

# **MSI.EAGLE User Manual**

Weljie Lab, University of Pennsylvania

GitHub page containing this manual, vignettes, and source code: [Link](#)

## **Table of Contents**

<i>Introduction</i> .....	<b>2</b>
<i>Installation</i> .....	<b>3</b>
<i>Data Setup</i> .....	<b>5</b>
<i>File Restore and Overview Analysis</i> .....	<b>6</b>
<i>Segmentation – UMAP</i> .....	<b>7</b>
<i>Segmentation – SSC</i> .....	<b>11</b>
<i>Phenotyping</i> .....	<b>12</b>
<i>Depth Analysis</i> .....	<b>14</b>
<i>Statistics</i> .....	<b>15</b>
<i>Heatmap</i> .....	<b>17</b>
<i>Colocalization Analysis</i> .....	<b>18</b>

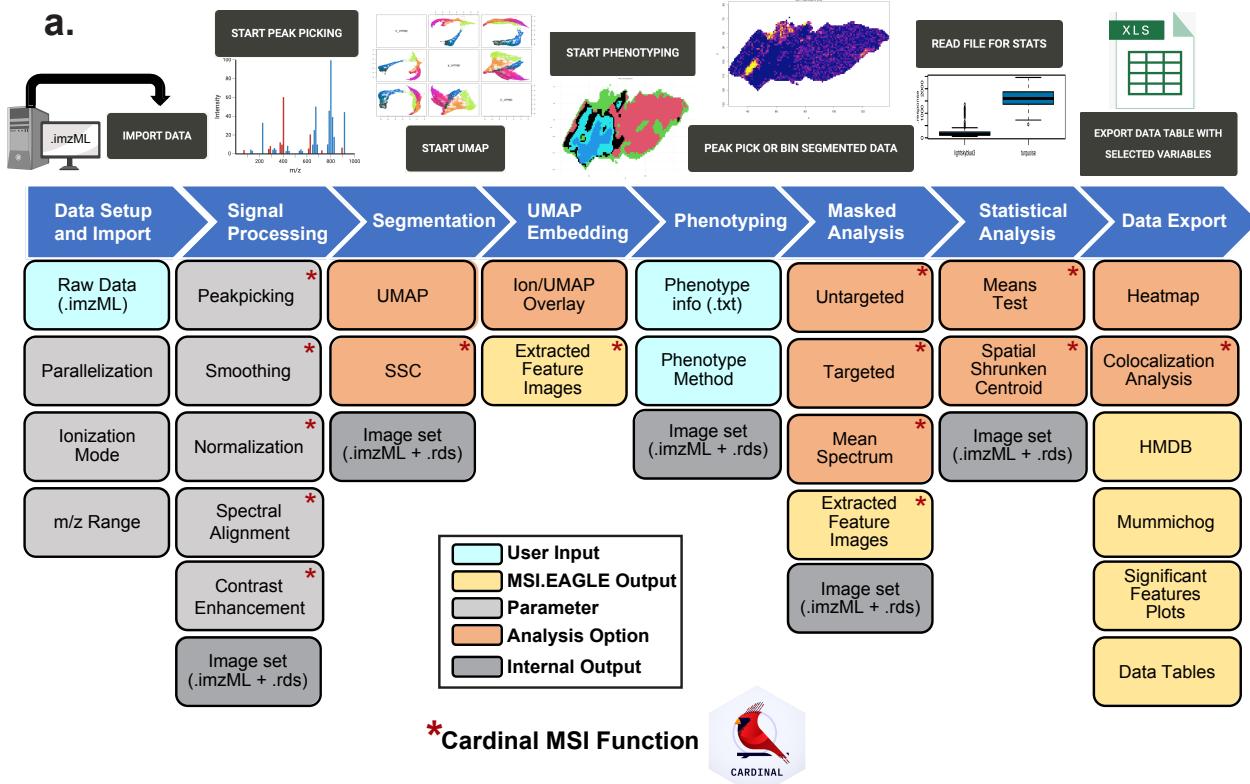
# Introduction

MSI.EAGLE is a modular R Shiny app designed to provide a user-friendly interface for processing of mass spectrometry imaging data. It relies heavily on the *Cardinal* R package for underlying spectra processing and certain data analysis functions, providing a number of other options for data analysis and visualization. A major advantage of the workflow and structure is processing of larger datasets which include multiple runs and the ability to easily manipulate whole datasets or subsets. The underlying architecture of the app is setup as individual modules accessed via tabs in the user interface, and the high-level major functions of each module are described in the diagram below. The modular layout also encourages further development of algorithms and methods which can be easily implemented. Within the modules there are a few workflow related dependencies to keep in mind:

- Data setup and import, Signal Processing / reading in files, Segmentation, and Phenotyping share data internally, and the data must be read in through the “File Restore and Overview Analysis” tab (corresponding the Signal processing arrow in the diagram below) or imported and processed as imzML data.
- The Depth Analysis tab relies on re-analysis of input raw data, and thus .imzML files must be available (and read into memory) from the Data Setup tab along with a set of coordinates from the same dataset to guide detailed analysis.
- The Heatmap and Colocalization tabs currently rely on having already performed a statistical analysis in the Statistics Tab.

Considerations for each tab will be outlined in subsequent further pages of the guide.

While running the app it is recommended that the user have access to the R console as there may be important information regarding dataset sizes, errors and other info.



# Installation

## Prerequisites

Before installing MSI.EAGLE, you need to install both R and R Studio: <https://rstudio-education.github.io/hopr/starting.html>. Then, install the required Bioconductor packages and Git as outlined below.

### Step 1: Install Required R Packages

```
# Install remotes package  
install.packages("remotes")  
  
# Install BiocManager if not already installed  
if (!require("BiocManager", quietly = TRUE))  
  install.packages("BiocManager")  
  
# Install required Bioconductor packages  
BiocManager::install(c("BiocParallel", "Cardinal"))  
  
# Optional: Install CardinalWorkflows for demo data  
BiocManager::install("CardinalWorkflows")
```

### Step 2: Install Git

If you don't have Git installed, download and install it from: <https://git-scm.com/book/en/v2/Getting-Started-Installing-Git>

### Step 3: Install MSI.EAGLE

```
remotes::install_github("https://github.com/Weljie-Lab-UPenn/MSI.EAGLE")
```

## Basic Setup

1. In R Studio, create a new R script, and copy and paste the start code below into it.
2. Add the file path to the folder containing your .imzML and .ibd raw data files (rawd), and the path to the folder where your results will be saved (wd).
3. Then, click "Source with Echo" to run the application.

The following start script allows the user to specify the raw file directory and working directory for saving processed files. If not specified, MSI.EAGLE will use the current R working directory for both raw files and data output.

The script also specifies the number of cpu cores for MSI.EAGLE to use (minus two reserved for background processes) for parallel processes. These parameters can also be adjusted in the app; however, it may be more convenient to set defaults if working from startup scripts specific to each project.

