

multiSLIDE User Manual

Version 2.0

Table of Contents

1.	<i>Run pre-existing demos</i>	3
2.	<i>Open a “.mslide” file</i>	5
3.	<i>Upload your own data</i>	6
3.1	Query using keyword search	9
3.2	Query by uploading pathways	10
3.3	Query by enrichment analysis.....	10
4.	<i>Functionalities</i>	11
5.	<i>Save visualizations</i>	12
6.	<i>Save workspace</i>	13
7.	<i>Upload externally curated network</i>	14

1. Run pre-existing demos

1. On the multiSLIDE homepage, clicking the “Demo” button displays three pre-loaded analyses. Click the “Open” to open the corresponding analysis. The details of the demos are available in the Case Studies section of our paper: <https://www.biorxiv.org/content/10.1101/812271v2>.

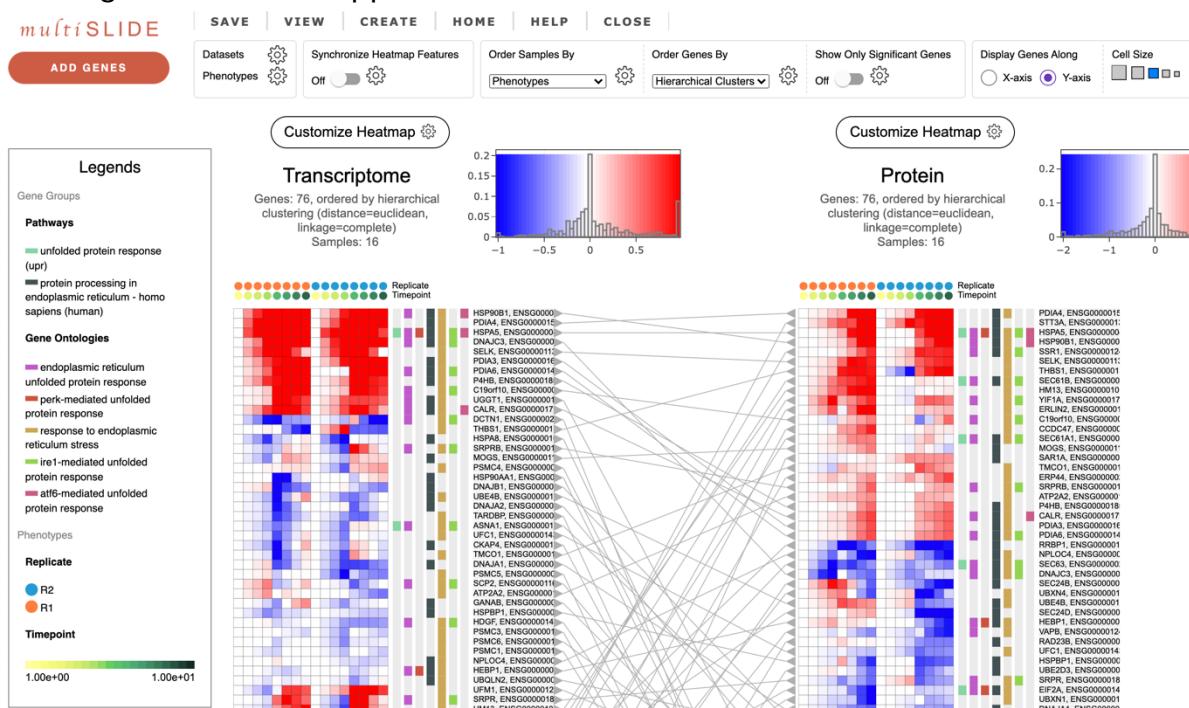
The screenshot shows the multiSLIDE homepage with a dark background. At the top center is the 'multiSLIDE' logo. Below it is the tagline 'Multi-omics Systems Level Interactive Data Exploration'. To the right are four buttons: 'CREATE', 'LOAD', 'OPEN', and 'DEMO'. Underneath these are three demo cards, each with a title, a small icon, and an 'OPEN' button:

- Unfolded Protein Response in Time-series mRNA and Protein data [1]
- Kinase-Substrate relationships in CPTAC Ovarian Cancer Data [2]
- Exploring microRNA to Protein targets using TargetScan Map [3]

