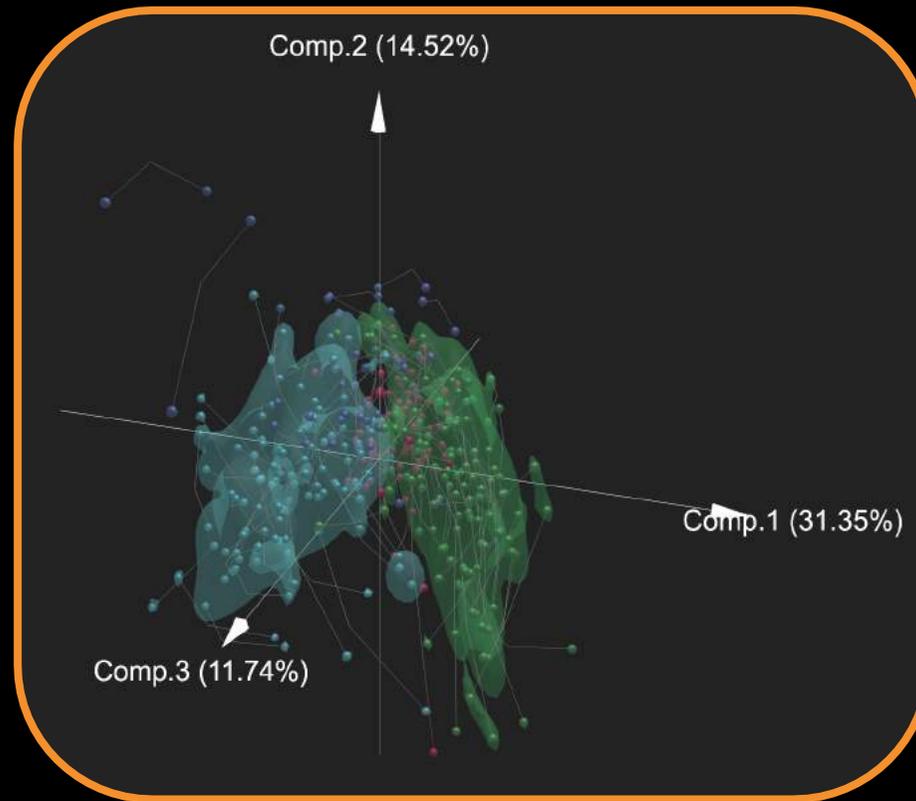


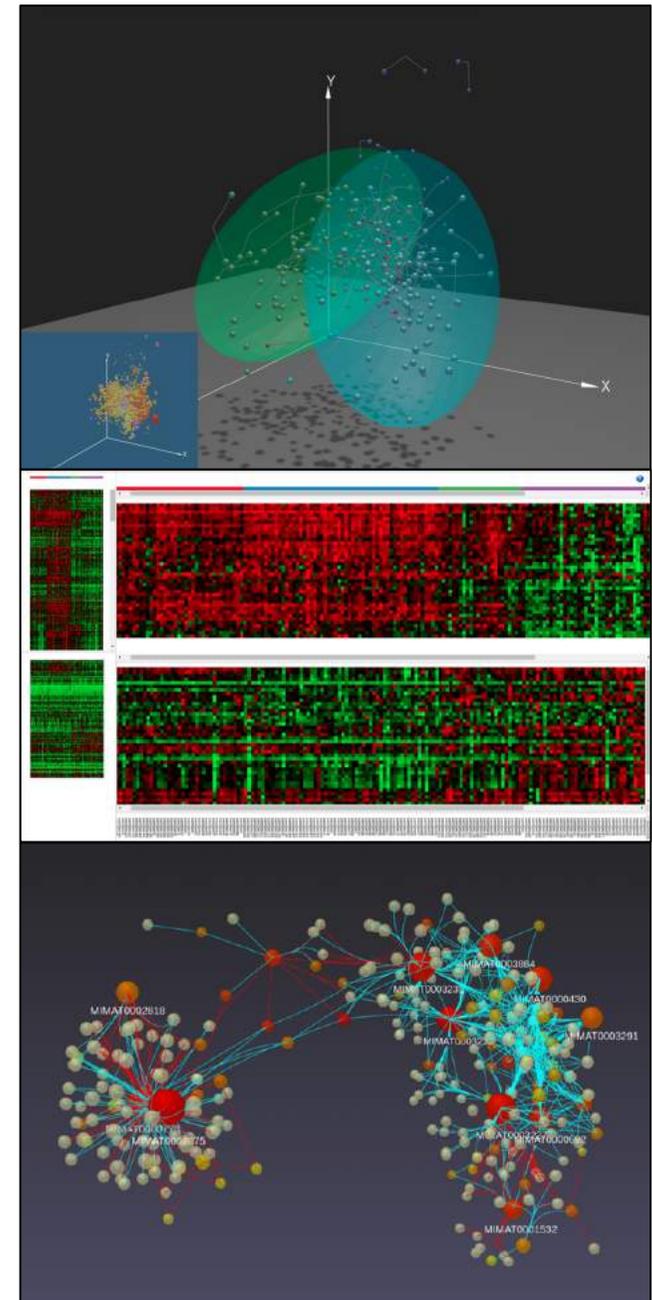
# TUTORIAL 2:

