Metaboverse Documentation

Release 0.11.0

Jordan A. Berg, et. al.

Jul 24, 2023

Contents

About

Metaboverse allows for the interactive and automated exploration of metabolic pathways and interesting patterns that emerge in the network, integrating transcriptomics, proteomics, metabolomics data and more.

To access the most recent version of our manuscript describing Metaboverse, please click here.

Table of contents

2.1 Overview

2.1.1 Background

Metaboverse is a cross-platform app built to aid users in contextualizing their data on their model's metabolic and global reaction network. Metaboverse is an interactive tool for exploratory data analysis that searches user data in the context of the metabolic network to identify interesting patterns and trends within the data. Metaboverse will aid users in interactively identifying interesting patterns and regulatory hotspots within their data for further experimental follow-up. You can find some of the interesting patterns we have discovered here.

2.1.2 Important Note

If you ever have any questions about using a particular feature, or what a particular button does, try hovering your mouse over the button or feature and often a short explainer will pop out. You can also search the Docs for more in-depth information or refer to the manuscript.

For information on symbols in Metaboverse, please refer to the `Shape Legend` and `Compartment Legend` references where applicable.

2.1.3 Data Inputs

Metaboverse is capable of handling several data types and structures. Users can input **any combination of paired transcriptomics and/or proteomics and/or metabolomics data** for their model. Metaboverse relies on Ensemble IDs, UniProt IDs, and ChEBI IDs for data mapping, so any data type that is able to map back to one of these data types can be used. For example, ribosome profiling translation efficiency data mapped to Ensembl gene IDs can be overlaid on the network. Data format consists of row names with the entities of interest, a column of $\log_2$ Fold Change data, and a column of a statistical value. An example for each datatype can be seen below, where (A) shows

single-condition data table examples, and (B) shows a single-timepoint proteomics dataset paired with a timecourse metabolomics dataset.

A

Transcriptomics

	log2FC	FDR
YHR043C	-2.9822449	1.25E-116
YOR221C	-3.6254778	6.61E-98
YLR303W	8.66431876	4.65E-60
YGL009C	-0.9819979	4.67E-54
YJL102W	-1.0826831	8.00E-46
YPR194C	-1.2207247	7.66E-45
YBL015W	-0.9913832	2.23E-44
YER001W	-1.1994954	1.22E-40
YER073W	-1.1782363	3.52E-39
YFL030W	-1.0443715	3.68E-34

Proteomics

	log2fc	bh_corrected
YHR033W	-1.5961476	0.00024614
GEP7	-1.3827167	0.00102381
PSD1	-0.6435572	0.00074099
MET17	6.33015829	0.00161708
MTG1	-1.3524357	0.00141944
COX13	-3.120406	0.00131153
BI2	-3.4097856	0.00122703
YJR149W	-1.4793181	0.00121765
YDR379C-A	-3.6457662	0.00127318
QCR7	-2.9436792	0.00120242

Metabolomics

	log2FC	p
D-Glucose	1.25862635	0.22716925
Glucose 6-phosphate	-0.4631542	0.24531688
Fructose-6-phosphate	-0.1895111	0.69037804
fructose-1,6-diphosphate	0.93995724	0.24073098
3-Phosphoglyceric acid	-1.1212842	0.12014577
Phosphoenolpyruvate	-1.0643069	0.03462956
Pyruvic acid	-0.2829846	0.46688602
L-Lactic acid	0.28519431	0.53999636
Citric acid	-2.8569398	0.00053647
cis-Aconitic acid	-2.2782394	0.21279282

B

Proteomics

	log2fc	bh_corrected
YHR033W	-1.5961476	0.00024614
GEP7	-1.3827167	0.00102381
PSD1	-0.6435572	0.00074099
MET17	6.33015829	0.00161708
MTG1	-1.3524357	0.00141944
COX13	-3.120406	0.00131153
BI2	-3.4097856	0.00122703
YJR149W	-1.4793181	0.00121765
YDR379C-A	-3.6457662	0.00127318
QCR7	-2.9436792	0.00120242

Metabolomics (Time Course)

	000min_log2FC	000min_p	015min_log2FC	015min_p	030min_log2FC	030min_p
D-Glucose	-1.212184779	0.04292119	1.53766345	0.13476555	-1.153392853	0.20414995
Glucose 6-phosphate	-0.14987541	0.55791308	-0.497647756	0.26324202	-0.15191735	0.75820291
Fructose-6-phosphate	0.187465355	0.58167208	-0.975317718	0.03748708	-0.104703812	0.8539282
fructose-1,6-diphosphate	-0.225557267	0.62759672	-0.576382163	0.24398156	-0.678747253	0.46773149
3-Phosphoglyceric acid	-1.119676775	0.00119412	-1.67109428	0.03376935	-0.645396824	0.3020287
Phosphoenolpyruvate	-0.763107235	0.04028404	-1.019296852	0.13709103	-1.296234656	0.02870063
Pyruvic acid	-1.091087055	0.00046675	-0.515642749	0.5060559	-0.655356778	0.01417618
L-Lactic acid	-0.467199485	0.30418128	0.417014075	0.089489	0.134197827	0.63847317
Citric acid	0.183446037	0.51261193	-1.753735688	0.00921643	-2.75307102	0.00046449
cis-Aconitic acid	-0.246479185	0.76665724	0.18960545	0.89540307	0.743625974	0.6405695

Along with single condition experiment set-ups, users can provide timecourse data by sequentially listing the fold change and statistical columns in temporal order in the input data table for that -omic type. Users can provide all paired -omics types for each time point, or provide one -omics table of timecourse data and a table for another -omics type with a single steady-state time point. In every situation, it is vital the user keeps in mind the caveats associated with comparing such mixed data types. Along with timecourse data, users can provide multiple conditions formatted as with the timecourse data.

The user may decide to select the desired statistical value based on numerous factors, including:

- How stringent of a multiple testing correction procedure they would like to use

- Whether or not multiple testing correction is necessary for an exploratory analysis

- Whether or not they would like to compare confidence intervals of comparison groups. Help with preparing confidence intervals can be found in the `Format Data` option. Metaboverse expectes confidence intervals for each comparison to be formatted as:

```
[
    [0.9,  [[21.5761801, 21.7571531], [24.3988067, 25.6011932]]],
    [0.95, [[21.6215126, 21.7118207], [24.6999961, 25.3000038]]],
    [0.99, [[21.6576415, 21.6756917], [24.9400369, 25.0599630]]]
]
```

This correlates with a list of lists, for the 90%, 95%, and 99% confidence intervals for each control and experimental group.

Users can also select the `Format Data` option. Doing so will launch a new window where an unformatted datatable can be uploaded and the user can select the relevant samples to experimental or control groups for a single 2-condition comparison or multiple comparisons for time-course or multi-condition datasets. A description of the usage of this module can be found in the Data Formatting Help section.

Variables and Data

Input Data

Input any of your omics data (optional)

See an example: (i)

Format dataset: [Launch]

Summary

- Each omic input should include fold change and statistal values.

- You may include **any combination** of corresponding transcriptomics AND/OR proteomics AND/OR metabolomics measurements.

- Timecourse data should be **repeated fold change and statistical values for each timepoint in order** (see example above).

- Multi-condition experiments can be provided as with timecourse data.

- Data names should correspond to Ensembl, UniProt, or ChEBI/KEGG/HMDB IDs and acceptable synonyms.

- Keep an eye out for measurements with weird characters in their names. This can often cause problems with the data mapping.

- Make sure the gene/protein/metabolite column name is blank, as shown in the examples.

> **Warning:** For most consistent behavior, you should only use tab-delimited tables with the suffix `txt` or `tsv`.
>
> For most consistent behavior, you should only symbolize a decimal with a period (.), NOT a comma (,).

> **Warning:** If duplicate entity names are provided within a given transcriptomics, proteomics, or metabolomics data table, Metaboverse will ignore these elements.

2.1.4 Unmapped Data

During network modeling of user data, any user-provided datapoints that are unable to be mapped with be output as a tab-delimited table in the same location as the original input files with the suffix `_unmapped.txt`. Data points could be unmapped for two reasons:

- The data point is not curated in a reaction within the selected organism's reaction network

- The provided data point name or ID is incompatible with the available synonyms for that entity within the network. Available synonyms are compatible: Ensembl gene ID or name; UniProt ID or name; ChEBI, KEGG, HMDB, JCBN, IUPAC, or MetaCyc IDs or names

- Try searching `Reactome + the name of the unmapped species` online to see if the species goes by another name in Reactome, or if it is missing from the organism curation.

2.1.5 Output File Types

There are two output file types associated with Metaboverse:

- `mvdb`: This is the organismal curation file, containing the underlying network structure for the selected organism, as well as important node mapping information for connecting user-provided gene, protein, and metabolite measurements to the appropriate locations in the network.

- `mvrs`: This is the completed organismal curation file, with your gene/protein/metabolite measurements mapped onto the network. You can load this file on the home page within the Metaboverse app to pick up where you left off in your analysis of the data.

2.1.6 Using Metaboverse

Metaboverse first asks users for the organism of interest and an output folder to store all output data and network representations. Afterwards, the user will be directed to provide any -omics data tables, along with some basic metadata, such as experimental setup, timecourse labels, etc. After this input is provided, the network will be modeled with the user's data overlaid. Two files will then be output to the location specified by the user. The first will be a file ending in `mvdb`, which contains the basic information needed to model the organism network without user data. This file can be saved for future usage if the user wishes to remodel data on that organism's network without have to wait for all the necessary source files to download. This file can be provided in leiu of the organism name and file output on the appropriate page.

The second file ends in `mvrs` and contains the final network with the user data and user metadata stored. This file will allow for all downstream visualizations and analyses and can be reloaded on the main page of Metaboverse to quickly re-analyze a dataset. Clicking on the `Session Data` tab in the menu on each page of Metaboverse will display the appropriate metadata for that particular model.