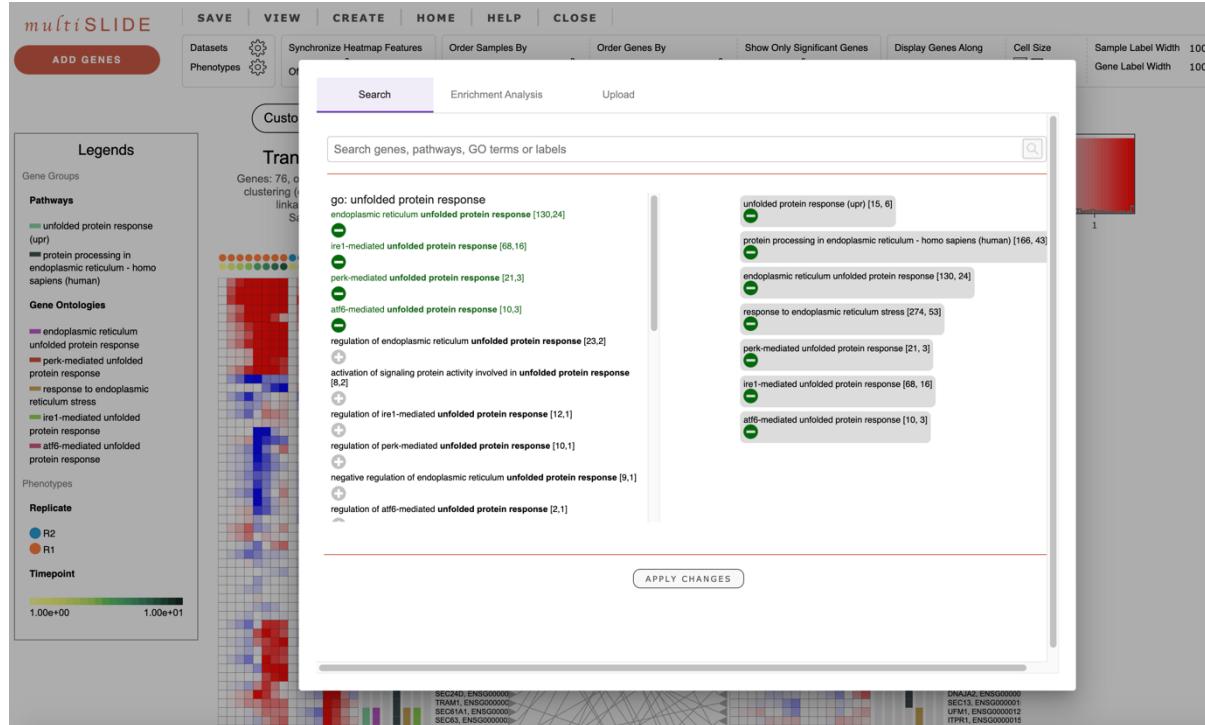
At the bottom of the page, there is a note with three references:

- [1] Cheng, Zhe, et al. "Differential dynamics of the mammalian mRNA and protein expression response to misfolding stress." *Molecular systems biology* 12.1 (2016): 855.
- [2] Zhang, Hui, et al. "Integrated proteogenomic characterization of human high-grade serous ovarian cancer." *Cell* 166.3 (2016): 755-765.
- [3] Choi, Hyungwon, et al. "Plasma protein and microRNA biomarkers of insulin resistance: A network-based integrative-omics analysis." *Frontiers in physiology* 10 (2019): 379.

2. Clicking ‘Open’ on multiSLIDE homepage leads to the visualization page. Here pathways, GO terms have been pre-selected and the clustering and feature ordering has also been applied.



3. Clicking ‘Add Genes’ on the top-left corner of the visualization page opens a panel showing the pre-selected pathways and GO terms. The ‘+’ and ‘-‘ buttons can be used to add or remove pathways for visualization, respectively. Selected (added for visualization) search results are added to the right-side panel. Finally click ‘Apply Changes’ for changes to take effect.



4. To customize visualizations, set global clustering and statistical filtering parameters and customize individual heatmaps techniques see Section 4.

