Data Visualization Portal Tutorial

Authors: Justin Davis and Ana Uriarte-Acuna

Link to Data Visualization Portal

Table of Contents

About OSDR

GeneLab is an open-access resource that enables scientists to upload, store, share, and analyze omics data from spaceflight experiments. It facilitates information sharing, fosters innovation, and accelerates scientific discovery in space biology. By studying the effects of microgravity and space environment on DNA, RNA, proteins, and metabolites, GeneLab contributes to our understanding of biology, and advances in genomics. It provides coordinated data sets and metadata, allowing users to gain comprehensive insights and make novel discoveries. GeneLab's impact lies in its ability to facilitate information sharing, drive innovation, and ultimately expand our knowledge of how space conditions affect the fundamental building blocks of life.

About Data Visualization

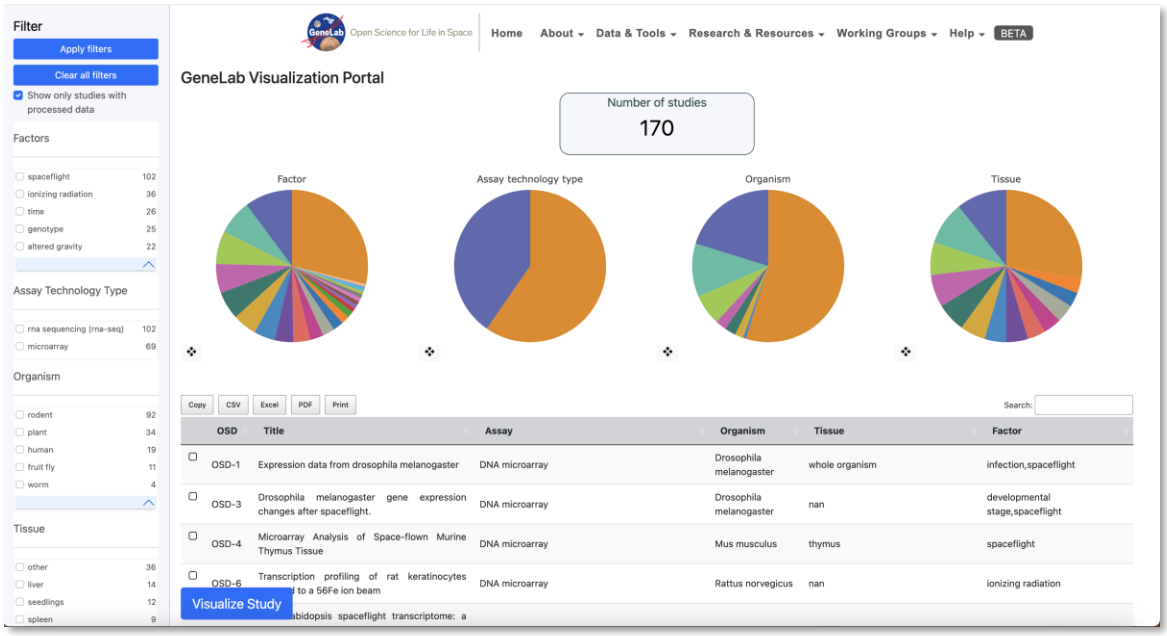
OSDR has a Data Visualization portal that currently provides users the ability to interact with the processed RNA sequencing data from the OSDR database. The portal encompasses various visualization types, including Gene Expression query tables, Dendrograms, Heatmaps, Gene Set Enrichment Analysis and a range of interactive plots including Principal Component Analysis (PCA) plots, Pair plots, and Volcano plots. Each tool offers researchers flexibility to adjust parameters and explore specific aspects of the data effectively.

Metadata Dashboard

Link to Metadata Dashboard

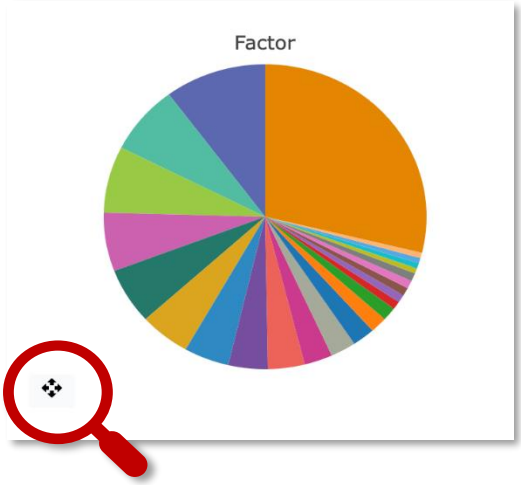
Filters

The metadata dashboard is designed to help users narrow search results for experimental data. It provides various tools for filtering and displaying results. The main tools for filtering the studies table's results are the pie charts and the filters on the left side of the dashboard. Each section of the pie chart acts as a separate section of filters, and when a filter from the pie chart is selected the results containing that category will automatically populate in the studies table below. A user can make one selection on each pie chart to narrow results in the studies table further.

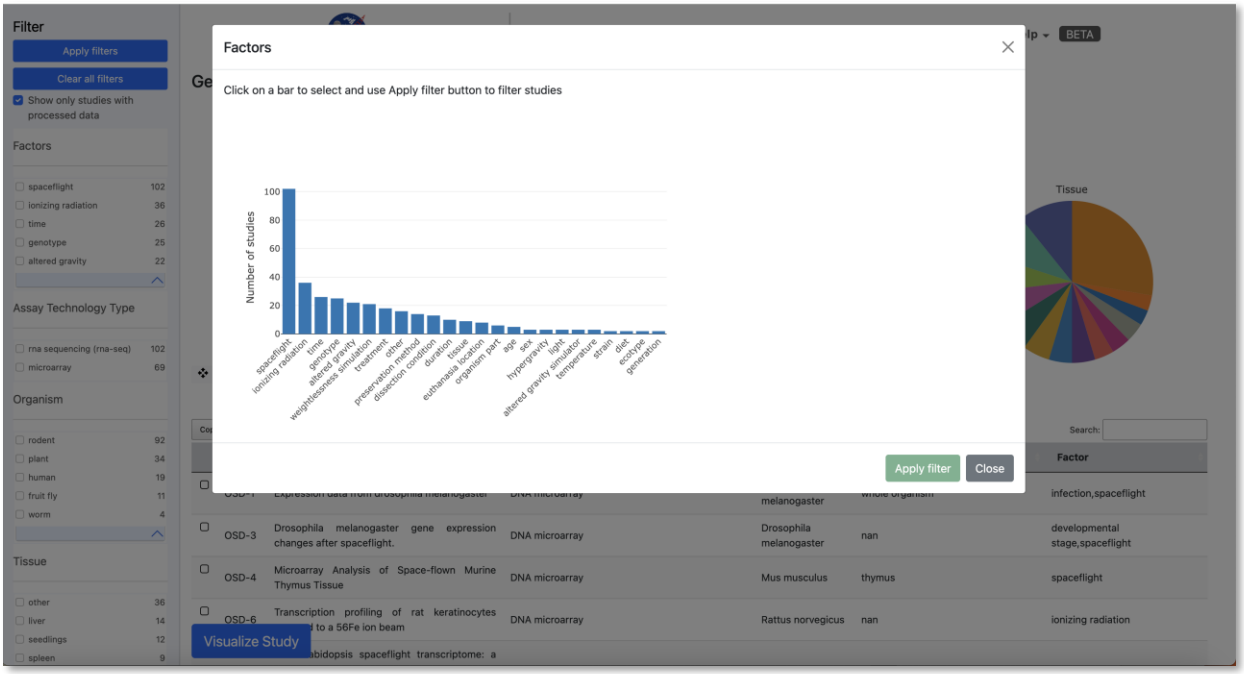


In addition to the pie charts, there are specific filters on the left side of the dashboard that can be selected to narrow down results. When selecting filters on the pie charts, both sections will be updated to show the selected filters and the studies table will be updated to show the relevant studies. When selecting filters on the sidebar, sections will be updated after clicking on the "Apply filters" button, this allows the user to select multiple filters.

