



Availability and implementation: The Odyssey software is freely available for non-profit academic use at <https://diagnose.shinyapps.io/odyssey/>. The source code, example datasets are available at <https://github.com/alptaciroglu/Odyssey>. Contact: taciroglu.alperen@metu.edu.tr

ODYSSEY

**A TOOL FOR MICRORNA-MRNA
EXPRESSION AND INTERACTION
VISUALIZATION**



**Bilkent University KONU Lab
METU Informatics ACAR Lab**



Table of Contents

Summary	3
Home Page	6
Expression Data Choice	7
Data Upload	9
Data Visualization	14
Data Information	15
Query Selection	18
Interaction database	18
Include Secondary Interactions	19
ID Checking	19
Differential Expression Analysis Visualization	23
Network Options and Network File	25
Network Options	26
Network File	27
Network Statistics and Visualization	28
Network Statistics	29
Network Visualization	32
Prioritization	35
Index	39

Summary

"Odyssey" has been built for interactive visualization the interaction networks of miRNAs with along with their target expressions for a dataset either user uploaded or found in the Odyssey as default datasets (please see the list below) (Figure 1). It is built using Shiny package of the R programming language leading to seamless online access and modularity. The aim is to provide users a user-friendly web-application which consists of modules that allows:

- Uploading of their own data or reanalyzing existing public mRNA/microRNA datasets;
- Performing differential expression (DGEx) analysis;
- Visualization of the network of "Odyssey" builds from either experimentally validated or predicted interactions;
- Run a prioritization algorithm for positive or negative associations between miRNA/mRNA pairs by the node ID or the node degree, to selectively highlight nodes and uncover emergent biological signatures.
- Draw scatterplot and analyze association between Correlation Score% \sim logFC quantitative variables for each miRNA / mRNA node in network

Interaction information between microRNAs and genes can be derived from experimentally validated interactions using miRNet [1] or predictions on TargetScan [2] or both.

Odyssey enables the user to filter selected nodes to create refined networks through;

- Using fold change cut-offs obtained in DGEx;
- Using degree value of the nodes;
- Highlighting negative-associated miRNA- mRNA pairs by phasing out positively associated pairs or vice versa.

Furthermore, the application has been demonstrated using two different public miRNA-mRNA expression datasets in the Example Case section of Tutorials. Odyssey has integrated 13 different public miRNA-mRNA expression datasets while it also allows for using user uploaded miRNA and/or mRNA datasets given the expression data platform is one of the below-mentioned ones.

Odyssey is available at <https://diagnose.shinyapps.io/odyssey/>

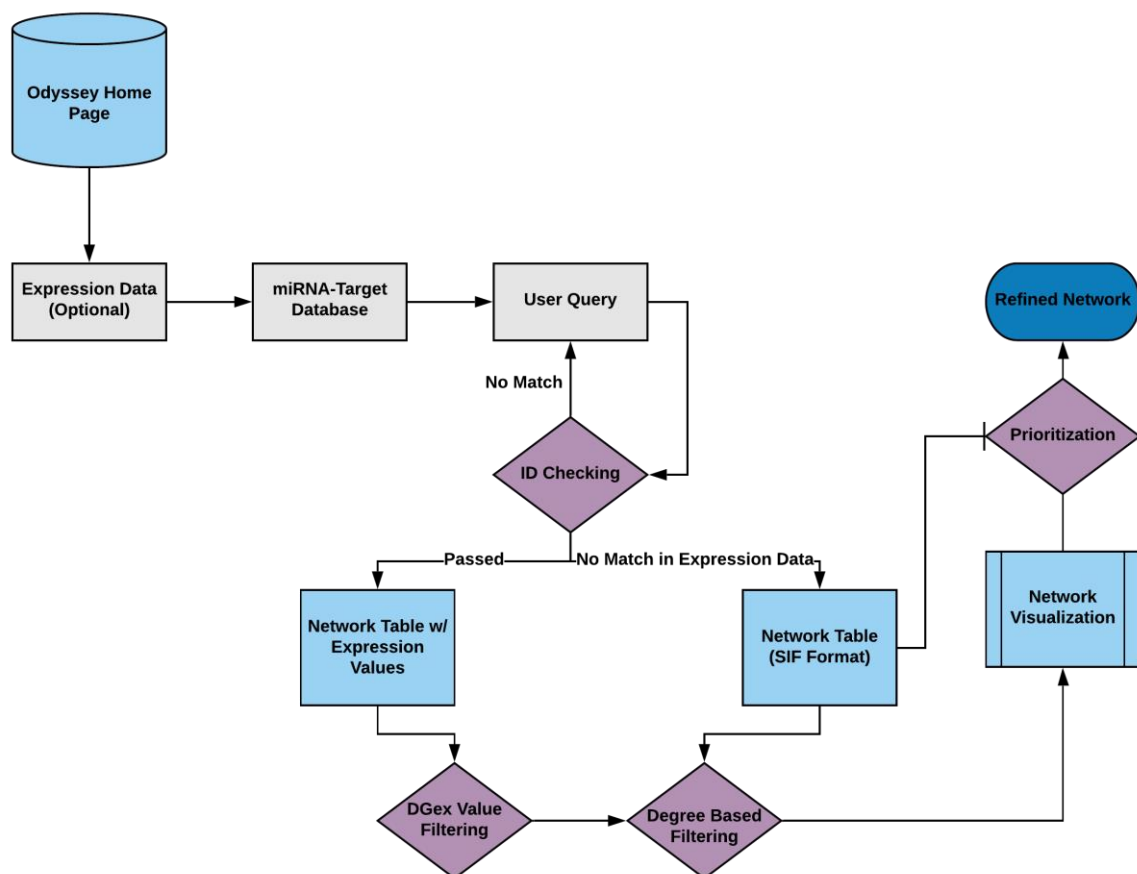


Figure 1. Flowchart of Odyssey v1.0.

Table 1: List of expression platforms supported in the current version:

mRNA expression data platforms	miRNA expression data platforms
GPL570 [HG-U133_Plus_2]	GPL8786 [miRNA-1]
GPL96 [HG-U133A]	GPL14613 [miRNA-2]
GPL97 [HG-U133B]	GPL16384 [miRNA-3]
GPL6244 [HuGene-1_0-st]	GPL21572 [miRNA-4]
GPL17692 [HuGene-2_1-st]	

Table 2: List of default expression datasets:

Dataset ID GEO	GEO URL	PubMed URL	Control Samples	Treatment Samples
GSE25402	GSE25402	PMID:22688341	Non-obese	Obese
GSE32539_1	GSE32539	PMID:23783374	Control	Idiopathic pulmonary fibrosis (IPF) / Usual interstitial pneumonia
GSE32539_2	GSE32539	PMID:23783374	Control	Nonspecific Interstitial Pneumonia
GSE32539_3	GSE32539	PMID:23783374	Control	Respiratory bronchiolitis-interstitial lung disease
GSE34681_1	GSE34681	PMID:22244333	Control	Ars2 siRNA-1
GSE34681_2	GSE34681	PMID:22244333	Control	DGCR8 siRNA-1
GSE38617	GSE38617	PMID:23723971	Control	oral lichen planus (OLP)
GSE39061_1	GSE39061	PMID:23590309	Confluent	Day 28 (differentiated)
GSE39061_2	GSE39061	PMID:23590309	Subconfluent	Confluent
GSE39061_3	GSE39061	PMID:23590309	Subconfluent	Day 28 (differentiated)
GSE40321	GSE40321	No PubMed URL	46 XY	47 XY, +8
GSE49697_1	GSE49697	PMID:25645730	US_48h (unstimulated) (poolBC removed)	S_48h (stimulated) (poolBC removed)
GSE49697_2	GSE49697	PMID:25645730	US_24h (unstimulated) (poolBC removed)	S_24h (stimulated) (poolBC removed)
GSE59702_1	GSE59702	PMID:25587024	Match normal of fusion negative tumor	Fusion negative tumor
GSE59702_2	GSE59702	PMID:25587024	Match normal of fusion positive tumor	Fusion positive tumor
GSE104268_1	GSE104268	No PubMed URL	Control	GSE
GSE104268_2	GSE104268	No PubMed URL	Control	TSA
GSE81867	GSE81867	No PubMed URL	Transwell Static	Chip
GSE90604	GSE90604	No PubMed URL	All Healthy Tissues	Glioblastoma
GSE35389_1	GSE35389	PMID:23056502	Normal melanocyte	Melanoma
GSE35389_2	GSE35389	PMID:23056502	Melanoma	Melanoma exosome
GSE88721	GSE88721	PMID:28327132	Meningial Cells	Meningioma

Home Page

The screenshot shows the Odyssey web application interface. At the top left is the Odyssey logo, a circular emblem with a network diagram. To its right is the word "Odyssey" in a large, bold, black font. Below the logo is a "Data Selection" button. To the right of the logo is a "Tutorial" link with a download icon. The main content area contains three red-bordered boxes. The first box, titled "Expression Data Choice", contains a drop-down menu with "Example Data" selected. The second box, titled "Which example data you want to use ?", contains a drop-down menu with "GSE38617" selected. The third box contains a table of phenodata for the selected dataset. A "Proceed" button is at the bottom. Three blue arrows point from text boxes on the right to the drop-down menus and the phenodata table.

