MetaboAnalyst By Du-Lab

Overview

MetaboAnalyst is a comprehensive platform dedicated for metabolomics data analysis via user-friendly, web-based interface. Over the past decade, MetaboAnalyst has evolved to become the most widely used platform (>300,000 users) in the metabolomics community. The current MetaboAnalyst (V5.0) supports raw MS spectra processing, comprehensive data normalization, statistical analysis, functional analysis, meta-analysis as well as integrative analysis with other omics data.
The objective is to enable high-throughput analysis for both targeted and untargeted metabolomics, and to narrow the gap from raw spectra to biological insights.

Website Link

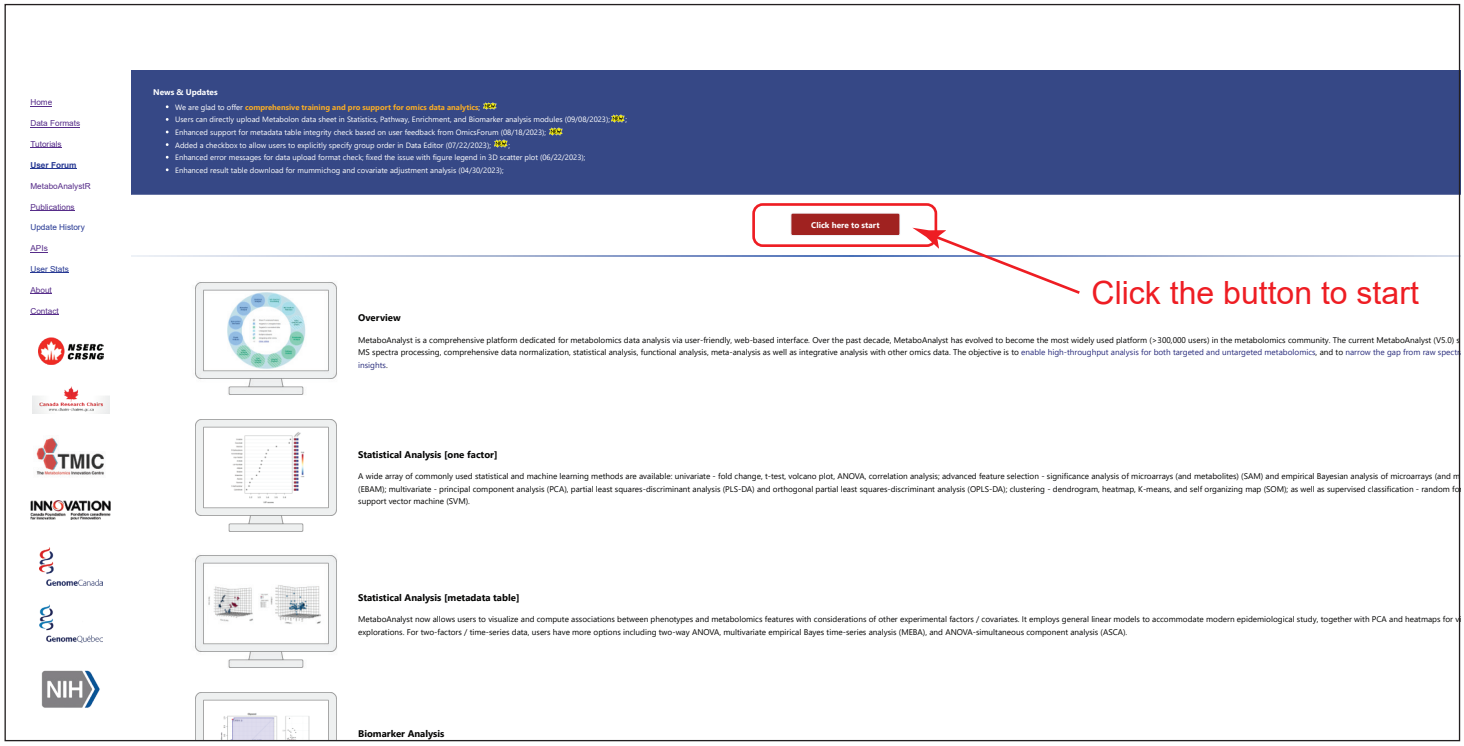
http://www.metaboanalyst.ca/

Operate Procedure

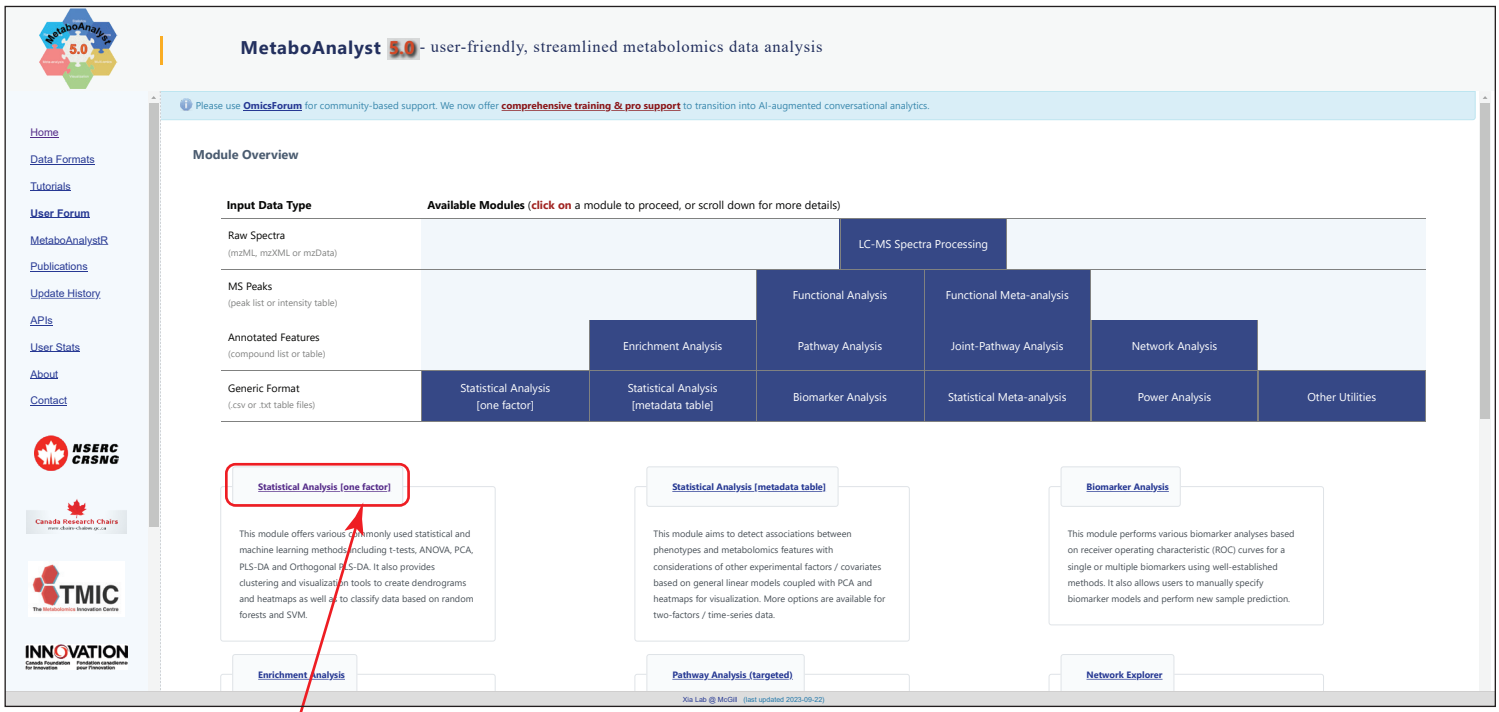
1. Format the data file
   a. Make a format in Excel as below (if two groups). For exemple, there are two groups (Apical and Basal)

Sample	M1 1A	M1 2A	M1 3A	M1 1B	M1 2B	M1 3B
Label	Apical	Apical	Apical	Basal	Basal	Basal
Proline	59844441	33791551	23474213	119517656	136735841	128762485
Allantoin	19858	34695	29799	71271	62978	48100
4-Hydroxybutyrate	2185475	1986435	2820447	3893676	2066648	3473576
Ribose-5-P	167070	197278	139778	195885	206918	130071
Acetylcarnitine	48707	62685	56021	51924	26663	41485
Homoserine	100250	102629	112778	151527	166205	156264
Folic Acid	15674	10762	15763	10833	21809	11802
Arachidonate	61729	66610	73376	79207	68003	54012
isoValeric Acid	796251	800411	767928	919306	1036717	872882
Homovanilate	731392	667279	585246	794879	765786	663936
F16BP/F26BP/G16BP	42954	40554	48960	54221	67125	55755
Choline	22743223	21776263	19723315	25825552	30454783	25465297
OH-Phenylpyruvate	63157	56977	51205	50963	69219	58001

b. Save the Excel file as CSV (Comma Delimited)