## Dimension reduction analysis



# Intro to OmicsAnalyst

- Web-based platform designed for data-driven multi-omics integration and visualization
- Designed to be accessible to bench scientists rather than bioinformatician
- Integrates well-established multivariate and univariate statistics with innovative visual analytics to support:
  - Integrative multi-omics analysis
  - Clustering and pattern discovery
  - Correlation analysis



# Requirements

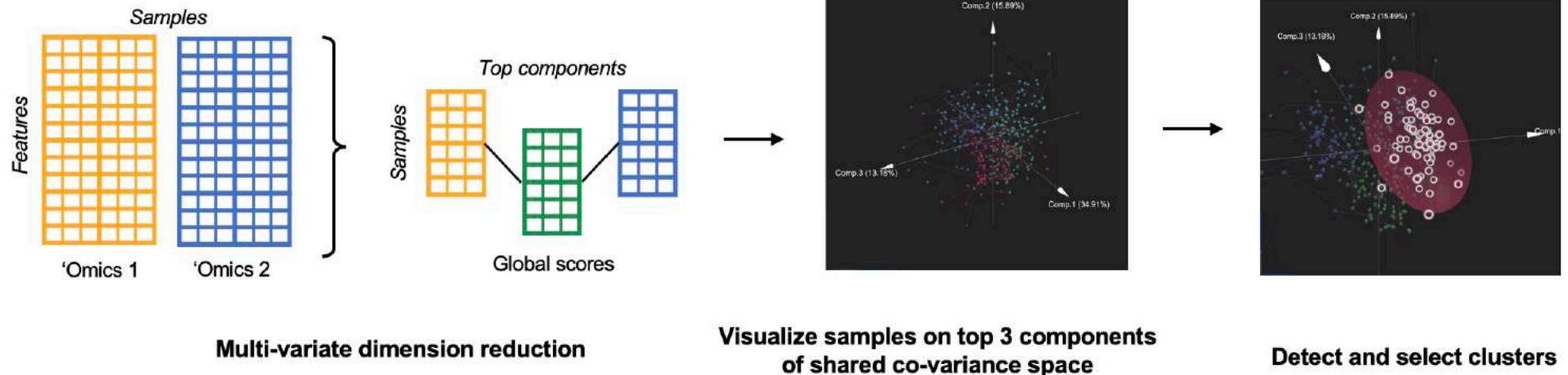
- Modern browser supporting WebGL.
- Ensure that WebGL is enabled in your browser!
  - Go to this page <https://get.webgl.org/> to verify your browser supports WebGL.
  - Refer to the FAQs for instructions on how to enable WebGL.
- For the best performance and visualization, use the latest version of Google Chrome.

# Goal for this tutorial

- A challenge for multi-omics analysis is to interpret highly complex and heterogeneous data without being overwhelmed by it
- Visualizing global patterns and clusters facilitate data interpretation
- Multivariate statistical methods integrate and transform omics data into dimensionally reduced data points in both feature space and sample space
- The goal of this tutorial is to:
  - Introduce OmicsAnalyst's tools for visualizing synchronized feature and sample space (3D Scatter Plot)
  - Perform clustering and functional analysis to explore overall patterns and biological functions
  - For a more detailed look at data processing, see Tutorial #1

# Dimension reduction + 3D scatter plot

- An overview of this analysis track is:



- To start the tutorial, click

**GET STARTED**

on the homepage

OmicsAnalyst currently supports transcriptomics, proteomics, metabolomics, microbiomics, and miRNA data. Human and mouse annotation are supported, but data from any species can be analyzed without annotation.

2  
Select 3<sup>rd</sup> dataset and click "Yes"

**Example Datasets**

Data	Description	Download
<input type="radio"/> Human pregnancy [2]	Human multi-omics data (proteomics, metabolomics) on modeling the chronology of these adaptations during full-term pregnancy. <a href="#">Multi-omics of pregnancy</a>	<a href="#">Proteomics</a> <a href="#">Metabolomics</a>
<input type="radio"/> Immune cells [3]	Mouse multi-omics data (transcriptomics, metabolomics, miRNA) on the effect of Ikaros transcription factor on B-cell differentiation from <a href="#">STATegRA</a>	<a href="#">Transcriptomics</a> <a href="#">Metabolomics</a> <a href="#">miRNA</a>
<input checked="" type="radio"/> Brain cancer [2]	Human multi-omics data (transcriptomics, miRNA) on glioblastoma multiforme of four different subtypes from <a href="#">TCGA</a> .	<a href="#">Meta-data</a> <a href="#">Transcriptomics</a> <a href="#">miRNA</a>

Yes Cancel

1  
Select "Try Example"

3  
Click "Proceed"

Try Examples

Proceed



Data Upload

Uploaded Data

tcga\_gene.csv

Feature: 1560  
Sample: 169  
DE #: 490  
Finished

tcga\_mirna.csv

Feature: 273  
Sample: 169  
DE #: 74  
Finished

### Processing Individual Data

Currently selected data:

tcga\_mirna.csv

See options for different datasets by changing this menu

If unsure whether the data are already normalized, click the eye icon to view boxplots of the data. If the data are not extremely right skewed, it is generally safe to assume that they are already normalized.

Processing Step			Action
Annotation	Data value type	Raw	Submit ✓
	Omics type	miRNA	
	Specify organism	H. sapiens (human)	
	ID type	miRBase ID (v15+)	
Comparison	Choose a method	Limma	Submit ✓
	Fold change cutoff	1.0 (for two groups only)	
	P-value (FDR) cutoff	5.0E-5	

The main form shows processing parameters that were used for each dataset. When uploading data, it is important to correctly specify whether the data are raw counts or continuous values, as different data transformations should be used prior to differential analysis. See the FAQs for more details on the processing methods.