Odyssey

[Data Selection](#) [Tutorial](#)

Expression Data Choice

Example Data

Which example data you want to use ?

GSE38617

Selection:	GSE38617
Description:	Disease-associated miRNA-mRNA networks in oral lichen planus
Sample Comparison:	Control vs. oral lichen planus (OLP)
GEO URL	GSE38617
PubMed URL	GSE38617

Proceed

Choice of the Expression Dataset is entered through this drop-down menu. Choices are "Example Data", "Upload Data", or "No Data".

Drop-down menu for selection from the Example Datasets. Selections are made from 13 different public miRNA-mRNA expression datasets

Phenodata of the Example Data choice is displayed here. Not available for user uploaded expression data.

Expression Data Choice

Expression Data Choice

Example Data

Which example data you want to use ?

GSE38617

GSE35389_1

GSE35389_2

GSE88721

GSE39061_1

GSE39061_2

GSE39061_3

GSE49697_1

GSE49697_2

Comparison.

GEO URL

GSE38617

PubMed URL

GSE38617

Odyssey currently utilizes 13 different public miRNA-mRNA expression datasets and phenodata of manually curated 22 total comparisons based on these datasets, which are indicated as GSExxx_1, GSExxx_2 etc. The comparisons are made by selecting different control and/or treatment samples to perform differential expression analysis with.

The screenshot shows a web form titled "Expression Data Choice". At the top is a dropdown menu with "Upload Data" selected. Below this is a note: "Uploading an expression file is optional". The form is divided into two sections. The first section is titled "Upload your miRNA expression file" and contains a "Browse..." button and a text box that says "No file selected". The second section is titled "Upload your mRNA expression file" and also contains a "Browse..." button and a text box that says "No file selected". At the bottom of the form is a "Reset!" button.

Selecting “*Upload Data*” from the “*Expression Dataset Choice*” drop-down menu creates two “fileInput” areas to upload the miRNA expression file and mRNA expression file respectively. Expected upload format is limited to “**Series Matrix File(s)**” in GEO database [3]. Odyssey checks whether the uploaded format is correct, and whether the uploaded expression data is log transformed by base 2. If the file is not log transformed, Odyssey will perform this automatically.


Users also have the option to upload either one of the expression files and leave the other “fileInput” area blank, e.g. miRNA data uploaded only with no mRNA data.

Data Upload

Users can upload their own expression data to analyze for any biological interest using Odyssey. There are two main limitations for the data upload; namely:

- Data format limitation
- Data platform limitation

Odyssey accepts expression data from platforms listed in the Home Page section of this tutorial document, in Series Matrix File(s) format downloaded and extracted from Gene Expression Omnibus database.


Download family		Format
SOFT formatted family file(s)		SOFT ?
MINiML formatted family file(s)		MINiML ?
 Series Matrix File(s)		TXT ?

Supplementary file	Size	Download	File type/resource
GSE32539_RAW.tar	831.2 Mb	(http)(custom)	TAR (of CEL)

This screenshot is taken from <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE32539>.

Users need to click on “Series Matrix File(s)” link to download the expression data.

Index of /geo/series/GSE32nnn/GSE32539/matrix/

 [\[parent directory\]](#)

Name	Size	Date Modified
 GSE32539-GPL6244_series_matrix.txt.gz	8.2 MB	10/26/19, 12:02:00 PM
 GSE32539-GPL8786_series_matrix.txt.gz	1.2 MB	10/26/19, 12:02:00 PM

When clicked, a directed webpage will be loaded. For the example dataset of GSE32539, there are two different series matrix files. GPL6244 [HuGene-1_0-st] is a mRNA expression array and GPL8786 [miRNA-1] is a miRNA expression array. Clicking on any of the links will start downloading the corresponding expression data file in compressed format (.gz). Users may use any of the file archiver / compressor supported by their Operating System to extract the uncompressed file.

Odyssey

[Data Selection](#) [Tutorial](#)

Expression Data Choice

Upload Data

Uploading an expression file is optional

Upload your miRNA expression file

Browse... GSE32539-GPL8786_series_matrix.txt

Upload complete

Upload your mRNA expression file

Browse... GSE32539-GPL6244_series_matrix.txt

Upload complete

Reset!

Proceed

After successfully extracting the Series Matrix Files, the obtained files should be in “.txt” format. In order to analysis to commence, “*Upload Data*” should be selected from Expression Data Choice drop-down menu and files should be uploaded to their respective fileInput zones. “*Upload complete*” sign will mark the successfully uploaded data, however successfully uploaded data can still be in wrong format and may not be accepted to proceed with the analysis. If there is need to remove uploaded data, “*Reset*” button may be clicked to do so.

After data is uploaded, to continue with the analysis “*Proceed*” is clicked. If at least one of the miRNA or mRNA fileInput zones are uploaded with data, a modal window will appear indicating the commencing of the analysis.

Your data is being processed... This process usually takes < 1 min, depending on the size of the data!

OK



Odyssey

Data Selection

[Tutorial](#)

Expression Data Choice

Upload Data

Uploading an expression file is optional

Upload your miRNA expression file

Browse... GSE32539-GPL8786_series_matrix.txt

Upload complete

Upload your mRNA expression file

Browse... GSE32539-GPL6244_series_matrix.txt

Upload complete

Reset!

Proceed

Data Information

[Tutorial](#)

miRNA Data Information

Number of samples in data: 14

Number of selected treatment samples: 7

Number of selected control samples: 7

mRNA Data Information

Number of samples in data: 14

Number of selected treatment samples: 7

Number of selected control samples: 7

miRNA Sample Selections

These control samples are selected:

- GSM946236, GSM946237, GSM946238, GSM946239, GSM946240, GSM946241, GSM946242

These treatment samples are selected:

- GSM946243, GSM946244, GSM946245, GSM946246, GSM946247, GSM946248, GSM946249

mRNA Sample Selections

These control samples are selected:

- GSM946253, GSM946254, GSM946255, GSM946256, GSM946257, GSM946258, GSM946259

These treatment samples are selected:

- GSM946260, GSM946261, GSM946262, GSM946263, GSM946264, GSM946265, GSM946266

Data Visualization

Expression Data

Your data is being processed... This process usually takes < 1 min, depending on the size of the data!

OK

Expression-miRNA

Show 2 entries

ID	GSM946236	GSM946237	GSM946238	GSM946239	GSM946240	GSM946241	GSM946242	GSM946243	GSM946244	GSM946245	GSM946246	GSM946247	GSM946248	GSM946249
IRS1														
LET	4.026	3.905	3.877	4.102	3.974	4.09	3.968	4.096	3.992	3.626	3.78	3.847	3.905	3.887
TA-SP														
IRS1														
LET	9.932	10.312	8.778	9.014	7.303	9.733	9.329	10.498	10.301	9.508	10.815	10.211	9.837	9.777
TA-SP														

Showing 1 to 2 of 437 entries

Previous 1 2 3 4 5 ... 219 Next

miRNA Controls miRNA Treatments

Expression-mRNA

Show 2 entries

ID	GSM946253	GSM946254	GSM946255	GSM946256	GSM946257	GSM946258	GSM946259	GSM946260	GSM946261	GSM946262	GSM946263	GSM946264	GSM946265	GSM946266
ATCF	3.897	3.877	4.132	3.969	3.697	3.84	4.252	3.606	4	3.648	3.943	3.631	3.785	3.913
A2M	9.389	10.827	9.854	9.96	9.652	8.957	8.73	10.794	10.426	9.48	8.881	9.999	10.78	8.235

Showing 1 to 2 of 20,001 entries

Previous 1 2 3 4 5 ... 10001 Next

Gene Controls Gene Treatments

Back Forward

If the uploaded data is in acceptable format, data will be displayed in the “*Main Panel*” as described in the Data Visualization section of the tutorial. However, uploading a data will prompt “Control” and “Treatment” sample selection for both miRNA and mRNA data. Using any integrated Example Data does not require Control and Treatment sample selection as these selections are already made for each dataset implemented.

The screenshot displays the Odyssey software interface. At the top, there are tabs for "Data Visualization" and "Expression Data". The "Expression Data" tab is active, showing two sections: "Expression-miRNA" and "Expression-mRNA".

Expression-miRNA Section:

- Search bar: []
- Table with columns: ID, GSM806421, GSM806422, GSM806423, GSM806424, GSM806425, GSM806426, GSM806427, GSM806428, GSM806429, GSM806430, GSM806431, GSM806432, GSM806433, GSM806434.
- Rows: HSA-LET-7A-3P, HSA-LET-7A-5P.
- Buttons: "miRNA Controls" and "miRNA Treatments".
- Navigation: Previous, 1, 2, 3, 4, 5, ..., 219, Next.

Expression-mRNA Section:

- Search bar: []
- Table with columns: ID, GSM806204, GSM806205, GSM806206, GSM806207, GSM806208, GSM806209, GSM806210, GSM806211, GSM806212, GSM806213, GSM806214, GSM806215, GSM806216, GSM806217.
- Rows: AADACL2, AADAT.
- Buttons: "Gene Controls" and "Gene Treatments".
- Navigation: Previous, 1, 2, 3, 4, 5, ..., 4532, Next.