The figure below shows an example of this process.

1) Choose organism and output location

Select the organism you would like to analyze:

Saccharomyces cerevisiae

Click on the folder to select output location:

Select Output File Name

Continue

	000min_mct1_lfc	000min_mct1_bh	015min_mc
Urea	-3.238686904	9.85E-05	-2.2272
L-Valine	1.236272279	0.001305057	0.4466
Putrescine	-1.981283414	7.02E-06	-1.5845
Methylmalonic	0.886017803	0.005095074	0.1653
Glycine	2.002589931	0.000133546	2.5797
Malonic acid	0.149937263	0.729091415	-0.0341
Fumaric acid	2.588709849	3.83E-06	1.0036
L-Alanine	2.322251455	5.26E-09	0.7780
Threonic acid	-0.382539465	0.050386726	-0.0852
D-Malic acid	2.728950429	1.97E-07	2.6067
L-Isoleucine	1.063415577	0.001035193	0.3891
Citric acid	0.183446037	0.438568557	-1.7537
L-Lysine	-2.473820035	0.003217549	-5.3311
Xanthurenic aci	0.905991059	0.163877202	0.3146
L-Glutamic acid	-0.626056458	0.203936688	-2.5790
Orotic acid	-3.96860657	0.000206079	-2.9325
L-Tyrosine	-2.309293197	0.000371556	-2.9632
Ornithine	-6.562107439	0.004785547	-8.0784
Phosphoenolpy	-0.763107235	0.013593318	-1.0192
Thymine	-0.719164516	0.000306792	-0.3706
Adenine	-1.898401034	3.13E-05	-1.3664
L-Aspartic acid	0.850815348	0.014706926	-1.8264
2-Hydroxygluta	2.388661515	0.00204618	2.2464

2) Load data

Load transcriptomics data:

Select File   No file selected

Load proteomics data:

Select File   No file selected

Load metabolomics data:

Select File   No file selected

3) Build network model and layer data

Building Network...

This may take several minutes...

Please check your output directory for any unmapped entities. If you would like these entities included in the analysis, please modify your original data table with the appropriate name and rebuild the network.

If you have a question about a particular feature or button in Metaboverse, try hovering your mouse over the item or the *i* icon to display a short description or check the documentation for more complete information.

100%

Pattern Analysis   Visualize   Perturbation Networks
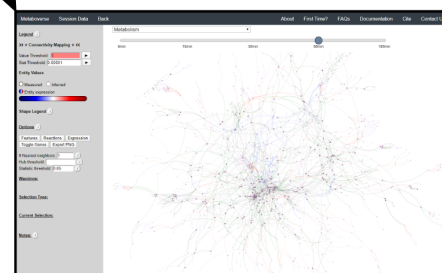
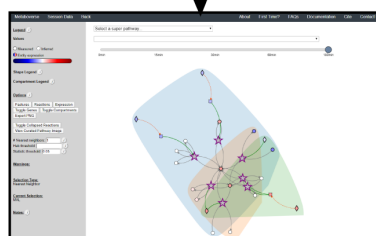4.1) Reaction Pattern Analysis
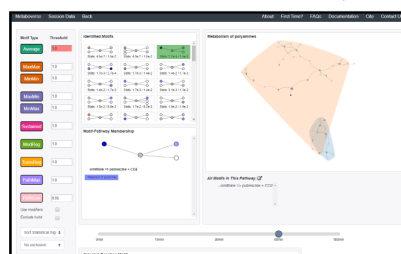
4.2) Pathway   Visualization
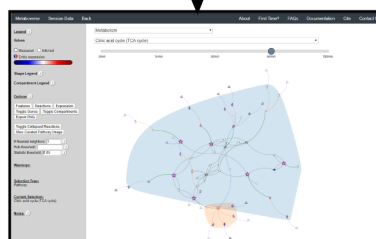
4.3) Perturbation Networks

4) Analysis

Nearest Neighborhood Searches

For more details on using each analytical feature within Metaboverse, please refer to the *Walkthrough* page.

TL;DR

* Users specify organism, output location, -omics files, and basic experimental metadata by following the prompts.

* User data is layered onto the reaction network.

* `mvdb` and `mvrs` files can be saved to quickly access the organism curation or data model. For `v0.3.0b` or earlier, these files ended in `pickle` and `json`, respectively.

* Once data is modeled on the network, the user can begin interactive analysis of their data.

2.1.7 Important Features of Metaboverse

While many tools currently available provide capabilities to analyze data on metabolic networks, we introduce several new or improved features in metabolic network analysis that aid in more robust, unbiased network analysis of biological data. These features are outlined below:

* **Collapsing data sparsity**: Due to the challenges in metabolomics revolving around missing values for metabolites in a reaction pathway, data sparsity can impede in the analytical process. We address this issue by introducing a reaction collapse scheme, that creates pseudo-reactions where up to 3 reactions have missing data points between the reactions, but the external ends of the reactions in this sub-path have measured values.

* **Dynamic Network Visualization**: The Metaboverse provides an interactive app that allows for dynamic exploration of regulatory hotspots, perturbation networks, and general visualization. Users can add or remove metabolites or other entities from being visualized, which is particularly useful in cases of highly connected metabolites, such as water, which clutter the visualization and do not contribute much to the interpretation of the data. Users can toggle on or off labels and features, drag and rearrange components of the networks, and more.

* **Identifying Regulatory Patterns**: In order to identify reactions where interesting regulatory events are occurring based on the provided -omic data, we introduce a regulatory hotspot search engine, which is based on concepts from activity motifs. Several hotspot patterns are pre-programmed into Metaboverse to allow the users to explore different regulatory patterns present in the data. In the future, we plan to include an interactive interface to allow users to design and implement custom search patterns.

* **Modeling Perturbation Networks**: Users can specify the threshold needed to consider a reaction perturbed and stitch together each of these reactions passing the threshold to reveal a perturbation network. This allows for the visualization of connected perturbations along a pathway and may indicate interesting characteristics of a condition that may complicate certain treatments if, for example, a drug targets a perturbed reaction, but the reaction downstream is also perturbed for independent reasons from the first and could thus reduce efficacy of the drug treatment.

* **Exploring global consequences of a change**: In the `Exploration` page, you can double-click on any reaction component and explore all connected reactions to this component across the total reaction network, not just the selected pathway. Additionally, you can select `All entities` from the `Select a super-pathway...` drop-down menu to select any component from the total network for this type of exploration.

2.1.8 Performance

Performance will vary network to network. For example, the curation and modeling of data on a yeast network will process several times faster (5-10min) than data modeled on a human network (20-30min).

For pre-curated reference files (where you do not force a fresh curation), this total process should only take a few minutes, even for a human network, and depends more on your internet conneciton speed.

You may experience a network taking longer to curate, but as long as no error messages appear, it is probably still working. Certain steps that are particular computationally intensive may pause the completion % for a couple minutes, but Metaboverse may still be working.

2.1.9 Technical Description

Metaboverse is currently segmented into two parts:

1. `Metaboverse` interactive app: The platform-independent app for visualizing and exploring data on the metabolic network. [code]

2. `metaboverse-cli`: The network curation and modeling utilities that form the back-end of the Metaboverse app. For each release of Metaboverse, OS-specific binaries are generated of the backend and incorporating into the GUI app. [code]

Metaboverse works by doing the following:

1. Generates an organism-specific reaction network using the Reactome Knowledgebase

2. References the ChEBI and HMDB databases to cross-reference metabolite synonyms.

3. Generates a network-based reaction model

4. Layers user data onto the global reaction network

5. Optionally broadcasts gene expression data to protein expression nodes where protein values are unavailable

6. Prevent the visualization of certain nodes to create simplified visualizations of pathways

7. Runs just-in-time searches of the global network for regulatory patterns of interest centered around a reaction

8. Generates just-in-time visualizations of global or super-pathway-specific perturbation networks

9. Generates just-in-time general visualization of canonical pathways

2.2 Installation

2.2.1 Basic requirements

* 64-bit macOS, Windows, or Linux operating system
* >= 4 GB RAM
* >= 5 GB of open storage space
* Internet connection (for curating step)

2.2.2 Installation Instructions

1) Go to https://github.com/Metaboverse/Metaboverse/releases
2) Click on the most recent release of Metaboverse

3) Under the Assets tab, click the appropriate file for your computer's operating system (`darwin` for macOS, `win32` for Windows, and `linux` for Linux-based operating systems)

4) Unzip the downloaded file

5) Double-click the Metaboverse icon (named `Metaboverse`)