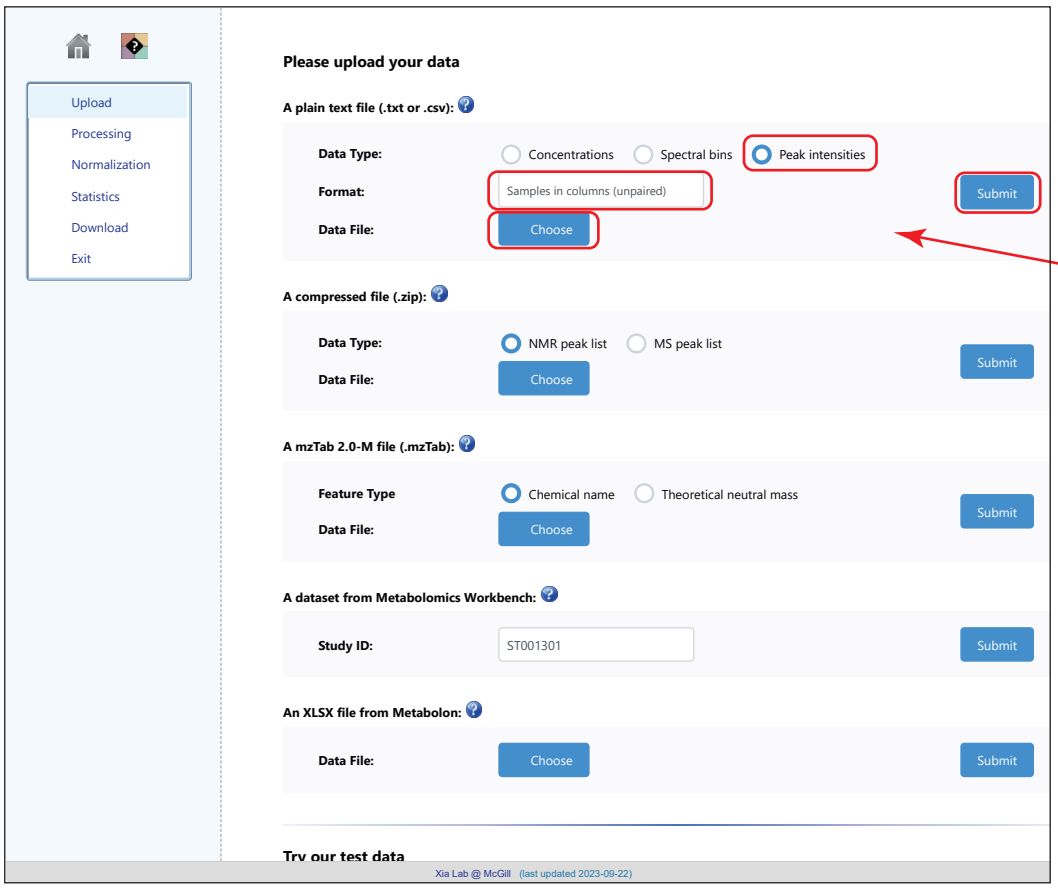


2. Select Statistical Analysis



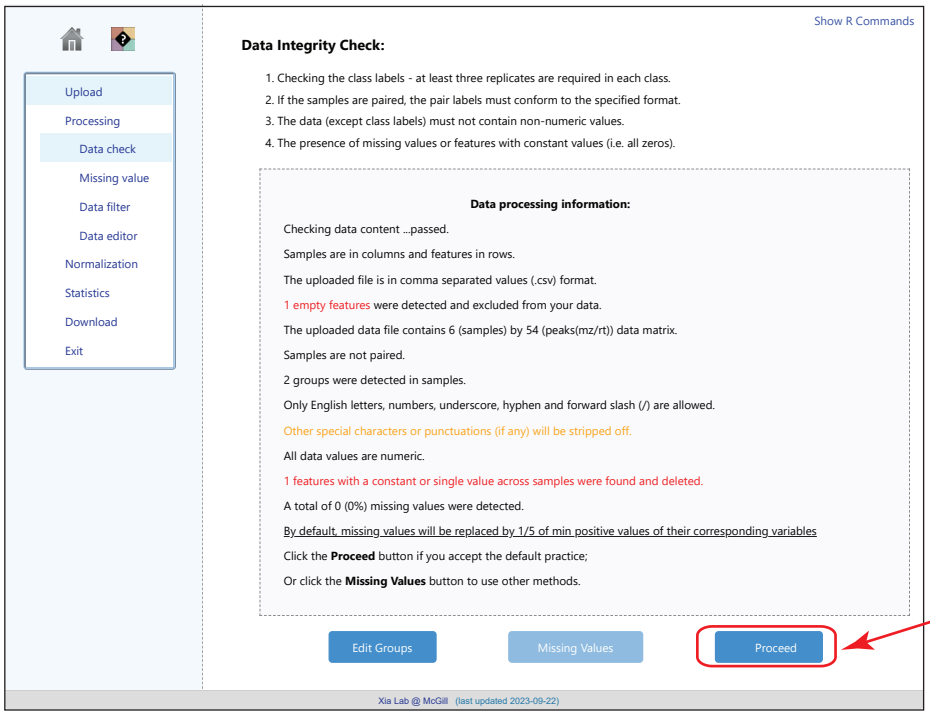
Click “Statistical Analysis” button

3. Upload the Data

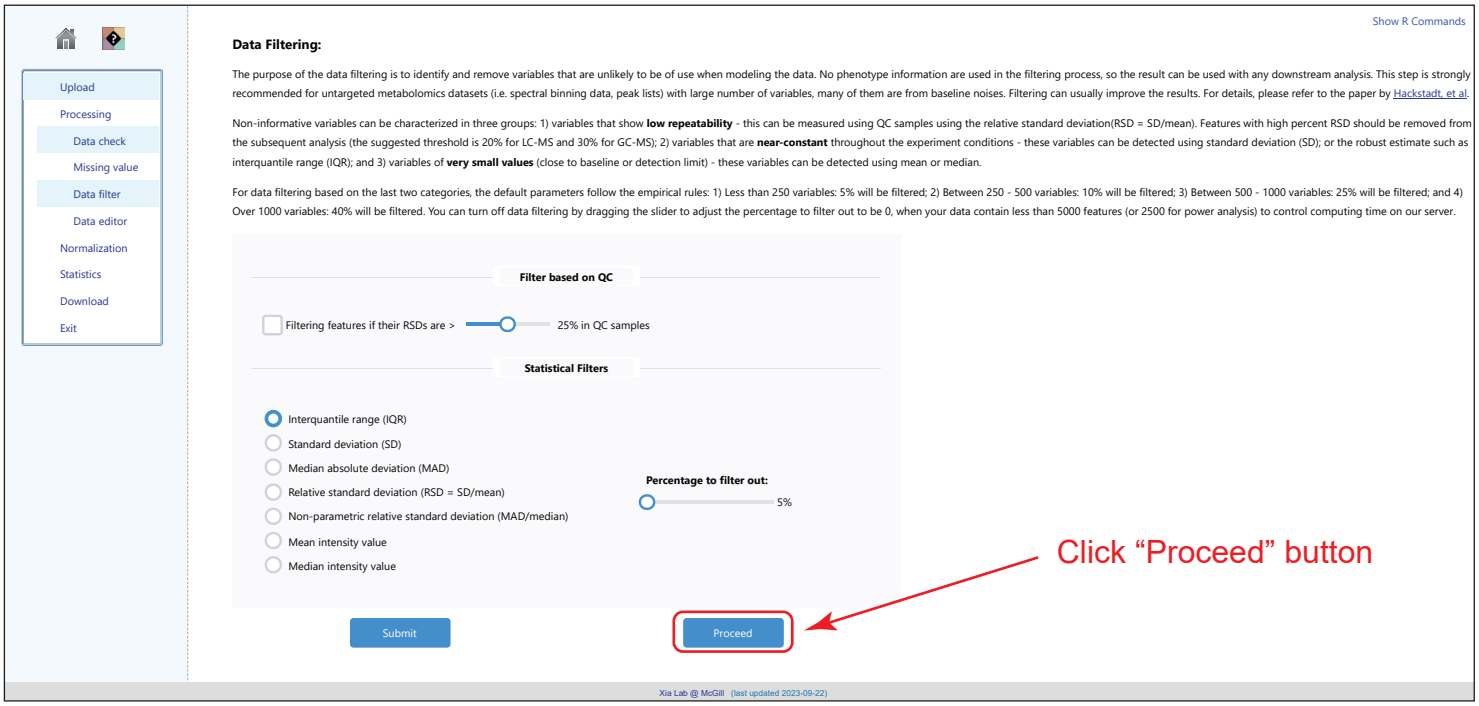


Choose“Peak intensity table” in data type, “Samples in columns (unpaired)” in format dropdown list, and find the .csv file by click the “Choose” button. Then click “Submit”