At the bottom, there are "Back" and "Forward" buttons.

Selection of the samples are handled by column-wise clicking on the samples on the data table. Odyssey allows multiple sample selections which means each clicked sample will be highlighted with blue color, shown in screenshot above. “*miRNA Controls*” and “*miRNA Treatments*” buttons mark the sample selections and the selected samples are displayed in an information box placed in the Side Panel.

miRNA Sample Selections

These control samples are selected:

- GSM806424, GSM806425, GSM806426, GSM806427

These treatment samples are selected:

- GSM806428

Expression-mRNA

Show entries

Search:

ID	GSM806204	GSM806205	GSM806206	GSM806207	GSM806208	GSM806209	GSM806210	GSM806211	GSM806212	GSM806213	GSM806214	GSM806215	GSM806216	GSM806217
AADACL2	3.153	2.874	2.879	3.448	3.313	2.777	4.094	2.897	3.128	3.328	2.742	3.455	5.938	2.90
AADAT	4.844	5.49	4.741	4.411	4.469	5.423	4.709	4.751	4.652	4.955	4.855	4.625	4.96	5.0

Showing 1 to 2 of 9,064 entries

Previous **1** 2 3 4 5 ... 4532 Next

Gene Controls Gene Treatments

Back Forward

Control and Treatment sample selection for mRNA data is displayed in the screenshot above. Clicking “*Forward*” button takes the analysis on to next step and clicking “*Back*” button at this step sends the application to Home Page.

Failing to select either Control or Treatment samples from either one of the miRNA or mRNA data prompts a modal window, preventing the analysis to go further without completing necessary selections.

Data upload requires Control and Treatment sample selection!

OK

Data Visualization

! Skip to next section if “No Data” option used

Upon making sure miRNA/mRNA ID is found without any irrecoverable errors, Odyssey loads and displays the expression data on the “*Data Visualization Tab*” inside “*Main Panel*”. Expression data are displayed in a nested manner using “*Expression Data Tab*”. Phenodata associated with selected dataset is also displayed fully in “*Phenodata Tab*” as stored in GEO database.

Phenodata Tab is specific to the integrated Example Data Selections in Odyssey and will not be visible on “Data Upload” option.

The screenshot displays the Odyssey web application interface. On the left is the 'Side Panel' with 'miRNA Data Information', 'mRNA Data Information', and 'miRNA Sample Selections'. The 'Main Panel' shows the 'Data Visualization' tab with 'Expression-miRNA' and 'Expression-mRNA' data tables. Each table has a search bar and a 'Show' dropdown set to 2 entries. The 'Expression-miRNA' table shows data for HSA-LET-7A-3P across various GSM IDs. The 'Expression-mRNA' table shows data for A1CF and A2M across various GSM IDs. At the bottom of the main panel are 'Back' and 'Forward' buttons.

ID	GSM946236	GSM946237	GSM946238	GSM946239	GSM946240	GSM946241	GSM946242	GSM946243	GSM946244	GSM946245	GSM946246	GSM946247	GSM946248	GSM946249
HSA-LET-7A-3P	4.026	3.905	3.877	4.102	3.974	4.09	3.968	4.098	3.992	3.626	3.78	3.847	3.903	3.887

ID	GSM946259	GSM946264	GSM946265	GSM946266	GSM946267	GSM946268	GSM946269	GSM946270	GSM946271	GSM946272	GSM946273	GSM946274	GSM946275	GSM946276
A1CF	3.897	3.677	4.132	3.969	3.697	3.84	4.252	3.606	4	3.648	3.943	3.631	3.785	3.911
A2M	9.169	10.627	9.854	9.96	9.652	8.957	8.73	10.794	10.426	9.48	8.881	9.999	10.78	8.255

Side Panel

Main Panel

“Back” and “Forward” buttons, which are displayed at the bottom of the main panel, need to be used to take the analysis one step backwards or one step forward, respectively.

Data Information



Odyssey

* [Data Information](#)

[Tutorial](#)

miRNA Data Information

Number of samples in data:	14
Number of selected treatment samples:	7
Number of selected control samples:	7

mRNA Data Information

Number of samples in data:	14
Number of selected treatment samples:	7
Number of selected control samples:	7

miRNA Sample Selections

These control samples are selected:

- GSM946236, GSM946237, GSM946238, GSM946239, GSM946240, GSM946241, GSM946242

These treatment samples are selected:

- GSM946243, GSM946244, GSM946245, GSM946246, GSM946247, GSM946248, GSM946249

mRNA Sample Selections

These control samples are selected:

- GSM946253, GSM946254, GSM946255, GSM946256, GSM946257, GSM946258, GSM946259

These treatment samples are selected:

- GSM946260, GSM946261, GSM946262, GSM946263, GSM946264, GSM946265, GSM946266

Information boxes display the number of samples in miRNA & mRNA expression data and how many samples are selected as treatment and control samples for the proceeding differential gene expression analysis.

Information boxes displays the names of the samples in miRNA & mRNA expression data that are selected as treatment and control samples for the proceeding differential gene expression analysis.

A close view of the *Data Visualization Tab* with *Expression Data Tab* selected (truncated) is as follows.

The screenshot displays the 'Data Visualization' tab interface. At the top, there are three tabs: 'Data Visualization' (selected), 'Expression Data', and 'PhenoData'. Below these, the 'Expression-miRNA' section is visible, showing a table of miRNA expression data. The table has columns for ID and GSM946236 through GSM946240. The first two rows are truncated. The first row shows HSA-LET-7A-3P with values 4.026, 3.905, 3.877, 4.102, and 3.974. The second row shows HSA-LET-7A-5P with values 9.932, 10.312, and 7.303. Below the table, it says 'Showing 1 to 2 of 437 entries'. The 'Expression-mRNA' section is also visible, showing a table of mRNA expression data. The table has columns for ID and GSM946253 through GSM946257. The first two rows are truncated. The first row shows A1CF with values 3.897, 3.677, 4.132, 3.969, and 3.697. The second row shows A2M with values 9.169, 10.627, 9.854, 9.96, and 9.652. Below the table, it says 'Showing 1 to 2 of 20,001 entries'.

Data Visualization Tab

PhenoData Tab

Expression Data Tab

Expression Data of miRNAs

Expression Data of mRNAs

ID	GSM946236	GSM946237	GSM946238	GSM946239	GSM946240
HSA-LET-7A-3P	4.026	3.905	3.877	4.102	3.974
HSA-LET-7A-5P	9.932	10.312			7.303

Showing 1 to 2 of 437 entries

ID	GSM946253	GSM946254	GSM946255	GSM946256	GSM946257
A1CF	3.897	3.677	4.132	3.969	3.697
A2M	9.169	10.627	9.854	9.96	9.652

Showing 1 to 2 of 20,001 entries

A close view of the *Data Visualization Tab* with *PhenoData Tab* selected (truncated) is as follows.

Data Visualization

Expression Data

PhenoData

phenoData-miRNA

Show 2 entries

ID	title	geo_accession	status	submission_date	last_update_date	type
GSM946236	K1_miRNA	GSM946236	Public on May 06 201...	Jun 08 2012	May 06 2013	RNA
GSM946237	K3_miRNA	GSM946237	Public on May 06 201...			RNA

Phenodata of miRNA samples

Showing 1 to 2 of 14 entries

Phenodata of miRNA samples

phenoData-mRNA

Show entries

ID	title	geo_accession	status	submission_date	last_update_date	type
GSM946253	K1_mRNA	GSM946253	Public on May 06 201...			RNA
GSM946254	K3_mRNA	GSM946254	Public on May 06 201...	Jun 08 2012	May 06 2013	RNA

Showing 1 to 2 of 14 entries

Phenodata of mRNA samples

Query Selection

The screenshot displays the Odyssey web interface. On the left, the 'Query Selection' panel includes a 'Gene or Micro-rna ID' input field, a 'Querying for' dropdown with 'miRNA' and 'Gene' options, and an 'Interaction database' dropdown with 'TargetScan' selected. A checkbox for 'Include Secondary Interactions' is also present. The main area shows 'Differential Expression Results - miRNA' with a table of 5 entries. The table columns are ID, logFC, AveExpr, t, P.Value, adj.P.Val, B, TargetScan, and miRNet. The entries are HSA-MIR-1238-3P, HSA-MIR-4721, HSA-MIR-6716-3P, HSA-MIR-4783-3P, and HSA-MIR-3780. Below the table is a 'Download' button. A second, identical panel is shown below the first, displaying results for a different set of miRNAs (TSPAN13, CA14, SERPINE1, CYP11A1, CTGF).