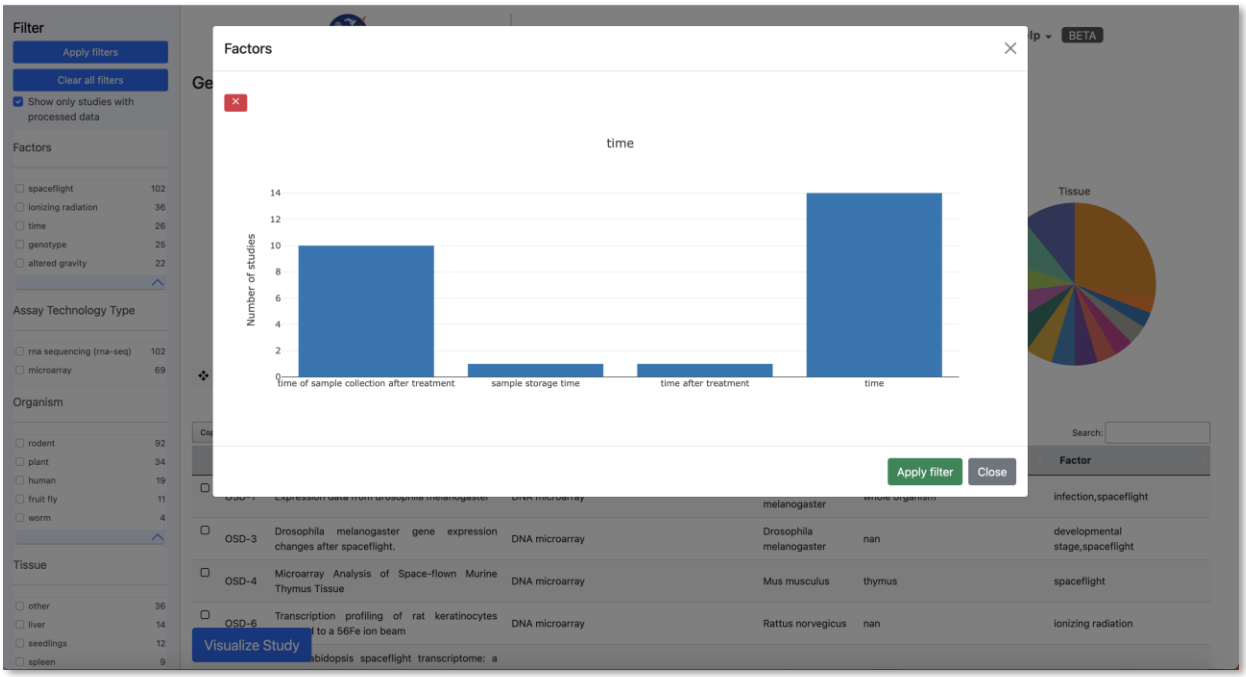
By default, the "Show only studies with processed data" filter is selected. Studies with processed data have the necessary data to display the visualization plots or combine multiple studies. The user can remove this filter to display information about all available studies in the repository.



Another tool that is provided with each individual pie chart is the crosshair located on the bottom-left of the chart. When you select the crosshair a bar graph displaying the different categories listed within the Pie chart will appear.

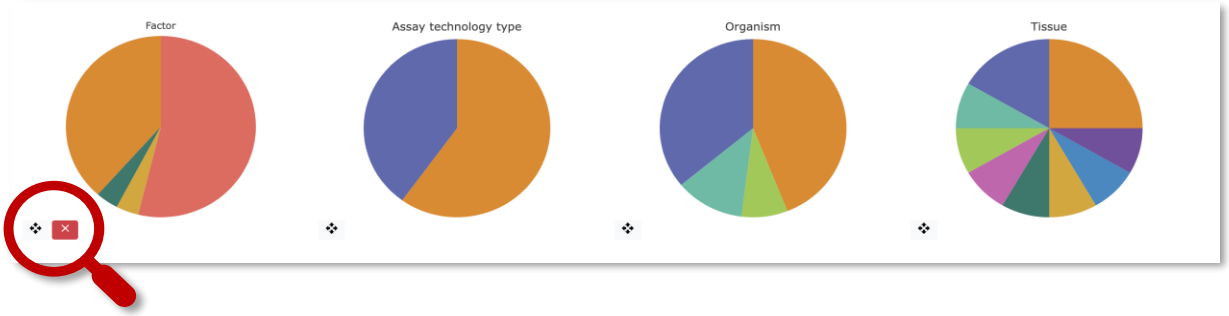


The bar graph is an additional way to display the number of results each category has, and when a bar is selected from the bar graph a new graph will be displayed with the subcategories included in the selected value. The below example shows the breakdown of choosing "Time" from the bar graph selection.



After selecting the desired filter from the bar graph, the user can press the “Apply filter” button to update the results within the studies table.

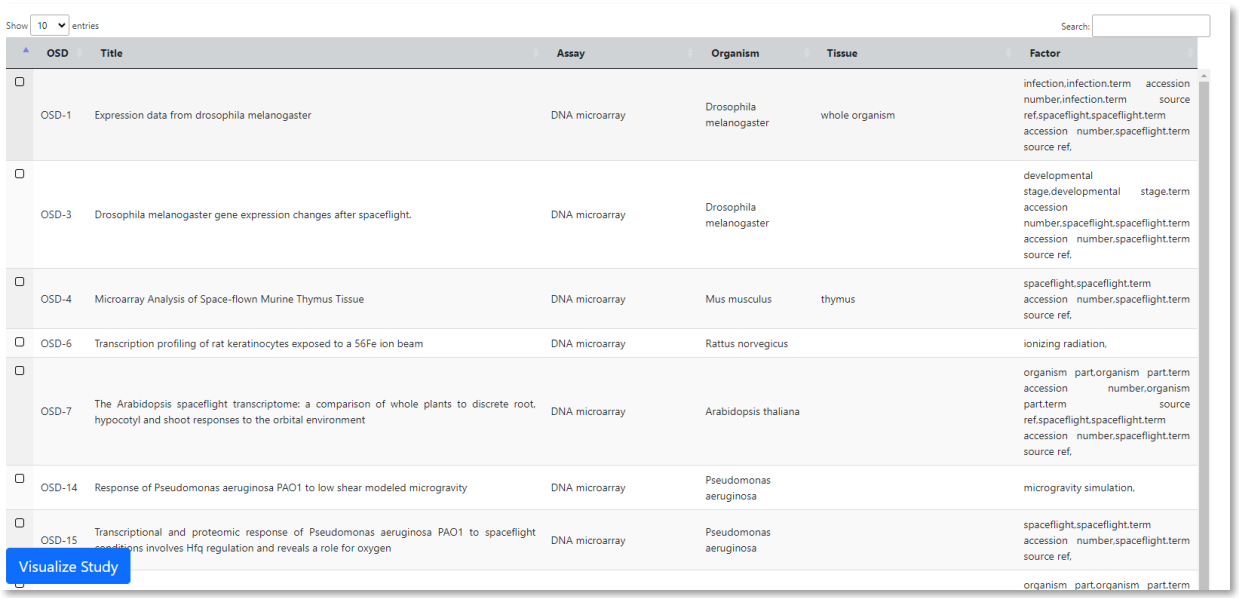
When a filter is selected, a red "X" will also appear next to the crosshair of the associated pie chart. Pressing the red "X" button will clear the selected filters.