2.2.3 Permissions Notices

In all foreseeable cases, you will likely see some sort of access warning when launching Metaboverse. This is due to the software being unsigned. This basically boils down to us (the developers) not being in the registered databases of Apple- or Windows-certified developers, as this cumulatively can cost several hundred dollars. This obviously comes with some inherent risks, but if Metaboverse sees considerable enough attention and use, we will look into investing in this signing ability. Please reach out to us in the Issues forum for any questions. But for now, please do the following when launching Metaboverse on the following platforms:
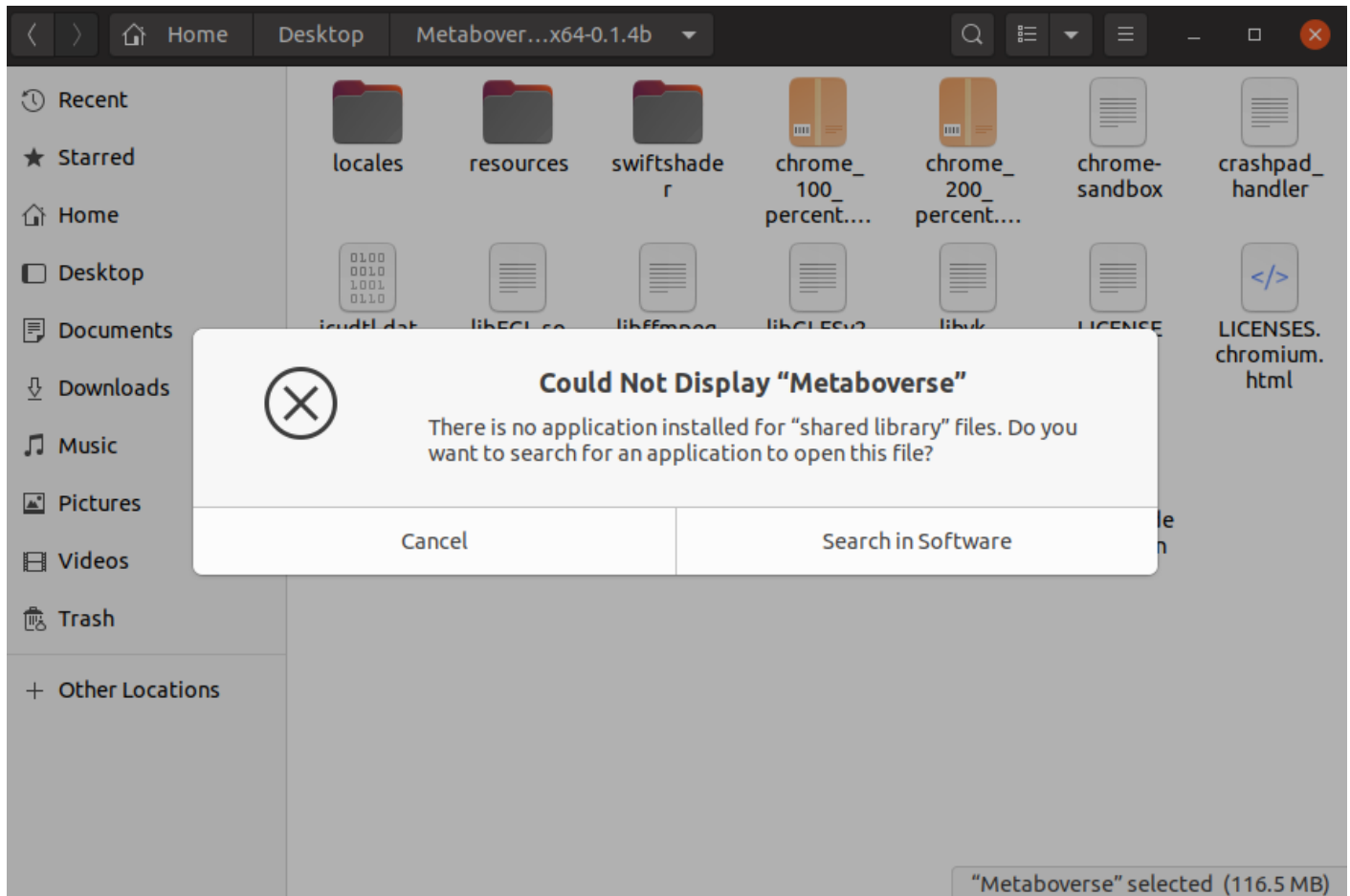
Windows
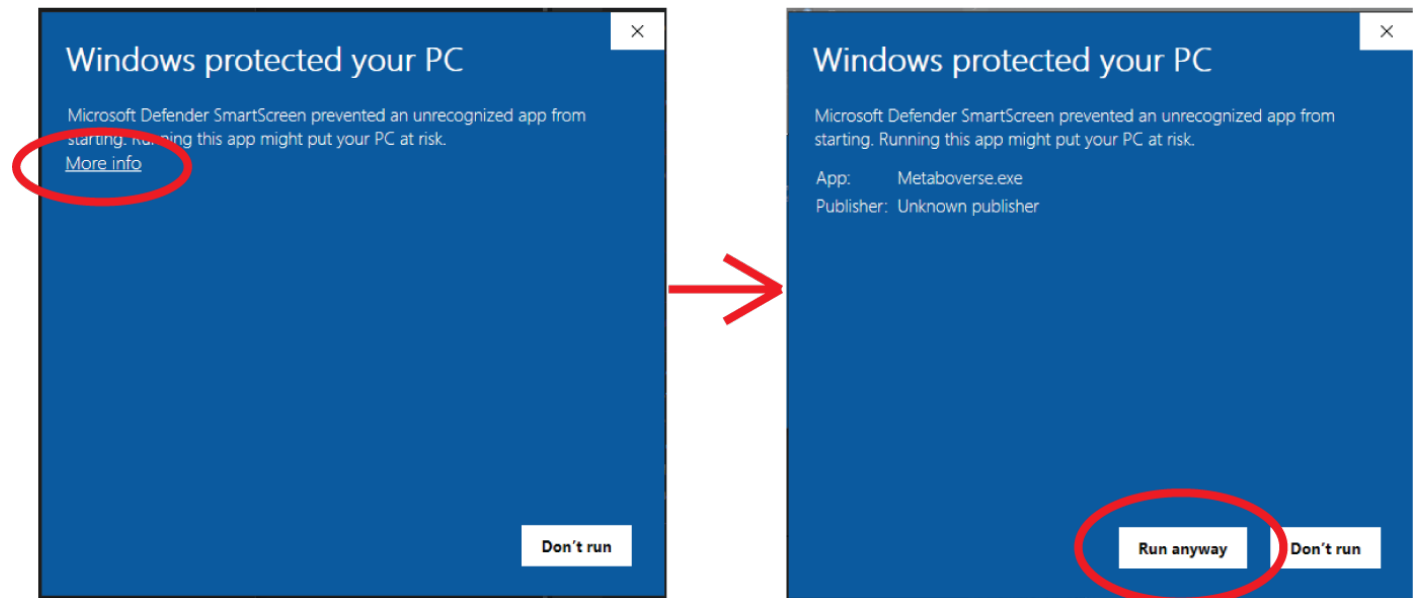
macOS
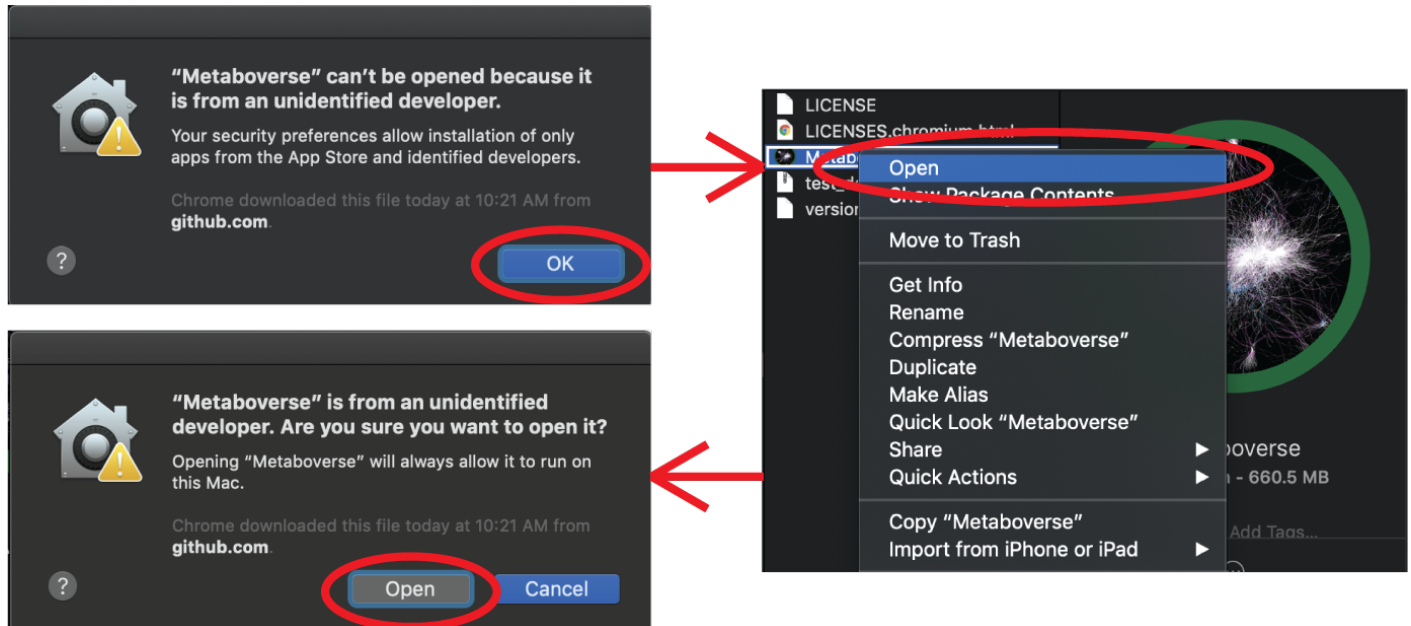
Linux

If you click on the Metaboverse app for Linux and you see the following error:

you should perform the following steps in the Terminal.

```
$ cd /path/to/unzipped/metaboverse/app/folder
$ chmod +x ./Metaboverse
$ chmod +x ./resources/app/python/metaboverse-cli-linux
```

and then launch the app by executing the following:

```
$ ./Metaboverse
```

2.3 Walkthrough

The majority of the following walkthrough was created using Metaboverse `v0.9.0` and previous versions. If using a later version, some aspects may look/behave slightly different. We will do our best to keep these walkthroughs updated as any relevant components change. If you have any questions, please let us know here.

2.3.1 Analysis Video Walkthrough

You can watch the Metaboverse walkthrough video here, or continue below for a more detailed walkthrough.

2.3.2 Downloading Metaboverse

You can access the most recent version of Metaboverse for your operating system at:
https://github.com/Metaboverse/Metaboverse/releases.