ID	logFC	AveExpr	t	P.Value	adj.P.Val	B	TargetScan	miRNet
HSA-MIR-1238-3P	1.12	1.304	6.419	0.001	0.998	-4.21	No	Yes
HSA-MIR-4721	0.963	4.409	4.869	0.004	0.998	-4.25	No	Yes
HSA-MIR-6716-3P	1.515	1.211	4.85	0.004	0.998	-4.25	No	Yes
HSA-MIR-4783-3P	1.472	2.334	4.817	0.004	0.998	-4.252	No	Yes
HSA-MIR-3780	1.146	1.569	4.635	0.004	0.998	-4.258	No	Yes

ID	logFC	AveExpr	t	P.Value	adj.P.Val	B	TargetScan	miRNet
TSPAN13	4.791	6.384	38.217	0	0	13.713	No	Yes
CA14	-3.386	6.832	-27.054	0	0	11.86	No	Yes
SERPINE1	4.733	5.137	25.485	0	0	11.489	Yes	Yes
CYP11A1	3.29	4.91	24.944	0	0	11.352	No	Yes
CTGF	3.164	4.836	24.24	0	0	11.167	Yes	Yes

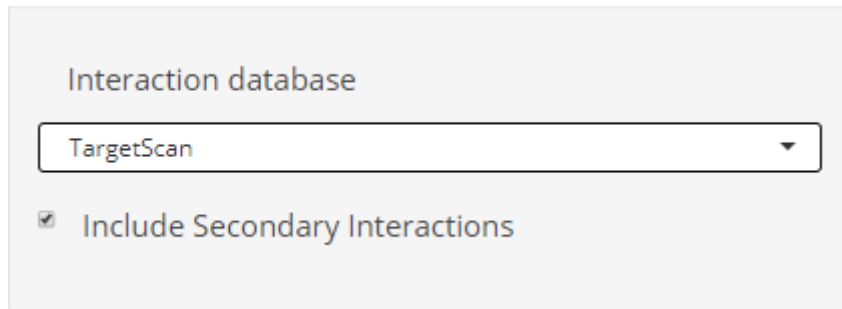
At this step, users are expected to input a gene / microRNA query. Also, differential expression gene analysis result on the previously selected expression data set (if selected) is displayed here.

Interaction database

The screenshot shows a dropdown menu for the 'Interaction database'. The menu is open, displaying three options: 'TargetScan', 'miRNet', and 'TargetScan&miRNet'. A tooltip above the dropdown states: 'TargetScan for predicted, miRNet for experimentally validated interactions'.

Selection of “*TargetScan&miRNet*” will reveal another drop-down menu for selection of “*Intersect*” or “*Union*” of interaction databases. Consequently, union of these two interaction databases is likely to harbor more miRNA-mRNA interactions.

Include Secondary Interactions



Interaction database

TargetScan ▼

☒ Include Secondary Interactions

Selection of “*Include Secondary Interactions*” expands the initial network by including second degree interactions on first degree interactions of the query.

ID Checking

Odyssey requires an ID for an mRNA or microRNA entered in the query box where it says enter gene/miRNA ID. If a query does not result in a hit in the Odyssey session the error handling module tries to handle these situations, depending on type of generated error.

There may not be a hit found in the Odyssey database due to using different identifiers than the allowed “Official Gene Name”, “miRBase v22” IDs. In other cases, no record of the query can be found in TargetScan and/or miRNet databases or in the platform of the expression dataset used. These types of queries generate an irrecoverable error and Odyssey directs user to select a different query.

The image shows the Odyssey web interface. At the top left is the Odyssey logo, a circular network diagram. Next to it is the word "Odyssey" in a large, bold, dark blue font. Below the logo is a button with a question mark icon and the text "Query Selection". To the right of the logo is a link with a download icon and the text "Tutorial".

The main form has two sections. The first section is for inputting a query. It has a label "Gene or Micro-rna ID:" followed by a text input field containing "Baby Yoda". To the right of this is a label "Querying for:" followed by two buttons: "miRNA" (with a checked checkbox) and "Gene" (with an unchecked checkbox).

The second section is for selecting an interaction database. It has a label "Interaction database" followed by a dropdown menu showing "TargetScan". Below this is a checkbox labeled "Include Secondary Interactions" which is checked.

A blue arrow points from the bottom of the form to a red error message box. The error message is in red text and says: "Query is not present in selected Interaction Database! Please try a different query...". At the bottom right of the error box is an "OK" button.

A query may be in the right format in terms of the class of identifiers used and they also have records of interactions in TargetScan and/or miRNet databases, yet the selected expression dataset might not have all of the counterparts of these identifiers of which Odyssey needs to color the network. These types of queries generate a recoverable error; and Odyssey warns the user if proceeded not all data are available.

The image shows the Odyssey web interface. At the top left is the Odyssey logo, a circular emblem with a network diagram. Next to it is the word "Odyssey" in a large, bold, dark blue font. To the right of the logo is a button with a question mark icon and the text "Query Selection". Further right is a link with a download icon and the text "Tutorial". Below these elements is a form with two main sections. The first section has a label "Gene or Micro-rna ID:" followed by a text input field containing "hsa-miR-495-3p". To the right of this is a label "Querying for:" followed by two buttons: "miRNA" (which has a checked checkbox) and "Gene" (which has an unchecked checkbox). The second section has a label "Interaction database" followed by a dropdown menu showing "TargetScan". Below the dropdown is a checkbox labeled "Include Secondary Interactions" which is checked. A blue arrow points from the "miRNA" button to a red error message box that appears in the foreground. The error message box has a white background and a black border. It contains the text "Query did not match up with expression data identifiers! Proceeding without full expression data benefits..." in red. At the bottom right of the error box is a black button with the text "OK".

Odyssey

[? Query Selection](#) [Tutorial](#)

Gene or Micro-rna ID:

Querying for: ☒ miRNA ☐ Gene

Interaction database:

☒ Include Secondary Interactions

Query did not match up with expression data identifiers! Proceeding without full expression data benefits...

OK

A query might have a typo(s) or does not include an exact match but a partial one. In this case, Odyssey reports an irrecoverable error and such cases are handled by a displayed report that contains the error cause and potential matches.



Odyssey

 [Query Selection](#)

 [Tutorial](#)

Gene or Micro-rna ID:

Querying for:

hsa-miX-495-3p

☒ miRNA

☐ Gene

Interaction database

TargetScan

☒ Include Secondary Interactions

Query is not present in selected Interaction Database! Select from similar identifiers or try a different query. Here are a few similar matches ...

HSA-MIR-490-3P
HSA-MIR-485-3P

OK

Differential Expression Analysis Visualization

Differential Expression

Differential Expression Results - miRNA

Show 5 entries

ID	logFC	AveExpr	t	P.Value	adj.P.Val	B	TargetScan	miRNet
HSA-MIR-342-3P	1.864	1.055	6.007	0.001	0.239	-1.45	No	Yes
HSA-MIR-29B-2-3P	-1.479	2.197	-4.028	0.005	0.601	-2.386	No	Yes
HSA-MIR-650	0.797	0.763	3.997	0.005	0.601	-2.405	No	Yes
HSA-MIR-548A-3P	0.851	0.424	3.873	0.006	0.601	-2.485	No	Yes
HSA-MIR-1202	0.683	0.484	3.787	0.007	0.601	-2.542	No	Yes

Showing 1 to 5 of 437 entries

Download

Differential Expression Results - mRNA

Show 5 entries

ID	logFC	AveExpr	t	P.Value	adj.P.Val	B	TargetScan	miRNet
CCR2	-1.191	5.372	-3.753	0.003	1	-4.259	No	Yes
ZDHHC14	-0.784	5.533	-3.206	0.008	1	-4.316	Yes	Yes
TAS2R43	1.118	4.335	3.034	0.011	1	-4.336	No	No
MS4A15	0.962	6.261	3.033	0.011	1	-4.336	Yes	No
NR4A2	1.361	5.879	2.97	0.012	1	-4.344	Yes	Yes

Showing 1 to 5 of 9,064 entries

Download

Back

Forward

Differential expression results of mRNA samples

Differential expression results of miRNA samples

A close view of the *Differential Expression Tab* selected (truncated) is as follows.

In the Differential Expression tab, differential expression analysis (DGEx) results are displayed. The analysis is conducted with limma package [4] and ordered descending with regards to adj. p. value. Correspondingly, most significant results are displayed at the top. Besides DGEx, the presence or absence of miRNAs / mRNAs are added to the tables, aimed to conveniently run Odyssey with different biological molecules that are significant in the dataset of choice according to limma results.

Network Options and Network File

Clicking on *Forward* button after the data selection part is completed (even if “No Data” is selected) takes the analysis to next step. At this step, information about the network is displayed, user can decide the strictness of the filtering criteria to be applied on to the prospective network.

Size of the network, i.e. number of nodes in the network and number of edges between these nodes are dynamically displayed as the filtering parameters changed, allowing user to have control over the magnitude of the network prior to its creation.

The screenshot displays the Odyssey web application interface, divided into two main sections: **Network Options** and **Network File**.

Network Options: This section includes a sidebar with a logo and a 'Tutorial' link. It features two filter options: 'Filter initial query nodes' (Number of nodes: 335) and 'Filter network by degree' (Number of edges: 539). Below these, there are two sliders for filtering miRNA nodes by logFC values, with a 'Filter using' dropdown set to 'Sliders'. Two histograms show the distribution of logFC values for miRNA nodes.