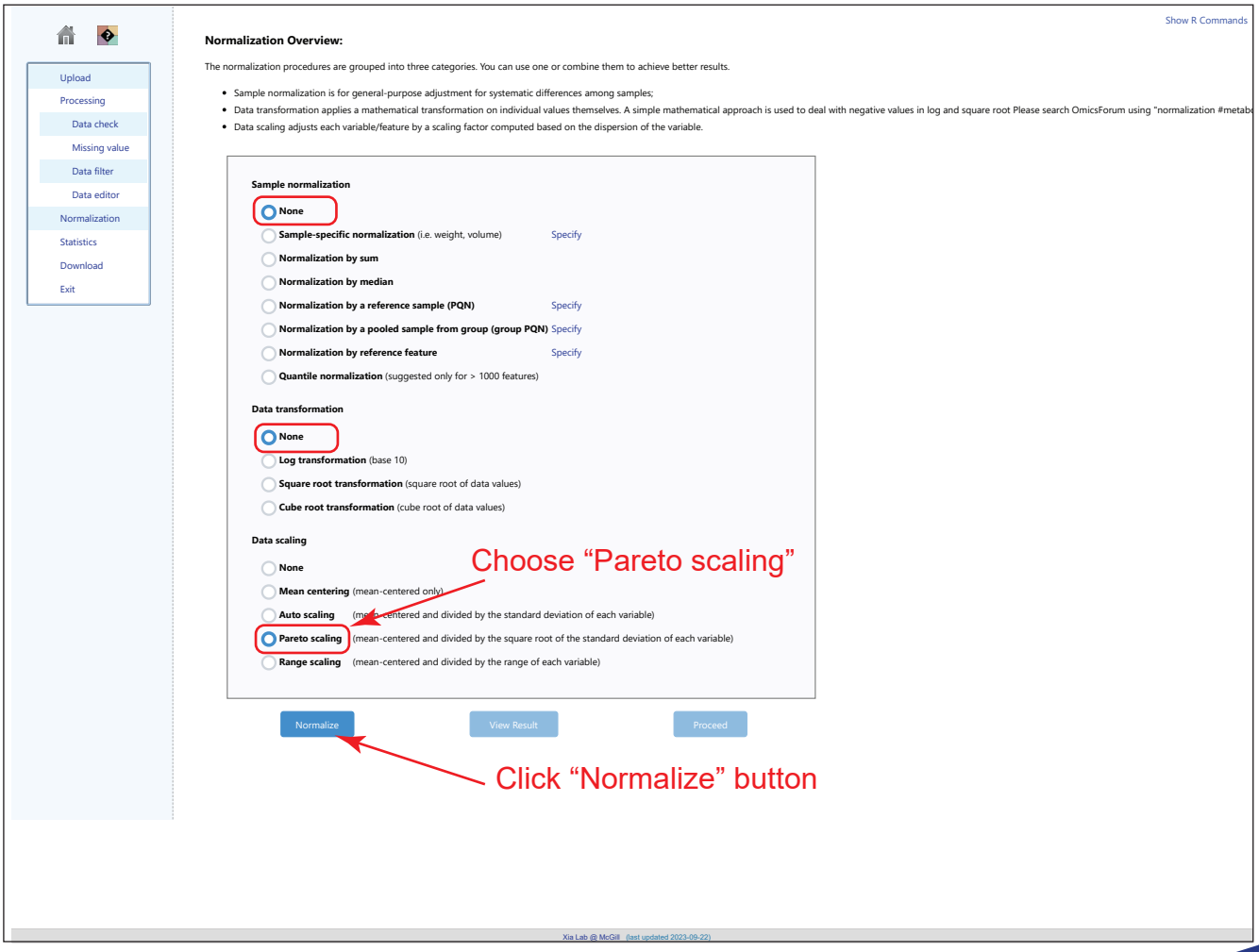
4. Click Proceed button if you accept the default practice.