2. Open a “.mslide” file

The analysis workspace can be saved as a ‘.mslide’ file. To reopen a “.mslide” file click ‘Load’ button in the multiSLIDE homepage. This opens a panel below where the file can be selected using ‘Choose file’. Clicking ‘Upload’ re-opens the saved workspace. Pre-generated example .mslide files for the demo visualizations are available in https://github.com/soumitag/multiSLIDE/tree/master/demo_data.



3. Upload your own data

1. To start uploading your data click ‘Create’ in multiSLIDE homepage. In the panel that opens up, provide an Analysis Name in the textbox and click ‘Create’.



Analysis Name

 CREATE

2. In the data upload page, first select the Species type from the dropdown list

1. Provide Species for Analysis 'test'

Species Type

- Homo sapiens
- Mus musculus
- Other

2. Please Upload a Single Clinical Information File for the Experiment

Upload Clinical Information File

Clinical Information File Delimiter

UPLOAD

RESET

3. Please Upload -omics Data File(s) using this module

Data Type

Data File Upload

Data File Delimiter

3. In the data upload page, select a sample/clinical information file using the ‘Choose file’ button, select the appropriate file delimiter from the dropdown list, and click ‘Upload’ button. Successful completion of file upload will prompt a success message. Example sample information files are available in https://github.com/soumitag/multiSLIDE/tree/master/demo_data.

The screenshot shows a dark-themed web application interface for data upload. At the top left, a message says "1. Provide Species for Analysis 'test'". Below it, a dropdown menu is open. To the right, a large central panel is titled "2. Please Upload a Single Clinical Information File for the Experiment". It contains a "Choose file" input field containing the text "ER_Stress_sa...ing_info.txt", a "Clinical Information File Delimiter" dropdown set to "Tab", and a prominent orange "UPLOAD" button. To the right of this panel, a message indicates the file was uploaded successfully: "File ER_Stress_sample_grouping_info.txt uploaded successfully". At the bottom left, another message says "3. Please Upload -omics Data File(s) using this module". A dropdown menu labeled "Data Type" is shown below this message.

4. Using the panel on the bottom left of the data upload page, upload each omics data file one at a time. To upload an omics file, select the Data Type from the dropdown list to indicate the type of ‘omics’ data. Then use ‘Choose file’ button to select the omics file, select the file delimiter from the dropdown list, and click the ‘Submit’ button. On successful file upload the file will be listed in the table on the right. Examples of omics quantitative expression files are available here: https://github.com/soumitag/multiSLIDE/tree/master/demo_data.

The screenshot shows the same dark-themed web application. The top section is identical to the previous screenshot, showing the clinical information file upload step. The bottom section shows the process for uploading omics data files. It includes a "Data Type" dropdown set to "Protein", a "Choose file" input field containing "protein_ER_st...rmalized.txt", a "Data File Delimiter" dropdown set to "Tab", and a "SUBMIT" button. To the right, a table lists the uploaded file: "Transcriptome mRNA_ER_stress_b6_sline_normalized.txt" with "Gene Expression (mRNA)" as the "Data Type", "Tab" as the "Delimiter", and "Linker(Gene) Column Mapping: Molecular-level Specific Identifier(s):" as the "Metadata Column Information". There are "Edit", "Preview", and "Delete" buttons for this entry. A "CREATE" button is also visible.

5. The table shows the details of uploaded omics files. Clicking ‘Edit’ will open the Metadata Column Selection and Identifier Mapping panel. All information requested in this panel has to be completed for each omics file before clicking ‘Create’ to create the analysis.

Display Name	Filename	Data Type	Delimiter	Metadata Column Information	Edit	Preview	Delete
Transcripome	mRNA_ER_stress_baseline_normalized.txt	Gene Expression (mRNA)	Tab	Metadata Columns: Linker(Gene) Column Mapping: Molecular-level Specific Identifier(s):			
Protein	protein_ER_stress_baseline_normalized.txt	Protein	Tab	Metadata Columns: Linker(Gene) Column Mapping: Molecular-level Specific Identifier(s):			

RESET **CREATE**

6. Identify all metadata columns (non-quantitative expression data) in the data file, by selecting from the dropdown list. To align omics data, specify any linker columns, if available. Linker columns can be mappable to one of the standard identifiers, such as Gene Symbol, Entrez etc. Next, if molecular level specific identifiers such as genomic coordinates, phosphosites are also available, identify these columns from the dropdown list. Click ‘Save Changes’ and return to the main page. Once the details for all datasets have been filled click ‘Create’ to create the analysis.