Network File: This section displays a table of network entries. The table has columns for ID, miRNA, Gene, DiffmiRNA, and DiffmiRNA. The first 10 entries are shown, with a 'Show 10 entries' dropdown and a 'Search' bar. The table lists interactions between miRNAs and genes, such as HSA-LET-7A-5P and SEMA4F. Below the table, there is a 'Download' button and a 'Generate Network' button. A 'Back' button is also present.

Annotations with blue brackets link the 'Filter using' dropdown to the 'Binary Interactions in network. miRNA_A -> Gene_X' box, the 'Generate Network' button to the 'Differential Gene Expression Value of the interacting parts' box, and the 'Network Options' section to the 'Network File' section.

ID	miRNA	Gene	DiffmiRNA	DiffmiRNA
1	HSA-LET-7A-5P	SEMA4F	0.264	0.33
2	HSA-LET-7A-5P	VEZT	0.907	0.33
3	HSA-LET-7A-5P	KAT5BL1	0	0.33
4	HSA-LET-7A-5P	MAP3K13	0.586	0.33
5	HSA-LET-7A-5P	ABHD17C	0	0.33
6	HSA-LET-7A-5P	ICMT	-0.016	0.33
7	HSA-LET-7A-5P	NEK9	0.192	0.33
8	HSA-LET-7A-5P	MRS2	0.305	0.33
9	HSA-LET-7A-5P	BIGN1	-0.152	0.33
10	HSA-LET-7A-5P	MARS2	0.582	0.33

Above is the example run from Odyssey with query “*PMAIP1*” as gene, “TargetScan” as interaction database and “GSE104268_2”. Filtering options are explained in detail on this example run in the next section.

Network Options

The screenshot displays the Odyssey Network Options interface. At the top left is the Odyssey logo. The main title is "Odyssey". Below it is a "Network Options" button. The interface is divided into several sections. The top section contains two checkboxes: "Filter initial query nodes" (checked) and "Filter network by degree" (unchecked). To the right of these checkboxes, the "Number of nodes" is 335 and the "Number of edges" is 539. Below this, there is a "Filter using:" section with two buttons: "Sliders" (checked) and "Numbers" (unchecked). Under the "Sliders" button, there are two sliders: "Filter miRNA nodes by logFC values" and "Filter mRNA nodes by logFC values". The miRNA slider has a range from -0.5 to 0.6, with a red bar indicating the selected interval from -0.3 to 0.5. The mRNA slider has a range from -2.6 to 4.7, with a red bar indicating the selected interval from -0.7 to 1. Below the sliders, there are two histograms: "logFC values (miRNA)" and "logFC values (mRNA)". The miRNA histogram shows a distribution of logFC values for miRNA, with a peak around 0.2. The mRNA histogram shows a distribution of logFC values for mRNA, with a peak around 0. The interface is annotated with several callouts explaining the functionality of different parts.

Applies filtering options to initial query and its primary interactions

Filters network by degree value of nodes

Histograms plotted with differential expression value (DGE_x) distribution for miRNA and mRNA datasets.

Display of number of nodes and edges between nodes in the network. Changing filtering settings, dynamically alters the display

Selecting "Numbers" hides slider bars and reveals numerical inputs for filtering UI module.

Filters network by logFC values according to the interval determined. On default, interval excludes logFC values that are in between 10 and 90 percentile values.

Ticking / unticking "Filter initial query nodes", "Filter network by degree" or changing either one of the intervals "Filter miRNA nodes by logFC values" or "Filter mRNA nodes by logFC values" dynamically changes number of nodes and edges between these nodes. These changes are displayed in Network Options.

Network File

The screenshot shows the 'Network File' interface. At the top, there is a 'Search:' input field. Below it, a table displays 10 entries. The table has columns: ID, miRNA, Gene, DiffmRNA, and DiffmiRNA. The first 10 rows of data are highlighted with red boxes. Below the table, there is a 'Download' button, a 'Back' button, and a 'Generate Network' button. Arrows point from the 'miRNA' column, 'Gene' column, 'DiffmRNA' column, and 'DiffmiRNA' column to labels: 'miRNA column', 'Gene column', 'logFC values of Genes', and 'logFC values of miRNAs' respectively.

ID	miRNA	Gene	DiffmRNA	DiffmiRNA
1	HSA-LET-7A-5P	SEMA4F	0.264	0.33
2	HSA-LET-7A-5P	VEZT	0.907	0.33
3	HSA-LET-7A-5P	KATNBL1	0	0.33
4	HSA-LET-7A-5P	MAP3K13	0.586	0.33
5	HSA-LET-7A-5P	ABHD17C	0	0.33
6	HSA-LET-7A-5P	ICMT	-0.016	0.33
7	HSA-LET-7A-5P	NEK9	0.192	0.33
8	HSA-LET-7A-5P	MRS2	0.305	0.33
9	HSA-LET-7A-5P	B3GNT1	-0.152	0.33
10	HSA-LET-7A-5P	MARS2	0.582	0.33

Showing 1 to 10 of 29,237 entries

Previous 1 2 3 4 5 ... 2924 Next

Download

Back

Generate Network

miRNA column

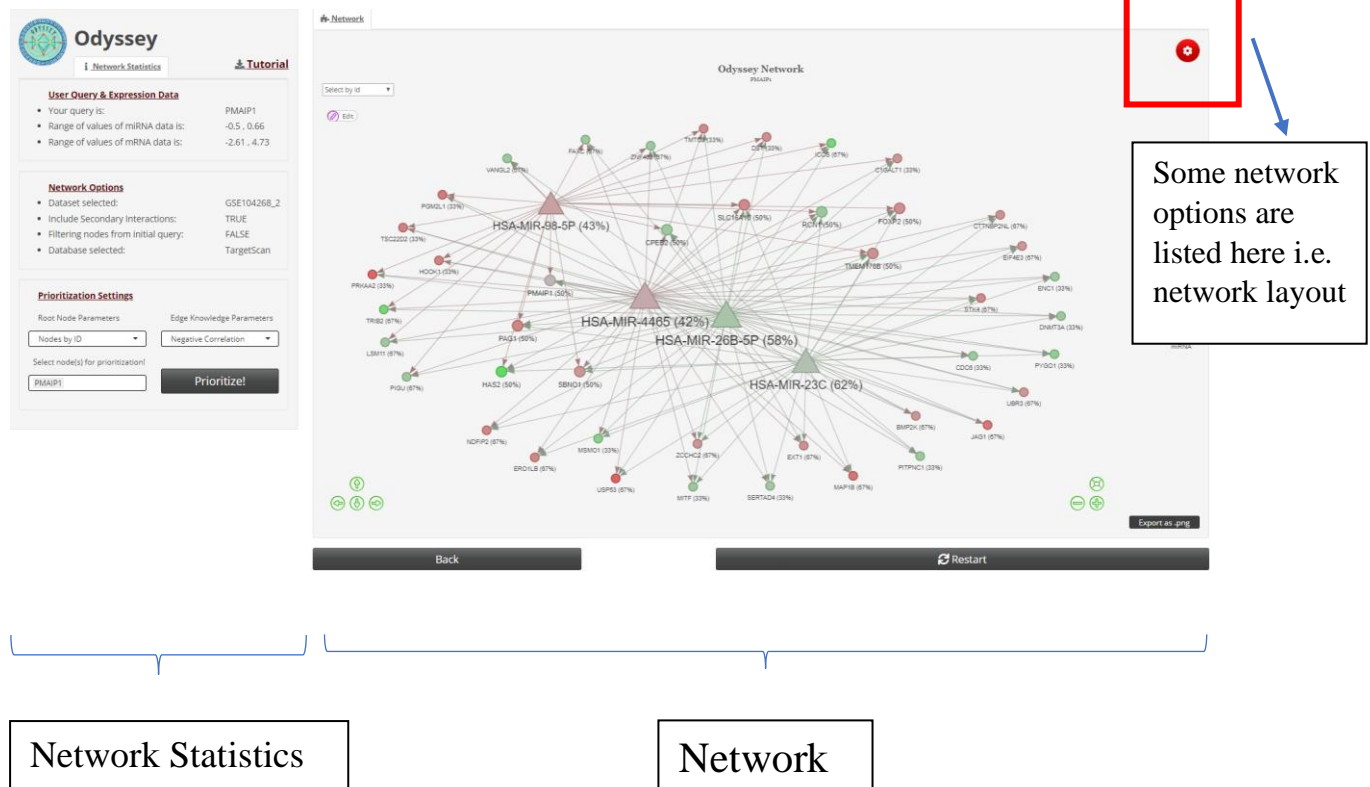
Gene column

logFC values of Genes

logFC values of miRNAs

Binary interactions are displayed row-wise with their respected differential expression values. Also, the search button placed on top of the data table conveniently allows searching individual miRNAs and /or Genes from a large network. After desired options for the network is determined, “*Generate Network*” button is clicked and the network is visualized in the next step.


Network Statistics and Visualization



Pressing “*Generate Network*” button in the previous step makes Odyssey visualize the network as shown in the example above. Parameters that were used in the Network’s creation is displayed in “Network Statistics” tab along with a few interactive settings that can be applied on to drawn network i.e. layout algorithm selection and prioritization (Explained in the Network Statistics section of Tutorial).