1

Click "Proceed"

Proceed

Try Examples

Uploaded Data

<input checked="" type="checkbox"/>	<b>tcga_gene.csv</b> Feature: 1560 Sample: 169 DE #: 490	
<input checked="" type="checkbox"/>	<b>tcga_mirna.csv</b> Feature: 273 Sample: 169 DE #: 74	

### Data Quality Check

The uploaded omics datasets are summarized below:

**Total number of samples:** 169  
**Group names:** Classical; Mesenchymal; Neural; Proneural  
**Individual datasets:** tcga\_gene.csv; tcga\_mirna.csv  
**Corresponding feature number:** 1560; 273

The density plot and PCA plots are generated to provide an overview of the omics datasets for quality control. The density plot shows the distributions across different omics layers. If the overall distribution seems to be in very different ranges, you should consider batch correction, accounting for batch effect prior to uploading your data to OmicsAnalyst, for example with the R package COMBAT.

**Current omics data:** tcga\_mirna.csv

**Sample normalization:** None

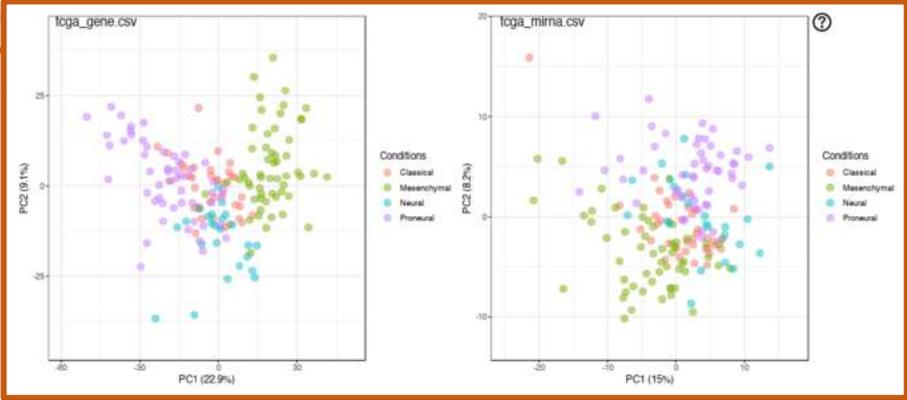
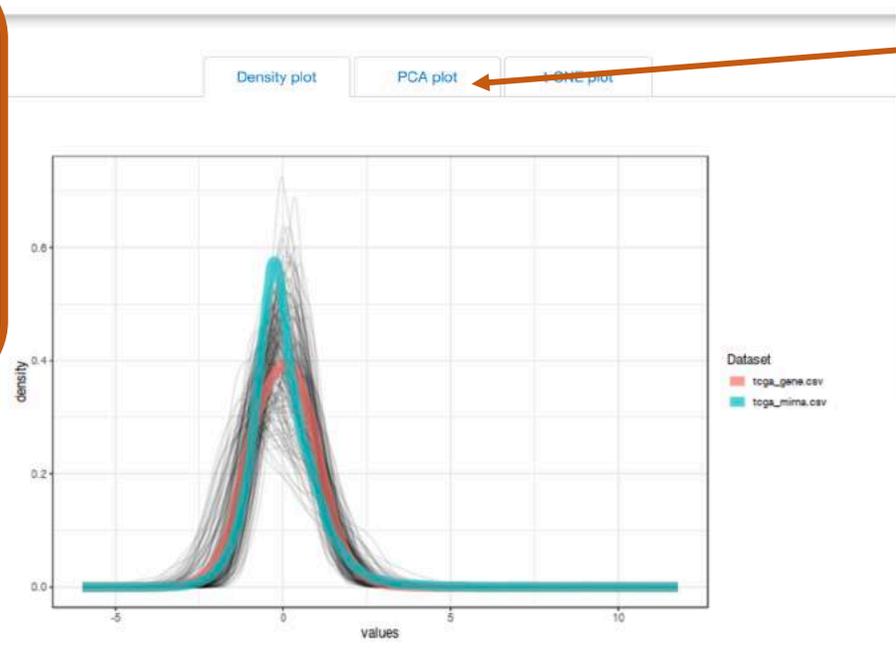
**Data transformation:** None

**Data scaling:** Auto scaling

The page provides graphics to ensure that the data has been properly normalized. Since different transformation may be preferred compared to differential analysis, you can update the method here. Scaling the datasets so that have comparable distributions across 'omics types is recommended.

See the number of differential features from each dataset. Click the icon to update thresholds.

