110	31052.59	25.90
MIMAT	MIMAT0	102	15531.09	13.66
MIMAT	MIMAT0	101	29418.61	26.88
MIMAT	MIMAT0	94	18554.98	19.84
MIMAT	MIMAT0	76	15600.46	28.89
MIMAT	MIMAT0	75	8710.568	20.12
MIMAT	MIMAT0	72	8225.622	40.51
MIMAT	MIMAT0	71	5099.787	40.37
MIMAT	MIMAT0	61	8682.606	26.93
MIMAT	MIMAT0	59	10224.21	7.857
MIMAT	MIMAT0	50	3167.33	23.43
MIMAT	MIMAT0	48	2710.692	21.89
MIMAT	MIMAT0	48	16501.53	15.48
MIMAT	MIMAT0	46	7811.515	14.39
MIMAT	MIMAT0	41	3914.091	7.151

Page 1 of 26

Current Selections

Function Explorer

Query: All nodes

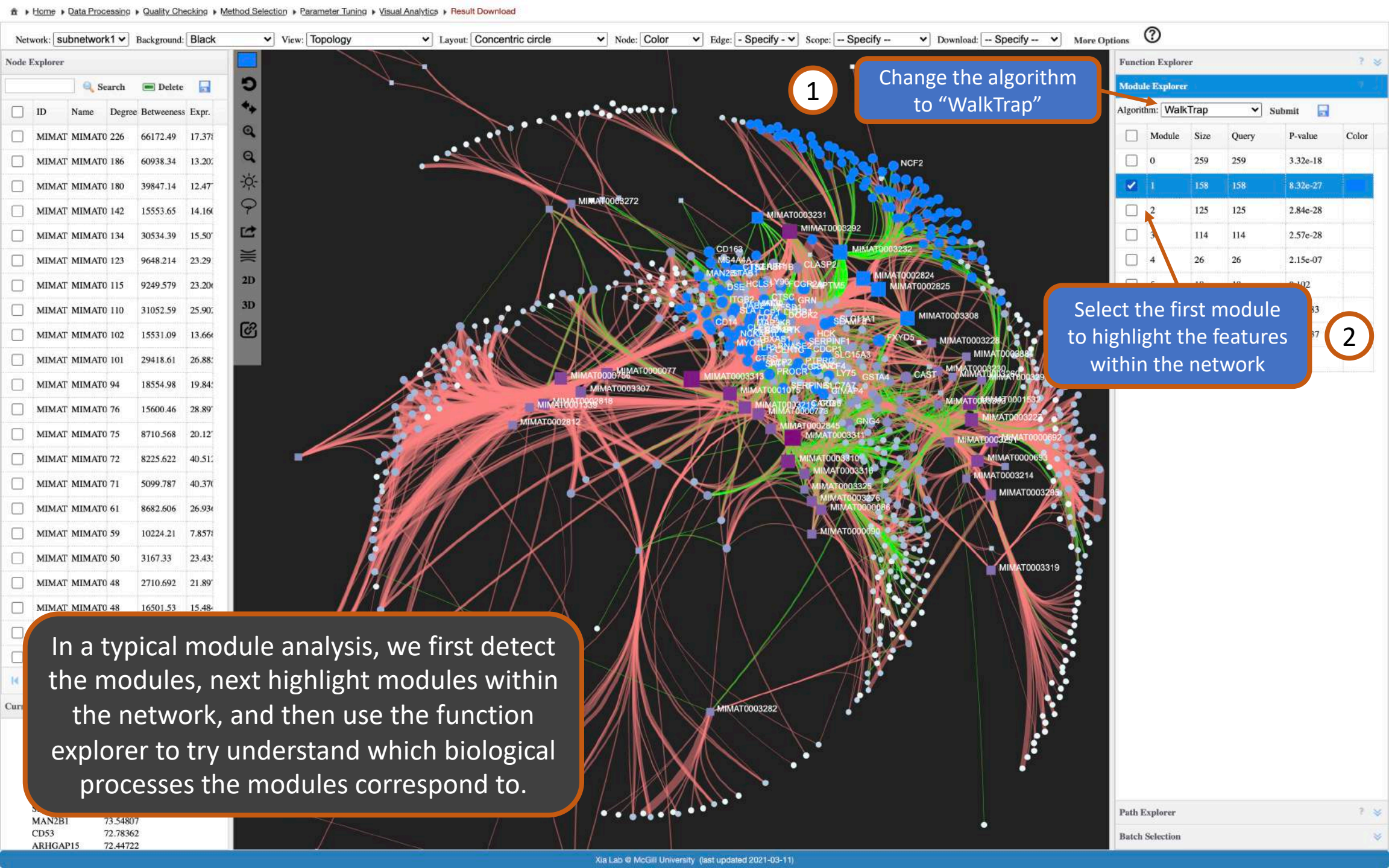
Database: KEGG (gene) | Submit | Save

Name	Hits	Pval	AdjP	Col

Module Explorer

Path Explorer

Batch Selection


Home ▸ Data Processing ▸ Quality Checking ▸ Method Selection ▸ Parameter Tuning ▸ Visual Analytics ▸ Result Download
Network: subnetwork1 Background: Black View: Topology Layout: Concentric circle Node: Color Edge: - Specify - Scope: -- Specify -- Download: -- Specify -- More Options
Node Explorer
Search Delete
ID Name Degree Betweeness Expr.
MIMAT MIMAT0 226 66172.49 17.37
MIMAT MIMAT0 186 60938.34 13.20
MIMAT MIMAT0 180 39847.14 12.47
MIMAT MIMAT0 142 15553.65 14.16
MIMAT MIMAT0 134 30534.39 15.50
MIMAT MIMAT0 123 9648.214 23.29
MIMAT MIMAT0 115 9249.579 23.20
MIMAT MIMAT0 110 31052.59 25.90