The network is visualized at the right section of the application, allowing interactive node selection and dragging and downloading the network in different formats (Explained in the Network Visualization section of Tutorial).

Network Statistics



Odyssey

i [Network Statistics](#)

 [Tutorial](#)

User Query & Expression Data

- Your query is: PMAIP1
- Range of values of miRNA data is: -0.5 , 0.66
- Range of values of mRNA data is: -2.61 , 4.73

Network Options

- Dataset selected: GSE104268_2
- Include Secondary Interactions: TRUE
- Filtering nodes from initial query: FALSE
- Database selected: TargetScan

Prioritization Settings

Root Node Parameters

Edge Knowledge Parameters

Nodes by ID

Negative Correlation

Select node(s) for prioritization!

PMAIP1

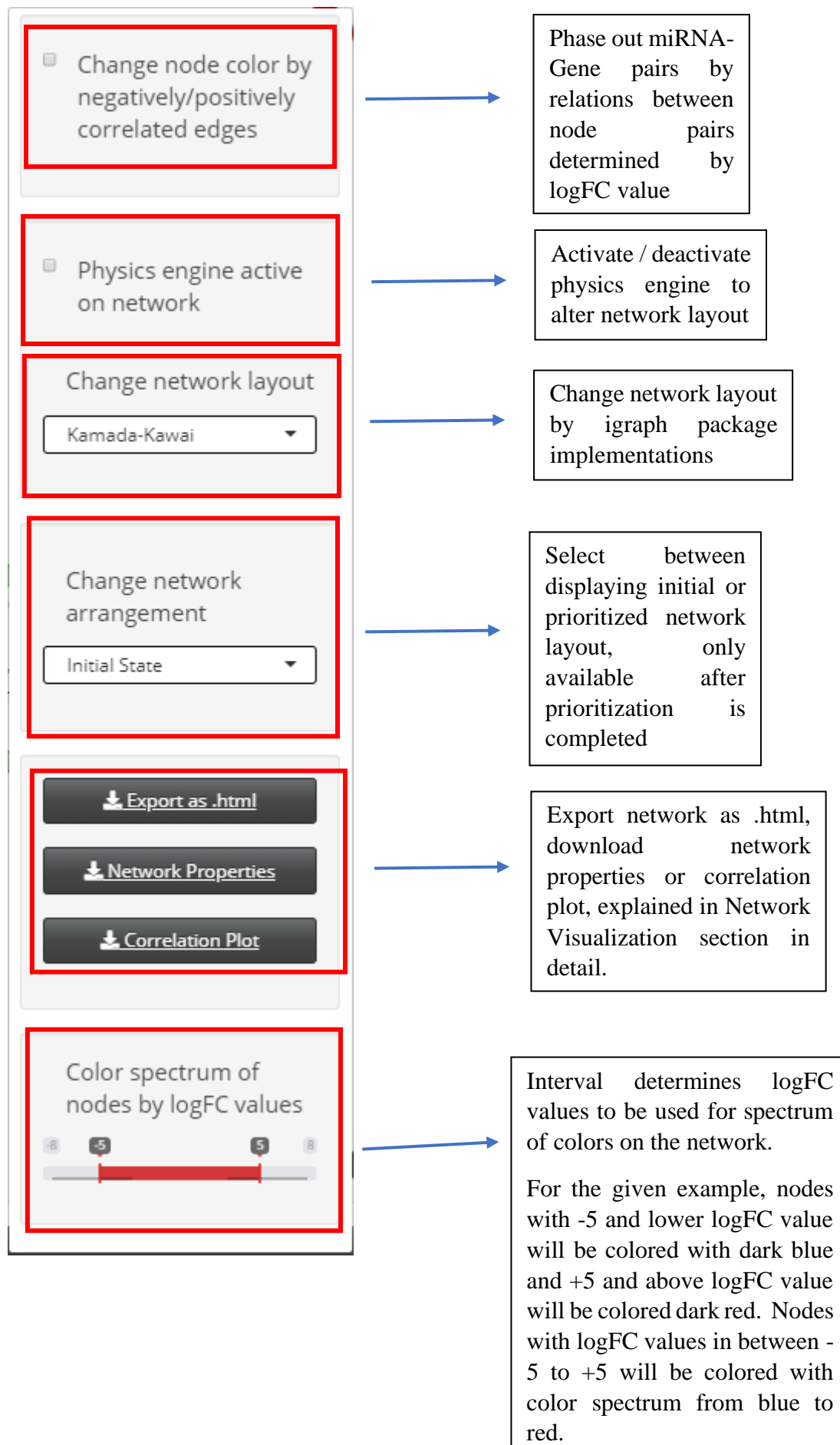
Prioritize!

Query and Expression Data

Parameters used for creating the network

Prioritize the network by node and edge parameter selections

Prioritization Settings contains the options to be used to employ the prioritization. Further details are given in the Prioritization section of this tutorial document.



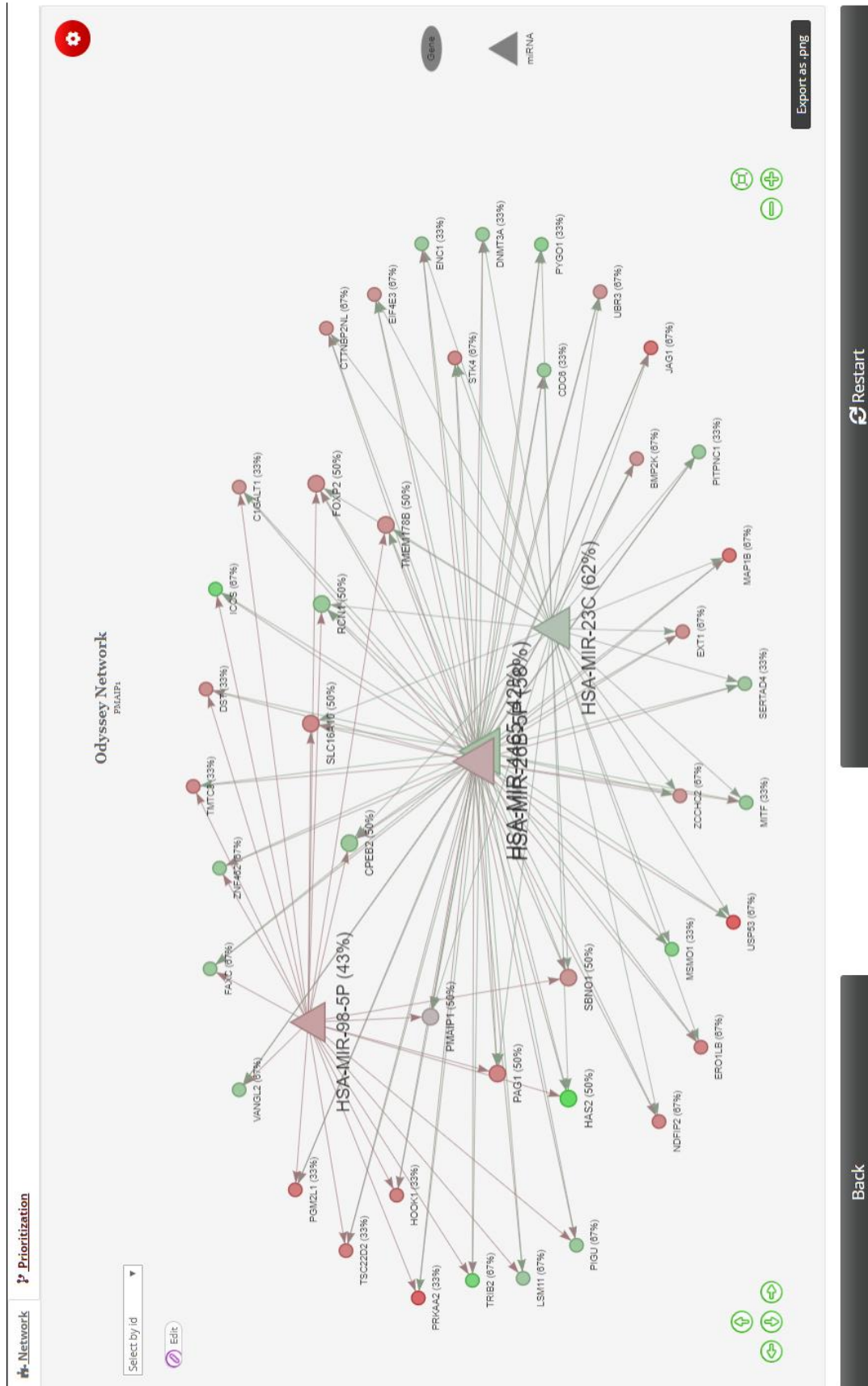
If expression data is used in the session, Odyssey draws networks based on miRNA / mRNA interactions with each node having a corresponding logFC value. logFC values can be positive (upregulation) or negative (downregulation) based on the comparison of Treatment vs Control samples listed in the expression dataset used. A percentage score of correlation is calculated for each node based direct interactions with other nodes shown in the network.