1 Look at the "PCA plot"



2 Click "Proceed"

Uploaded Data	
<input checked="" type="checkbox"/>	<b>tcga_gene.csv</b> Feature: 1560 Sample: 169 DE #: 490
<input checked="" type="checkbox"/>	<b>tcga_mirna.csv</b> Feature: 273 Sample: 169 DE #: 74

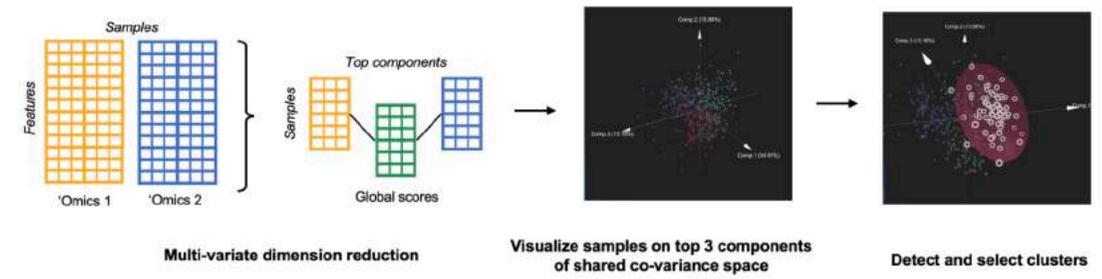
Please choose a method to proceed

- Correlation Network Analysis
- Cluster Heatmap Analysis
- Dimension Reduction

Change to "Dimension Reduction" tab

1

The objective of this analysis is to perform dimension reduction, and then visually explore corresponding scores, loadings and biplots in interactive 3D scatter plots to understand the common trends and underlying patterns. The multivariate dimension reduction techniques are kind of like parallel versions of PCA, where we try to find sets of multi-dimensional components that both reduce redundant information within individual datasets, and are related to each other across datasets. These sets of components are related to each other through some global scores, which are the dimensions that we use to visualize the sample space. The different methods are mainly distinguished by the way that they optimize similarity of component sets across the 'omics datasets. Select an individual method to see more details on its unique statistical features.



dimension reduction method

Proceed

2

Select "MCIA"

Click "Proceed"

3

OmicsAnalyst supports four unsupervised dimension reduction methods, and one supervised one. Select each method from the drop down menu to read a short description.

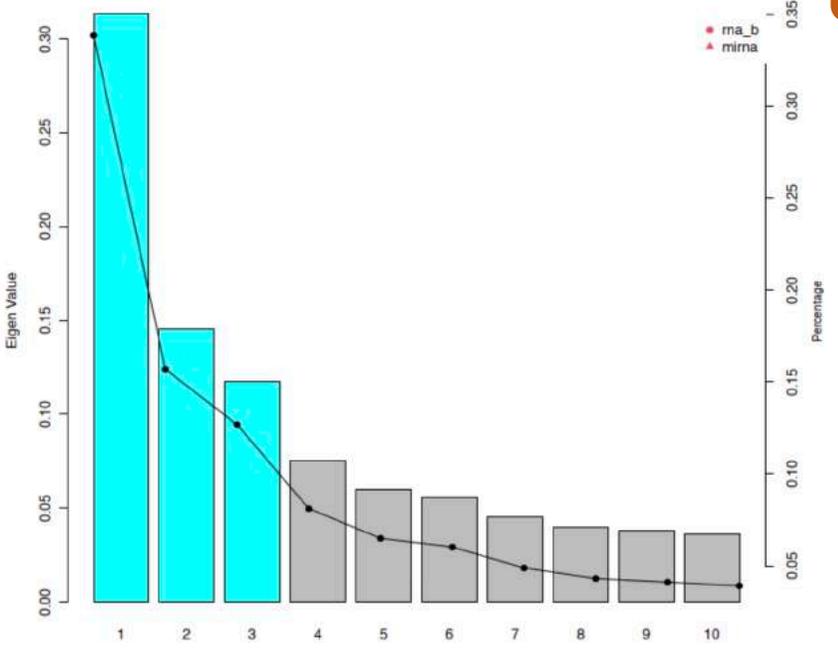
This page allows you to refine key tuneable parameters from the selected statistical method (if any – MCI A does not have any exposed parameters). Sample and feature space plots are provided to help determine whether the parameters should be changed.

### Analysis Overview & Parameter Tuning

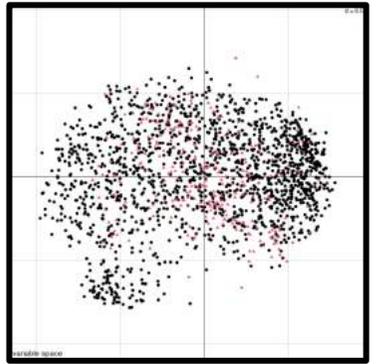
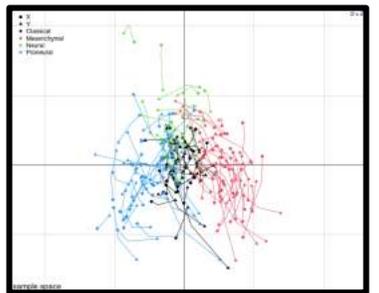
- **Diagnostic** includes a figure that helps in deciding the
- **Sample space** visualizes overall distribution of samples
- **Feature space** focuses on the contributions of features

Diagnostic   Sample space   Feature space

Eigenvalue plot displaying its relationships to eigenvectors (components). The bars represent absolute eigenvalues while the linked dots represented variance percentage. Please refer to [omicade4](#) for more details.