## Start Script – Copy to R studio, then click “Source with Echo” to run MSI.EAGLE

```
#MSI.EAGLE Start  
  
library(MSI.EAGLE)  
  
# Set path to raw data files  
rawd = 'path/to/rawfilesdirectory'  
  
# Set working directory (can be same as raw data directory, wd = rawd)  
wd = 'path/to/workingdirectory'  
  
# Configure number of CPU cores to use  
# Option 1: Use half of available processors  
ncores <- as.integer(parallel::detectCores()/2)  
  
# Option 2: Use all but 2 processors (recommended)  
ncores <- as.integer(parallel::detectCores()) - 2  
  
# Launch the application  
MSI.EAGLE()
```

## Documentation

For more instructions, tutorials, and examples, see our [Manuals and Vignettes](#).

## Requirements

- R version 3.6 or higher
- Bioconductor packages: Cardinal, BiocParallel
- Git (for installation)
- Sufficient RAM for mass spectrometry imaging data processing

# Data Setup

The purpose of this tab is to import initial imaging data for processing. The interface provides a set of parameters for processing files in .imzML format. Currently only data in centroid mode is fully supported, although continuous data can be read in (not fully tested). Files in imzML format are read from the Raw files directory and a selectable table presented to the user. Once one or more files are selected, import parameters are selected to be used by [Cardinal 3](#) and the data imported via the “Import Data” button. After Import, a small subset of the data can be optionally processed directly in order to quickly check spectral quality via the “Extract Sample and MS Sample Plot”. Full processing is available in the next tab.

Setup data directories, some options, and import .imzML files

```
folder= /Volumes/Woolfork/MALDI/02-10-2023_files/
wd= /Volumes/Woolfork/MALDI/02-10-2023_files/
```

Show 10 rows

Select files for further processing

**files**

1	03012023_DESI vs MALDI Fly 100um_Slide_A_Analyte 2AFAMM
2	regions-total ion count

Showing 1 to 2 of 2 entries

**Image visualization ions**

- All
- Custom

**Contrast enhancement** **Colorscale**

suppression	cividis
-------------	---------

**Smoothing options** **Normalize image settings**

none	linear
------	--------

Draw color key?

**Image plot width (px)** 800 **Image plot height (px)** 600

Extended font options?

Raw files directory (ie. .imzML file location)  
/Volumes/User/Woolfork/Desktop/DESI

Working directory (ie. where results are stored)  
/Volumes/User/Woolfork/Desktop/DESI

**SELECT DIRECTORY**

Regex syntax to extract files if needed.  
(.\*)imzML\$

**Parallelization mode**

- None
- MulticoreParam
- SnowParam

**Number of cores to use**  
6

**Peak picking params**  
qToF1

**resolution for peakpicking (ppm)**  
15

**tolerance for spectral alignment (ppm)**  
50

**m/z min for import**  
50

**m/z max for import**  
1700

**IMPORT DATA**  
Limits for plotting spectrum to ensure MS looks reasonable

**% of pixels to randomly sample for plot**  
1

**EXTRACT SAMPLE AND MS SPECTRUM PLOT**

**Choose sample dataset**  
Human Renal Cell Carcinoma (RCC)

**LOAD DEMO DATA**

<b>Parallelization Mode Options</b>	See <a href="#">Intro to BiocParallel</a> for Parallelization mode options.
<b>Select Number of Cores to Use</b>	Number of cores to use is typically $\frac{1}{2}$ the available processors.
<b>Select Peak Picking Parameters</b>	Peak picking parameters are a way to set default parameters for peak picking resolution and spectral alignment tolerance.
<b>Select Resolution for Peak Picking (ppm)</b>	
<b>Select Tolerance for Spectral Alignment (ppm)</b>	
<b>Select m/z Min for Import</b>	
<b>Select m/z Max for Import</b>	
<b>Select Percentage of Pixels to Randomly Sample for Initial Visualization Plot</b>	For large datasets, it's best to randomly sample only 1% of pixels.
<b>Choose Sample Dataset</b>	This is an option to load demo data for following the workflow of the MSI.EAGLE app. The data from this option will be used in the remainder of the user manual. Suggested parameters are located in images. To load this package: > BiocManager::install("CardinalWorkflows")

# File Restore and Overview Analysis

The purpose of this tab is to manipulate datasets either by de novo peak picking, restore previously processed files, adding or subsetting existing datasets. 1-100 files / runs can be manipulated at once.

The screenshot shows the 'File Restore and Overview Analysis' tab. On the left, under 'Peakpicking operation', there are several options: 'Open previously peakpicked file (.rds)' (selected), 'Peakpick rawfiles (imzML loaded in Data Setup tab)', 'Add two imagesets (same coordinates)', 'Add two imagesets (same peaklist)', and 'Demo data from CardinalWorkflows package'. Below this is a section 'Start from an existing peak\_picked file?' with 'Yes' selected. A dropdown menu labeled 'Imageset with peaked peaks' is shown, and a 'RESTORE SAVED FILE' button is at the bottom. On the right, there are sections for 'Image visualization ions' (radio buttons for 'All' or 'Custom'), 'Contrast enhancement' (dropdown), 'Colorscale' (dropdown), 'Smoothing options' (dropdown), 'Normalize image settings' (dropdown), and 'Draw color key?' (checkbox checked). Below these are 'Image plot width (px)' and 'Image plot height (px)' input fields, and an unchecked checkbox for 'Extended font options?'. A horizontal line separates this from the table below.

**Open previously peakpicked file (.rds)** opens a previously processed dataset using a peakpicked .rds file from the directory.

**Peakpicked raw files (imzML loaded in Data Setup tab)** opens a new dataset using raw data imzML files that must be already available in memory from the previous tab.

**Add two image sets (same coordinates)** combines two imagesets containing the coordinates. The coordinates must have been peak picked already, but usually with different peak lists (i.e., same x,y coordinates, but different m/z lists). The features selected option/filtered allows for uses of specific features to create subsets.

**Add two image sets (same peaklist)** combines two image sets containing the same peak list. The two sets must come from the same original peakpicked file (i.e., adding two datasets with different x,y coordinates but the same peak list).

Select Files for Further Processing and Import Data	R Console will give object summary containing feature, pixel, and spectral information.
Extract Sample and MS Spectrum Plot to Confirm Proper Data Import	The plotting speed will depend on the size of the dataset and will initially need to be calculated and decompressed as a MSProcessedImagingExperiment object.
Other Features	
Image Visualization Ions	All ions or a custom m/z can be visualized for all selected files.
Contrast Enhancement, Colorscale, Smoothing Options, and Normalization Image Settings	- Draw Color Key: Places the key within the visualization plot. - Image Plot Width and Height: Change the size of the plot.

# Segmentation – UMAP

The purpose of this tab is to segment a dataset to summarize the MS imaging experiment define regions of interest and assign phenotypes to the pixels. MSI.EAGLE implements two methods for this, UMAP in the first tab, and shrunken spatial centroids in the neighboring tab.