2.3.3 Using a Test Dataset in Metaboverse

If you would like to practice using Metaboverse with a test dataset, several are included with each distribution of Metaboverse. You will need to unzip the `test_data` file within the Metaboverse folder you just downloaded and unzipped.

This vignette using the *mct1* Δ timecourse dataset published with the original Metaboverse manuscript. This dataset is included with each distribution of Metaboverse in the `test_data/yeast-multiomics-timecourse` directory.

2.3.4 Getting Started in Metaboverse

To launch the Metaboverse desktop app, open the Metaboverse app within the unzipped folder you downloaded.
1. If this is your first time analyzing a given dataset, you should press `Skip` on the opening page.

2. Select the organism of interest you would like to model your data.
3. Select an output location for files generated by Metaboverse.

4. Provide the following variables for your model.
a) Provide the appropriate data files for your -omics data. Please see the *Data Inputs* sub-heading for more information.
b) Provide a name for your experiment/model.

---

c) Provide the appropriate experimental type. If you are providing timecourse or multi-condition data, a second entry space will open to input sample names.

d) Select these modification options if desired.

  - *Broadcast gene expression to proteins*: Check to broadcast gene expression values to proteins when protein values are not available (checked by default).

  - *Broadcast metabolite values to protein complexes*: Include metabolite measurements in inferring protein complex level.

  - *Consider modifiers in reaction collapsing*: Check to include modifiers in reaction collapsing. Catalysts are included as outputs, inhibitors are included as inputs. Please refer to documentation for more information.

e) Add or remove any items you wish to not be displayed in the network visualizations.

5. Build your network model. After your network model has been built, you will have the option to perform regulatory hotspot identification, model perturbation networks, or perform general visualization of pathways and components by clicking on the appropriate buttons at the bottom of the screen that will appear once Metaboverse has finished processing your data.

2.3.5 Reaction Pattern Analysis

Regulatory patterns can be searched for in the global reaction network using this page. Users can select different reaction pattern types by clicking on the appropriate buttons and adjusting their thresholds as necessary.

Currently, these reaction pattern types are available:

1. **Average**: A reaction whose absolute difference between the average measured values of reactants and the average measured values of products is greater than or equal to the threshold.

2. **Sustained**: Find instances of sustained perturbation along a reaction where there are both an input and an output that are not the same molecule with changes above the selected threshold.

3. **ModReg**: Find instances with one regulated modifier and one core component in the reaction.

4. **TransReg**: Find instances where a component is the same for input and output, is regulated, along with a modifier being regulated.

5. **Enzyme**: Find instances of two neighboring reactions, both with at least one non-metabolite component that passes the specified threshold.

6. **Metabolite**: Find instances of two neighboring reactions, both with at least one metabolite component that passes the specified threshold.

7. **MaxMax**: A reaction whose absolute difference between the maximum measured value of reactants and the maximum measured value of products is greater than or equal to the threshold.

8. **MaxMin**: A reaction whose absolute difference between the maximum measured value of reactants and the minimum measured value of products is greater than or equal to the threshold.

Identified reaction patterns can be sorted for the user using multiple methods:

1. **Sort FDR**: e multiplied by the geometric mean of the relevant p-values to the reaction pattern components are calculated, which roughly leads to a generalized version of a false discovery rate procedure [ref].

---

2. **Sort Statistical Significance**: Either the relevant or maximum reactant, product, and/or modifier p-values are considered in sorting.

3. **Sort Magnitude Change**: Either the relevant or maximum reactant, product, and/or modifier fold changes and the net change across the reaction are considered in sorting.

4. **Sort Number of Pathways**: Reaction patterns annotated in more pathways are prioritized.

Users can then select a given reaction pattern, view the pathways that reaction is present in across the global reaction network, select a pathway to view, and see all other patterns of that type available in that pathway. In the `Reaction-Pathway Membership` box, a simplified glyph of the reaction is drawn, with inputs (substrates) on the left of the reaction node and outputs (products) on the right of the reaction node. Modifiers and input/output types are not shown in this simplified glyph.

If the model includes time-course or multi-condition data, the patterns present at each timepoint or condition can be viewed.

Users can also choose to include the following modifications in the reaction pattern search:

- **Collapse reactions**: Search for reaction patterns using collapsed reaction representations for instances where intermediate reaction components are not measured.

- **Use modifiers**: Include modifiers in reaction pattern analysis. Catalysts are included as outputs, inhibitors are included as inputs. You will need to re-run the reaction pattern analysis to include modifiers.

- **Exclude hubs**: Exclude high-hub components from consideration in the reaction pattern search. This will remove hubs with more than 100 connections. You will need to re-run the reaction pattern analysis to include modifiers.

- **Infer complexes**: Infer protein complex levels using complex component measurements during pattern searching.

- **Statistic threshold**: Define the statistical threshold to bold nodes in visualization. If provided values are confidence intervals, nodes will be bolded if selected confidence intervals values do not overlap.

- **Dropdown 1 - Sorting type**: Select how to sort reaction patterns.

- **Dropdown 2 - Exclusion**: Select a time-point or condition to exclude. In other words, if we were analyzing time-point 1, we could exclude any of the patterns that were also present in time-point 0.

Additionally, users can choose how to sort identified reaction patterns, or choose to not return reaction patterns for the selected time-point or condition that were also found in another time-point or condition. For time-course or multi-condition experiments, a pane will appear that will display the behavior of all reaction components across all time-points or conditions.

For time-course and multi-condition data, a line plot is drawn for a selected reaction with the behavior of that reaction's components across all time-points or conditions.

2.3.6 General Pathway Exploration

While the emphasis of Metaboverse is at the reaction level, where we do not want to bias our analysis towards a specific, familiar pathway, there may be cases where we are interested in a particular pathway and the patterns that emerge. We can access pathway-level visualization by clicking the `Explore` button.

a) At the top of the legend, reference to the node types are displayed. Grey nodes are reactions, nodes with solid outlines are measured, and nodes with dashed outlines are inferred using the gene broadcasting feature. Components with significant statistical values based on the threshold (defined in section d) are bolded. Users can hover across the color scale to see what color corresponds to what value.

Note: Identified reactions with a reaction pattern will have enlarged nodes and will be outlined in bold purple.

b) Relationship types between nodes are shown by hovering over the `Shape Legend` icon. Compartment shadings for the selected pathway are shown by hovering over the `Compartment Legend` icon.

c) Users can toggle component, reaction, and expression labels on and off using the appropriate buttons. Users can also toggle gene nodes and compartment shading on and off with the appropriate buttons. Users can also press the `Export PNG` to select the viewed network to a `png` file. Users can decide whether to show collapsed reactions, or to show the full, non-collapsed pathway of reactions that would have been collapsed. Collapsed reactions are displayed by default. Users can also open the selected pathway in Reactome to view a more classical representation of the pathway.

d) Users can modify the number of neighbors to plot when double-clicking on a node to expand its nearest reaction neighbors. Users can also select a threshold to not display any node with more than the defined number of connections. By modifying the statistic threshold, users can change the minimum statistical value needed to bold the component node for easy discrimination of significant measurements in the network.

e) Metadata about pathways, reactions, and components, along with warnings or errors, will be displayed in this section of the legend.

f) Users can select from the following super-menus:

- *All entities*: A list of all components (metabolites, proteins, etc.). This will draw the nearest neighborhood graph for the selected component.

- *All pathways*: A list of all pathways across all super-pathways.

- *Other*: A list of selected super-pathways that will group pathways in the next drop-down menu. For example, if selecting `Metabolism`, only pathways related to Metabolism will be shown in the next menu.

Users can easily toggle between timepoints or conditions by moving the slider.

2.3.7 Perturbation Network Modeling

Users can explore the role and extent of perturbation within the reaction network in their model by navigating to the appropriate page. Users can modify the perturbation threshold for the fold change or statistical value.

2.3.8 Investigating Upstream and Downstream Effects

Users interested in exploring a components reaction neighborhood can double-click the component to expand this view. The number of reaction neighbors can be selected, and limits on the number of connections a connecting component can have to be displayed can be modified. This hub selection is particularly helpful for connected components that are involved in several reactions.

Users can also target an entity for exploration directly by selecting `All entities` in the `Explore` page from the `Select a super-pathway...` drop-down menu.

2.3.9 Reaction Collapsing

Missing quantifications along the metabolic network are frequent with metabolomics data (Do, 2018; Chen, 2021). As Metaboverse is designed to identify patterns across reactions, missing data can greatly inhibit the ability to identify these patterns. Within Metaboverse, we implemented a reaction collapsing algorithm that can bridge multiple (up to 3) subsequent reactions with missing data at intermediate steps.

Exhaustive Reaction Collapsing

Partial Reaction Collapsing
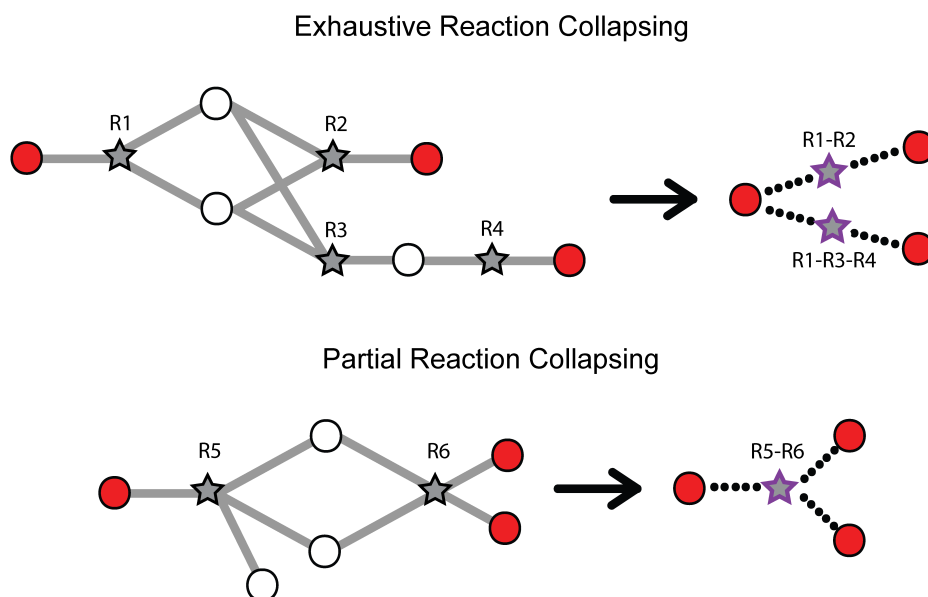
Within Metaboverse, you can toggle back and forth between networks with or without collapsed reaction representations.