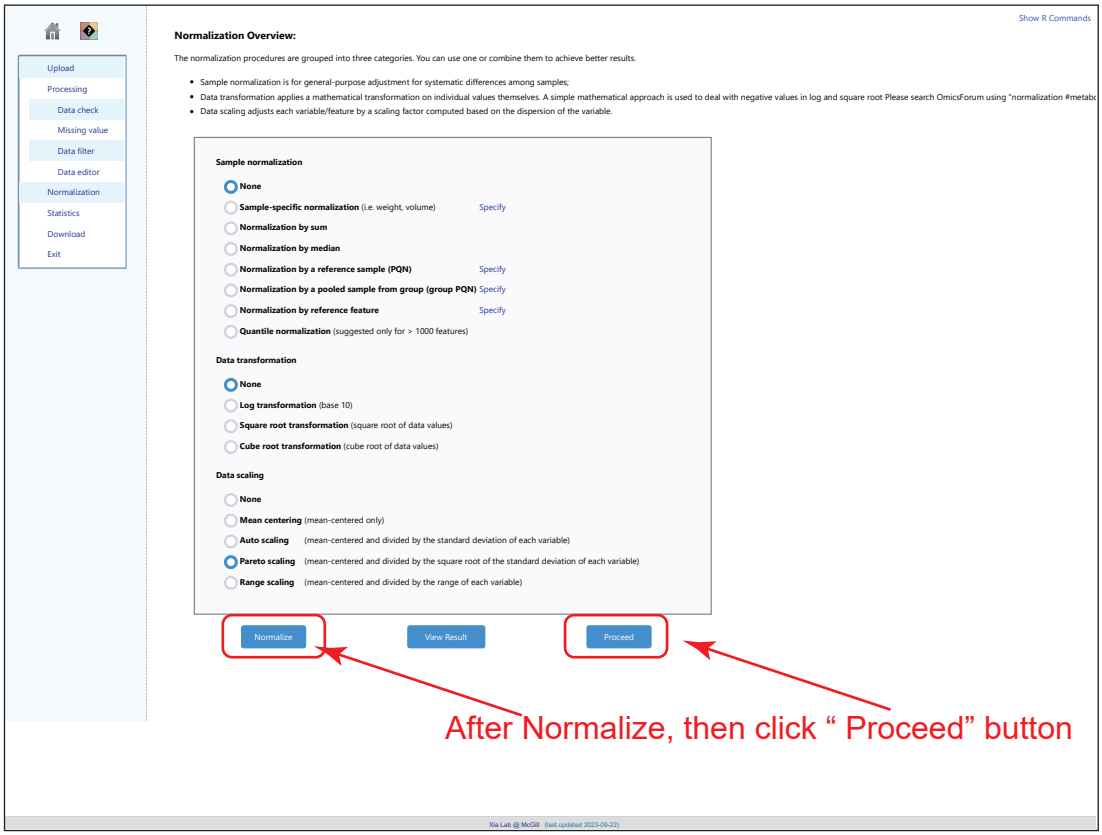


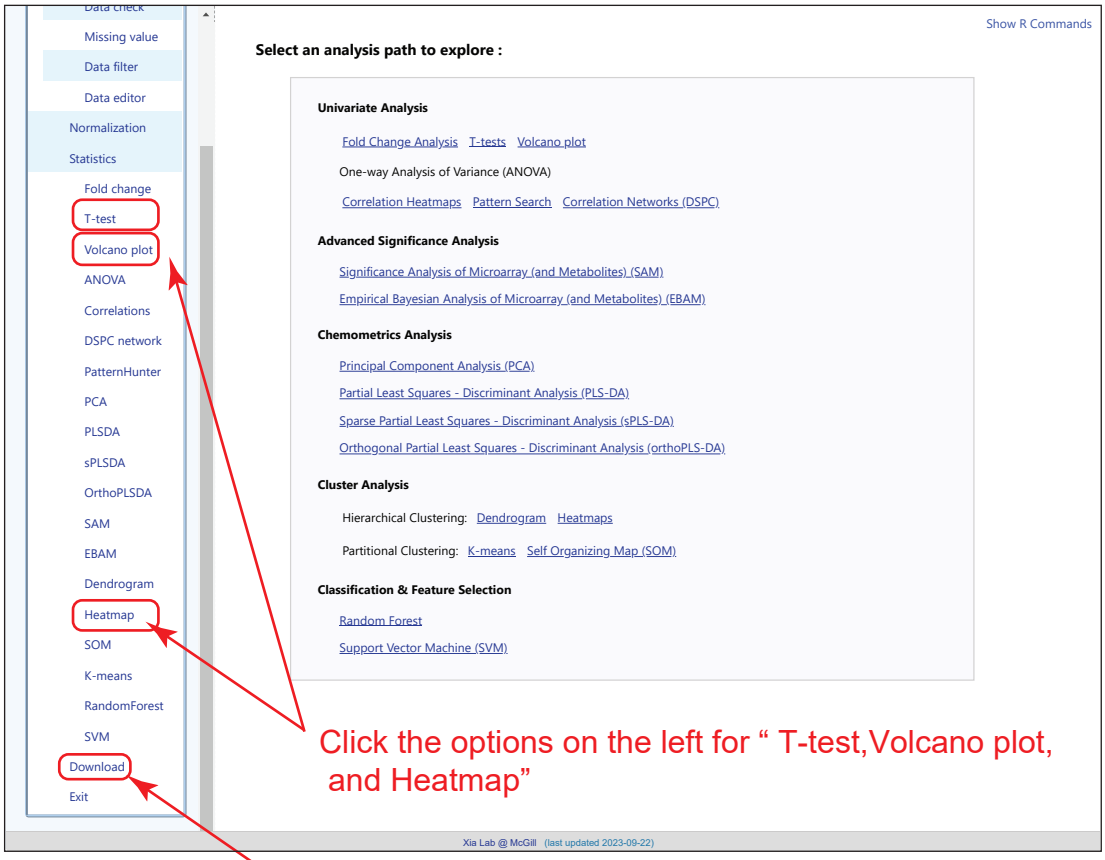
Click “Proceed” button

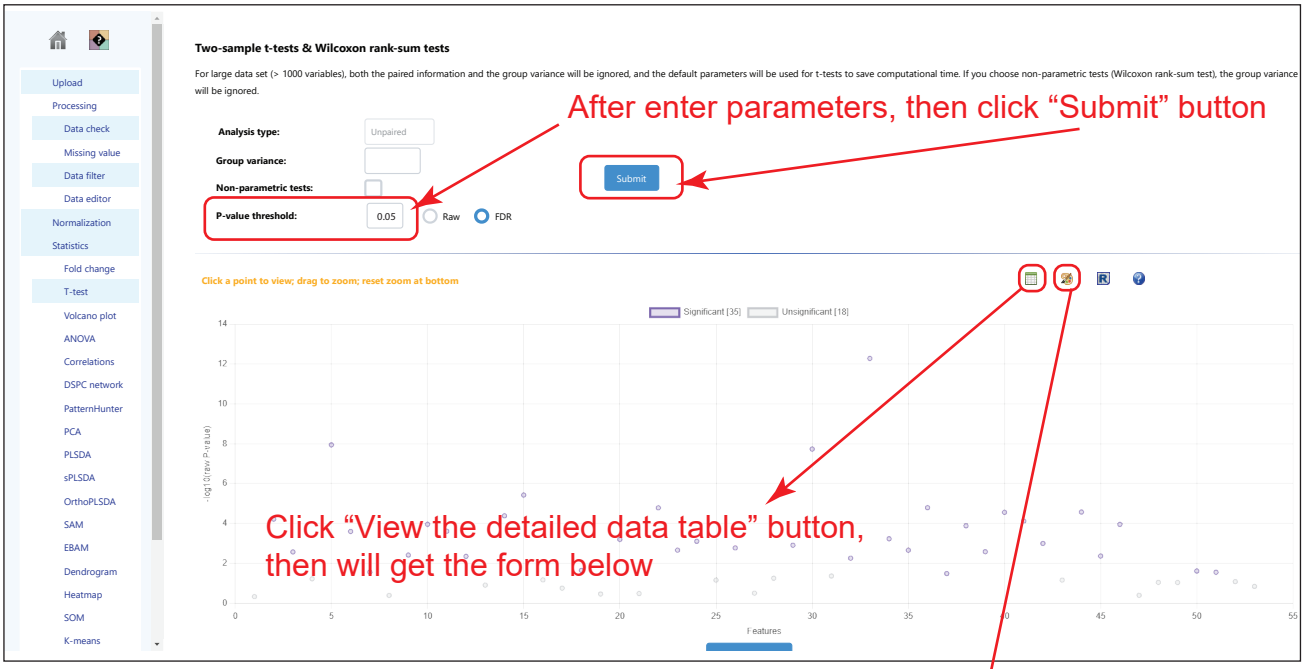


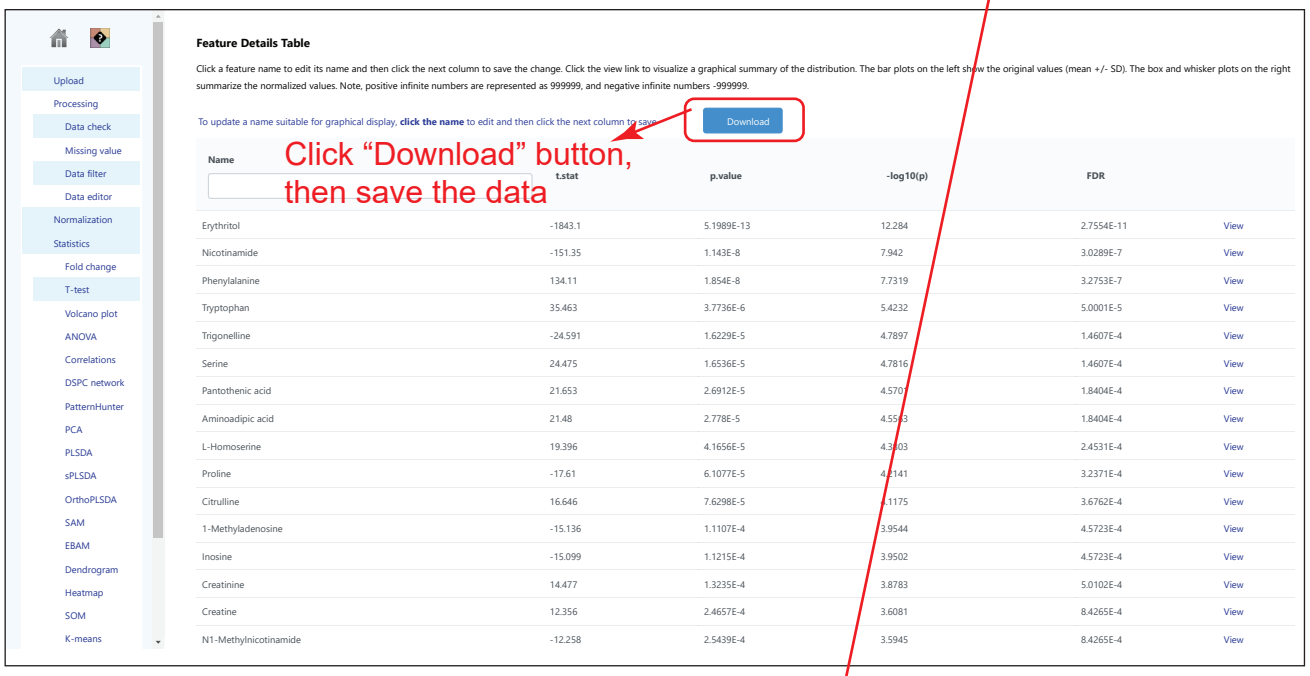
5. Normalization overview

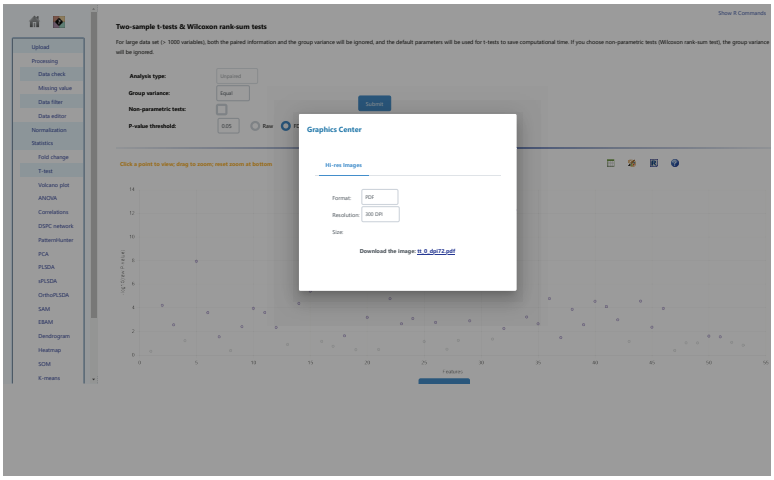


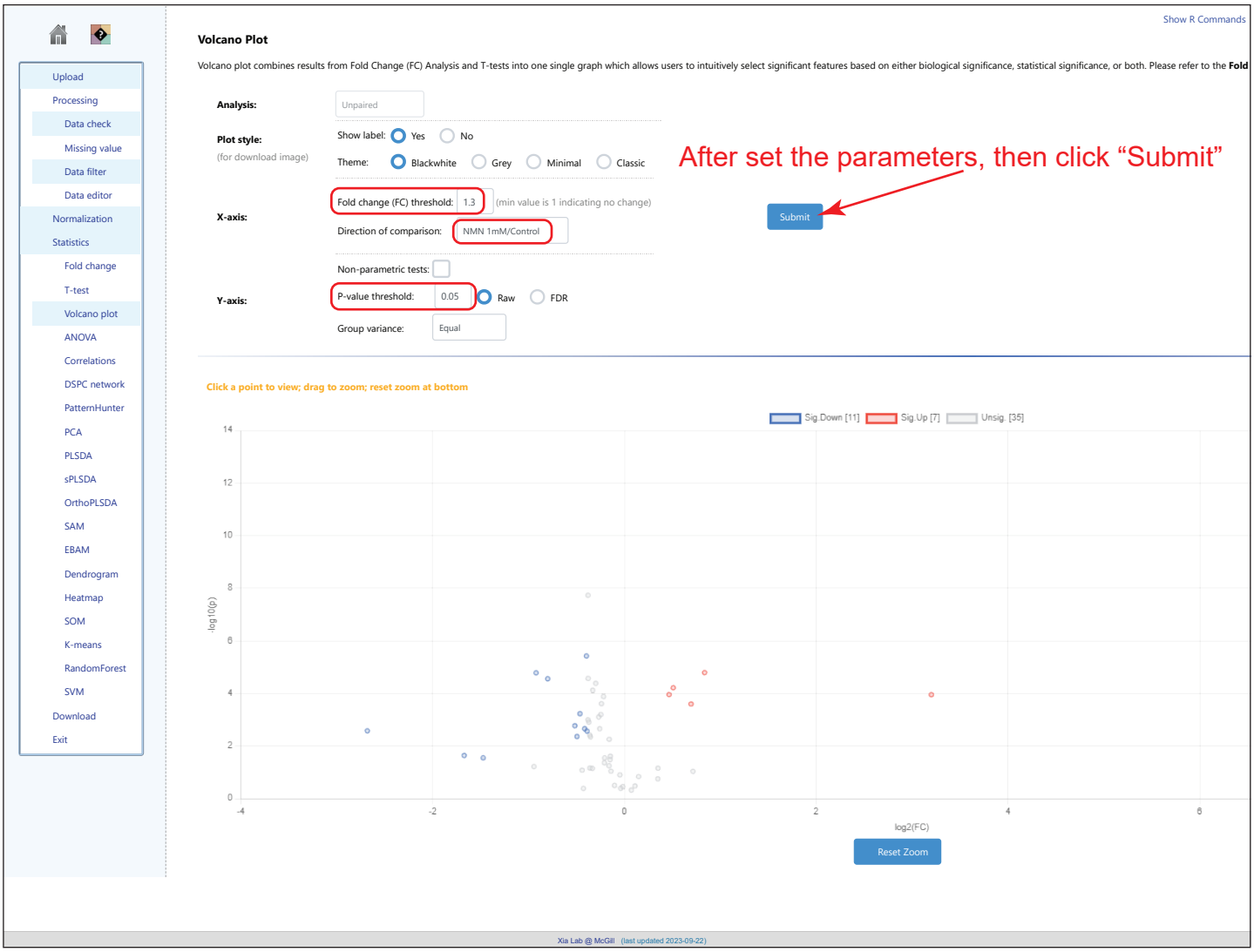

Show R Commands
Upload
Processing
Data check
Missing value
Data filter
Data editor
Normalization
Statistics
Download
Exit
Normalization Overview:
The normalization procedures are grouped into three categories. You can use one or combine them to achieve better results.
Sample normalization is for general-purpose adjustment for systematic differences among samples;
Data transformation applies a mathematical transformation on individual values themselves. A simple mathematical approach is used to deal with negative values in log and square root Please search OmicsForum using "normalization #metabo
Data scaling adjusts each variable/feature by a scaling factor computed based on the dispersion of the variable.
Sample normalization
None
Sample-specific normalization (i.e. weight, volume) Specify
Normalization by sum
Normalization by median
Normalization by a reference sample (PQN) Specify
Normalization by a pooled sample from group (group PQN) Specify
Normalization by reference feature Specify
Quantile normalization (suggested only for > 1000 features)
Data transformation
None
Log transformation (base 10)
Square root transformation (square root of data values)
Cube root transformation (cube root of data values)
Data scaling
None
Mean centering (mean-centered only)