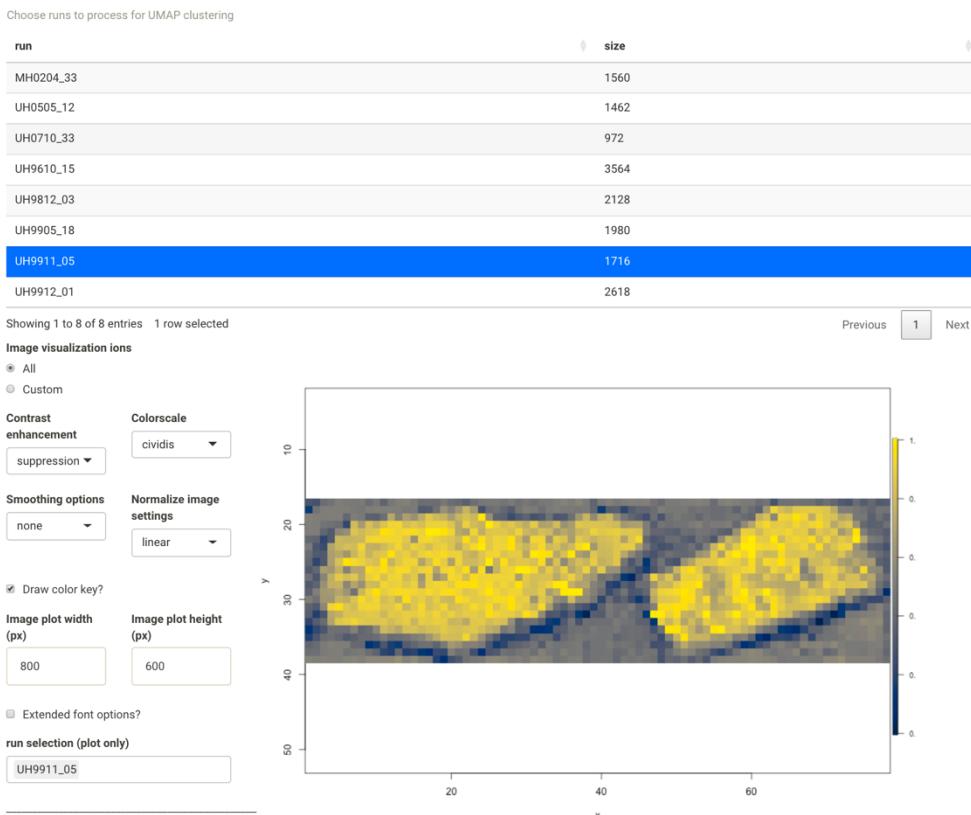
There are different goals for segmentation in this tab, both based on the Uniform Manifold Approximation and Projection for Dimension Reduction (UMAP) technique. This technique scales a dataset in size and dimension in a way that preserves the structure of the data, giving a visual representation of distinct data clusters that allows for assumptions to be made about the data. The implementation here relies on the '[uwot](#)' library in R and further information about specific parameters can be found in the package documentation.

Background removal can be used to remove the background of the slide used for imaging and anatomical segmentation is used to further segment tissue images for anatomical annotation. Both make use of UMAP parameters based on the nearest neighbors, the number of nearest neighbors used to construct the cluster and minimum distance between the points in the space. As the number of nearest neighbors increase, more neighboring points are connected and constructs a representation of the data that closely aligns with the global structure of the data. As the minimum distance increases, the UMAP spreads out the projected points and decreases the clustering of the data and reduces the global structure. The difference between the two implementations (background vs anatomical) is that the anatomical goal allows for access to a greater set of features for assigning and visualizing phenotype information assigned to the pixels, along with some slight differences in default parameters.

The UMAP PCA components runs a PCA on the dataset prior to UMAP, which is useful when the dataset is very large. The NULL feature runs on the entire dataset and gives more depth in colors and clustering. Tissue images use a reduced dataset due to image size.

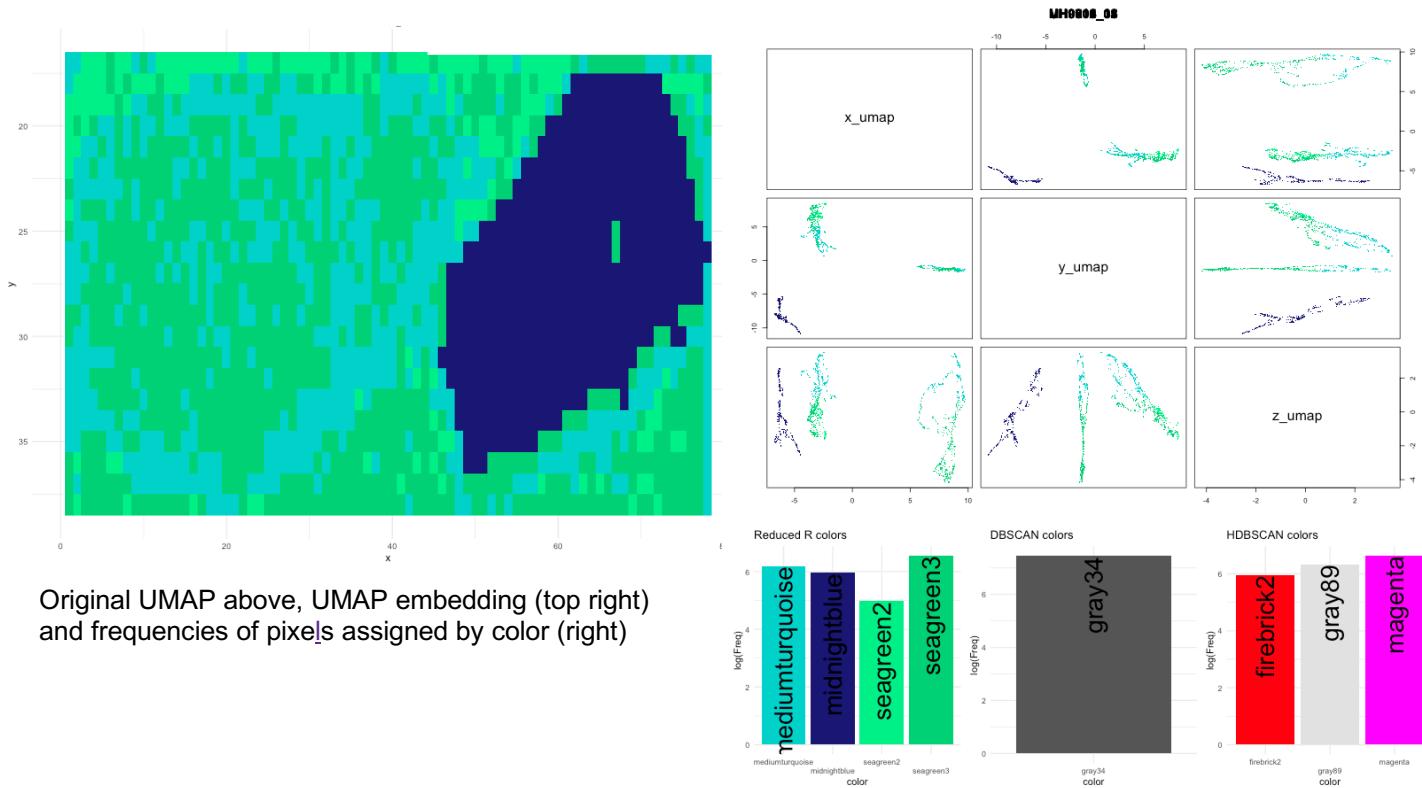
Color visualization of the UMAP embeddings can occur using R DBSCAN (via the 'dbSCAN' package or via method to color the UMAP space via mapping of x, y and z coordinates to rgb color space [1]. The R color reduction quantile allows for control of the color distribution of the clusters. A higher value should be used initially to accurately remove the background of the slide. With each successive UMAP, the value can be decreased to give more defined colors of background regions to be removed. The background can be segmented by removing colors one at a time while observing the original image as well as visualizing the UMAP space. The store processed data function should be used before beginning another UMAP of the image.