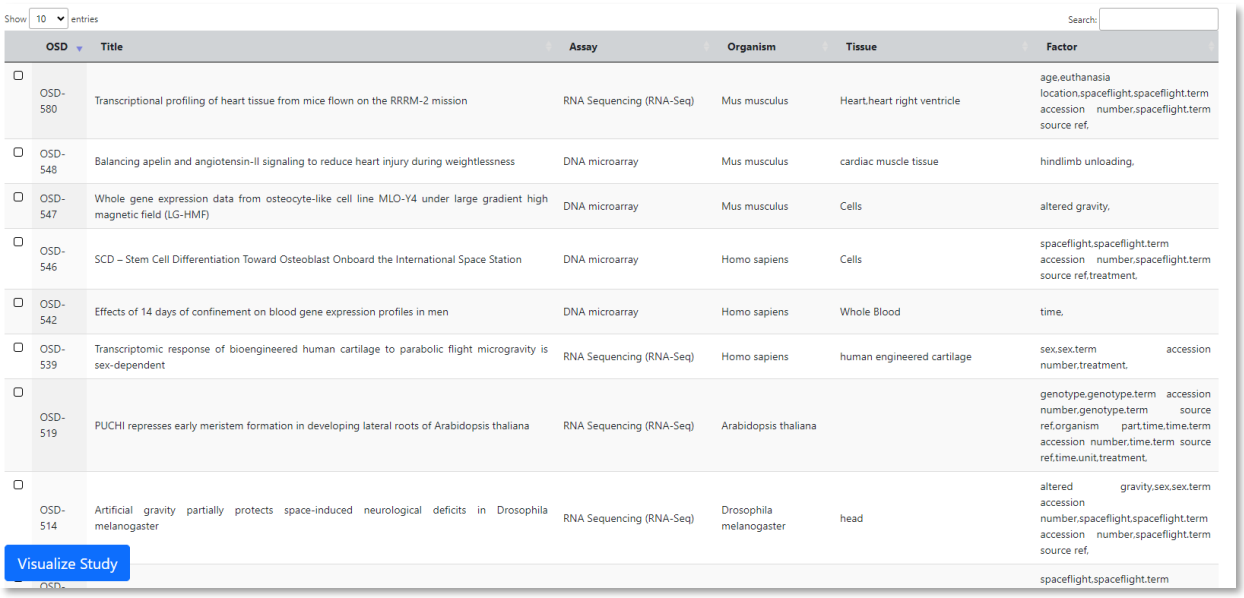
The “Clear all filters” button at the top left can be used to remove all selected filters at once.

Studies Table

Below the pie charts is a table that lists the studies resulting from the selected filters from above. The table includes the following information for each study: OSD, Title, Assay, Organism, Tissue, and Factor.



By default, the studies will be listed in order of OSD-# from smallest to largest, but the order can be flipped based on each information category by clicking the title of the category twice.

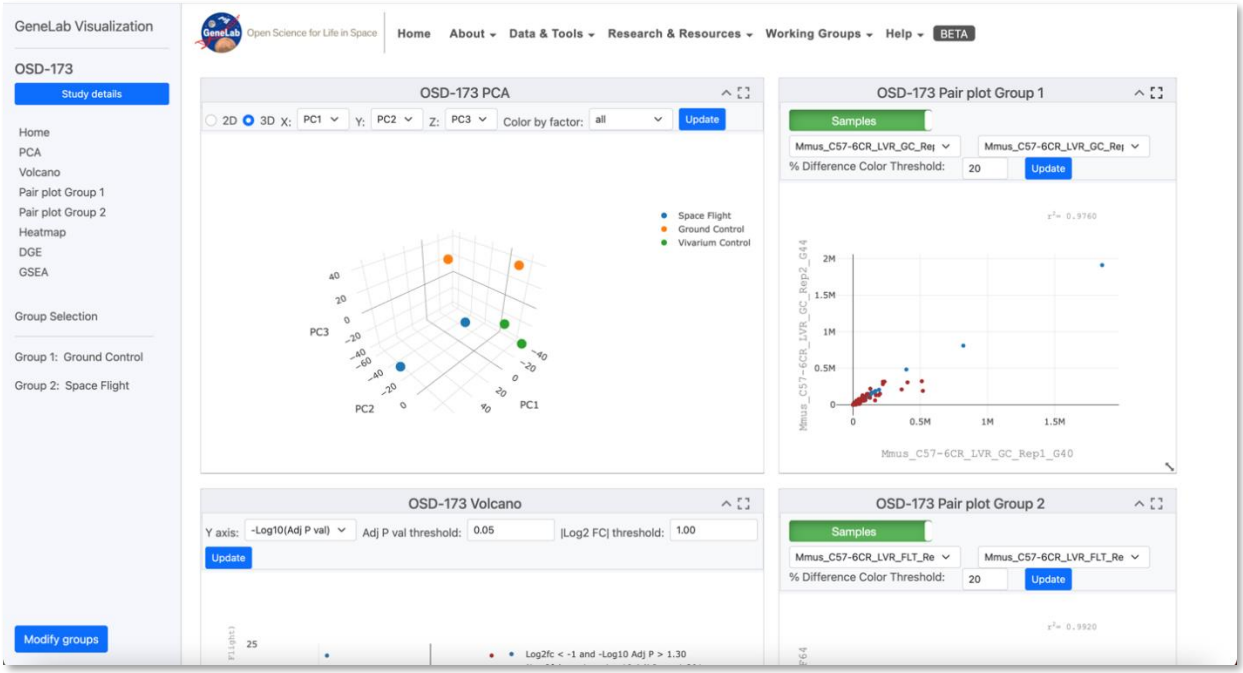


Finally, once a user has selected a study or multiple studies, they can press the "Visualize Study" button to be directed to the data visualization tools. A user may also select multiple studies to visualize simultaneously in which case a user will be directed to a Multi-Study preview page before being directed to the data visualization tools.

Single Study Visualization Portal

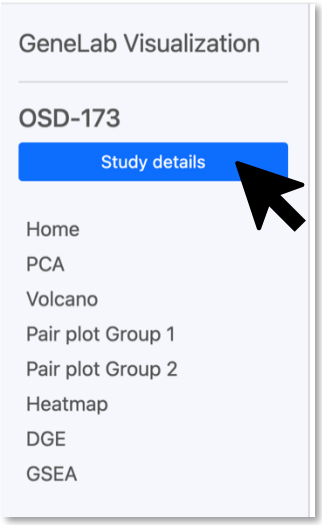
When a user has selected one study to visualize, they will be directed to the data visualization portal. The following plots are available for each study:

- PCA plot
- Pair plots
- Volcano plot
- Heatmap
- Interactive Differential Gene Expression (DGE) table
- Gene Set Enrichment Analysis (GSEA) plots:
    - Normalized Enrichment Score (NES) Table
    - Enrichment Score Plot
    - Normalized Enrichment Score (NES) Plot
    - Dot Plot
    - Ridge Plot
    - Network Plot