MIMAT MIMAT0 102 15531.09 13.66
MIMAT MIMAT0 101 29418.61 26.88
MIMAT MIMAT0 94 18554.98 19.84
MIMAT MIMAT0 76 15600.46 28.89
MIMAT MIMAT0 75 8710.568 20.12
MIMAT MIMAT0 72 8225.622 40.51
MIMAT MIMAT0 71 5099.787 40.37
MIMAT MIMAT0 61 8682.606 26.93
MIMAT MIMAT0 59 10224.21 7.857
MIMAT MIMAT0 50 3167.33 23.43
MIMAT MIMAT0 48 2710.692 21.89
MIMAT MIMAT0 48 16501.53 15.48
MAN2B1 73.54807
CD53 72.78362
ARHGAP15 72.44722
2D
3D
1
Change the algorithm to “WalkTrap”
Function Explorer
Module Explorer
Algorithm: WalkTrap Submit
Module Size Query P-value Color
0 259 259 3.32e-18
1 158 158 8.32e-27
2 125 125 2.84e-28
114 114 2.57e-28
4 26 26 2.15e-07
Select the first module to highlight the features within the network
2
Path Explorer
Batch Selection
In a typical module analysis, we first detect the modules, next highlight modules within the network, and then use the function explorer to try understand which biological processes the modules correspond to.
Xia Lab @ McGill University (last updated 2021-03-11)

Network: subnetwork1 | Background: Black | View: Topology | Layout: Concentric circle | Node: Color | Edge: - Specify - | Scope: -- Specify -- | Download: -- Specify -- | More Options

3

Click "Submit"

1

Go to the "Function Explorer"

Change to "Highlighted nodes"