## Background Removal



Example from demo data derived from the 'CardinalWorkflows' package. Select the data from the table to analyze via UMAP

UMAP may be performed as a group of data files or as individual data files. If performed as individual data files, the data must be restored and saved before unselecting the current file and selecting a new file to UMAP. Once all the data files are individually processed, then, the overall UMAP data can be saved as one file.



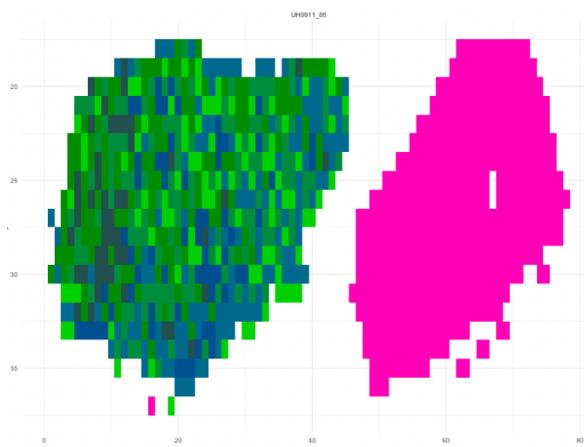
### Colors for UMAP visualization

R reduced

Choose colors

- mediumturquoise
- seagreen2
- seagreen3
- midnightblue

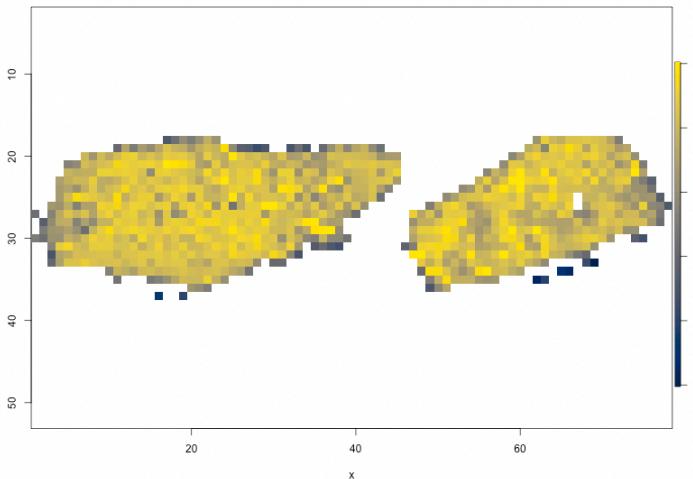
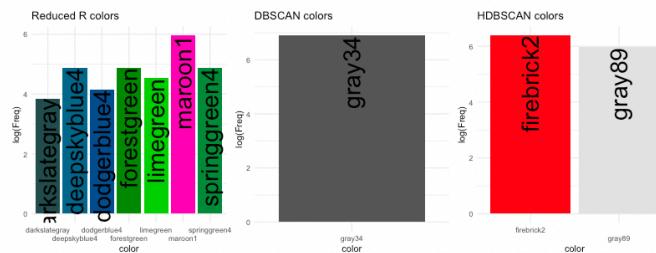
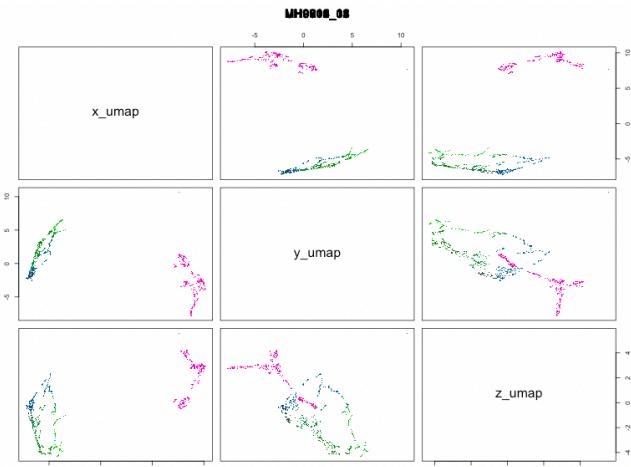
**STORE PROCESSED DATA** **SAVE PROCESSED DATA**



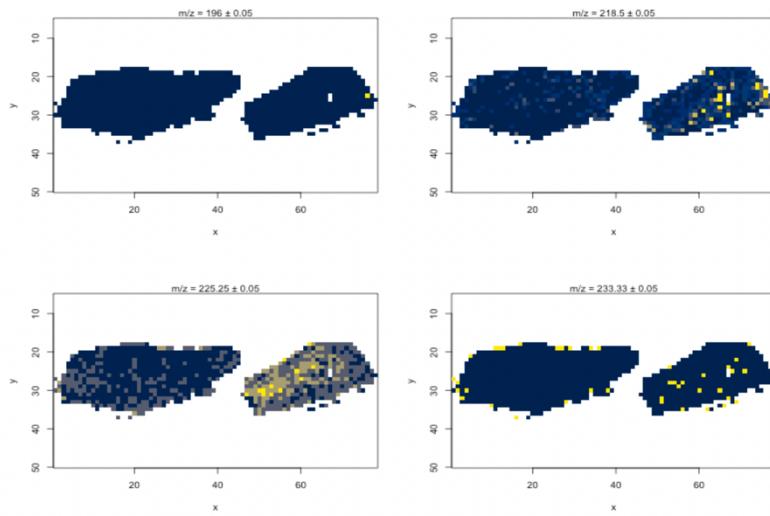
After successive UMAPs for background removal

Uncheck the colors that corresponds to the background to be removed during the next UMAP.

Store the processed data before reducing the R color reduction quantile and beginning the UMAP process again.



Check your MS image to make sure the background is being removed.



Once the background is successfully removed, UMAP segmentation can also be for visualizing a specific ion using *custom image visualization of ions*.

This option can also perform *ion math* which can give the sum, min, max, standard deviation for the custom ions (current selection: none).

## pData/UMAP Editor

Under this option, the pData field for visualization must be set and Values to show must be defined or a new field must be added. Clusters and other pData variables can be re-named for downstream analysis.