Auto scaling (mean-centered and divided by the standard deviation of each variable)
Pareto scaling (mean-centered and divided by the square root of the standard deviation of each variable)
Range scaling (mean-centered and divided by the range of each variable)
Normalize
View Result
Proceed
After Normalize, then click " Proceed" button
Xia Lab @ McGill (last updated 2023-09-22)

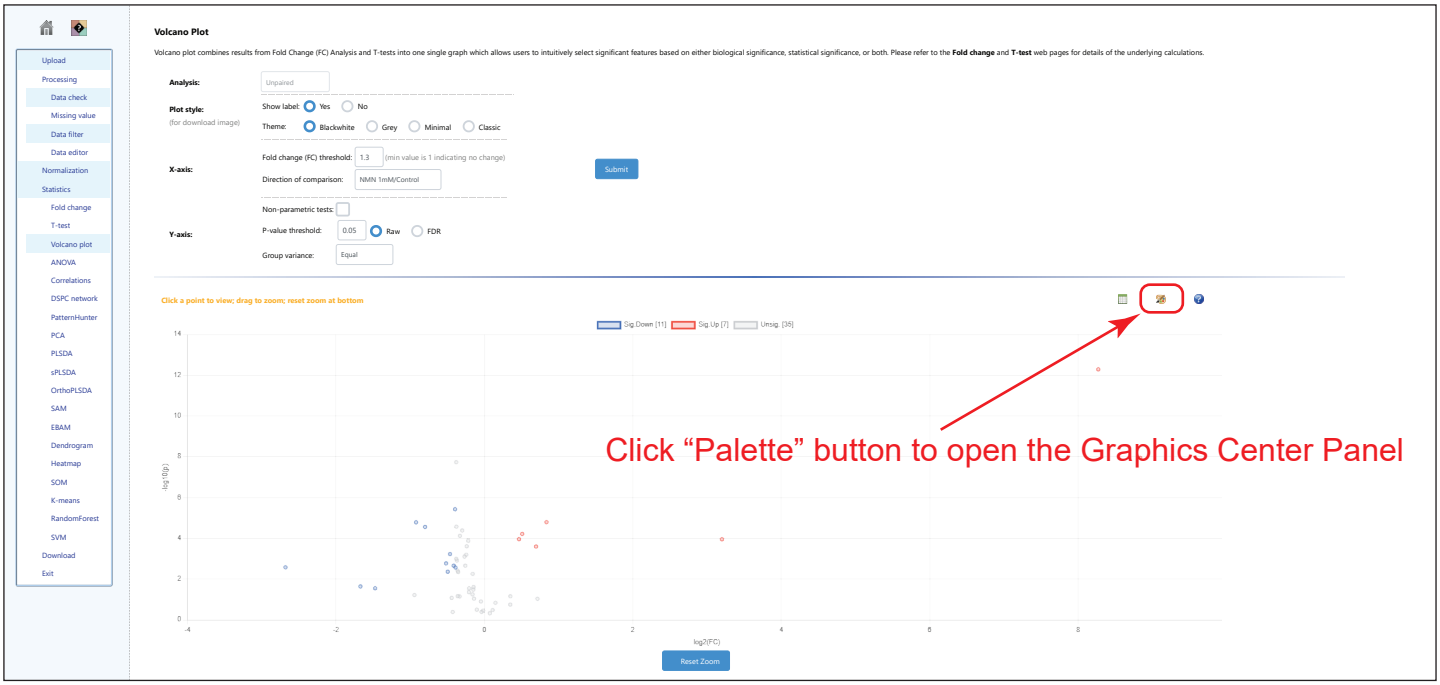

Missing value
Data filter
Data editor
Normalization
Statistics
Fold change
T-test
Volcano plot
ANOVA
Correlations
DSPC network
PatternHunter
PCA
PLSDA
sPLSDA
OrthoPLSDA
SAM
EBAM
Dendrogram
Heatmap
SOM
K-means
RandomForest
SVM
Download
Exit
Show R Commands
Select an analysis path to explore :
Univariate Analysis
Fold Change Analysis T-tests Volcano plot
One-way Analysis of Variance (ANOVA)
Correlation Heatmaps Pattern Search Correlation Networks (DSPC)
Advanced Significance Analysis
Significance Analysis of Microarray (and Metabolites) (SAM)
Empirical Bayesian Analysis of Microarray (and Metabolites) (EBAM)
Chemometrics Analysis
Principal Component Analysis (PCA)
Partial Least Squares - Discriminant Analysis (PLS-DA)
Sparse Partial Least Squares - Discriminant Analysis (sPLS-DA)
Orthogonal Partial Least Squares - Discriminant Analysis (orthoPLS-DA)
Cluster Analysis
Hierarchical Clustering: Dendrogram Heatmaps
Partitional Clustering: K-means Self Organizing Map (SOM)
Classification & Feature Selection
Random Forest
Support Vector Machine (SVM)
Click the options on the left for " T-test,Volcano plot, and Heatmap"
or Click "download" button to download .zip for all the report.
Xia Lab @ McGill (last updated 2023-09-22)