2

Node Explorer

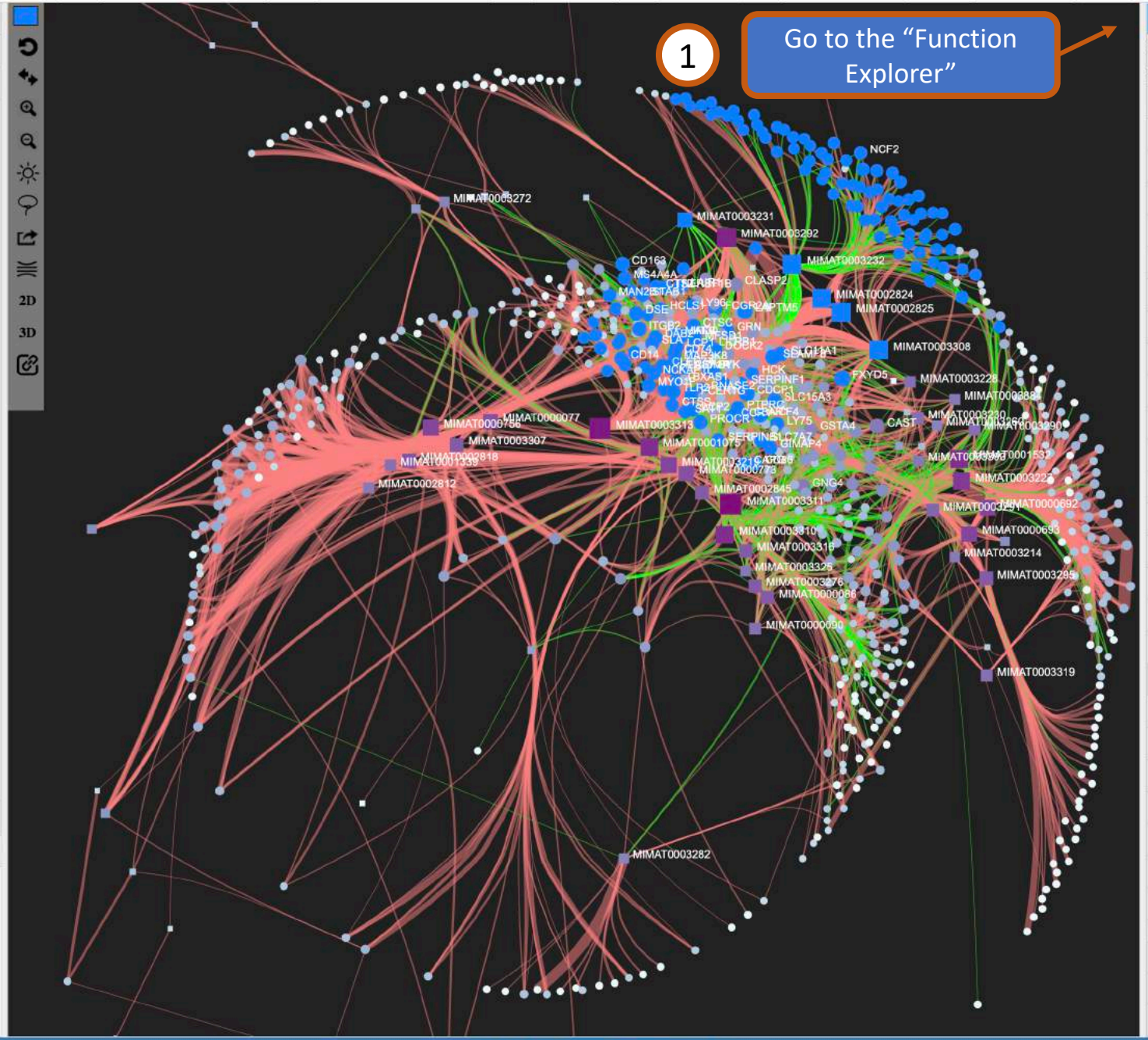
Search | Delete

ID	Name	Degree	Betweeness	Expr.
MIMAT	MIMAT0	226	66172.49	17.37
MIMAT	MIMAT0	186	60938.34	13.20
MIMAT	MIMAT0	180	39847.14	12.47
MIMAT	MIMAT0	142	15553.65	14.16
MIMAT	MIMAT0	134	30534.39	15.50
MIMAT	MIMAT0	123	9648.214	23.29
MIMAT	MIMAT0	115	9249.579	23.20
MIMAT	MIMAT0	110	31052.59	25.90
MIMAT	MIMAT0	102	15531.09	13.66
MIMAT	MIMAT0	101	29418.61	26.88
MIMAT	MIMAT0	94	18554.98	19.84
MIMAT	MIMAT0	76	15600.46	28.89
MIMAT	MIMAT0	75	8710.568	20.12
MIMAT	MIMAT0	72	8225.622	40.51
MIMAT	MIMAT0	71	5099.787	40.37
MIMAT	MIMAT0	61	8682.606	26.93
MIMAT	MIMAT0	59	10224.21	7.857
MIMAT	MIMAT0	50	3167.33	23.43
MIMAT	MIMAT0	48	2710.692	21.89
MIMAT	MIMAT0	48	16501.53	15.48
MIMAT	MIMAT0	46	7811.515	14.39
MIMAT	MIMAT0	41	3914.091	7.151

Page 1 of 26

Current Selections

module 1	
ZEB1	93.93268
SERPINF1	90.69636
SLCO2B1	86.06108
SELL	79.91001
SAMSN1	75.34214
SDC2	73.94697
MAN2B1	73.54807
CD53	72.78362
ARHGAP15	72.44722