Select pData field for visualization

Values to show

 cancer  normal

Choose pData field to annotate:

 Replace
 Add new field

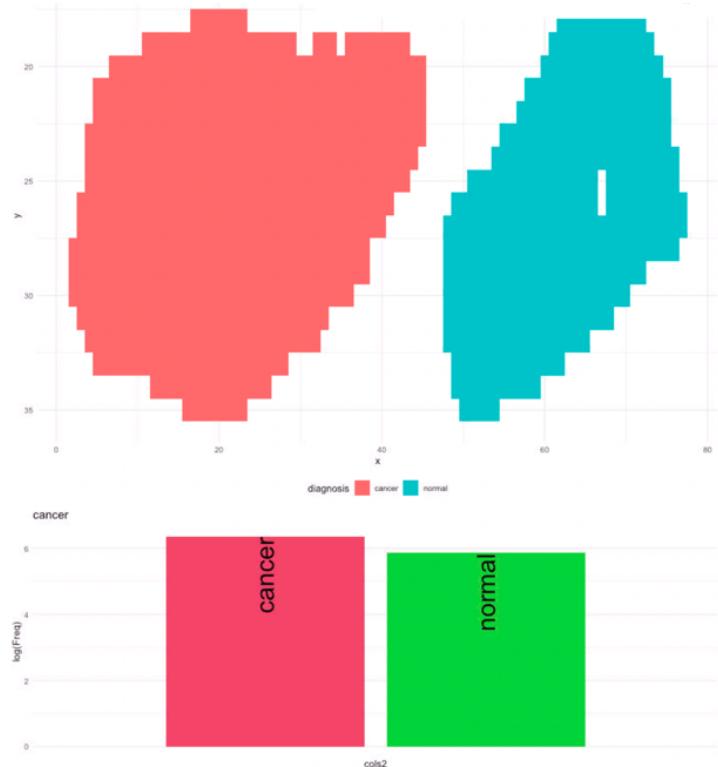
Choose field to annotate

Choose existing value

OR

Input new annotation value

**APPLY ANNOTATION**



# Segmentation – SSC

The purpose of this tab is to segment a dataset to summarize MS imaging experiment and define regions of interest.

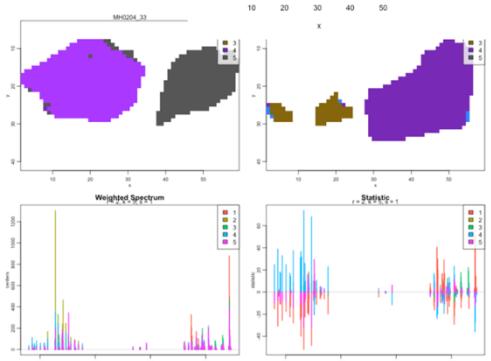
While PCA components is a popular way for exploring a dataset, segmentation can also occur with spatial shrunken centroids (SSC) [as described by Bemis et al \(2016\)](#) [2]. This performs spatially aware nearest shrunken centroid clustering or classification a dataset. SSC is based on four parameters: method, r, s, and shrinkage (c or s). Method determines the spatial weight (Gaussian or adaptive). The r parameter is the smoothing radius, ranging from 2 and 5. The k parameter is the initial number of clusters, where a large value is higher than the expected number of clusters and the algorithm will automatically drop empty segments (ranges from 5 and 15). The c parameter default is the shrinking parameter ranging between 5 and 10, where higher values result in fewer contributing clusters to segmentation.

Because of the potential time-burden of performing SSC, a file is created which saves the analysis in the working directory. The user can re-access the file by inputting the same parameters used to generate the initial file. Each model is given as a separate option to visualize. Once a model is selected, three panels are plotted below the image plot. This is an image of the ssc clusters themselves, a weighted spectrum showing which masses are contributing to each cluster, and a plot of the statistic to visualize the highest contributing m/z values. Further analysis to extract specific m/z values can be performed by opening the saves ssc model file in R.

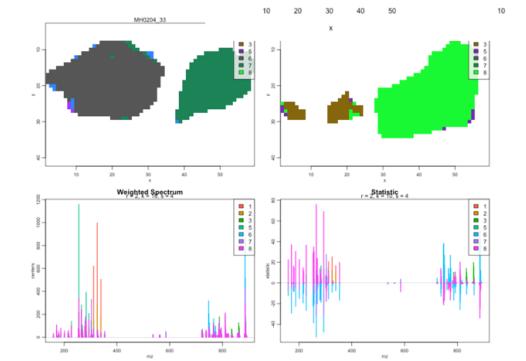
The SSC models may not work if there are pixels with few neighbors. As a result, the “Show fix stray pixel options?” check box is offered which will remove these stray pixels and can in some (but not all cases) allow the ssc to run smoothly. Note that at the moment one would need to perform the same stray pixel fix (same parameters) in order to restore a SSC model previously run otherwise a size mismatch will occur in the number of pixels. The SSC models are automatically saved to the working directory, but can be deleted if one wants to run SSC with the same parameters more than once.

To reproduce the demo data as produced in this manual, below are the following parameters:

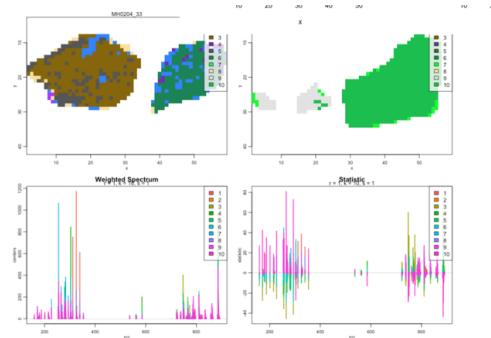
- Under the Data Setup Tab:
  - Load the RCC without background data
  - Change the Tolerance for spectral alignment to 1200
- Under the File Restore and Overview Analysis
  - Select Peakpick rawfiles
  - Change the S/N for overview peakpicking to 3
  - Change Peak picking method to mad
  - Start peak picking



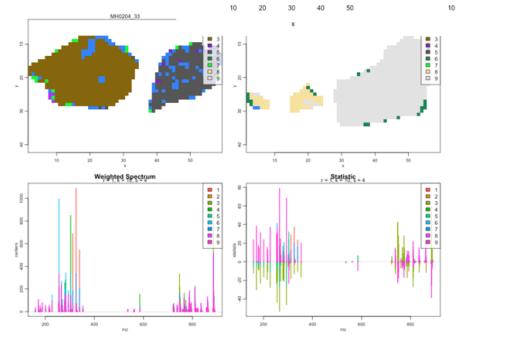
SSC segmentation for adaptive method  
r = 2, k = 5, s = 1; shrinkage c(1,4)



SSC segmentation for adaptive method  
r = 2, k = 10, s = 4; shrinkage c(1,4)



SSC segmentation for adaptive method  
r = 1, k = 10, s = 1; shrinkage c(1,4)



SSC segmentation for adaptive method  
r = 1, k = 10, s = 4; shrinkage c(1,4)

# Phenotyping

The purpose of this tab is phenotype MSI data. This is a critical step in being able to assign meaning of the data through statistics. Each pixel can have multiple phenotypes assigned. The approach used in MSI.EAGLE relies on removing the background pixels to then assign phenotypes based on areas with and without sample information. The automated method also relies on a grid-like structure of the input data.