Metadata Column Selection and Identifier Mapping

Metadata Columns

- GeneSymbol
- Ensembl GeneSymbol
- GeneSymbol

Does this dataset contain a 'Gene' column or a 'Linker' variable?
Linker variable are used to align omics identifiers across molecular levels. Learn more.

Yes No

Map Linker Column -> Standard Identifier

Linker Column

GeneSymbol

Standard Identifiers

Gene Symbol

GeneSymbol → Gene Symbol

Does this dataset have molecular-level specific identifiers?
(for e.g. genomic coordinates, transcript identifiers etc.)

Yes No

SAVE CHANGES

3.1 Query using keyword search

Clicking ‘Add Genes’ on the top-left corner of the visualization page opens a panel with three tabs: ‘Search’, ‘Enrichment Analysis’, and ‘Upload’. Use the ‘Search’ tab for keyword search-based selection. Specify the search query in the search text box and click the search icon. The search results are indicated in the panel below. Scroll through the results and use the “+” and “-“ icons to add or remove selections from visualization. Selected (added for visualization) search results are added to the right-side panel.

The screenshot shows the multiSLIDE search interface. At the top, there are three tabs: 'Search' (which is selected), 'Enrichment Analysis', and 'Upload'. Below the tabs is a search bar containing the query 'pathway: stress'. To the right of the search bar is a magnifying glass icon. The main area displays a list of search results. On the left side of the results, each item has a '+' or '-' button next to it, indicating whether it is selected for visualization. The results are as follows:

- pathway: stress
cellular responses to stress [393,58]
+
cellular response to heat stress [100,27]
+
fas ligand (fasl) pathway and stress induction of heat shock proteins (hsp) regulation [43,10]
+
oxidative stress induced senescence [129,9]
+
activation of atr in response to replication stress [37,8]
+
dna damage/telomere stress induced senescence [28,6]
+
oxidative stress [30,5]
+
cellular response to heat stress [16,5]
+
oxidative stress induced gene expression via nrf2 [19,4]
+
simplified interaction map between loxl4 and oxidative stress pathway [18,4]
- unfolded protein response (upr) [15, 6]
-
protein processing in endoplasmic reticulum - homo sapiens (human) [166, 43]
-
endoplasmic reticulum unfolded protein response [130, 24]
-
response to endoplasmic reticulum stress [274, 53]
-
perk-mediated unfolded protein response [21, 3]
-
ire1-mediated unfolded protein response [68, 16]
-
atf6-mediated unfolded protein response [10, 3]
-

At the bottom center is a 'APPLY CHANGES' button. At the very bottom of the interface is a small link 'Info about Search'.

The table below shows three example queries that can be used in multiSLIDE. In the three examples, multiple searches are combined into a single query by specifying a semi-colon separated list of individual search queries. The individual queries can either take the form keyword=terms, for exact search, or keyword:terms, for inexact search. Here terms indicate the search terms, which can be a single search term or a comma separated list of search terms. Given that the user has to explicitly specify the keyword to search, multiSLIDE uses extensive name mapping, to support multiple syntaxes. For instance, in the examples in the table, “go”, “go-term”, and “Term” all imply the keyword GO term. Similarly, in the first and third examples, “path” and “PATHNAME” are equivalent and imply the keyword pathway name. To ensure composite searches containing multiple queries are not any slower than the individual searches these are performed parallelly.

Examples of composite search queries in multiSLIDE

- 1) path = estrogen signaling pathway; go:chemotaxis; gene: cdk1, ESR1,traf2
- 2) path-id = 04915; go-term:chemotaxis; genesymbol: cdk1, ESR1,traf2
- 3) PATHNAME=estrogen signaling pathway; Term:chemotaxis;entrez: 983, 2099, 7186

3.2 Query by uploading pathways

Clicking ‘Add Genes’ panel on the top-left corner of the visualization page opens a panel with three tabs. Click on the third tab ‘Upload’ to upload your own pathways or subsets of data to visualize.