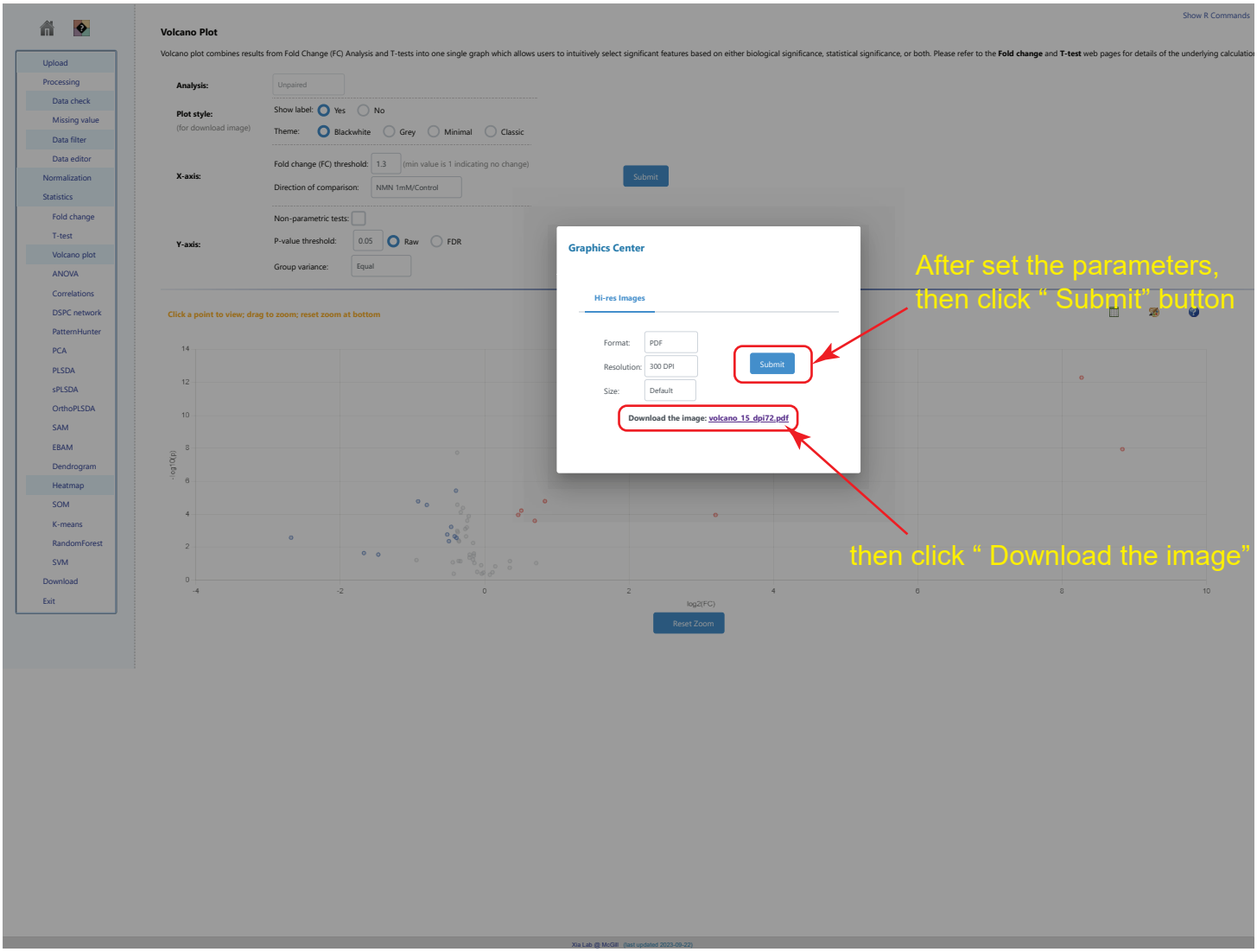

Two-sample t-tests & Wilcoxon rank-sum tests
For large data set (> 1000 variables), both the paired information and the group variance will be ignored, and the default parameters will be used for t-tests to save computational time. If you choose non-parametric tests (Wilcoxon rank-sum test), the group variance will be ignored.
After enter parameters, then click “Submit” button
Analysis type:
Unpaired
Group variance:
Non-parametric tests:
P-value threshold:
0.05
Raw
FDR
Submit
Click a point to view; drag to zoom; reset zoom at bottom
Significant [35]
Unsignificant [18]
Click “View the detailed data table” button, then will get the form below
Upload
Processing
Data check
Missing value
Data filter
Data editor
Normalization
Statistics
Fold change
T-test
Volcano plot
ANOVA
Correlations
DSPC network
PatternHunter
PCA
PLSDA
sPLSDA
OrthoPLSDA
SAM
EBAM
Dendrogram
Heatmap
SOM
K-means

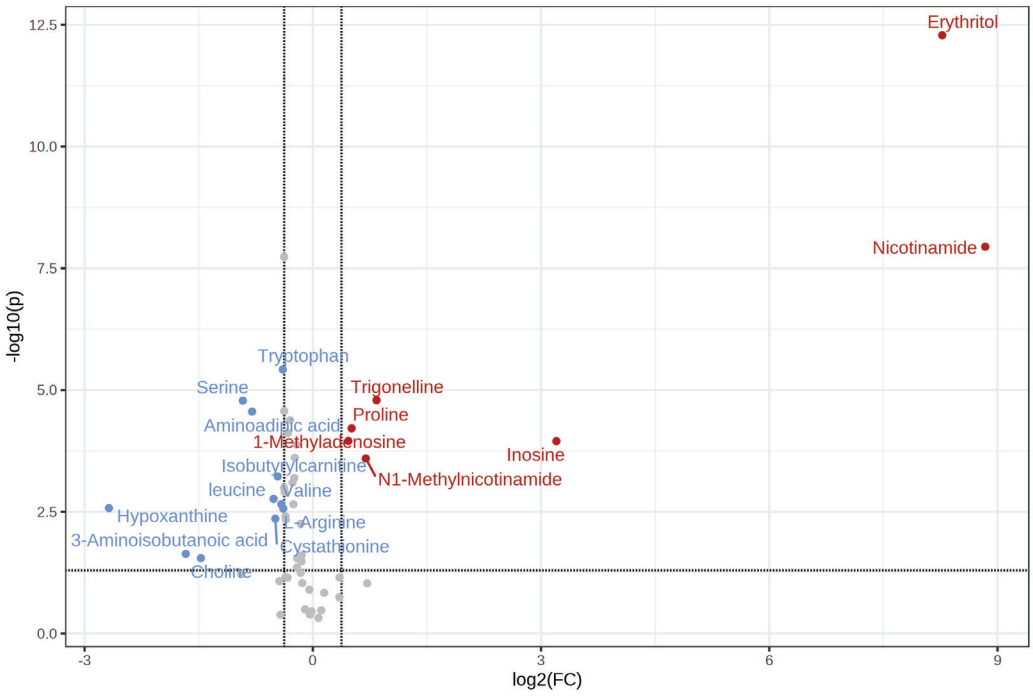

Feature Details Table
Click a feature name to edit its name and then click the next column to save the change. Click the view link to visualize a graphical summary of the distribution. The bar plots on the left show the original values (mean +/- SD). The box and whisker plots on the right summarize the normalized values. Note, positive infinite numbers are represented as 999999, and negative infinite numbers -999999.
To update a name suitable for graphical display, click the name to edit and then click the next column to save
Download
Click “Download” button, then save the data
Name
t.stat
p.value
-log10(p)
FDR
Erythritol -1843.1 5.1989E-13 12.284 2.7554E-11 View
Nicotinamide -151.35 1.143E-8 7.942 3.0289E-7 View
Phenylalanine 134.11 1.854E-8 7.7319 3.2753E-7 View
Tryptophan 35.463 3.7736E-6 5.4232 5.0001E-5 View
Trigonelline -24.591 1.6229E-5 4.7897 1.4607E-4 View
Serine 24.475 1.6536E-5 4.7816 1.4607E-4 View
Pantothenic acid 21.653 2.6912E-5 4.570 1.8404E-4 View
Aminoadipic acid 21.48 2.778E-5 4.553 1.8404E-4 View
L-Homoserine 19.396 4.1656E-5 4.3803 2.4531E-4 View
Proline -17.61 6.1077E-5 4.2141 3.2371E-4 View
Citrulline 16.646 7.6298E-5 4.1175 3.6762E-4 View
1-Methyladenosine -15.136 1.1107E-4 3.9544 4.5723E-4 View
Inosine -15.099 1.1215E-4 3.9502 4.5723E-4 View
Creatinine 14.477 1.3235E-4 3.8783 5.0102E-4 View
Creatine 12.356 2.4657E-4 3.6081 8.4265E-4 View
N1-Methylnicotinamide -12.258 2.5439E-4 3.5945 8.4265E-4 View


Graphics Center
Hi-res Images
Format:
Resolution:
Size:
Click “Palette” button, then will get “Customize the graphics output”


Show R Commands
Volcano Plot
Volcano plot combines results from Fold Change (FC) Analysis and T-tests into one single graph which allows users to intuitively select significant features based on either biological significance, statistical significance, or both. Please refer to the Fold
Upload
Processing
Data check
Missing value
Data filter
Data editor
Normalization
Statistics
Fold change
T-test
Volcano plot
ANOVA
Correlations
DSPC network
PatternHunter
PCA
PLSDA
sPLSDA
OrthoPLSDA
SAM
EBAM
Dendrogram
Heatmap
SOM
K-means
RandomForest
SVM
Download
Exit
Analysis:
Unpaired
Plot style:
(for download image)
Show label: Yes No
Theme: Blackwhite Grey Minimal Classic
X-axis:
Fold change (FC) threshold: 1.3
(min value is 1 indicating no change)
Direction of comparison: NMN 1mM/Control
Non-parametric tests:
Y-axis:
P-value threshold: 0.05
Raw FDR
Group variance: Equal
After set the parameters, then click “Submit”
Submit
Click a point to view; drag to zoom; reset zoom at bottom
Sig.Down [11] Sig.Up [7] Unsig. [35]
log10(p)
log2(FC)
Reset Zoom
Xia Lab @ McGill (last updated 2023-09-22)