Spot	Plate	Sample Name	Time Point
A1	P1	AGW#1	AM0800
A2	P1	AGW#3	AM400
A3	P1	AGW#4	AM2400
A4	P1	AGW#6	AM0400
A5	P1	AGW#7	AM1600
A6	P1	AGW#8	AM0800
A7	P1	AGW#13	PM2400
A8	P1	QC	NA
A9	P1	AGW#14	AM400
A10	P1	AGW#15	AM2400
A11	P1	AGW#16	AM0800
B1	P1	AGW#20	AM400
B2	P1	AGW#21	AM2400
B3	P1	AGW#24	AM400
B4	P1	AGW#25	AM400
B5	P1	AGW#27	AM0800
B6	P1	AGW#28	AM2400
B7	P1	QC	NA
B8	P1	AGW#29	AM1600
B9	P1	AGW#31	AM400
B10	P1	AGW#32	AM0800
B11	P1	AGW#33	AM0800

Sample data to insert into a .txt file for spotted samples

Spot	Plate	Sample Name	Phenotype
A1	P1	MH0204_33	Cancer
A2	P1	MH9294_33	Normal
A1	P2	UH0505_12	Cancer
A2	P2	UH0505_12	Normal
A1	P3	UH0710_33	Cancer
A2	P3	UH0710_33	Normal
A1	P4	UH9610_15	Cancer
A2	P4	UH9610_15	Normal
A1	P5	UH9812_03	Cancer
A2	P5	UH9812_03	Normal
A1	P6	UH9905_18	Cancer
A2	P6	UH9905_18	Normal
A1	P7	UH9911_05	Cancer
A2	P7	UH9911_05	Normal
A1	P8	UH9912_01	Cancer
A2	P8	UH9912_01	Normal

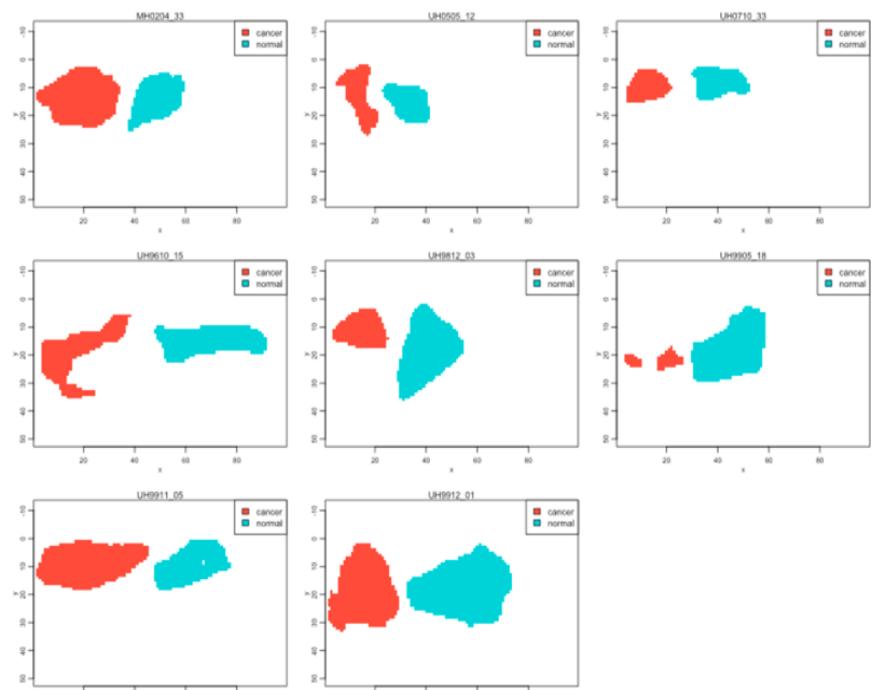
Sample data to insert into a .txt file for tissue samples

If starting from an existing peak picked file, the file must be restored from the file restore tab. If continuing data processing from segmentation, choose the 'Stored data' option. In addition, a text file with the phenotype data needs to be uploaded in the phenotyping tab. The text file must be tab-delimited with a .txt ending. The input text file for all methods must have minimally three columns, the first two being titled "Spot" and "Plate". The third and subsequent columns are the phenotype information (eg. Biological grouping, Dose, Time etc...). The phenotype name will match the column header name. "Spot" must be specified at the grid position (A1, A2... B1, B2...etc.) of the sample. If there are only tissues, a single row (A1, A2, A3 etc.) can be used. When analyzing tissues, automated phenotyping in the app is dependent on collecting tissues in a grid-like pattern. The "Plate" variable defines the run to which the spot/sample must be contained in the run name (i.e., the name displayed on the table during import or display in the previous tabs). Furthermore, the "Plate" variable must be unique to each run. For if the runs were called 'run\_P1' and 'run\_P2', samples from the former could have plate values of 'run\_P1' or simply 'P1'.

There are four phenotyping methods: spectral density, periodicity, breaks between samples, and manual.

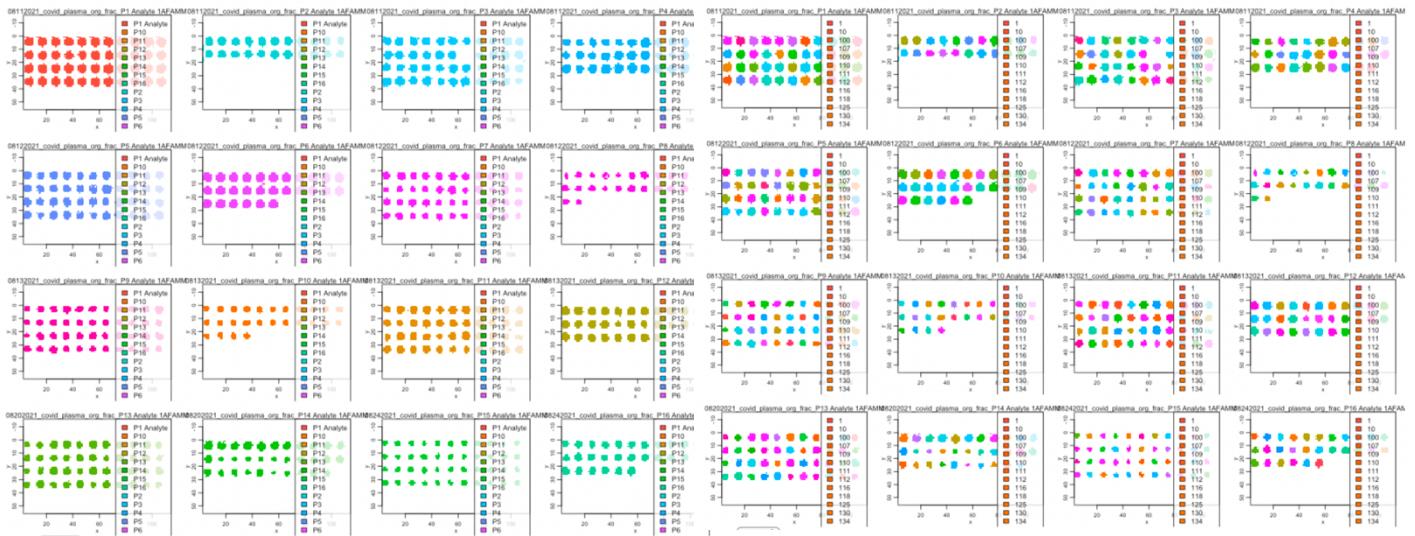
Spectral density and periodicity will attempt to use the frequency of the pixel density along the x and y axis to estimate the minimum to find breaks between samples. The breaks method is a simpler version which looks for areas with no pixels to assign phenotype.