Search Enrichment Analysis **Upload**

Upload Functional Group Information

Functional Group Information Upload

Pathways_36...8082020.txt

File Pathways_36_kinases_corresponding_substrates_28082020.txt uploaded successfully

Functional Groups

3.3 Query by enrichment analysis

Clicking ‘Add Genes’ panel on the top-left corner of the visualization page opens a panel with three tabs. Click on the second tab ‘Enrichment Analysis’ and specify the dataset and parameters to run enrichment analysis and list the enriched pathways. These pathways can be added in the visualization

Search **Enrichment Analysis** Upload

Differential Analysis

Dataset

Phenotype

Test Type

Significance Level

Multiple Testing Correction

False Discovery Rate

% (Enter a number between 1-100)

Enrichment Analysis

Enrichment Type

Pathway Gene Ontology Terms

Significance Level

4. Functionalities



The above image shows the global settings available in multiSLIDE. The numbered panels can be used to:

1. select the omics datasets and phenotypes to visualize
2. turn on/off synchronized and unsynchronized (independent) clustering
3. select the sample ordering scheme
4. select the molecule ordering scheme
5. turn on/off statistical filtering of the data.
6. change the orientation of the heatmaps
7. adjust heatmap cell sizes to one of the five preset sizes
8. adjust the size heatmap labels

The settings buttons besides panels 1-5 can be used to specify parameters specific to the selected customizations.

Individual omics heatmaps can be customized (by setting the number of histogram bins, the heatmap color scheme, range of the data etc.) by clicking the ‘Customize Heatmap’ settings button available beside each heatmap. This opens the panel shown below. Click ‘Apply Changes’ to close and apply the selected changes to the heatmap.

Aggregate by Linker

Aggregator Function

Feature Identifier(s)

gene_symbol

Number of Colors (Valid Values: 0-255)

Binning Range

- Use Min and Max of Data (Min=-5.59, Max=6.25)
- Use Symmetric Bins (about 0)
- Use Range

Start and End

Heatmap Color Scheme

Diverging colormaps

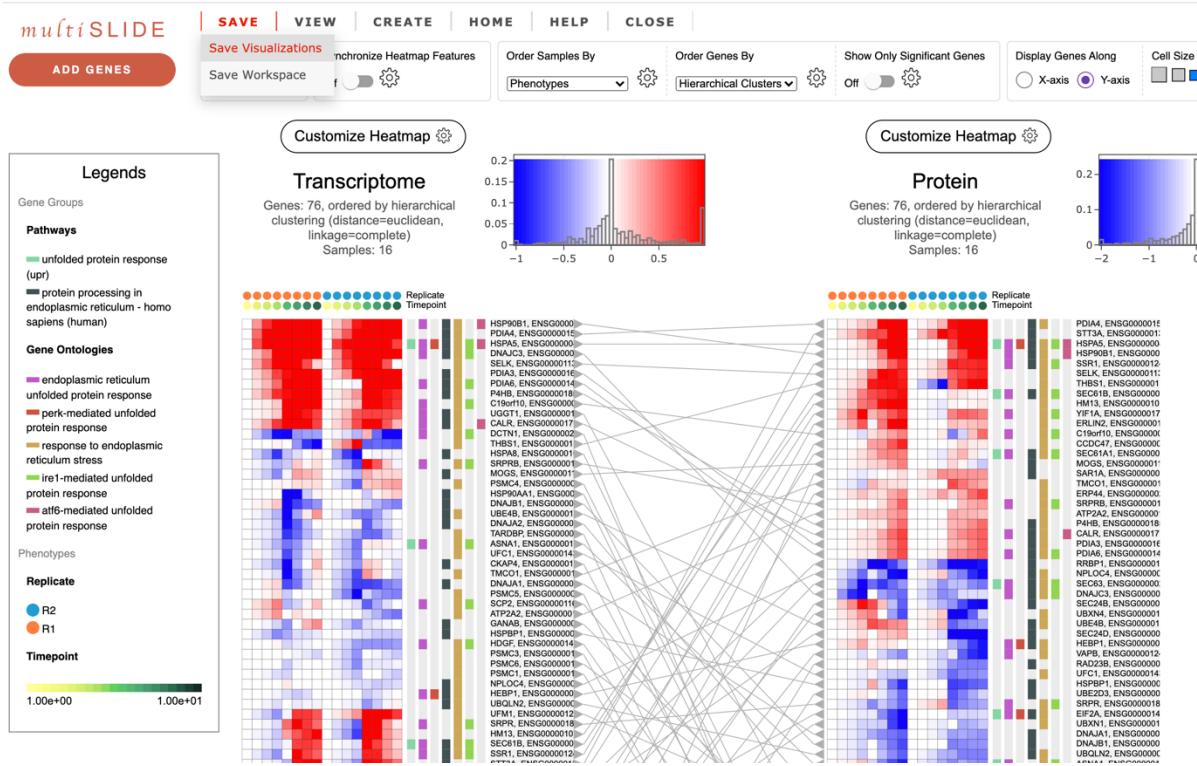
- BBKY
- BLWR
- BYR
- GBR
- SEISMIC
- COOLWARM
- BRBG
- PRGN
- PIYG
- RDYLGN
- RDBU
- RDYLBU
- SPECTRAL