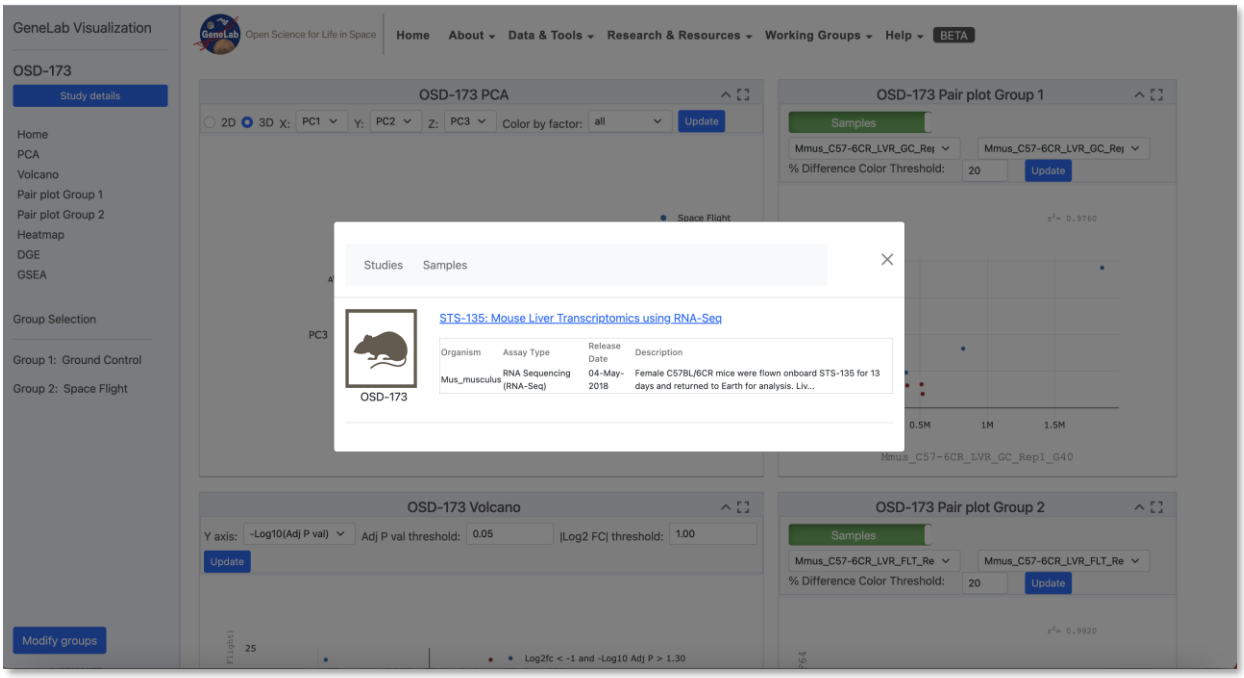


Sidebar Functions

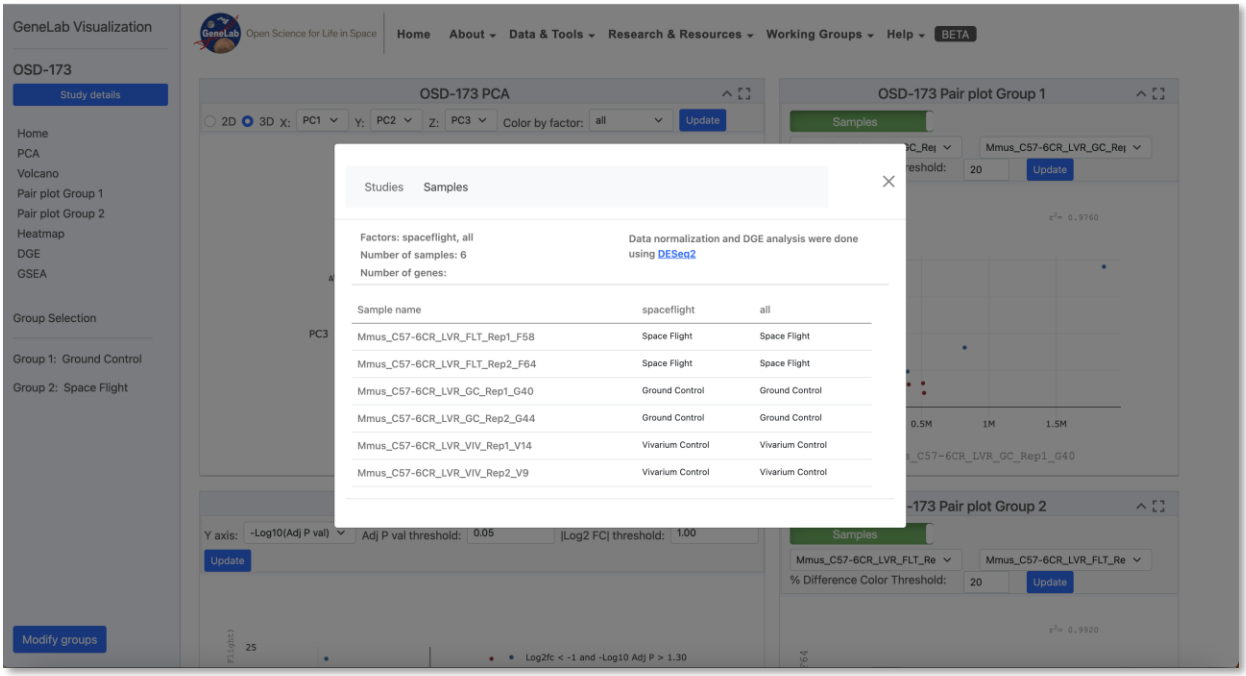
A sidebar of helpful tools is provided on the left side of the screen.

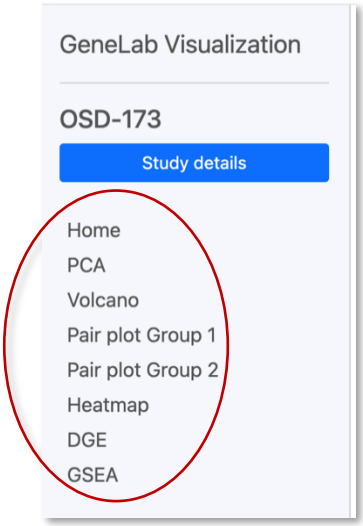


The "Study details" button is located at the top of the sidebar. This button pulls up a display with the study information including a small description.

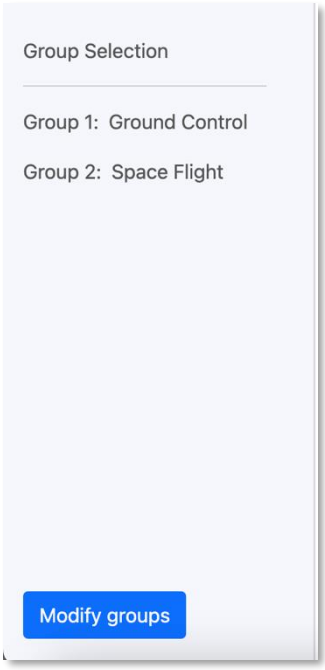


The display also includes a tab labeled “Samples” that a user can press to see the individual samples and additional information for the study.





Below the study details button is a label for each individual plot provided for a user within the data visualization tool. Clicking these labels will automatically direct the user to the plot associated with the label.