1  
View other plots

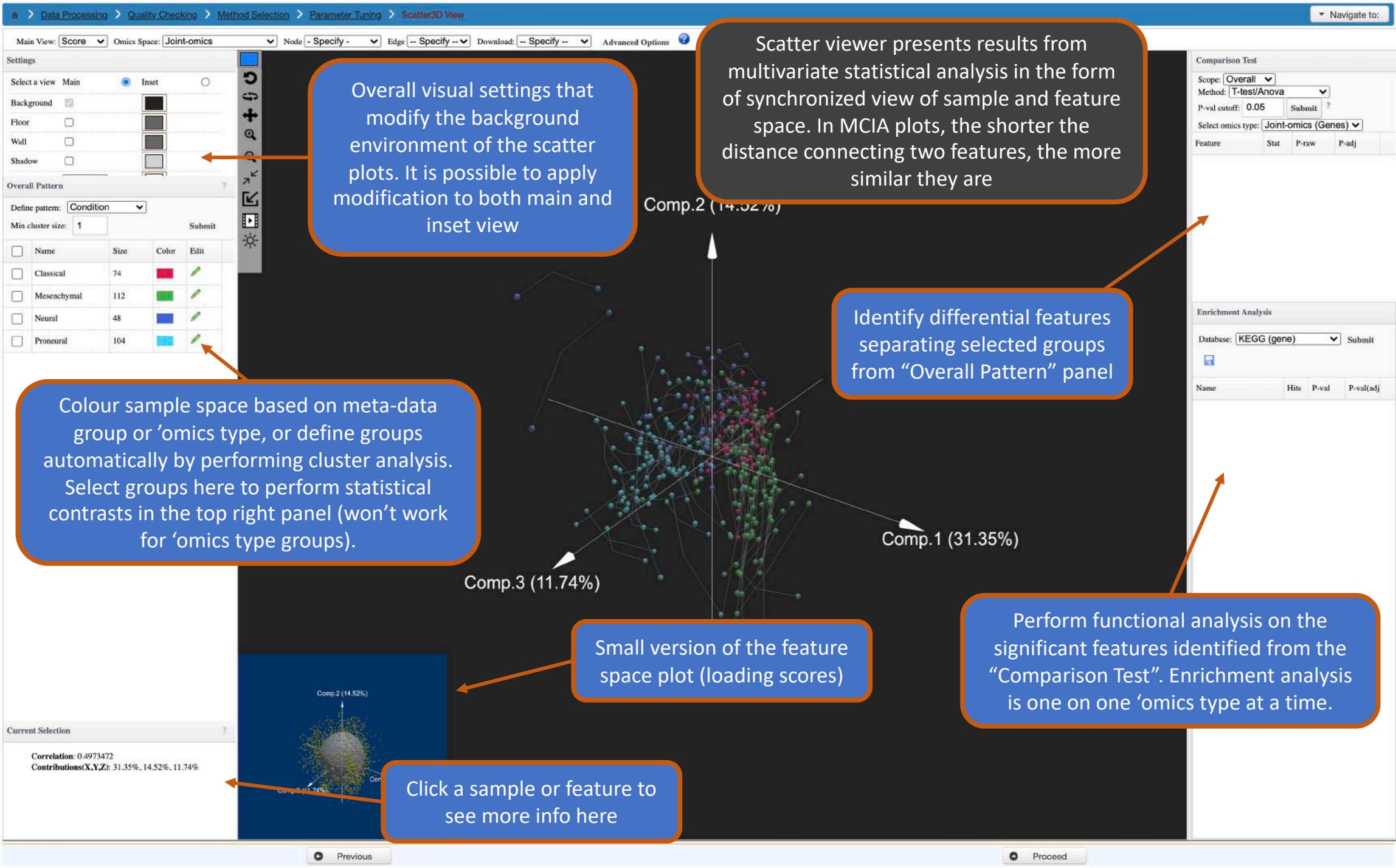


Click "Proceed"

2

Previous

Proceed



Overall visual settings that modify the background environment of the scatter plots. It is possible to apply modification to both main and inset view

Scatter viewer presents results from multivariate statistical analysis in the form of synchronized view of sample and feature space. In MCA plots, the shorter the distance connecting two features, the more similar they are

Colour sample space based on meta-data group or 'omics type, or define groups automatically by performing cluster analysis. Select groups here to perform statistical contrasts in the top right panel (won't work for 'omics type groups).

Identify differential features separating selected groups from "Overall Pattern" panel

Small version of the feature space plot (loading scores)

Perform functional analysis on the significant features identified from the "Comparison Test". Enrichment analysis is one on one 'omics type at a time.