If a microRNA is listed as upregulated and shown to be interacting with 3 downregulated and 1 upregulated gene in the network, corresponding score of correlation for that particular microRNA is 75%. If another microRNA in the network is downregulated and shown to be interacting with 9 upregulated genes and 1 downregulated genes, corresponding score of correlation for that microRNA is 90% and so forth. Odyssey uses the following formula for calculation of correlation for each node in the network.

$\text{Correlation Score \%} = \text{Number of inversely interacting nodes} / \text{Number of total interactions}$
--

According to this formula, “*Change node color by correlation percentage*” option greys out positively or negatively correlated nodes based on selection based on calculated correlation score.

Network Visualization



Odyssey generates an interactive network where the nodes and correspondingly connected edges can be moved to obtain a user-tailored view. Users can use keyboard arrow buttons or mouse drag to move around the network. Mouse wheel or implemented minus and plus signs at the bottom-right corner of the network can be utilized to zoom-in to or out.

There are two options to download the visualized network in Odyssey. First option is to download the network in .png format specified as “*Export as.png*”. This option yields a non-interactive network where as “*Export as .html*” option yields an interactive version of the network as is possible while using the application.

Two more download buttons are implemented directly below the network. First of these download button is “*Network Properties*”.

Row.names	Degree	Correlation	logFC	AveExpr	t	P.Value	adj.P.Val	B	TargetScan	miRNet
BMP2K	3	67%	1.029	6.904	6.028	0	0.004	0.547	Yes	Yes
C1GALT1	3	33%	1.006	6.859	6.22	0	0.004	0.779	Yes	Yes
CDC6	3	33%	-0.802	6.522	-6.096	0	0.004	0.63	Yes	Yes
CPEB2	4	50%	-0.855	4.846	-3.988	0.004	0.027	-2.241	Yes	Yes
CTTNBP2NL	3	67%	1.148	5.931	8.245	0	0.001	2.949	Yes	Yes
DNMT3A	3	33%	-0.85	5.63	-5.444	0.001	0.007	-0.191	Yes	Yes
DST	3	33%	1.184	7.48	7.949	0	0.001	2.661	Yes	Yes
EIF4E3	3	67%	1.027	5.313	5.857	0	0.005	0.336	Yes	Yes
ENC1	3	33%	-0.815	4.655	-5.158	0.001	0.009	-0.568	Yes	Yes
ERO1LB	3	67%	1.406	4.974	6.827	0	0.002	1.481	Yes	Yes
EXT1	3	67%	1.068	7.702	8.646	0	0.001	3.325	Yes	Yes
FAXC	3	67%	-0.845	5.138	-4.569	0.002	0.015	-1.386	Yes	Yes
FOXP2	4	50%	1.165	3.11	7.955	0	0.001	2.666	Yes	Yes
HAS2	4	50%	-2.293	5.785	-14.647	0	0	7.521	Yes	Yes
HOOK1	3	33%	1.563	2.944	9.609	0	0.001	4.167	Yes	Yes
HSA-MIR-23C	28	61%	-0.326	1.203	-0.572	0.59	0.998	-4.65	Yes	Yes
HSA-MIR-26B-5P	42	57%	-0.504	2.019	-1.185	0.284	0.998	-4.575	Yes	Yes
HSA-MIR-4465	42	43%	0.528	1.941	1.91	0.109	0.998	-4.474	Yes	Yes
HSA-MIR-98-5P	22	45%	0.657	1.879	1.002	0.358	0.998	-4.6	Yes	Yes
ICOS	3	67%	-1.706	4.878	-10.939	0	0	5.206	Yes	No

Network Properties file is downloaded as displayed above (truncated). Apart from the differential expression results of the nodes on the network, degree and correlation score of the nodes are also added in this file to supply a comprehensive info file.

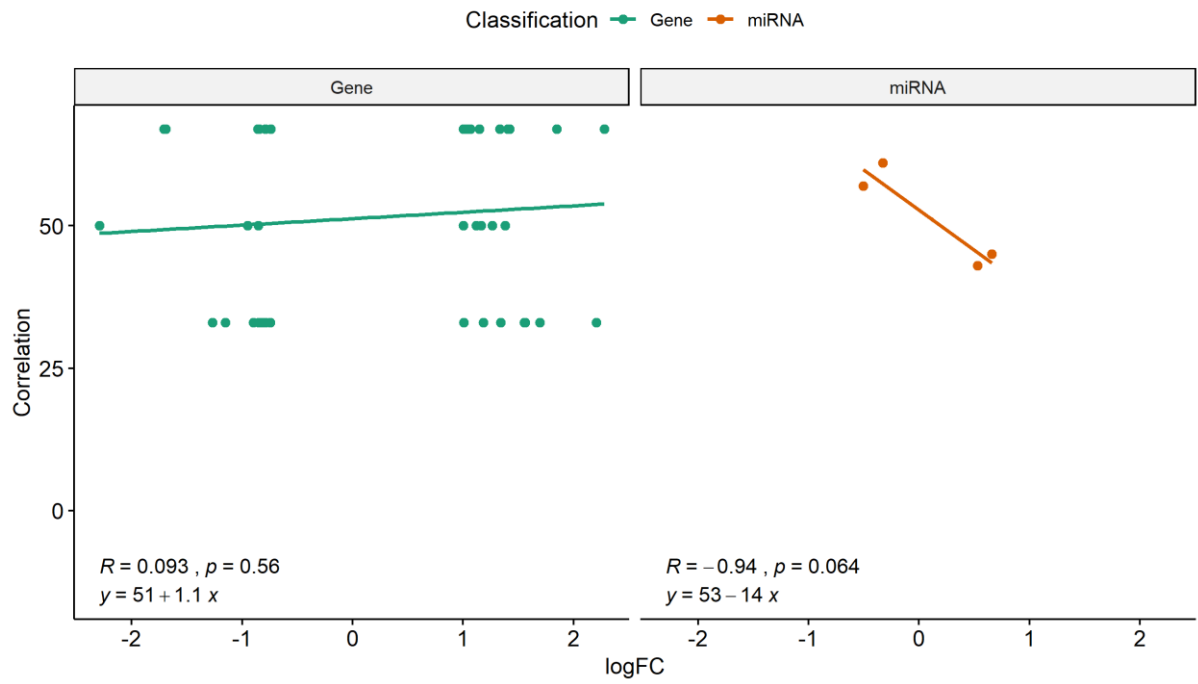
Correlation plot draws a scatterplot between the quantitative variables correlation score% ~ logFC separately for miRNA and mRNA data. We aim to provide insight into experimental settings where miRNA – gene interaction is plotted with parameters mentioned.

	Upregulated Node #	Downregulated Node #
Correlation Score >= 50%		
Correlation Score < 50%		

Fisher’s exact test is conducted separately for miRNA and gene data according to the contingency table above. Therefore, scatterplot and the fisher’s exact test is conducted to reveal any significant association between correlation score and expression levels of the nodes.

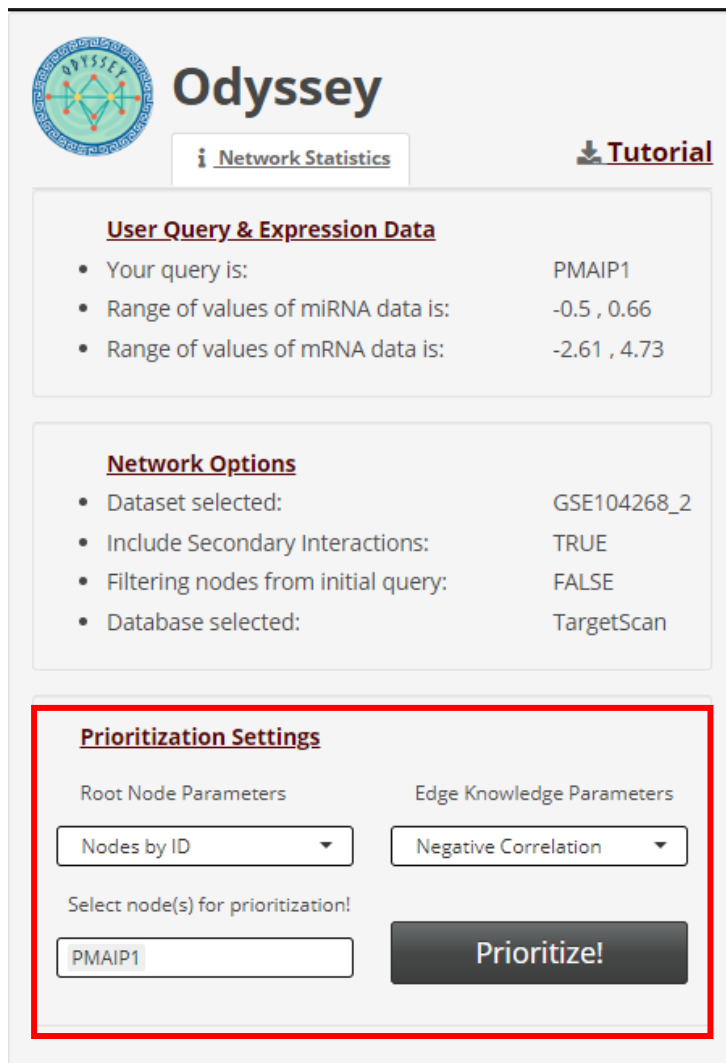
Fisher's test p value for Genes: 0.542 Fisher's test p value for miRNAs: 0.333

PMAIP1 GSE104268_2



Prioritization

Odyssey uses prioritization to highlight nodes that are selected as important according to Odyssey's implementation of KNGP algorithm [5]. KNGP algorithm is a Random Walk algorithm that requires root node parameter, node parameter and edge parameter defined for each network. Root nodes are defined as start point(s) of the random walk that can be selected manually from node ID table or automatically by node degree in the network. Node parameter is weight of nodes defined as absolute logFC value calculated in the differential expression analysis step. Finally, edge parameter is interaction weight between nodes that can be selected to favor negatively correlated node pairs or positively correlated node pairs. Together, these parameters define the start point, direction of random walk and finally the list of results that contain prioritization score of each node within the network.