2.3.10 Analyzing Complexes

During reaction pattern analysis, protein complexes consisting of multiple elements are aggregated to generate a pseudo fold change and statistical value by calculating the median of all magnitude changes and Euhler's constant (**e**) x the geometric mean of the component statistical values. If the resulting aggregated statistical value is greater than 1, 1 is substituted.

We aimed to take a flexible approach that biases the representative fold change and statistical value towards the most frequent, so if most complex components are statistically significant with high fold changes, the resulting values will be so as well.

By unchecking the `Infer complexes` box on the Reaction Pattern page, you can alternatively analyze complex components individually within that reaction, where Metaboverse will find the optimal element that will give the best magnitude difference to compare against the selected threshold.

2.3.11 Analyzing Previous Models

Users can load a previously modeled network with their data on it by loading the `mvrs` file output by Metaboverse when originally generated. This is done on the Home page.

Users can load a previously modeled organism network by loading the `mvdb` file output by Metaboverse when originally generated. This is done on the Curation page.

2.3.12 Publishing Data Analyzed with Metaboverse

When publishing analyses that used Metaboverse, we recommend attaching the appropriate `mvrs` file that contains the network with your data overlaid as a supplementary file. This will allow for others to easily reproduce and explore your data.

Users can explore the metadata related to their Metaboverse model by clicking on the `Session Data` tab in the menu once the `mvrs` file is loaded.

You can find more information on citing this software on the *Citing Metaboverse* page.

2.4 Other Features

2.4.1 Data Formatting Help

1. On the `Variables` page of Metaboverse, launch the `Format dataset` module.

2. Upload your datatable. The datatable's first column should be a blank cell, followed by each of the measured entities' names. Each subsequent column should start with the sample name, followed by the corresponding measurements for each measured entity.

3. By default, calculated p-values will use the Benjamini-Hochberg p-value adjustment procedure for multiple hypothesis testing. This is a less conservative adjustment procedure ideal for exploratory data analysis. If you do not wish to use a p-value correction procedure, uncheck the appropriate box.

Note: The statistical procedures used by this module assume data are normally distributed, such as is often the case with proteomics and metabolomics data. However, for transcriptomics data, which follow a negative binomial distribution, a package, such as DESeq2 or limma should be used. The resulting fold change and adjusted p-values should then be isolated from the results, exported into a tab-delimited file, and uploaded for use in Metaboverse. This formatting tool will display the distributions of each of the samples in the uploaded dataset to assist in verifying the distributions of the underlying data.

> **Warning:** This Format Dataset tool from Metaboverse v0.10.1 onward will display the sample distributions within the uploaded datasets. Users should examine the samples to ensure their values follow normal distributions. If the data are not normally distributed, the user will need to utilize other statistical testing in preparing their dataset for Metaboverse.

4. Select the experiment type used.

5. Provide a label for the comparison, or use the default name, and select the contol and experimental samples from the columns and assign to their appropriate group by clicking the group button (`Control` or `Experiment`).

6. For multi-condition or time-course experiments, continue to add additional groups.

7. For metabolomics data, click the `Check Names` button to cross-reference the names you provided with MetaboAnalyst to improve the chances that the metabolite correctly maps to the metabolic network.

8. Export your formatted datatable.

2.4.2 Tracing Metabolomics Data

Currently, Metaboverse does not contain any integrated methods for handling tracing/flux metabolomics data. We chose to do so for a couple of reasons:

1. Coupling tracing metabolomics with reaction pattern searches, reaction collapses, and the other features of Metaboverse would add more dimensionality that would reduce the effectiveness of what Metaboverse has to offer.

2. Tracing metabolomics can be difficult to automate pattern searches across. For example, M+4 vs M+5 citrate could imply drastically different metabolic outcomes and is better suited for manual analysis.

3. Escher-Trace is a publicly available visualization tool that already provides the capabilities to analyze tracing data.

We strongly suggest users interested in analyzing tracing data in conjunction with Metaboverse to check out Escher-Trace. For such an analysis, users might consider analyzing the different

ionization products with Escher-Trace and analyzing derived steady-state metabolomics data (i.e. M+0) with Metaboverse's reaction pattern search engine. Cross-referencing the outputs of these two tools may then provide biological clues for their system, such as to the downstream outcomes of differential metabolite behavior.

2.5 Data Vignettes

Test datasets are included with each release of Metaboverse. Once you have opened the executable, unzip the `test_data.zip` file and explore the provided datasets. Each sub-directory is explained below.

2.5.1 `test_data/human-metabolomics`

Source: Metabolomic Markers of Altered Nucleotide Metabolism in Early Stage Adenocarcinoma (Wikoff, et al., Cancer Prev Res (Phila), 2015).

Contents:
- `lung_tumor_vs_normal_measurements.txt`: GC-TOF metabolomics for paired tumor and normal lung tissue of the LUAD dataset (Wikoff, 2015) as used in the Metaboverse manuscript vignette. This dataset provides log2 Fold Change and Benjamini-Hochberg adjusted p-values for each measured entity at each time-point, and should be provided on the VARIABLES AND DATA page of Metaboverse.

1) On the Curation page of Metaboverse, select *Homo sapiens* as the model organism.
2) On the VARIABLES AND DATA page of Metaboverse, you would specify the experiment type as "Default (2-condition)".

2.5.2 `test_data/mouse-multiomics`

Source: STATegra, a comprehensive multi-omics dataset of B-cell differentiation in mouse (Gomez-Cabrero, et al., Sci Data, 2019).

Contents:
- `proteomics_formatted.txt`: Proteomics for the timecourse of B-cell differentiation in mouse. This dataset provides log2 Fold Change and Benjamini-Hochberg adjusted p-values for each measured entity at each time-point, and should be provided on the VARIABLES AND DATA page of Metaboverse.
- `metabolomics_combined.txt`: A combination of GC- and LC-MS metabolomics data for the timecourse of B-cell differentiation in mouse. This dataset provides log2 Fold Change and Benjamini-Hochberg adjusted p-values for each measured entity at each time-point, and should be provided on the VARIABLES AND DATA page of Metaboverse.

1) On the CURATION page of Metaboverse, select Mus musculus as the model organism.
2) On the VARIABLES AND DATA page of Metaboverse, you would specify the experiment type as "Time-Course", and would provide labels such as "0hr, 2hr, 6hr, 12hr, 18hr, 24hr".

2.5.3 `test_data/fish-proteomics`

Source: Comprehensive and quantitative proteomic analyses of zebrafish plasma reveals conserved protein profiles between genders and between zebrafish and human. (Li, et al., Sci Rep, 2016).

Contents:
- `proteomics_male_female_log.txt`: Proteomics for male vs female zebrafish (Li, 2016). This dataset provides the log10 transformed quantifications for each biological replicate between the two measured conditions. These data should be used as practice for the "Format dataset" tool on the VARIABLES AND DATA page of Metaboverse.

1) On the Curation page of Metaboverse, select *Danio rerio* as the model organism.

2) On the VARIABLES AND DATA page of Metaboverse, you would launch the "Format dataset" tool and upload the file to generate the log2 Fold Change and p-values for your comparison conditions.

3) On the VARIABLES AND DATA page of Metaboverse, you would specify the experiment type as "Default (2-condition)".

2.5.4 `test_data/yeast-multiomics-timecourse`

Source: Metaboverse: Automated discovery and visualization of diverse metabolic regulatory patterns (Berg, et al., bioRxiv, 2020).

Contents:
- `sce_mct1_03hr_counts_diffx.txt`: DESeq2-processed RNA-seq data for the 3 hr time-point of the MCT1 dataset as used in the Metaboverse manuscript vignette. This single time-point provides log2 Fold Change and FDR values for each measured entity, and should be provided on the VARIABLES AND DATA page of Metaboverse.
- `metabolomics_timecourse_mct1.txt`: GC-MS metabolomics for the 0min, 15min, 30min, 60min, and 180min time-points of the MCT1 dataset as used in the Metaboverse manuscript vignette. This time-series provides log2 Fold Change and Benjamini-Hochberg adjusted p-values for each measured entity at each time-point, and should be provided on the VARIABLES AND DATA page of Metaboverse.

1) On the CURATION page of Metaboverse, select *Saccharomyces cerevisiae* as the model organism.

2) On the VARIABLES AND DATA page of Metaboverse, you would also specify the experiment type as "Time-Course", and would provide labels such as "0min, 15min, 30min, 60min, 180min".

2.5.5 `test_data/yeast-multiomics-singletimepoint`

Source: Metaboverse: Automated discovery and visualization of diverse metabolic regulatory patterns (Berg, et al., bioRxiv, 2020).

Contents:
- `sce_mct1_12hr_counts_diffx.txt`: DESeq2-processed RNA-seq data for the 12 hr time-point of the MCT1 dataset as used in the Metaboverse manuscript vignette. This single time-point provides log2 Fold Change and FDR values for each measured entity, and should be provided on the VARIABLES AND DATA page of Metaboverse.
- `proteomics_mct1_12hr.txt`: Quantitative proteomics for the 12 hr time-point of the MCT1 dataset as used in the Metaboverse manuscript vignette. This single time-point provides log2 Fold Change and Benjamini-Hochberg adjusted p-values for each measured entity, and should be provided on the VARIABLES AND DATA page of Metaboverse.