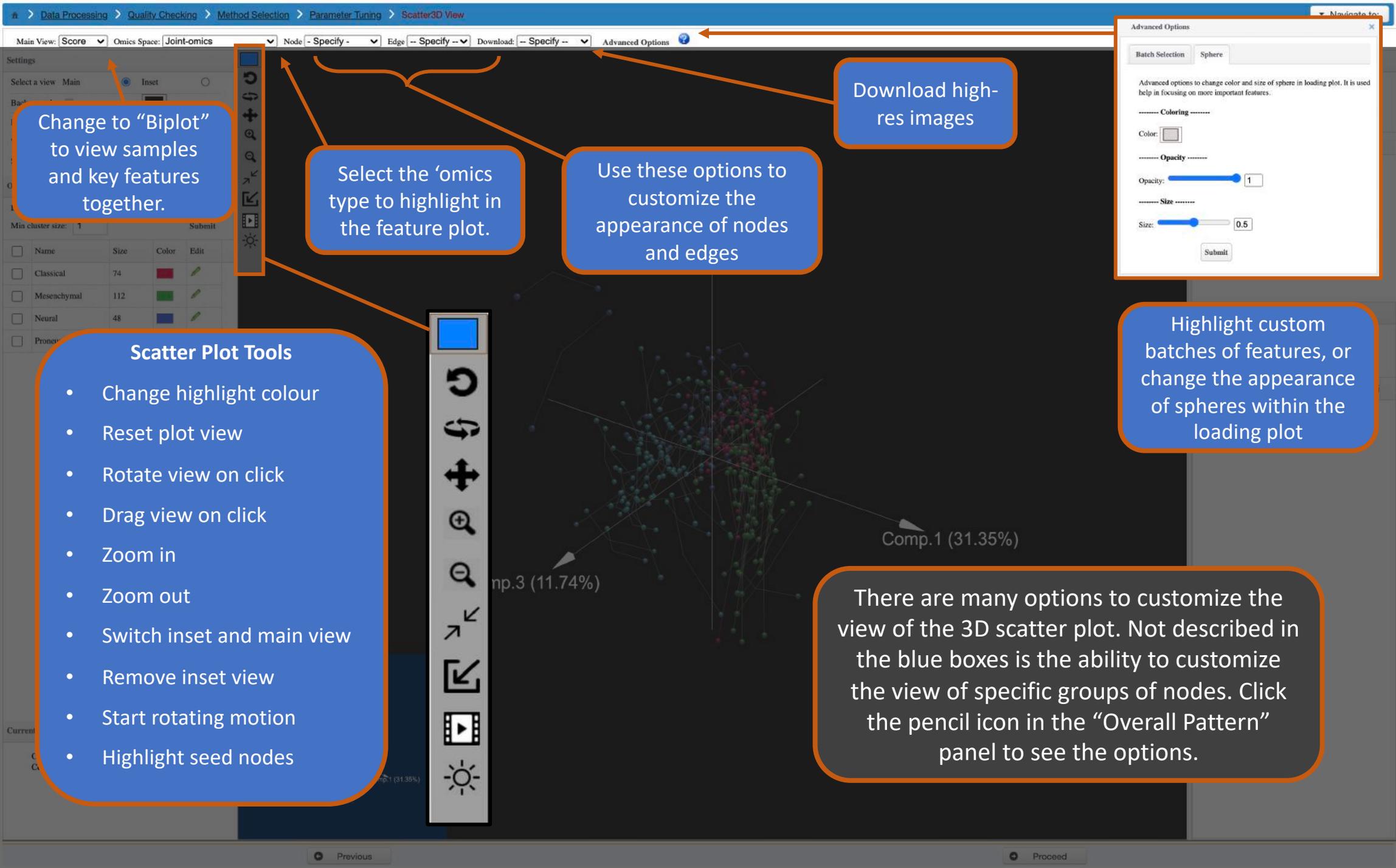
Click a sample or feature to see more info here

Current Selection  
Correlation: 0.4973472  
Contributions(X,Y,Z): 31.35%, 14.52%, 11.74%

Comparison Test  
Scope: Overall  
Method: T-test/Anova  
P-val cutoff: 0.05  
Submit  
Select omics type: Joint-omics (Genes)

Enrichment Analysis  
Database: KEGG (gene)  
Submit

Name	Hits	P-val	P-val(adj)
------	------	-------	------------



Change to "Biplot" to view samples and key features together.

Select the 'omics type to highlight in the feature plot.