The manual method will assign phenotypes based on user specified x- and y- start and end positions. These are specified in the input phenotype text file through four columns (xstart, xend, ystart, yend) added to the phenotyping .txt file for each sample.



Plotting of phenotype data by diagnosis (cancer vs normal) tissue samples

The phenotyping data (once applied) can also be plotted based on one of the columns. Interaction terms between variables can be added, which may be useful for downstream statistics (for example to define grouping variables). This is often used to define replicates of a condition: For example, if there is a “Genotype” column containing “WT” and “KO” conditions with three replicates each, one should create an interaction term between “Genotype” and “Replicate” columns. Based on these choices, various phenotyping data can be added into different or combined files by clicking the save phenotype data button.



## Interaction data added for phenotyping

# Masked Analysis

The purpose of this tab is to analyze specific pixels in the data without the burden of analyzing the complete dataset. This may be useful in a number of situations, such as analyzing only tissue pixels once background is removed or analyzing only certain tissue structures. By refining the analysis, targeted methods benefit by removing noise from pixels not relevant to the analysis which can be particularly important for low-abundance peaks which may get removed by a frequency filter (i.e., the requirement that a peak must appear in 1% of the data to not be considered noise). Another advantage of this approach is the ability to use a mean spectrum for peak picking which is reflective of tissue / pixels of interest which will also reduce spurious noise.

Because the Masked Analysis relies on re-analysis of the data, the source imzML files must be available and open in memory from the 'Data Setup' tab prior to any analysis.

The 'Imageset with segmented coordinates' file is a processed file which has some subset of pixels in the data (for example, tissue only, background removed). These sample coordinates will be used to specifically extract and perform the analysis on the raw data. Note that the run names in the input coordinate file and the .imzML file must match exactly since we are assuming this is a reanalysis of the same data. This can be a bit tricky if the imzML files are being read from a sub-directory, in which case the names may not be an exact match.

The other options are analogous to those from the 'File Restore and Overview Analysis tab'

# Statistics

The purpose of this tab is to perform statistical analysis on MSI data. There are three methods from Cardinal (means test, supervised SSC, and DGMM). A description of these and specific use cases are fully described in the Cardinal manual.

To begin statistics, the file generated from phenotyping must be read in for analysis. The simplest test is a means test from Cardinal which compares the mean of a variable in a group with the same variable in other groups across all pixels of interest. The variable for testing is then chosen and then members of this variable can be chosen along with interaction for grouping. The grouping variable should define the 'biological' replicate of the sample. As a result, all pixels and technical replicates from the same biological replicate will be considered together. After running the test, a notification will appear, and a table of the generated significant features will appear at in the 'Output table and ion plots' subtab in the main window. The FDR threshold can be adjusted based on needs and a value of 1 will give the entire list. The data table can be exported with selected variable as a CSV, excel, or PDF file.

To reproduce the demo data as produced in this manual, below are the following parameters:

## Means Test (Cardinal)

Imageset for analysis  
background\_Users\_ashwoo/Desktop\_0(Temp\_MSI-  
proc\_hi\_SN-peak\_picked-subset-2023-09-14.rds

READ FILE FOR STATS

Type of test  
 Means test (Cardinal)  
 Spatial shrunken centroids (Cardinal)  
 Spatial DGMM (Cardinal)  
 ANOVA (Experimental)

FDR threshold to include in results  
0.05

# Peaks (debugging); set before reading .rds file  
10

LOAD DEMO RCC DATA (CARDINALWORKFLOWS)

SETUP AND RUN STATISTICS   OVERVIEW PLOTS   OUTPUT TABLE AND ION PLOTS   DATA EXPORT

[1] "working filename= background\_Users\_ashwoo/Desktop\_0.Temp\_MSI-  
proc\_hi\_SN-peak\_picked-subset-2023-09-14.rds"

Choose variable to test  
diagnosis

Choose members to include  
cancer normal

Choose interaction for grouping  
run diagnosis

RUN TEST

Choose variables for modeling  
diagnosis

SAVE MEANS TEST MODEL   RESTORE MEANS TEST MODEL

The Imageset for analysis is a saved RDS file with background from the demo data

## Select plots to view

SETUP AND RUN STATISTICS   OVERVIEW PLOTS   OUTPUT TABLE AND ION PLOTS   DATA EXPORT

Show 10 entries

Choose runs to visualize

run	V2
MH0204_33	run
UH0505_12	run
UH0710_33	run
UH9610_15	run
UH9812_03	run
UH9905_18	run
UH9911_05	run
UH9912_01	run

Showing 1 to 8 of 8 entries

Image visualization lens  
 All  
 Custom

Contrast enhancement  
suppression

Smoothing options  
none

Draw color key?

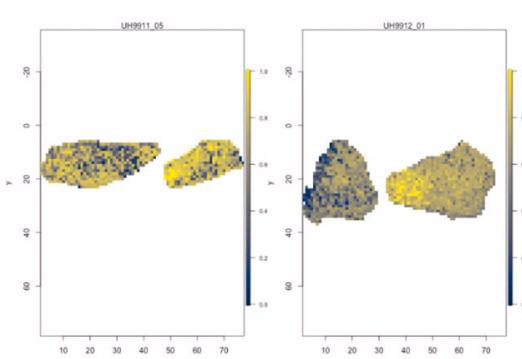
Image plot width (px)  
800

Image plot height (px)  
600

Normalize image settings  
linear

Extended font options?

Select individual runs for plotting only?