-`mct1_12hr_metabolomics.txt`: LC-MS metabolomics for the 12 hr time-point of the MCT1 dataset as used in the Metaboverse manuscript vignette. This single time-point provides log2 Fold Change and Benjamini-Hochberg adjusted p-values for each measured entity at each time-point, and should be provided on the VARIABLES AND DATA page of Metaboverse.

1) On the CURATION page of Metaboverse, select *Saccharomyces cerevisiae* as the model organism.
2) On the VARIABLES AND DATA page of Metaboverse, you would also specify the experiment type as "Default (2-condition)".

2.6 FAQs

If you have questions, requests, or bugs to report, please use the Metaboverse issues forum.

2.6.1 Reporting problems during model building

If you encounter an error during the build of your model, you should see a notification alerting you that a error log has been output. This file will be output to your selected `output` folder. Error logs are output to ensure all relevant details are visible to the user and developers. You may read through the log to see if there is a solution you can find, or you can submit an issue to us at the Metaboverse issues forum and attach the log file there for us to review.

2.6.2 Model building not showing any progress

Most likely, if Metaboverse does not show any progress on the `Build` page, this is due to a permissions issue. This can often be remedied by running Metaboverse as an administrator. Do so by right-clicking on the Metaboverse app icon and select `Run as administrator`. If a error log is output, please submit it to the Metaboverse issues forum.

2.6.3 Model building freezes

A frozen progress bar during model building is often due to internet connection issues. Please check you internet connection and connection speed in case this is the issue. If your internet connection seems fine, please submit an issue to the Metaboverse issues forum.

2.6.4 Not accepting a particular input

Try substituting spaces with underscores in input fields. Please also attach any errors or logs you receive when this happens and report on the Metaboverse issues forum.

2.6.5 Linux app not launching

If you click on the Metaboverse app for Linux and you see the following error:
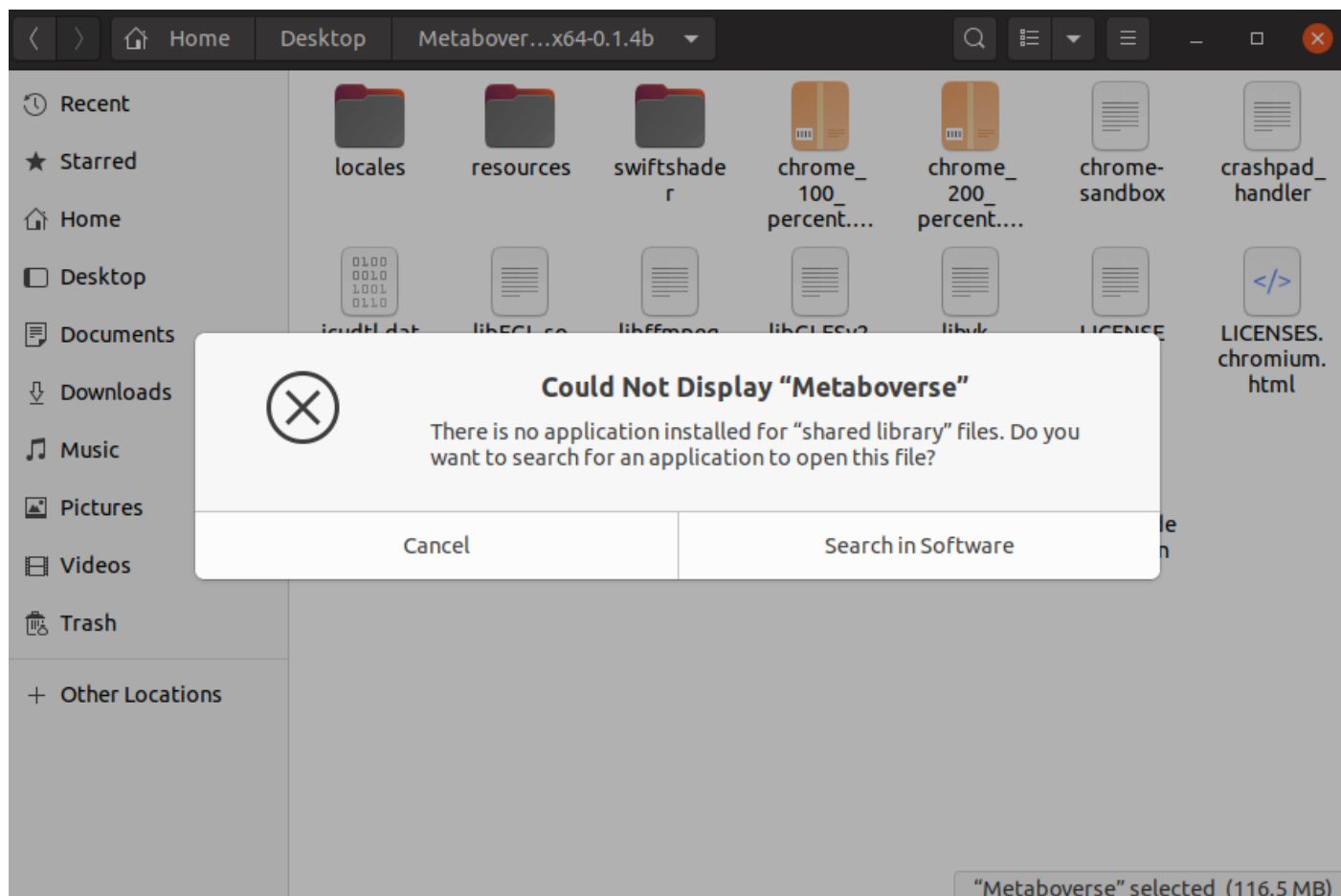
you should perform the following steps in the Terminal.

```
$ cd /path/to/unzipped/metaboverse/app/folder
$ chmod +x ./Metaboverse
$ chmod +x ./resources/app/python/metaboverse-cli-linux
```

and then launch the app by executing the following:

```
$ ./Metaboverse
```

2.6.6 Error: `AttributeError: 'float' object has no attribute 'lstrip'`

If the error, `AttributeError: 'float' object has no attribute 'lstrip'`, appears this is likely caused by missing values in one of the provided data tables. While current procedures should now handle this issue, if this error persists, you might try deleting any trailing whitespace from the gene/protein/metabolite names in your data tables.

2.7 Citing Metaboverse

2.7.1 Metaboverse

When reporting results obtained using Metaboverse, please include the following language or similar in your text:
"Network regulatory analysis was performed using Metaboverse v#.#.# [ref]."

[ref] Berg JA, Zhou Y, Ouyang Y, Cluntun AA, Waller TC, Conway ME, Nowinski SM, Van Ry T, George I,
       Cox JE, Wang B, Rutter J.
       Metaboverse enables automated discovery and visualization of diverse metabolic regulatory patterns.
       Nat Cell Biol. (2023) DOI: 10.1038/s41556-023-01117-9.

2.7.2 Reactome

Please include the following citations when referencing Reactome:
Bijay et al. 2020 PMID: 31691815.
Fabregat et al. 2018 PMID: 29145629.

2.8 Updates

2.8.1 v0.11.0

- *#132 <https://github.com/Metaboverse/Metaboverse/issues/132>'_*: Fixes issue with p-value/FDR calculation in interactive datatable formatting module
- *#134 <https://github.com/Metaboverse/Metaboverse/issues/134>'_, #136 <https://github.com/Metaboverse/Metaboverse/issues/136>'_*: Moves curated files to Sourceforge, adds automated scripts for release building
- *#128 <https://github.com/Metaboverse/Metaboverse/issues/128>'_, #129 <https://github.com/Metaboverse/Metaboverse/issues/129>'_, #130 <https://github.com/Metaboverse/Metaboverse/issues/130>'_, #131 <https://github.com/Metaboverse/Metaboverse/issues/131>'_*: Returns more information for values not being mapped to networks
- Updates to front-end calls to improve performance

We are trying to get a better sense of how we can best continue to develop Metaboverse for the community. If you are using Metaboverse, we would appreciate it if you took a few moments to fill out this survey:
https://forms.gle/YdM1caZ7NhLAjxCGA. Thank you!

2.8.2 v0.10.1

- Updated license to MIT
- Fixed version check when launching GUI to make sure it sorts version numbers correctly
- Updated documentation (expanded walkthrough, added dev notes)
- Added more test datasets distributed with each release
- Added sample distribution display for Data Formatting tool
- Fixes issue where collapsed reactions in other timepoints were not being highlighted

2.8.3 v0.10.0

- Introduces **Co-factor Selection**, a drop-down menu available in Pattern Analysis. Selected a specific factor will limit the Pattern Analysis results to only the reactions containing the selected entity.

- Implements a Pattern Analysis **Table Export** utility, whereby the user can export the Reaction Patterns of a given type.

- Fixes an issue where some non-collapsed reactions were not showing up in Pattern results.

2.8.4 v0.9.0

Note: Files curated and analyzed using Metaboverse v0.9.0 will not be backwards compatible with files generated using earlier versions of Metaboverse.

Major

- Integration of confidence values during data formatting, curation, and analysis.

- Fixed issues during curation arising from updates to the formatting of Reactome source files.

Minor

- Streamlined option buttons available during `Pattern Analysis`.

- Force directed networks now have sticky nodes once a user a dragged them to a new position.

- Added statistical thresholding visualization to `Pattern Analysis` module.

- Output session file with each successful build - will have experiment name appended to file.

- Sort `Average` reaction pattern by best statistic on each side of reaction

- Fix reaction filtering to more strictly and flexibly remove transport reactions from reaction pattern categories, and to only display transport reactions in the `TransReg` reaction pattern category.

- Node labels will default to the user-provided names if not the same as the curated default name.