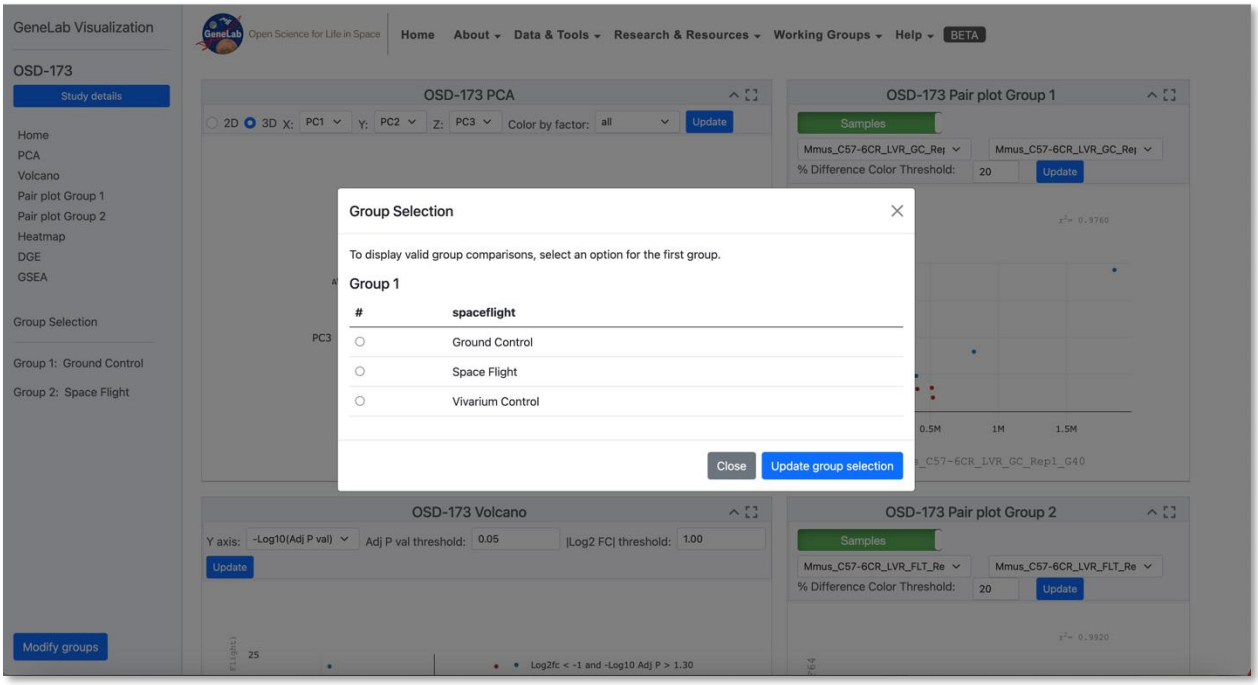


At the bottom of the sidebar is the default Group selection that is utilized for each plot. This groups represent the factor values used on the pairwise comparisons displayed on, or used to filter the Volcano plot, Pair plots, Heatmap, DGE table and GSEA.

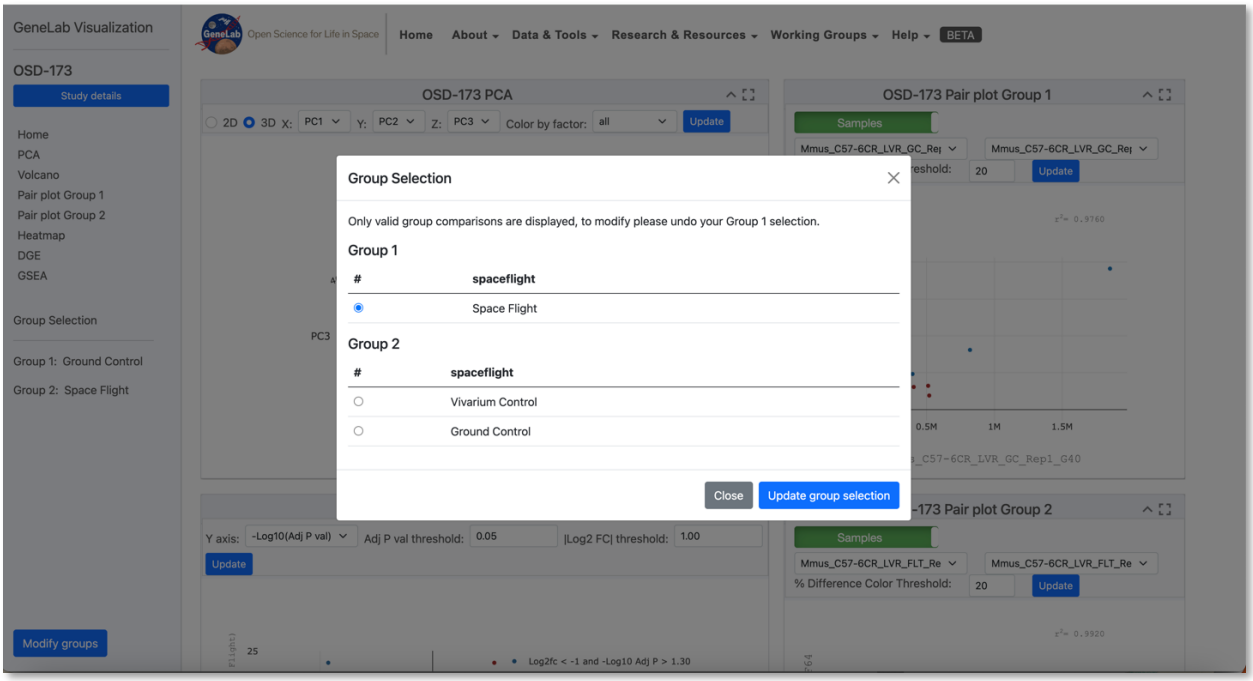
For example, if Group 1 is Ground Control and Group 2 is Space Flight:

- Volcano plot will display the Log2 Fold Change and P value between Ground Control samples and Space Flight samples.
- Pair Plot Group 1 will only display Ground Control samples.
- Pair Plot Group 2 will only display Space Flight samples.
- Heatmap will use Log2 Fold Change and Adjusted P value between Ground Control samples and Space Flight samples to filter genes.
- DGE table will display Log2 Fold Change, P value and Adjusted P value between Ground Control samples and Space Flight samples.
- GSEA will be performed using only Ground Control and Space Flight samples.
- PCA will not change based on group selection because it is calculated using data from all samples.

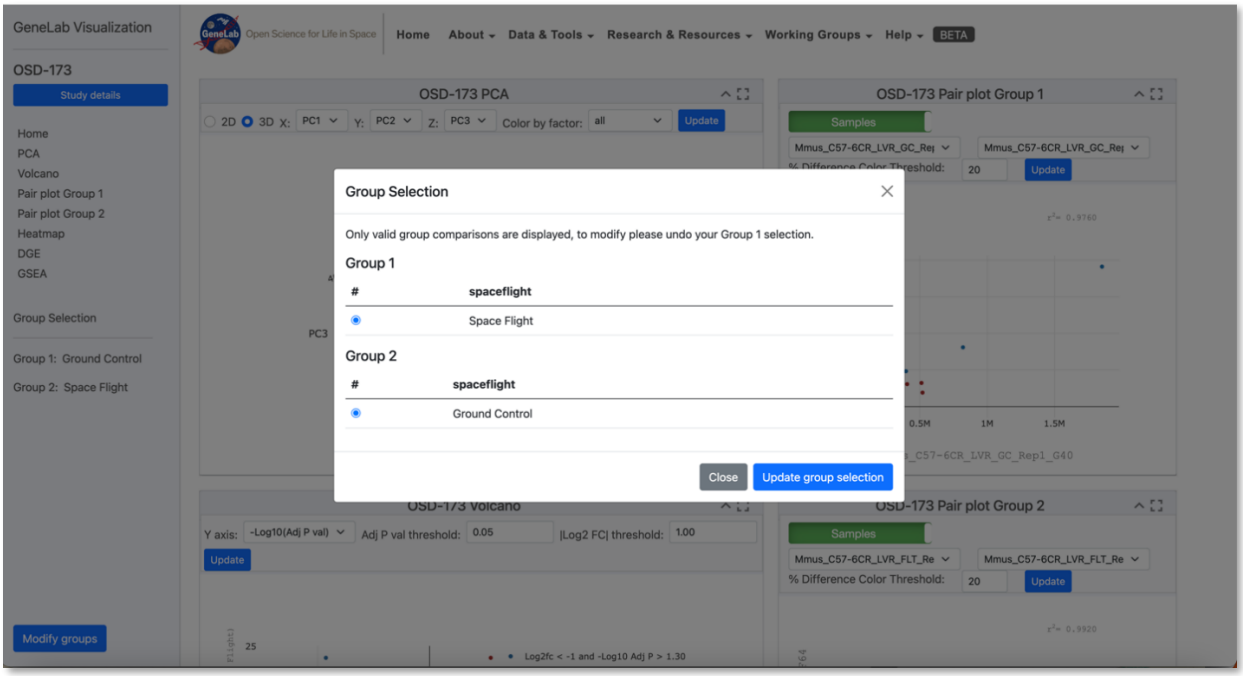
A user can modify the groups that are selected by pressing the "Modify Groups" button. This button will open a modal displaying the options for the first group.