Function Explorer

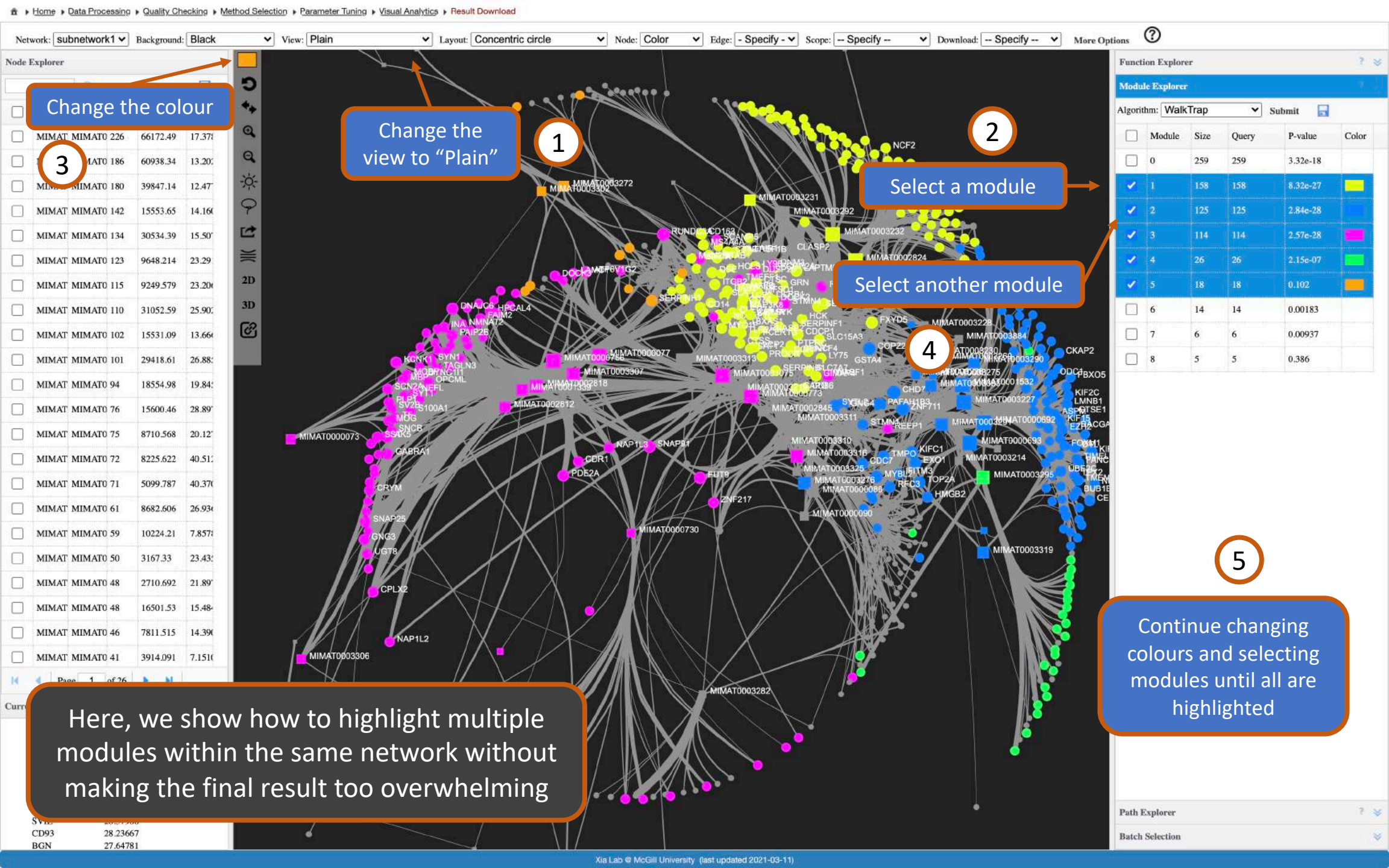
Query: Highlighted nodes | Database: KEGG (gene) | Submit | Save

Name	Hits	Pval	AdjP	Col
Apoptosis	18	4.28e-13	1.44e-10	
Influenza A	18	7.27e-12	9.37e-10	
Amoebiasis				
Human cytomeg				
Complement and				
Arrhythmogenic				
T cell receptor signa	10	3.98e-7		
Dilated cardiomyop	7	4.04e-7		
Cytosolic DNA-sens	9	5.45e-7	0.0000203	
Viral myocarditis	7	6.98e-7	0.0000235	
Proximal tubule bica	7	9.8e-7	0.0000274	
Allograft rejection	6	0.0000021	0.0000543	
Graft-versus-host di	7	0.00000422	0.000101	
Cell adhesion molec	10	0.0000047	0.000105	
NOD-like receptor s	8	0.00000651	0.000131	
Hypertrophic cardio	10	0.00000664	0.000131	
Primary immunodef	7	0.0000168	0.000314	
Synaptic vesicle cyc	6	0.0000331	0.000586	
African trypanosom	7	0.0000472	0.000792	
JAK-STAT signaling	8	0.0000495	0.000792	
Epstein-Barr virus in	10	0.0000565	0.000863	
Malaria	6	0.0000716	0.00105	
Hepatitis B	7	0.000538	0.00753	
RIG-I-like receptor	7	0.000987	0.0131	
Chemical carcinoge	9	0.00101	0.0131	
Fc gamma R-mediat	6	0.00109	0.0135	