2.8.5 v0.8.0

Major

- Added data formatting aid page (see issue #86).

- Added ability to cross-reference metabolite names with MetaboAnalyst and find more compatible names (see issue #74).

Minor

- Fixed issue where Metaboverse would not populate selected organism from drop-down if previously selected and the user returns to the page (see commit).

- Fixed issue where Metaboverse would not show the `Continue` button on the `Curation` page if necessary inputs were previously selected and the user returns to the page (see commit).

2.8.6 v0.7.1

Important Note

Many of the changes introduced in `v0.7.0` to session and intermediate file metadata will likely be incompatible with previous versions of Metaboverse.

Minor

- Fix Session Data page to format variables, file paths better (see commit)

- More flexible blocklist to capture all components with the same name, even if they have different species IDs (see commit1 ; commit2)

- Find largest change possible for modifier regulation patterns (see commit)

- Sort by best p-value (previously had taken a more conservative approach by using the worst p-value on each side of the reaction) (see commit)

- Add button and capabilities to switch between inferred complex values or to compare each complex component individually within the reaction pattern (commit1 ; commit2)

> - Will still display the complex as inferred value, but evaluated as each individual component during reaction pattern search

- Protein complex inference/aggregation

> - mean -> median for generating aggregate magnitude value for protein complex from component parts (see commit)
> - Aggregate statistic calculated using `e * gmean(p-array)` (see commit)
> - Max aggregate p-value set to 1 (see commit)

- Allow exporting line plots for timecourse and multi-condition datasets (see issue #89)

- Use user-provided names in labeling (see issue #87)

- Toggle analyte labels on by default (see commit)

- Allow flexibility for input data where commas are used in place of decimals (see issue #92)

- Remove duplicate rows from input data (interactive input will warn about these) (see commit)

- Fix **Sustained** reaction pattern to not identify is input and output value being compared are identical (see commit)

- Fix issue with `parseComponents()` function where usage of modifiers in pattern determination was pushing all modifiers (catalysts and inhibitors) to reactants list (see commit)

- For upregulated sustained reactions, get max of inputs and outputs (previously was getting min) (see commit)

- Fixed global motif searching for pathway and perturbation visualization to search non-collapsed reaction dictionary too. (see commit)

- Added reaction pattern tooltips on button to show a graphical example of each (see commit)

- Move some shared utilities to the *motif-utils.js* file (see commit)

2.8.7 v0.7.0

Important Note

Many of the changes introduced in `v0.7.0` to session and intermediate file metadata will likely be incompatible with previous versions of Metaboverse.

Major

- `Enzyme`/`Metabolite` reaction patterns added: The `Enzyme` reaction pattern evaluates for two neighboring reactions both with perturbed enzymes matching the given threshold. This will allow for better pattern identification, especially in RNA-seq/proteomics-only datasets. The `Metabolite` reaction pattern looks for neighboring reactions both with perturbed metabolites matching the given threshold (see issue #81).

Minor

- Removed pathway-specific pattern detection: In our testing, these seemed to be minimally helpful.

- Option added for users to define percentage of matching nodes between two reaction to allow for a collapse (see

issue #82).

- SVG export option (Full support for Inkscape, partial support for Illustrator) (see issue #83).

- Improved and more explicit session data for all intermediate files (see issue #78).

- Migrated source files to rutter.chpc.utah.edu/Metaboverse/source. This change should allow for faster downloads of pre-curated intermediate source files (see issue #80).

- Loading icon in reaction pattern page to let user know patterns are loading, especially in cases where many reaction patterns are discovered and the software may take some time to display them all.

2.8.8 Previous versions

v0.6.0

Major

- New database integration: First supported release with the ability to overlay data on BiGG and BioModels network models and enable reaction pattern searching across a wider array of organisms. Note: Network models from these sources can be less uniform as Reactome sources, so users should exercise some caution when using these capabilities and perform some sanity checks (see issue #73).

- kNN visualization improved to allow for more stable NN building without error (see commit 2395cd6).

- Neighbors dictionary backend curation is simplified and sped up (see commit 355abd4).

- Improved security policies. Specifically, external websites are opened in an isolated browser window and explicitly are context isolated and unable to access node integration. Enabled GitHub and Reactome URLs (`connect-src`) are more specific (see commit 96b1c9f).

Minor

- Progress bar during network build now update with more incremental steps for longer processes (see issue #77).

- New variables for more unified backend processing. Metaboverse v0.6.0 and later will not be compatible with files curated using Metaboverse v0.5.0b or earlier.

v0.5.0-beta

Major

- Addresses issue #66 , by hosting curated reference and template files for each organism per version of Metaboverse. Also provided user argument options to directly include already downloaded or curated files (`--organism_curation_file`, `--neighbor_dictionary_file`, `--graph_template_file`). Using pre-downloaded files, this reduced processing time for curating data on the human network from ~30-40 min to ~2 min. These pre-curated files will be processed with each release of Metaboverse and are hosted on SourceForge currently.

Minor

- Fixes an issue where an empty unmapped dataframe would cause an error (fixed by #32e9283)

- Fixes an issue the working path would be appended to the organism ID (fixed by #91a490d)

- Addresses issue #67 , where experiment name inputs with spaces would cause an error.

- Bump required version of Electron to >=9.4.0 (see pull request #68).

- Removed some unused user arguments from command-line interface.

- Fixed an issue where the backend argument parse checker would try to append a file path to the organism ID.

- Fixes internal warning for UI when CLI did not output blocklist or labels.

- Updated copyright info.

- Removed deprecated files.

- Migrated from Travis-CI to GitHub Actions.

v0.4.0-beta

Major

- Partial collapse: Addresses #51 , which introduces partial collapsing to the reaction collapsing utility within Metaboverse. Previously, perfect matches were required between two reactions to collapse the reactions. However, this can be overly stringent in key metabolic pathways where a metabolite that is output by one reaction may not be required for the subsequent reaction (perhaps ATP is produced by reaction A but is not required for reaction B). To perform a partial collapse, Metaboverse operates by largely the same scheme as before, but now checks for a perfect match from each neighboring reaction, and if a perfect match is not available, checks for partial matches by filtering out high-degree nodes (quartile 98 of all non-reaction node degrees) and then checking if at least 30% of the nodes match with its neighbor.

- Improvements to nearest neighbor searches where all iterations of a species are included in the graphing. Previously, it would only use the literally selected node to search for neighbors, but Reactome provides separate species IDs for a metabolite's different organelle-localizations, which was complicating these searches.

Minor

- Displays a preview of the user-selected reaction in an interactive format on the Pattern Search Analysis page until the user selects a Pathway to visualize. If a reaction is collapsed and spans two pathways, no pathways will be shown and instead a note that the reaction spans two pathways is displayed.

- The Pattern Search Analysis page now allows users to filter out collapsed reactions from the search results. By default, collapsed reactions will be displayed until the checkbox is unchecked by the user.

- Minor updates to Pattern Search Analysis page to make better usage of blank space

- Fixes #60 , where the `mvrs` file extension would not be automatically added to the user-provided output file name in Linux.

- Addresses #62 , where the some time-course/multi-condition slider bars would be improperly formatted.

v0.3.3-beta

Minor

- Closes #63 by applying `safestr()`` function to all user input encodings to make sure no errors arise.

v0.3.2-beta

Minor

- Closes #59 where non-ascii characters in reaction names would break the info extraction. Added a safestring conversion utility to prevent ascii-character issues.

v0.3.1-beta

Minor

- Fixes path separator for motif page name identification to allow for including modifiers in motif ID and exclusion of hubs (#55)

- Fixes CHEBI mapping so that CHEBI IDs provided as input data are more reliably used as mapping IDs if it cannot match the metabolite by name (#58)

- Fixes issue that arose in `v0.3.0b` where some motif stamps could not be clicked on for viewing for timecourse/multi-condition data where it could not identify the shape for an unknown component type (#54)

- Addresses #59 where non-ascii characters in reaction names would break the info extraction. Was not able to recapitulate the error, but this fix, where relevant reaction metadata is forced to a string data-type, should allow for some flexibility here.

- Updates walkthroughs and documentation to address (#31) and update formatting

v0.3.0-beta

Major

- Allows for more flexible gene/protein mapping with Reactome-formatted node names. For example, Reactome will label a gene or protein with its isomer coordinates. Metaboverse now ignores those coordinates during attribute mapping of the user's data.

- New naming of modules: `Motif Search` is now called `Pattern Analyis`, `Visualize` is now called `Explore`, and `Connectivity` is now called `Perturbation Networks`. Changed to be more descriptive and accessible to all users from broader backgrounds

Minor

- Fixed nearest neighbors capabilities in Perturbation Network visualization. In a previous release, a change had interfered with its function.

- Use of outdated version will now direct user to the download page for the most recent version

- The `Back` button from any of the analysis modules will now redirect back to the index page

- Removed compartment visualization from the `Perturbation Networks` page as these often would just clutter the visualization and would not actually be helpful

- Fixed reaction node formatting to turn off motif symbols/highlighting when timepoint/condition changed as this had been disabled by a previous change

- Fixed collapsed reactions to ensure they were included in all reaction and motif formatting events

- General formatting changes

- Updated documentation

v0.2.0-beta

Major

- Fixes issues with missing metabolites during the network mapping stage (#37). Addressed by re-working the metabolite synonym scheme to provide the same coverage of synonym look-up as before, but with more robustness so that some that were difficult to map would start mapping (i.e., Fructose)

- Added dynamic line-plots of a selected reaction motif for time-course and multi-condition data (#15). When exploring motifs on the Motif page for time-course and multi-condition experiments, a new panel appears at the bottom of the page which, for a selected motif, will show those reaction motif's component's behavior across all the time-points or conditions.