After the user has selected an option, valid options for the second group will be displayed.



Once the user has selected an option for both groups, the Update group selection button can be used to update the plots.

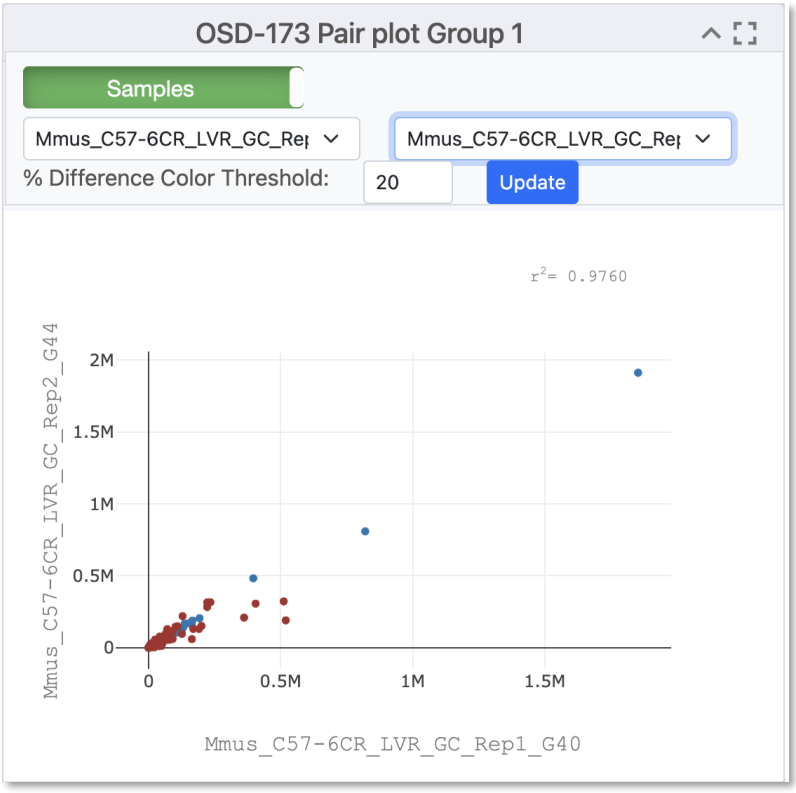


A feature exclusive to multi-study visualization is the option to download the combined Differential Gene Expression table.

DGE stands for Differential Gene Expression and represents the number of sequence reads that originated from a particular gene. The higher the number of counts, the more reads associated with that gene, and the assumption that there was a higher level of expression of that gene in the sample.

Plotly

Plotly is a third-party software that is using data provided by GeneLab to create the interactive visualizations displayed. At the top-right corner of each plot will be options to help a user better visualize the data.

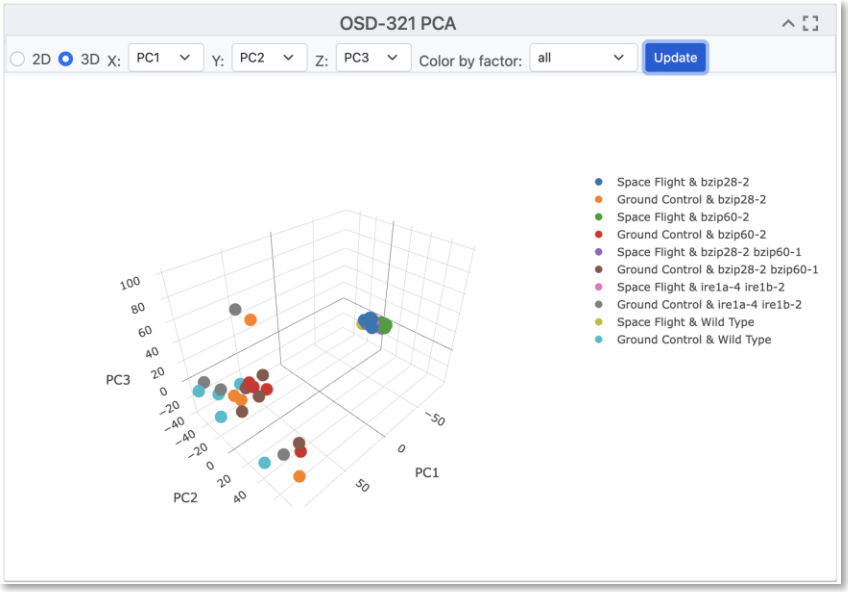


The house icon within the options will reset the axes of the plots back to default. Users are also provided the option to zoom in/out on each plot as well as auto-scale the graphic. There are two tools provided for data point selection, which are the lasso tool and box tool. Each of these tools provides a shape that will select any data points that fall within them. Lastly, there is a download button in the shape of a camera that will let you download the plot as a PNG file.

PCA Plots

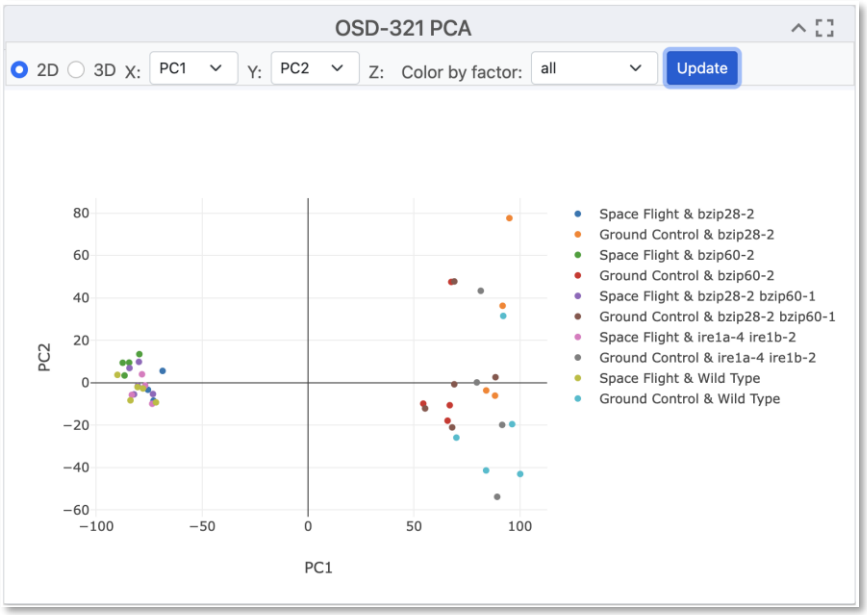
[Link to read about PCA plots]

PCA stands for Principal Component Analysis, and this type of plot is used to reduce the dimensionality of large sets of data to simplify the process of analyzing the data points.