The screenshot shows the Odyssey web application interface. The top section includes the Odyssey logo, a "Network Statistics" link, and a "Tutorial" link. Below this, there are three main sections: "User Query & Expression Data", "Network Options", and "Prioritization Settings". The "Prioritization Settings" section is highlighted with a red box. It contains two columns of settings: "Root Node Parameters" and "Edge Knowledge Parameters". Under "Root Node Parameters", there is a dropdown menu set to "Nodes by ID" and a text input field containing "PMAIP1". Under "Edge Knowledge Parameters", there is a dropdown menu set to "Negative Correlation". Below these, there is a "Select node(s) for prioritization!" label and a "Prioritize!" button.

Odyssey

[Network Statistics](#) [Tutorial](#)

User Query & Expression Data

- Your query is: PMAIP1
- Range of values of miRNA data is: -0.5 , 0.66
- Range of values of mRNA data is: -2.61 , 4.73

Network Options

- Dataset selected: GSE104268_2
- Include Secondary Interactions: TRUE
- Filtering nodes from initial query: FALSE
- Database selected: TargetScan

Prioritization Settings

Root Node Parameters Edge Knowledge Parameters

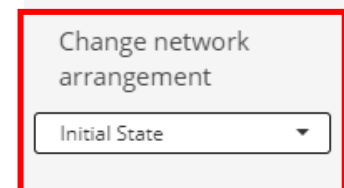
Nodes by ID Negative Correlation

Select node(s) for prioritization!

PMAIP1 **Prioritize!**

Odyssey aims to highlight biologically “important” nodes from a large network that is otherwise challenging to analyze. After prioritization, prioritization scores are obtained for each node, and the scores are ordered ascending. From the list of ordered scores, quantile values are calculated and nodes with prioritization score above 90 percentile are displayed in the network in square shape. Other nodes are grouped and merged in clusters according to quantile values. Nodes with prioritization score within 75th percentile and 90th percentile are clustered within “Medium” cluster and the rest of the nodes are clustered in “Low” cluster.

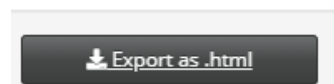
“Change network arrangement” menu is used to switch between prioritized network view and the initial network view.



The screenshot shows a dropdown menu titled "Change network arrangement". The menu is open, showing a single option: "Initial State".

Change network arrangement

Initial State



The screenshot shows a button labeled "Export as .html" with a download icon.

Export as .html

Prioritization Results

Show 10 entries

ID	Node_ID
1	HSA-MIR-26B-5P
2	HSA-MIR-4465
3	HSA-MIR-23C
4	HSA-MIR-98-5P
5	HAS2
6	USP53
7	PRKAA2
8	MAP1B
9	JAG1
10	ICOS

List of nodes that are in the network. A star (*) sign designates the root nodes used for prioritization.

Involvement_Score	Node Rank #
0.121	1
0.091	2
0.08	3
0.054	4
0.026	5
0.025	6
0.024	7
0.021	8
0.021	9
0.02	10

Involvement score ranges from 0-1. Involvement score of 1 means the node is very highly prioritized in the network.

Node Ranks

Prioritization Info

Root node selection:

- Nodes by ID

These are the root nodes:

- PMAIP1

Edge Knowledge Parameter:

- Negative Correlation

Results were obtained for bestf value of:

- 0

Prioritization information box contains parameters used for prioritization algorithm.

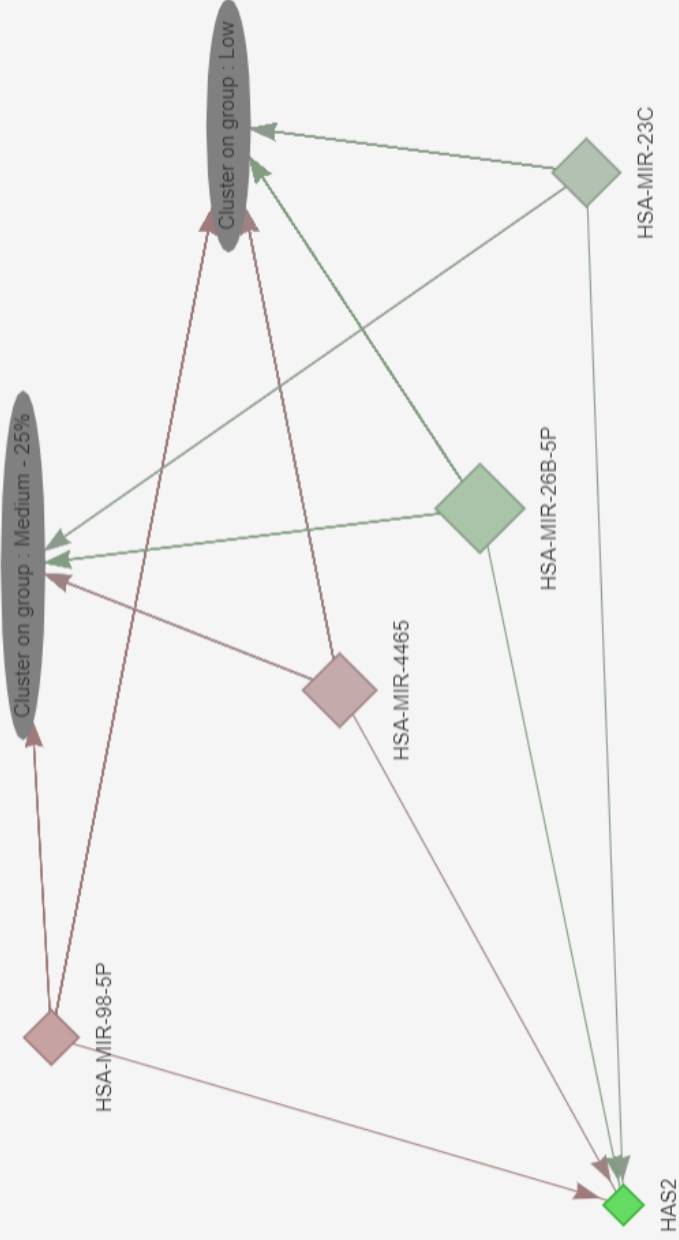
Back

Odyssey Network

FMAIP4

Select by id

Edit



Reinitialize clustering

Export as .png

Back

Restart



Low

High - 10%

Medium - 25%

References

1. Fan, Y. et al. miRNet — dissecting miRNA-target interactions and functional associations through network-based visual analysis. *Nucleic Acids Res.* 1, W135–W141 (2016).
 2. Agarwal, V., Bell, G. W., Nam, J. W. & Bartel, D. P. Predicting effective microRNA target sites in mammalian mRNAs. *eLife* 4, e05005 (2015).
 3. Barrett, T., Suzek, T.O., Troup, D.B., Wilhite, S.E., Ngau, W.C., Ledoux, P., Rudnev, D., Lash, A.E., Fujibuchi, W., Edgar, R. 2005NCBI GEO: mining millions of expression profiles—database and tools *Nucleic Acids Res.* 33D562–D566
 4. M. E. Ritchie, B. Phipson, D. Wu, Y. Hu, C. W. Law, W. Shi, and G. K. Smyth. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic acids research*, 43(7):e47, 2015.
 5. Kimmel C, Visweswaran S: An algorithm for network-based gene prioritization that encodes knowledge both in nodes and in links. *PLoS One* 2013, 8: e79564. 10.1371/journal.pone.0079564
-

Index

D

[Data Information](#) · 15
[Data Upload](#) · 9
[Data Visualization](#) · 14
[Differential Expression](#) · 23

E

[Expression Data Choice](#) · 7

H

[Home Page](#) · 6

I

[ID Checking](#) · 19
[Include Secondary Interactions](#) · 19
[Interaction database](#) · 18

N

[Network File](#) · 27

[Network Options](#) · 26
[Network Options and Network File](#) · 25
[Network Statistics](#) · 29
[Network Statistics and Visualization](#) · 28
[Network Visualization](#) · 32

P

[Prioritization](#) · 35

Q

[Query Selection](#) · 18

R

[References](#) · 38

S

[Summary](#) · 3