Module Explorer

Path Explorer

Batch Selection


Home
Data Processing
Quality Checking
Method Selection
Parameter Tuning
Visual Analytics
Result Download
Network: subnetwork1
Background: Black
View: Plain
Layout: Concentric circle
Node: Color
Edge: - Specify -
Scope: -- Specify --
Download: -- Specify --
More Options
Node Explorer
Function Explorer
Module Explorer
Algorithm: WalkTrap
Submit
Module
Size
Query
P-value
Color
Change the colour
Change the view to "Plain"
Select a module
Select another module
Continue changing colours and selecting modules until all are highlighted
Here, we show how to highlight multiple modules within the same network without making the final result too overwhelming
Path Explorer
Batch Selection
Xia Lab @ McGill University (last updated 2021-03-11)

Network: subnetwork1 | Background: Black | View: Plain | Layout: Concentric circle | Node: Color | Edge: - Specify - | Scope: -- Specify -- | Download: -- Specify -- | More Options

Node Explorer

Search | Delete

ID	Name	Degree	Betweeness	Expr.
MIMAT	MIMAT0	226	66172.49	17.37
MIMAT	MIMAT0	186	60938.34	13.20
MIMAT	MIMAT0	180	39847.14	12.47
MIMAT	MIMAT0	142	15553.65	14.16
MIMAT	MIMAT0	134	30534.39	15.50
MIMAT	MIMAT0	123	9648.214	23.29
MIMAT	MIMAT0	115	9249.579	23.20
MIMAT	MIMAT0	110	31052.59	25.90
MIMAT	MIMAT0	102	15531.09	13.66
MIMAT	MIMAT0	101	29418.61	26.88
MIMAT	MIMAT0	94	18554.98	19.84
MIMAT	MIMAT0	76	15600.46	28.89
MIMAT	MIMAT0	75	8710.568	20.12
MIMAT	MIMAT0	72	8225.622	40.51
MIMAT	MIMAT0	71	5099.787	40.37
MIMAT	MIMAT0	61	8682.606	26.93
MIMAT	MIMAT0	59	10224.21	7.857
MIMAT	MIMAT0	50	3167.33	23.43
MIMAT	MIMAT0	48	2710.692	21.89
MIMAT	MIMAT0	48	16501.53	15.48
MIMAT	MIMAT0	46	7811.515	14.39
MIMAT	MIMAT0	41	3914.091	7.151

Page 1 of 26

Current Selections

- **Name**: MIMAT0003232

Select a node from the "Node Explorer" to zoom in on and highlight in the network

1

2D | 3D

Function Explorer

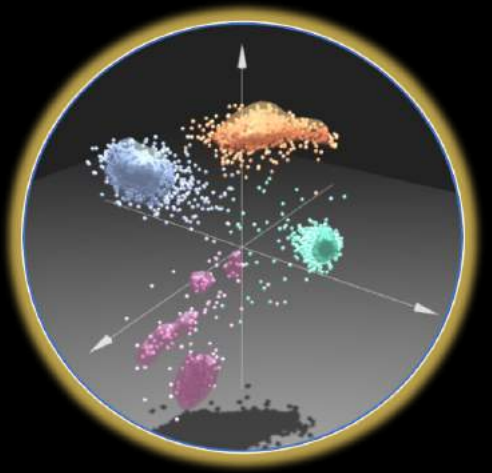
Module Explorer

Algorithm: WalkTrap | Submit

Module	Size	Query	P-value	Color
0	259	259	3.32e-18	
1	158	158	8.32e-27	
2	125	125	2.84e-28	
3	114	114	2.57e-28	
4	26	26	2.15e-07	
5	18	18	0.102	
6	14	14	0.00183	
7	6	6	0.00937	
8	5	5	0.386	

Path Explorer

Batch Selection



The End

For more information, visit the **FAQs, Tutorials, Resources**

and **Contact** pages on www.omicsanalyst.ca