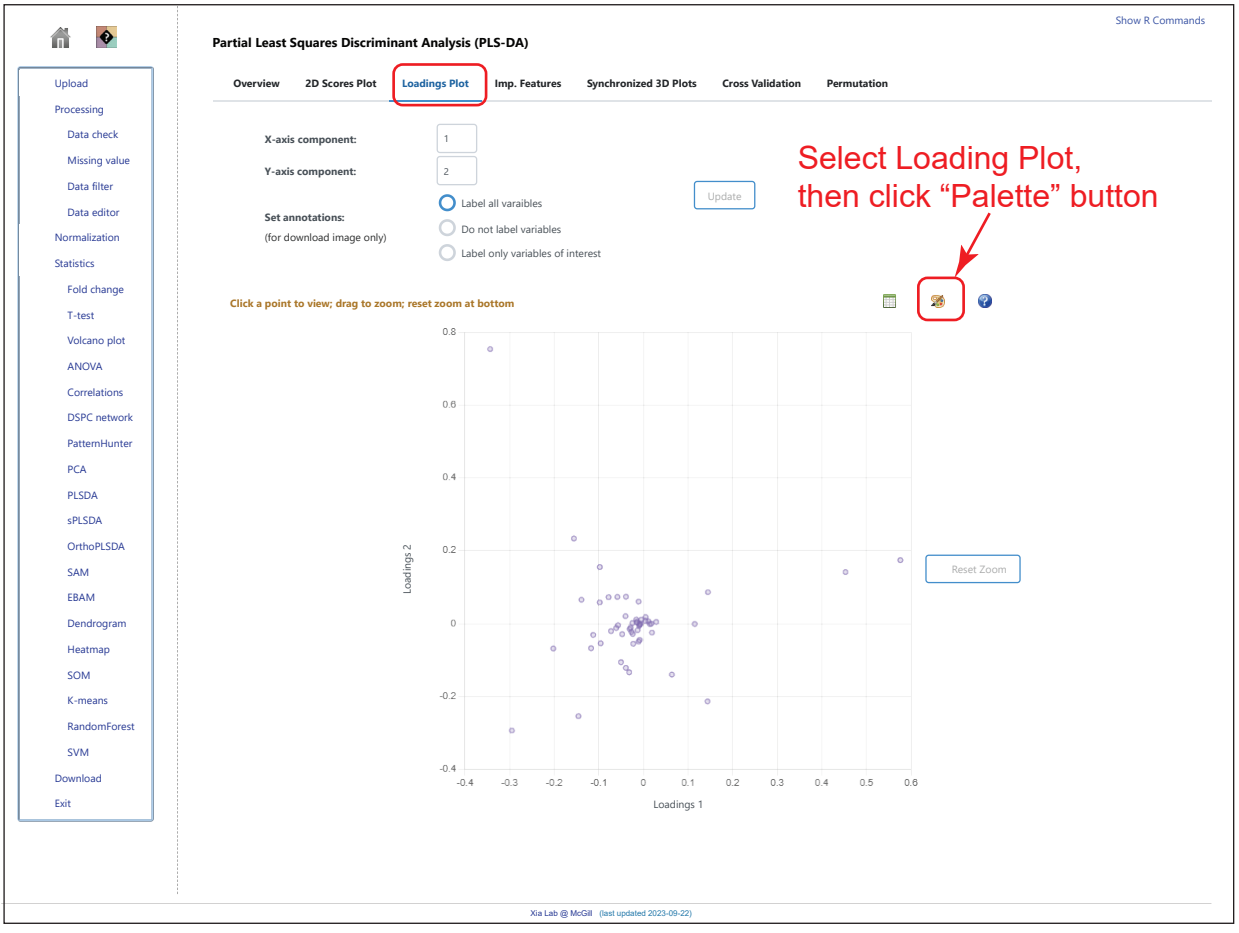

Volcano Plot
Volcano plot combines results from Fold Change (FC) Analysis and T-tests into one single graph which allows users to intuitively select significant features based on either biological significance, statistical significance, or both. Please refer to the Fold change and T-test web pages for details of the underlying calculations.
Analysis:
Plot style:
X-axis:
Y-axis:
Submit
Click a point to view; drag to zoom; reset zoom at bottom
Click “Palette” button to open the Graphics Center Panel
Reset Zoom