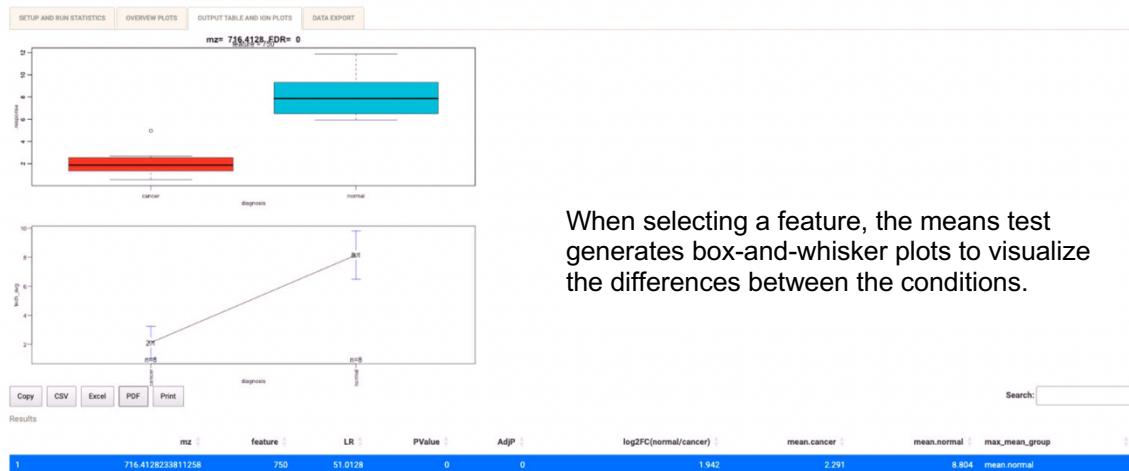


	mz	feature	LR	PValue	AdjP	log2FC(normal/cancer)	mean.cancer	mean.normal	max_mean_group
1	716.4128233811258	750	51.0128	0	0	1.942	2.291	8.804	mean.normal
2	809.311505251934	890	48.7747	0	0	1.623	4.068	12.531	mean.normal
3	808.2567210424116	889	39.8979	0	0	1.934	4.058	15.509	mean.normal
4	740.3572055267042	794	38.5411	0	0	1.832	5.102	18.168	mean.normal
5	714.471743751577	747	34.2426	0	0	2.337	2.281	11.524	mean.normal
6	714.941319115965	748	34.3553	0	0	1.384	2.155	5.626	mean.normal
7	742.444688428813	796	32.6884	0	0	1.757	5.298	17.901	mean.normal
8	215.2334970328965	54	31.1804	0	0	2.381	13.48	70.206	mean.normal
9	715.5781471790017	749	29.6014	0	0.00001	1.325	1.805	4.524	mean.normal
10	717.0909265107317	751	29.1577	0	0.00001	1.052	2.135	4.427	mean.normal

Showing 1 to 10 of 70 entries Show 10 entries

Previous 1 2 3 4 5 6 7 Next

Table of significant features with FDR threshold: 0.05. To generate a table with all results, change the FDR threshold to 1.



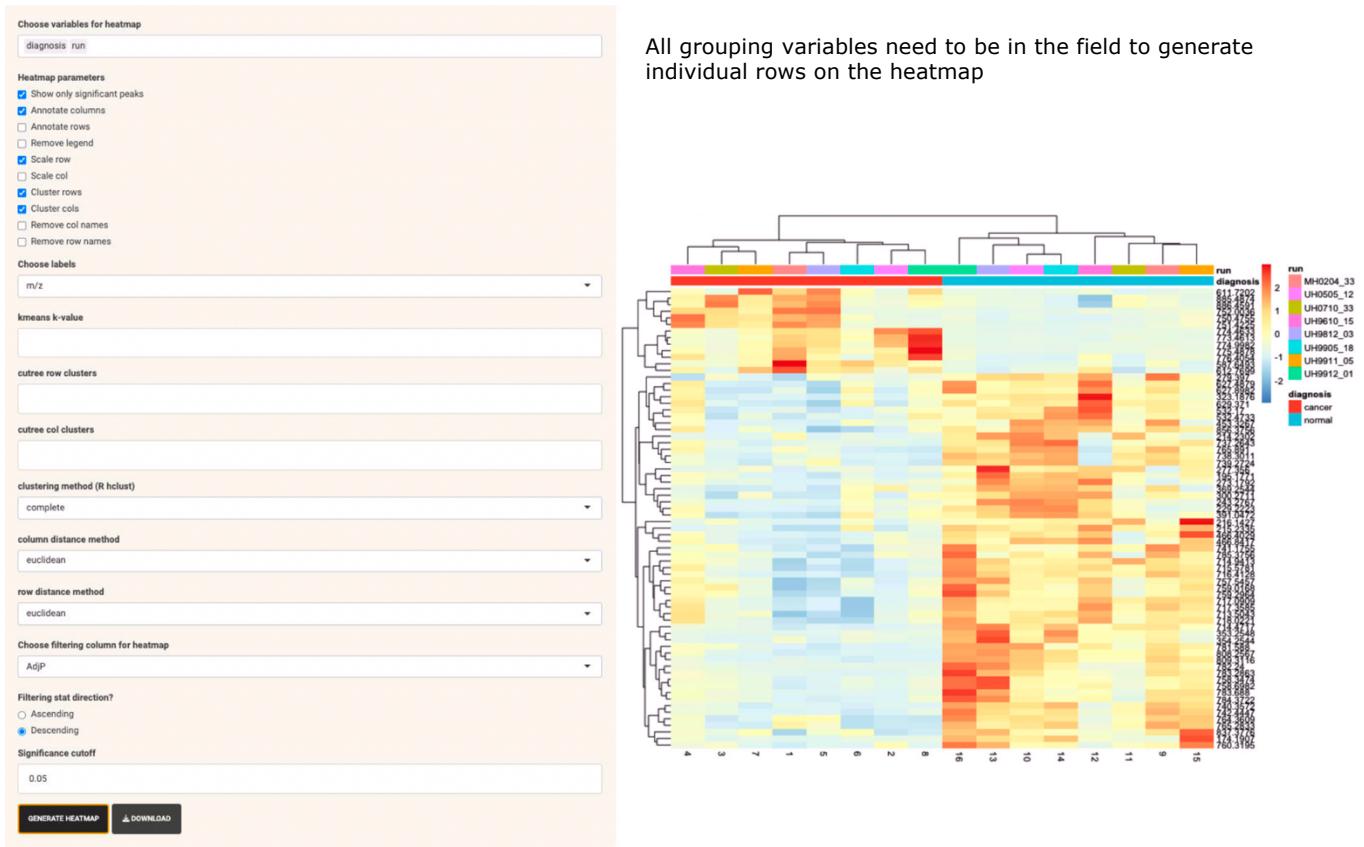
When selecting a feature, the means test generates box-and-whisker plots to visualize the differences between the conditions.

The means test model (and all other models in this tab) can be saved using the save button and reopened from a saved R.Data file in this tab.

# Heatmap

The purpose of this tab is to generate heatmaps based on statistical analysis performed on MSI data from the Statistics tab.

To reproduce the demo data as produced in this manual, below are the following parameters:



# Colocalization Analysis

The purpose of this tab is to visualize colocalization between features.

To reproduce the demo data as produced in this manual, below are the following parameters:



The console will generate a table of the top ten colocalized ions.

## Colocalized features:

	mz	correlation	M1	M2
1	716.4128	1.0000000	1.0000000	1.0000000
2	760.3195	0.4391020	0.6237656	0.6239710
3	742.4447	0.4334985	0.6171824	0.6173856
4	764.3609	0.4249884	0.6247531	0.6247531
5	782.2400	0.4236315	0.6323239	0.6323239
6	738.3011	0.4180463	0.6237656	0.6237656
7	714.4717	0.3943332	0.6263989	0.6263989
8	808.2567	0.3757609	0.6369322	0.6369322
9	740.3572	0.3661958	0.6175115	0.6175115
10	353.2548	0.3487755	0.6520737	0.6520737

## References

- [1] Anal. Chem. 2019, 91, 9, 5706–5714. Publication Date: April 2019. <https://doi.org/10.1021/acs.analchem.8b05827>
- [2] MCP. 2016, 15, 5, 1761-1772. Publication Date: May 2016. <https://doi.org/10.1074/mcp.O115.053918>