Perceptually uniform colormaps

- VIRIDIS
- CIVIDIS
- INFERNO
- MAGMA
- PLASMA

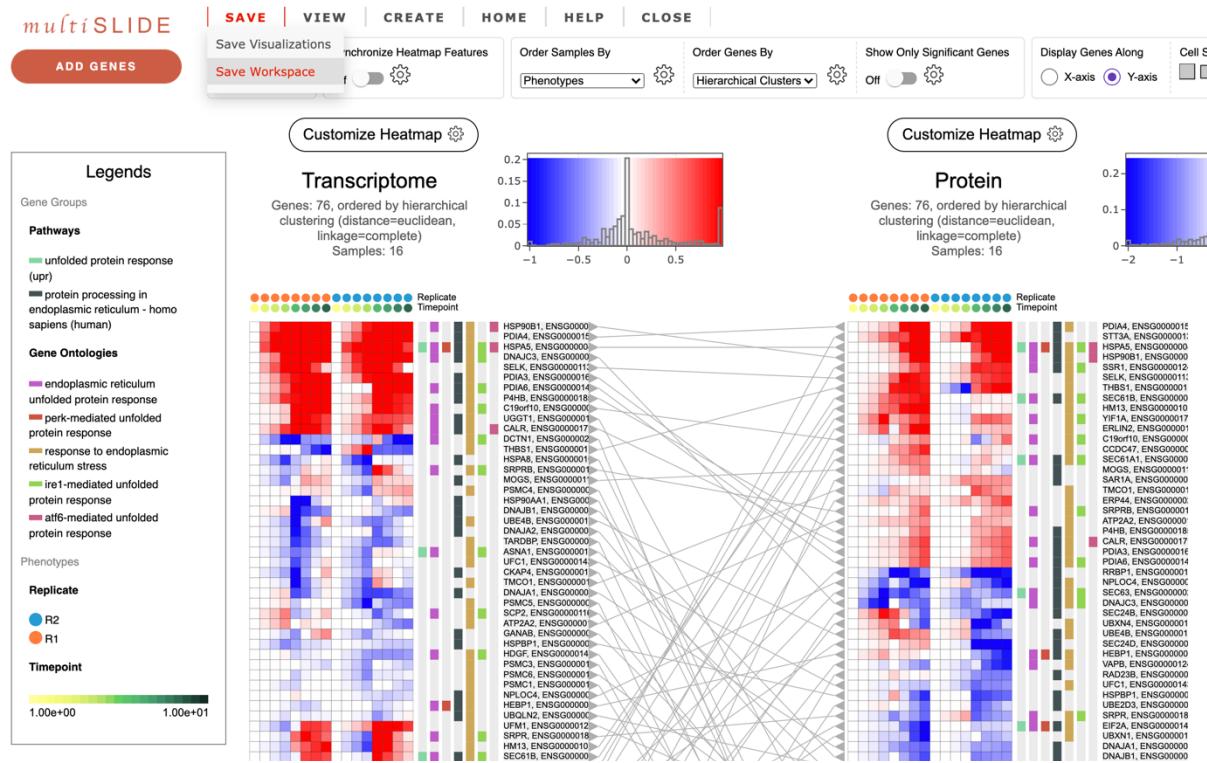
5. Save visualizations

In the top menu panel of the visualization page, mouse-over the ‘Save’ option opens a drop-down list. Clicking ‘Save Visualizations’ opens a new tab with the visualization. Right-click anywhere on the page and select ‘Print’ from the Context menu. In the Print dialog that opens select Destination ‘Save as PDF’ and using the ‘Scale’ option in the Print Settings Panel, scale the visualization to fit it in a single page.