Use these options to customize the appearance of nodes and edges

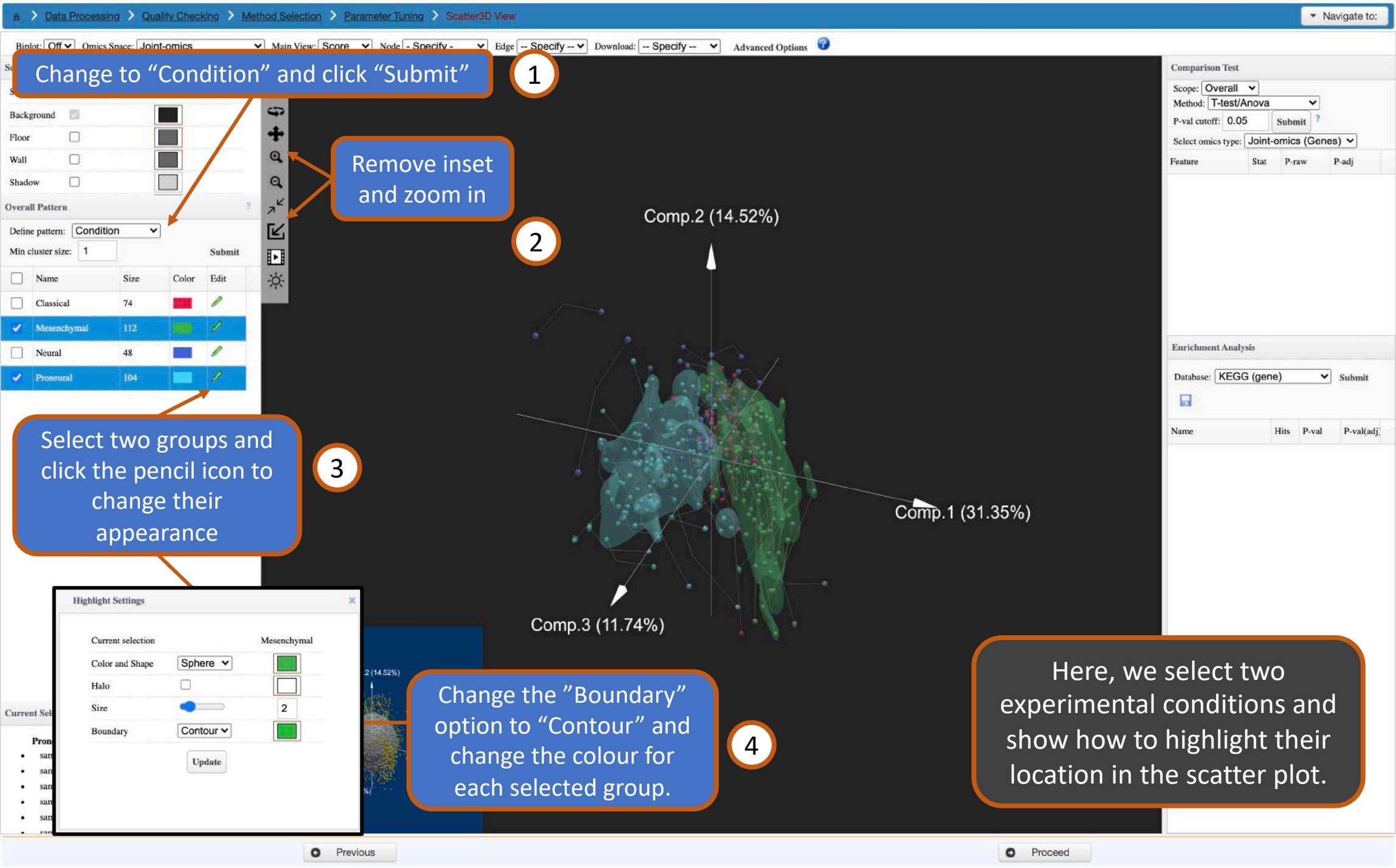
Download high-res images

Highlight custom batches of features, or change the appearance of spheres within the loading plot

### Scatter Plot Tools

- Change highlight colour
- Reset plot view
- Rotate view on click
- Drag view on click
- Zoom in
- Zoom out
- Switch inset and main view
- Remove inset view
- Start rotating motion
- Highlight seed nodes

There are many options to customize the view of the 3D scatter plot. Not described in the blue boxes is the ability to customize the view of specific groups of nodes. Click the pencil icon in the "Overall Pattern" panel to see the options.



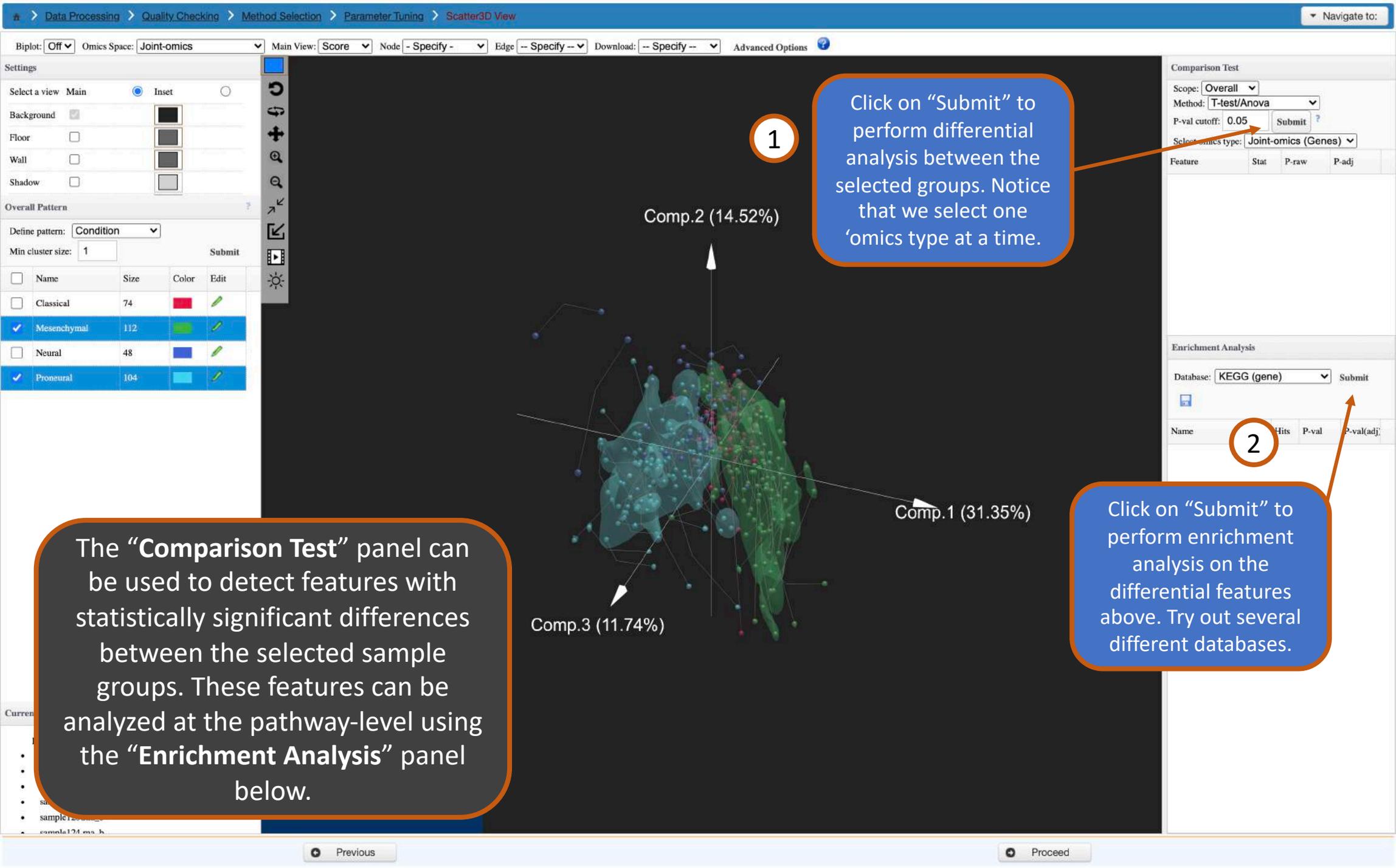
1 Change to "Condition" and click "Submit"