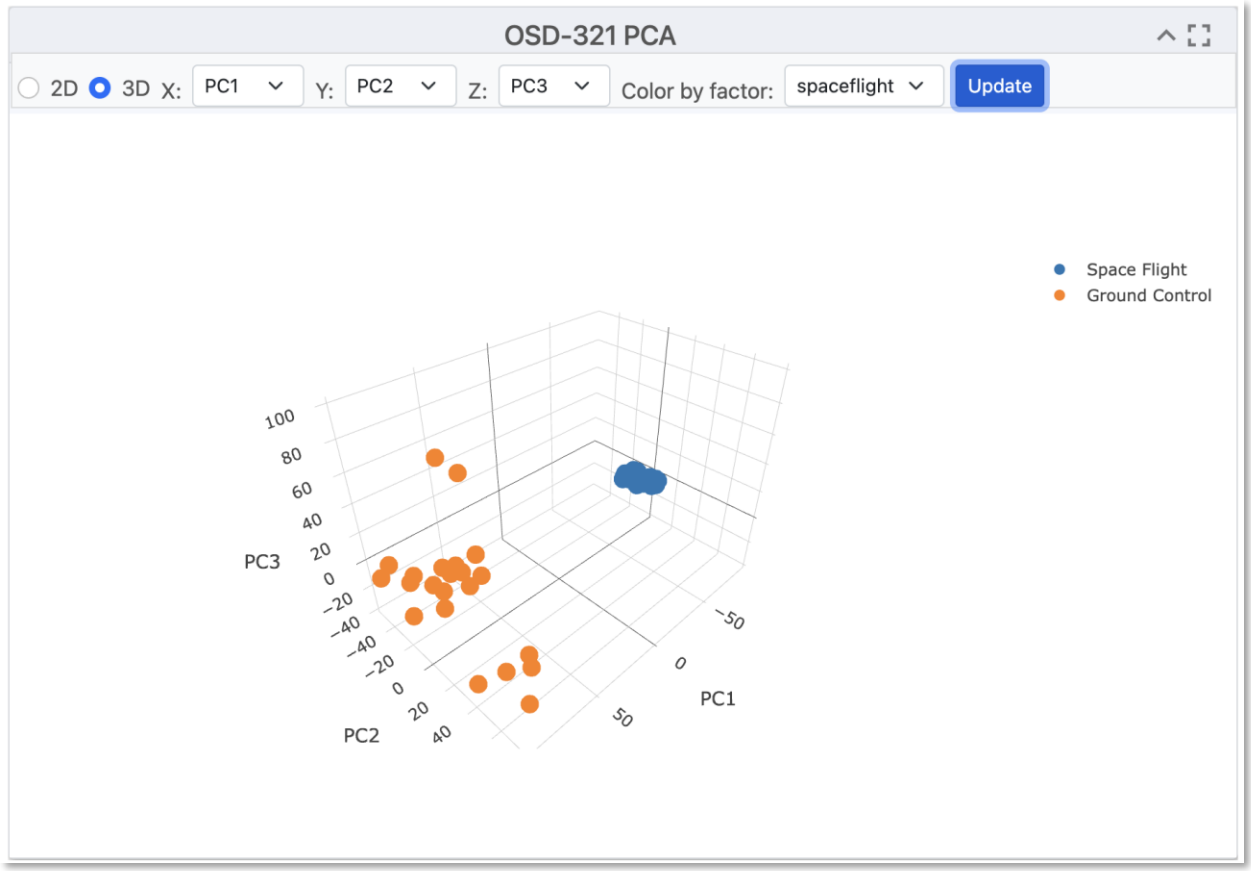
Each PCA plot will include options for a 2D and 3D representation of the data. The default selection is a 3D representation on an "X", "Y", and "Z" axis.

- In the upper left corner of the plot area select the "2D" button and then press "Update." The graph will update to display the data on an "X", and "Y" axis only.



The "Color by Factor" feature allows users to select a specific factor from the study for representation on the graph to allow for an easier comparison between differences in the data.

- Select the "Color by Factor" drop down menu.
- Within the drop-down menu select one factor, then press the "Update" button. In this example, the "spaceflight" factor was selected from the drop down.

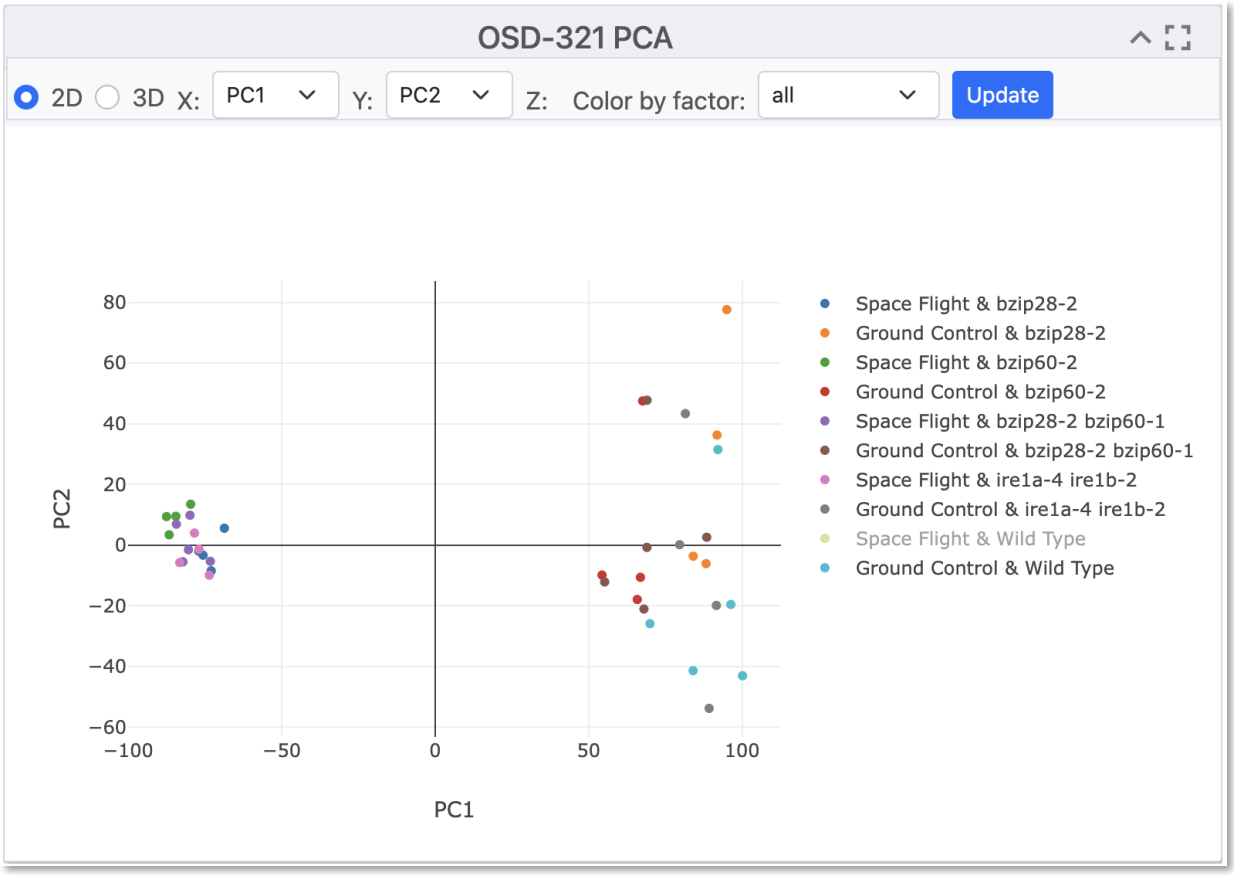


The results will now be represented by colors matching the factor that was selected.

In this example, (OSD-321) the colors are representing the different spaceflight conditions from the experiment and clearly shows how it could be a factor in the differences between the data points.

Another feature within the PCA plot tool allows users to hide factors by selecting the label located on the right side of the plot.