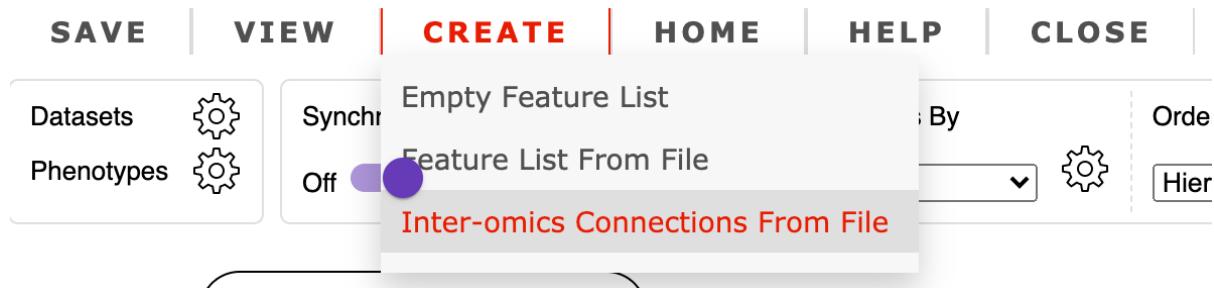
6. Save workspace

In the top menu panel of the visualization page, mouse-over the ‘Save’ option opens a drop-down list. Clicking ‘Save Workspace’ opens the Save Workspace panel. Click the ‘Download’ button to download the analysis workspace as .mslide file.



7. Upload externally curated network

In the top menu panel of the visualization page, mouse-over the ‘Create’ option opens a drop-down list. Clicking ‘Inter-omics Connections From File’ opens a panel where you can upload externally curated networks to the visualization. Provide a name to the network and select the file and its delimiter and click ‘Upload’



The screenshot shows the MultiSLIDE visualization interface. At the top, there is a horizontal menu bar with buttons for 'SAVE', 'VIEW', 'CREATE' (which is highlighted in red), 'HOME', 'HELP', and 'CLOSE'. Below the menu, there are two sections: 'Datasets' and 'Phenotypes', each with a gear icon. A 'Syncrh' button is shown with a 'Off' switch. The main area has a title 'Empty Feature List' and a sub-section 'Feature List From File'. A red box highlights the 'Inter-omics Connections From File' option in the dropdown menu. Below this, there are two panels: 'Upload Inter-omics Connections' on the left and 'Available Inter-omics Connections' on the right. The 'Upload' panel contains fields for 'Give This Set of Connections a Name' (set to 'kinase_substrate'), 'Choose a File' (set to 'kinase_substrate_network_28082020.txt'), and 'File Delimiter' (set to 'Tab'). It also has an 'UPLOAD' button. The 'Available' panel lists one item: 'kinase_substrate (kinase_substrate_network_28082020.txt)' with a green minus sign icon.

Upload Inter-omics Connections

Give This Set of Connections a Name
kinase_substrate

Choose a File
Choose file kinase_substrate...28082020.txt

File Delimiter
Tab

UPLOAD

Available Inter-omics Connections

kinase_substrate (kinase_substrate_network_28082020.txt)

DONE

7.1 Externally curated network file structure

Network files are delimited text files, with each row in the file specifying one interaction. The file can have up to four columns; the first two columns specify the interacting molecules, while the optional third and fourth columns specify colors and names/sources for the interaction, respectively. The following example network files are available in the GitHub repository: https://github.com/soumitag/multiSLIDE/tree/master/demo_data

Demo	Filename	Has Color and Source?
2	kinase_substrate_network_28082020.txt	No
2	kinase_substrate_network_with_sources_annotated.txt	Yes
3	HS_mirFamilyToTargets_with_filename_headers.txt	No

The figure below shows the kinase substrate network in Demo 2 visualized using the “kinase_substrate_network_with_sources_annotated.txt” file, which has color and source information. Note that the annotations are summarized in the Legends Panel.

multiSLIDE