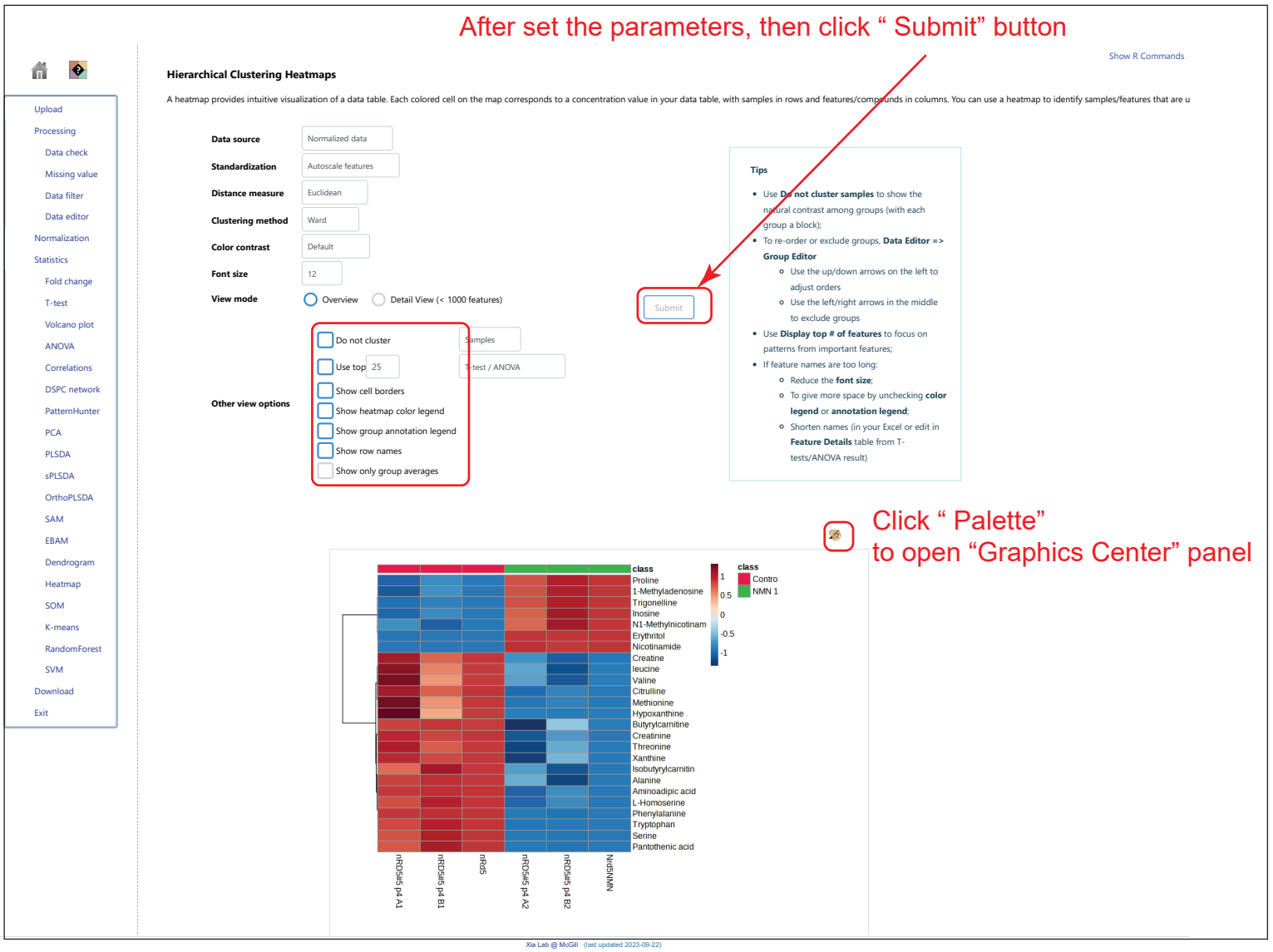


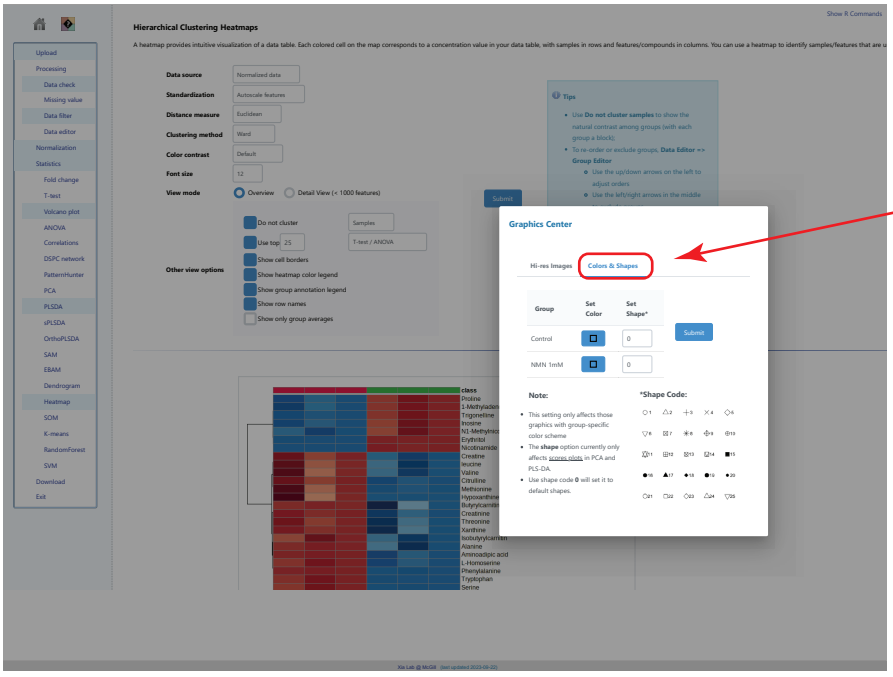


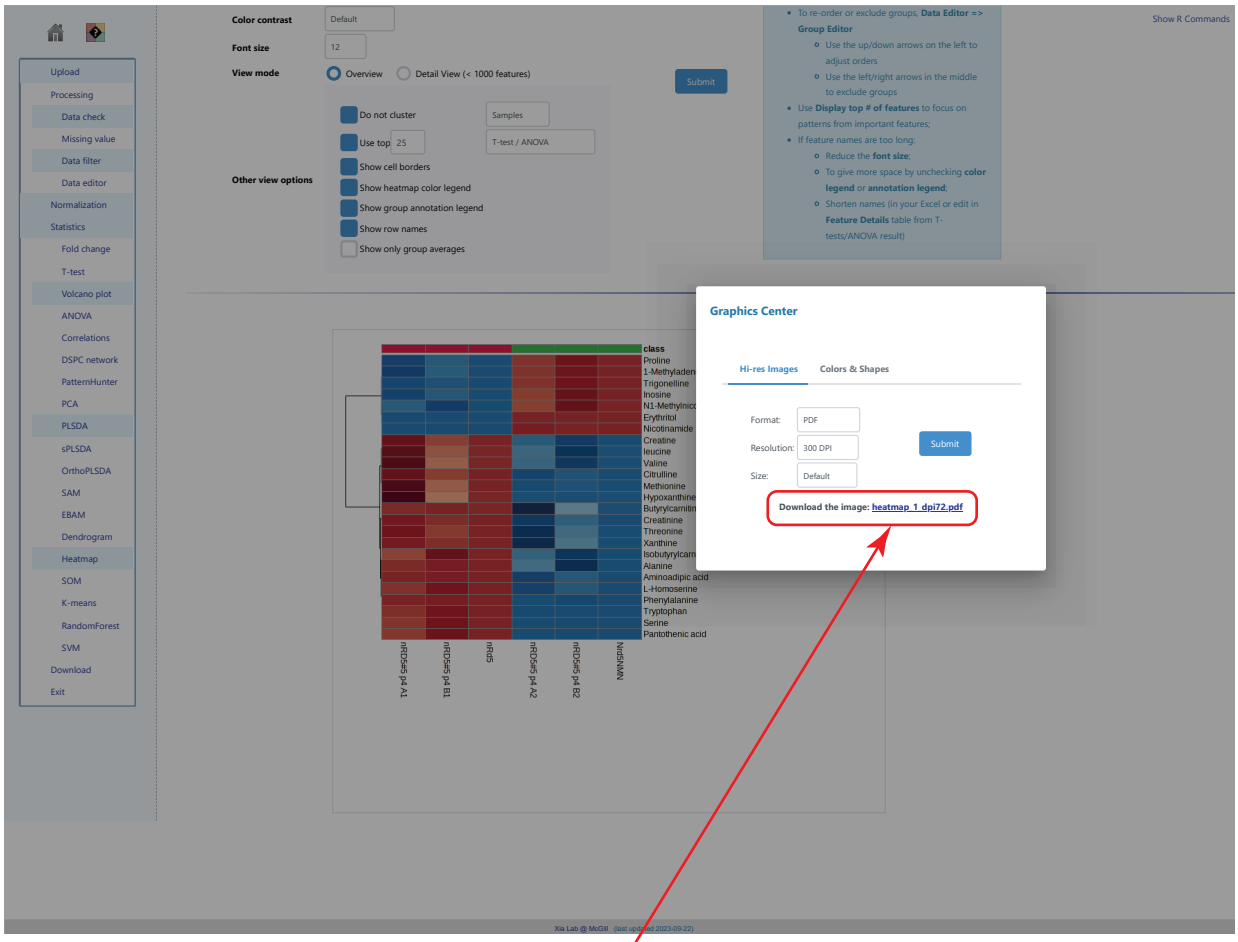
Volcano plot image


Show R Commands
Partial Least Squares Discriminant Analysis (PLS-DA)
Upload
Processing
Data check
Missing value
Data filter
Data editor
Normalization
Statistics
Fold change
T-test
Volcano plot
ANOVA
Correlations
DSPC network
PatternHunter
PCA
PLSDA
sPLSDA
OrthoPLSDA
SAM
EBAM
Dendrogram
Heatmap
SOM
K-means
RandomForest
SVM
Download
Exit
Overview
2D Scores Plot
Loadings Plot
Imp. Features
Synchronized 3D Plots
Cross Validation
Permutation
Select component for X-axis:
Select component for Y-axis:
Display 95% confidence region:
Display sample names:
Use grey-scale colors:
Update
Select 2D Scores Plot,
then click “Palette” button
Scores Plot
Control
NMN 1mM
Component 2 (7.6 %)
Component 1 (88.2 %)
Note, PLS-DA maximizes the covariance between X (data) and Y (group). The variance displayed in the plot above is the explained variance for X. Covariance and x-variance may r
2nd component.
Xia Lab @ McGill (last updated 2023-09-22)


Show R Commands
Partial Least Squares Discriminant Analysis (PLS-DA)
Upload
Processing
Data check
Missing value
Data filter
Data editor
Normalization
Statistics
Fold change
T-test
Volcano plot
ANOVA
Correlations
DSPC network
PatternHunter
PCA
PLSDA
sPLSDA
OrthoPLSDA
SAM
EBAM
Dendrogram
Heatmap
SOM
K-means
RandomForest
SVM
Download
Exit
Overview
2D Scores Plot
Loadings Plot
Imp. Features
Synchronized 3D Plots
Cross Validation
Permutation
X-axis component:
Y-axis component:
Set annotations:
(for download image only)
Label all varaibles
Do not label variables
Label only variables of interest
Update
Select Loading Plot,
then click “Palette” button
Click a point to view; drag to zoom; reset zoom at bottom
Reset Zoom
Loadings 2
Loadings 1
Xia Lab @ McGill (last updated 2023-09-22)


After set the parameters, then click " Submit" button
Show R Commands
Hierarchical Clustering Heatmaps
A heatmap provides intuitive visualization of a data table. Each colored cell on the map corresponds to a concentration value in your data table, with samples in rows and features/compounds in columns. You can use a heatmap to identify samples/features that are u
Upload
Processing
Data check
Missing value
Data filter
Data editor
Normalization
Statistics
Fold change
T-test
Volcano plot
ANOVA
Correlations
DSPC network
PatternHunter
PCA
PLSDA
sPLSDA
OrthoPLSDA
SAM
EBAM
Dendrogram
Heatmap
SOM
K-means
RandomForest
SVM
Download
Exit
Data source
Normalized data
Standardization
Autoscale features
Distance measure
Euclidean
Clustering method
Ward
Color contrast
Default
Font size
12
View mode
Overview
Detail View (< 1000 features)
Do not cluster
Samples
Use top
25
T-test / ANOVA
Other view options
Show cell borders
Show heatmap color legend
Show group annotation legend
Show row names
Show only group averages
Submit
Tips
Use Do not cluster samples to show the natural contrast among groups (with each group a block);
To re-order or exclude groups, Data Editor => Group Editor
Use the up/down arrows on the left to adjust orders
Use the left/right arrows in the middle to exclude groups
Use Display top # of features to focus on patterns from important features;
If feature names are too long:
Reduce the font size;
To give more space by unchecking color legend or annotation legend;
Shorten names (in your Excel or edit in Feature Details table from T-tests/ANOVA result)
Click " Palette"
to open "Graphics Center" panel
class
Proline
1-Methyladenosine
Trigonelline
Inosine
N1-Methylnicotinam
Erythritol
Nicotinamide
Creatine
leucine
Valine
Citrulline
Methionine
Hypoxanthine
Butyrylcarnitine
Creatinine
Threonine
Xanthine
Isobutyrylcarnitin
Alanine
Aminoadipic acid
L-Homoserine
Phenylalanine
Tryptophan
Serine
Pantothenic acid
1
0.5
0
-0.5
-1
Contro
NMN 1
nRD5#5 p4 A1
nRD5#5 p4 B1
nRd5
nRD5#5 p4 A2
nRD5#5 p4 B2
NRd5NMN
Xia Lab @ McGill (last updated 2023-09-22)


Graphics Center
Hi-res Images
Colors & Shapes
Group
Set Color
Set Shape*
Control
NMN 1mM
Submit
Note:
*Shape Code:
At Graphics Center panel,
Colors & Shape can be edited.



Click “ Download the image”

6. Back to the home page

Click the “House” icon on the left top and go the home page.

For enrichment and pathway analysis, just paste the changes in T-test.