- Added the option to exclude reaction motifs at a given time-point or condition that appear in another selected time-point or condition. (#16). For example, if a user has selected to view motifs at a terminal time-point, but they want to know which reactions are motifs at this time-point but not at the initial time-point, they can exclude the motifs that show up at both time-points using the appropriate drop-down menu on the motif page.

Minor

- Metaboverse now outputs a table of unmapped metabolites (#35).

- Exploration pages now have pop-out bubbles with all information for compartments and node/link types 7d17d34.

- Metaboverse new remembers and provides session info for experiment name, experiment type, labels, etc. and automatically fills those out for the user if returning to a page within the session 172d21a.
- Updated minor page formatting to make display more stable between Windows/Linux/Mac 52a100d.
- Added test cases to CI for new/updated features
- Updated package dependency information
- Updated docs and FAQs

v0.1.4-beta

- Fixes #26, where an error log is output if build fails
- Removes direct Matplotlib imports in metaboverse-cli modules to prevent unnecessary errors and incompatibilities

v0.1.3-beta

- Fixes bug where user paths with spaces were unable to be used (#26)

v0.1.2-beta

- Fixes bug that prevented the curation from running without a blocklist (#19)
- Fixes bug during data mapping that caused protein or gene values to occasionally map to metabolites (#20)

v0.1.1-beta

- Fixes minor run-time issues with the Metaboverse interactive app
- Fixes version alert to let users know if there is a newer version of Metaboverse available

v0.1.0-beta

Initial beta release

v0.0.1-beta

Demo pre-release with included human network data file for network visualization and exploration. Currently only available for MacOS.

How to run:

1. Download attached `zip` demo file.
2. Double-click on `zip` file to uncompress Metaboverse and the accompanying test file
3. Within the uncompressed folder, right-click on Metaboverse to launch the app
4. Drag and drop the `json` file to the appropriate load icon and click the Visualize button.

2.9 Backend Development

Note: This page includes information about the underlying scripts of Metaboverse and Metaboverse-cli and how to contribute to and modify the software.

2.9.1 Expectations

It is essential before getting started that you read through the Contributing and Code of Conduct documents. Harassment, bullying, or other inappropriate behavior of any form will absolutely not be tolerated. Those violating these standards will be banned from contributing to the Metaboverse project.

2.9.2 Communication

We highly recommend making use of the Issues, Discussions, and Projects pages to communicate developments, issues, questions and more.

2.9.3 First Steps

The code base for Metaboverse is currently housed in two primary Github repositories. metaboverse-cli handles the backend processing of metabolic networks and preparing user data for analysis, and Metaboverse handles front-end processing and visualization of user data.

If you would like to modify either of these repositories, you should be begin by forking a branch of the repository. You should reference the documentation: https://docs.github.com/en/get-started/quickstart/fork-a-repo.

Once you have made the desired changes, you can submit a Pull Request, which will be reviewed and incorporated into the code base for future deployment if the changes meet the required standards. You should reference the documentation: https://docs.github.com/en/pull-requests.

2.9.4 `metaboverse-cli`

`metaboverse-cli` handles a large portion of the backend processing, primarily to curate the metabolic network for the organism of interest and to layer the user's data onto said network. The source code is generally organized as follows:

```
metaboverse-cli
├── setup.py:              The command line tool setup file
├── requirements.txt:      The depencency file used for distribution
├── metaboverse-cli.spec:  The Pyinstaller file to compile the distributable␣
↪executable
└── /metaboverse_cli
    ├── arguments.py:      A file to set and parse user-provided command line␣
↪arguments
    ├── __main__.py:       The main execution script for the backend
    ├── /mapper:           This submodule generates the metabolite synonym mapper␣
↪that assists
    │                         in mapping user data to the metabolic network
```
(continues on next page)

```
    ├── /curate:              This submodule builds the organism specific reaction␣
→network and
    │                         stores any relevant metadata about the reaction␣
→network component
    └── /analyze:             This submodule handles layering the user's data onto the␣
→metabolic
                              network. This includes implementing the metabolite␣
→synonym mappper.
                              This submodule also handles generating the reaction␣
→network with
                              reaction collapsing
```

2.9.5 `Metaboverse`

`Metaboverse` handles the front-end user interface and data visualization, and performed reaction pattern recognition in real-time. The source code is generally organized as follows:

```
Metaboverse
├── build.sh:                A helpful script for compiling Metaboverse for␣
→distribution
├── /docs                    Source files for metaboverse.readthedocs.io
└── /app
    ├── /css:                CSS style files, including both custom and distributed␣
→files
    ├── /data:               Test data that is distributed with each release of␣
→Metaboverse,
    │                         along with icons and images used within the GUI
    ├── /html:               HTML files for each of the pages used with in the␣
→Metaboverse GUI
    ├── /js:                 Javascript files for the Metaboverse GUI. Selected␣
→important files
    │   │                     are listed below:
    │   ├── index.js:        Handles the home screen interactions
    │   ├── curate.js:       Handles the Curate screen interactions, where users␣
→provide
    │   │                     information about their model organism and more
    │   ├── variables.js:    Handles the Variables and Data screen interactions,␣
→where users
    │   │                     provide input data and other experimental data
    │   ├── build.js:        Handles the Build screen interactions, where Metaboverse␣
→compiles
    │   │                     the information provided by the user and performs back-
→end
    │   │                     processing with metaboverse-cli
    │   ├── visualize.js:    Handles network visualization
    │   ├── timecourse.js:   Handles formatting and toggling between time-course or␣
→multi-condition
    │   │                     samples during visualization
    │   ├── js-colormaps.js: A Javascript implementation of the "seismic" colormap␣
→from Matplotlib
    │   ├── motif-script.js: Front handler for the Reaction Pattern screen
    │   ├── motif-graph.js:  Handles interactions and visualization of Reaction␣
→Patterns
```

```
    |    ├── motifs.js:      Handles the actual, real-time Reaction Pattern searching␣
→across all
    |    |                       reactions in the network
    |    ├── motif-global.js: Searches for every possible Reaction Pattern during␣
→Pathway visualization
    |    ├── perturbations.js:Handles the Perturbation screen interactions and network␣
→generation of
    |    |                       a network consisting of all perturbed reactions within␣
→a given pathway
    |    └── datatable.js     Handles the Format Data screen interactions
    ├── main.js:              This handles initializing the Electron app window
    ├── package.json          This handles initializing dependencies and other␣
→settings for an
    |                           an Electron app
    └── __version__.txt:      This supplies the current software version displayed␣
→within the
                                software
```

2.9.6 Documentation and Testing

Any changes made should be supported by intuitive and descriptive comments throughout the code, documentation for analysis features for metaboverse.readthedocs.io, and with test cases.

2.9.7 Distribution

Distribution of a new release of Metaboverse requires several steps, outlined below.

1) Compile `metaboverse-cli`

    a) Update the version number in `metaboverse_cli/__init__.py`

    b) You will then need to do the following for each operating system you are distibuting Metaboverse on.

```
$ conda create --name pyinstaller
$ conda activate pyinstaller
$ conda config --add channels conda-forge
$ conda install python=3.8
```

On Windows:

```
$ conda install pyinstaller
$ conda install --file requirements.txt
```

On Mac/Linux:

```
$ pip install pyinstaller
$ pip install -r requirements.txt
```

Then, on each operating system

```
$ cd /path-to/metaboverse-cli
$ pyinstaller metaboverse-cli.spec
```

From here, you can archive each operating system executable on the Github release page for `metaboverse-cli`.

2) Compile `Metaboverse`

a) You will need to download the `metaboverse-cli` executables for each operating system. These should be stored in the directory, `/path-to/Metaboverse/app/python/`

b) Update version numbers in `/path-to/Metaboverse/docs/conf.py`, `/path-to/Metaboverse/build.sh`, `/path-to/Metaboverse/app/__version__.txt`, `/path-to/Metaboverse/app/package.json`, and `/path-to/Metaboverse/CITATION.cff`

c) Compile the Metaboverse distributions:

```
$ cd /path-to/Metaboverse
$ bash build.sh
```

> **Warning:** You will likely need to compile the Metaboverse package for Mac on a Mac. If the Mac-version `metaboverse-cli` executable is packaged into Metaboverse via Electron on another operating system, we have consistently received an error in executing `metaboverse-cli`.

2.10 Contact Us

Contact the Metaboverse team using one of the following methods:

2.10.1 Submit and issue or feature request on GitHub

The best method to contact us to through submitting an issue on the Metaboverse GitHub page.

- You will need to log-in to your GitHub account, or create a GitHub account here.

- On the issues page, click the `New issue` icon.

- You can choose whether to submit a **bug** or **question** using the `Bug report` template, or submit a **feature request** using the `Feature request` template.

- Please clearly describe the problem, what you have tried, as well as screenshots of any error information.

- If you are willing to include the datasets used during thei Metaboverse session, please include them in the issue.

- Generally, for any errors occurring during network building, a file named `metaboverse_session.log` will be output to your specified Output folder. If you receive this file, please upload it to your GitHub Issue.

2.10.2 Email

If you would like to reach us directly via email, please contact Jordan Berg at
`jordan<dot>berg<at>biochem<dot>utah<dot>edu` (replace the symbols within `<>` with their respective symbol).

Note: After July 31, 2022, please email jordanberg17<at>gmail<dot>com. Please also include an informative email subject line, such as "Metaboverse issue".

CHAPTER 3

Analysis Video Walkthrough

You can watch the Metaboverse walkthrough video here, or go the *Walkthrough* page for a more detailed walkthrough.

CHAPTER 4

Questions and Feedback?

- Check out the *FAQs* page.
- Learn how to *Contact Us*.
- You can also find a general purpose discussion forum here.

License

Metaboverse is developed and maintained by Jordan Berg in the Rutter Lab @ the University of Utah, along with other collaborators. We welcome pull requests if you would like to contribute to the project.

Metaboverse is perpetually open-source under an MIT license.