2 Remove inset and zoom in

3 Select two groups and click the pencil icon to change their appearance

4 Change the "Boundary" option to "Contour" and change the colour for each selected group.

Here, we select two experimental conditions and show how to highlight their location in the scatter plot.



1

Click on "Submit" to perform differential analysis between the selected groups. Notice that we select one 'omics type at a time.

2

Click on "Submit" to perform enrichment analysis on the differential features above. Try out several different databases.

The "Comparison Test" panel can be used to detect features with statistically significant differences between the selected sample groups. These features can be analyzed at the pathway-level using the "Enrichment Analysis" panel below.

Biplot: **On** Omics Space: **Joint-omics** Main View: **Score** Node: **- Specify -** Edge: **-- Specify --** Download: **-- Specify --** Advanced Options

Settings

Select a Background

Floor

Wall

Shadow

Overall Pattern

Change view to "Biplot"

1

**Biplot Options**

This feature aims at overlaying significant features on top of the sample space to easily visualize their contribution in terms of sample separation.

- Rank-based: display top features based on differential expression analysis performed on data upload page.
- Custom: enter a list of features (up to 50) of interest.

Color:

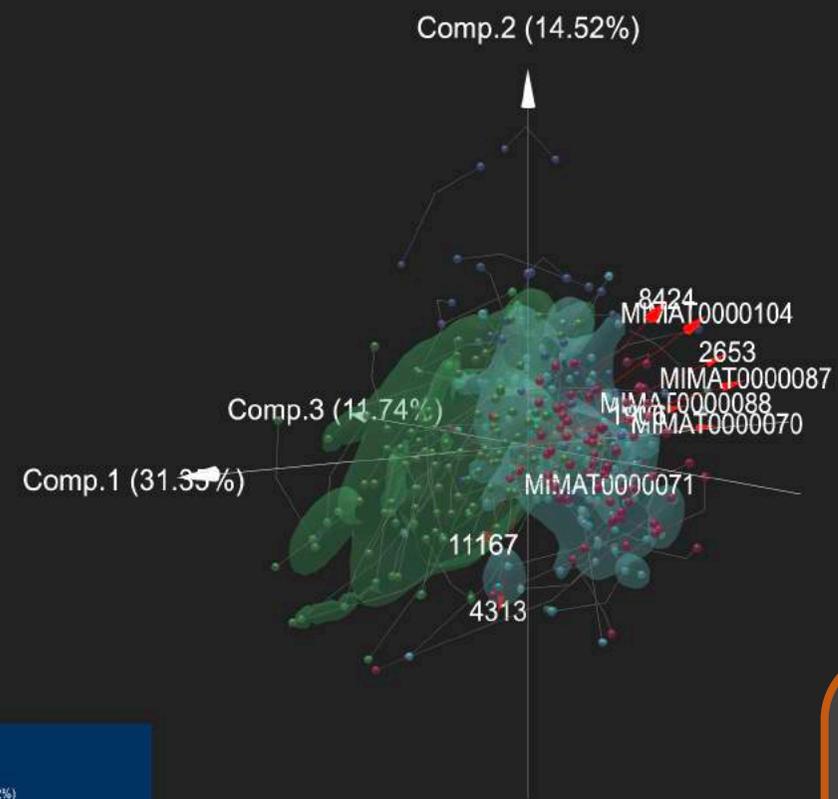
Selection: **Rank-based**

Omics: **Both**

Number of features:

2

Click "Submit"



Comparison Test

Scope: **Overall**

Method: **T-test/Anova**

P-val cutoff: **0.05**

Select omics type: **Joint-omics (Genes)**

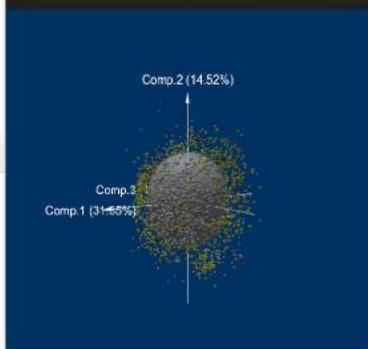
Feature	Stat	P-raw	P-adj
---------	------	-------	-------

Enrichment Analysis

Database: **KEGG (gene)**

Name	Hits	P-val	P-val(adj)
------	------	-------	------------

- Current Selection
- Proneural**
- sample1.ma\_b
  - sample100.ma\_b
  - sample118.ma\_b
  - sample12.ma\_b
  - sample120.ma\_b
  - sample124.ma\_b



Use the "Biplot View" to overlay key features on top of the sample space. Features to display are selected based on the statistical contrasts performed during initial data processing.



# THE END

For more information, visit the **FAQs, Tutorials, Resources**  
and **Contact** pages on [www.omicsanalyst.ca](http://www.omicsanalyst.ca)