The labels provided represent the factor values corresponding to each sample. Click on the label “Space Flight & Wild Type” and the data points will be hidden as shown below:



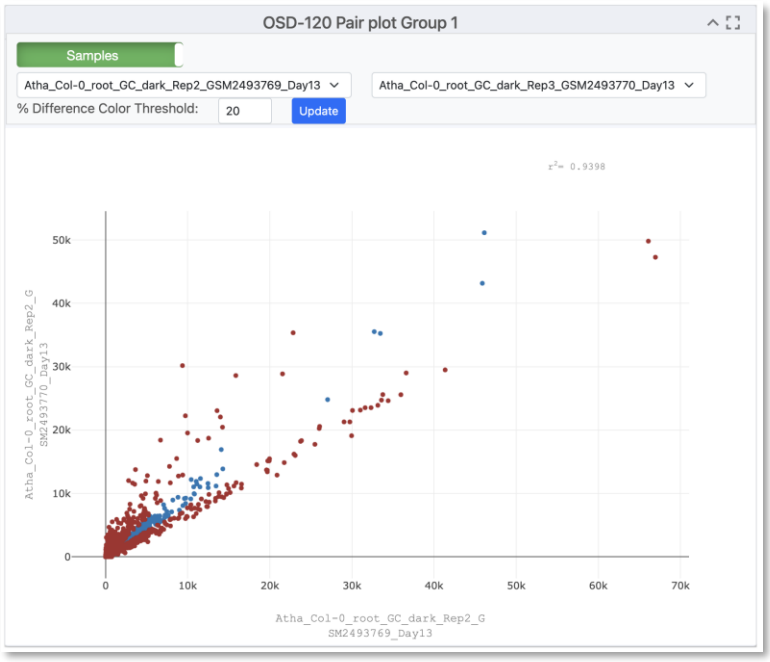
- Click on the label a second time and the data will return.

Pair Plots

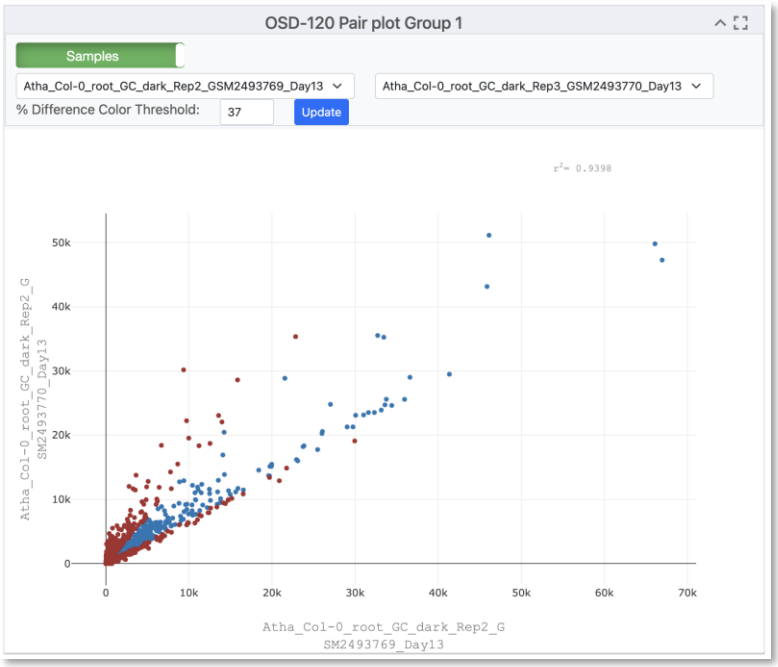
[Link to read about Pair Plots]

Pair plots are used for Exploratory Data Analysis, where the plot visualizes the data in order to find a relationship between variables that can be continuous or categorical. A Pair plot is used to understand the best set of features to explain a relationship between two variables or to form the most separated clusters. It also helps to form some simple classification models by drawing some simple lines or make linear separation in a dataset.

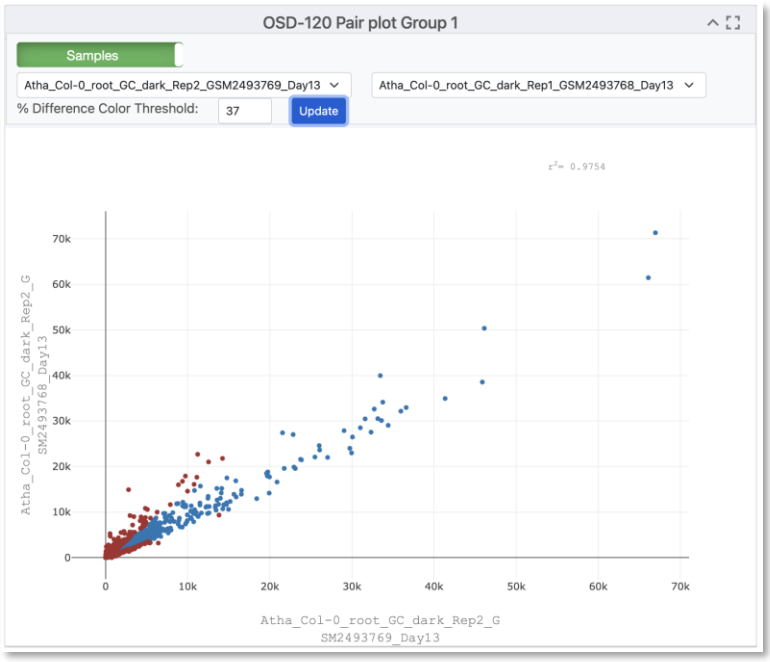
The default display for the pair plot will be the comparison between two sets of data with a % difference color threshold of 20%. Two plots will be displayed on the dashboard for the ability to compare multiple sets of data simultaneously.



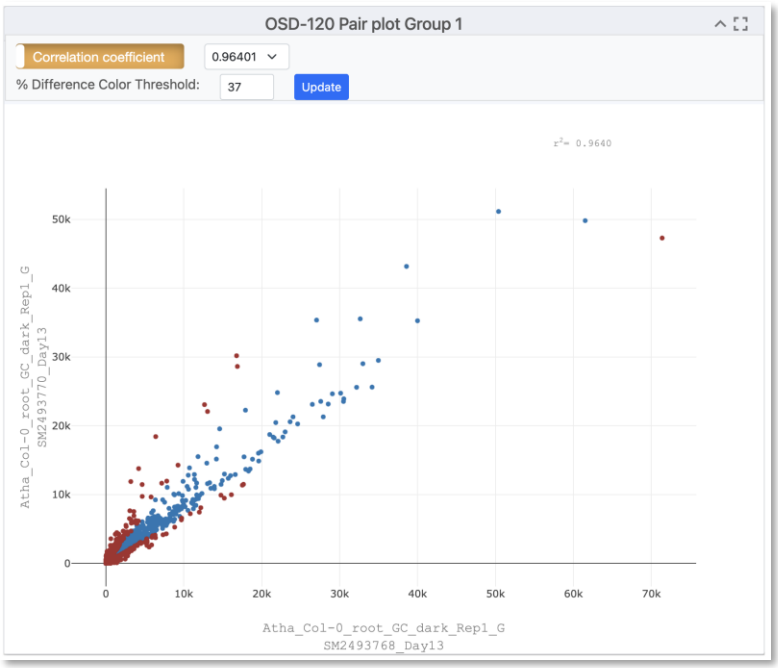
Within the plot, users have the ability to change the % difference color threshold. Below is an example of the color threshold being altered to 37%.



Clicking each of the drop-down menus will allow user to change the samples displayed. By default, the pair of samples with the lowest correlation coefficient is displayed.



Users also have the capability to view different data correlations by clicking the green "Samples" button at the top of the plot. Clicking this button will change the dropdown to show multiple correlation coefficients for a set of data. When clicking “Update” the plot will show the pair of samples with the selected correlation coefficient.

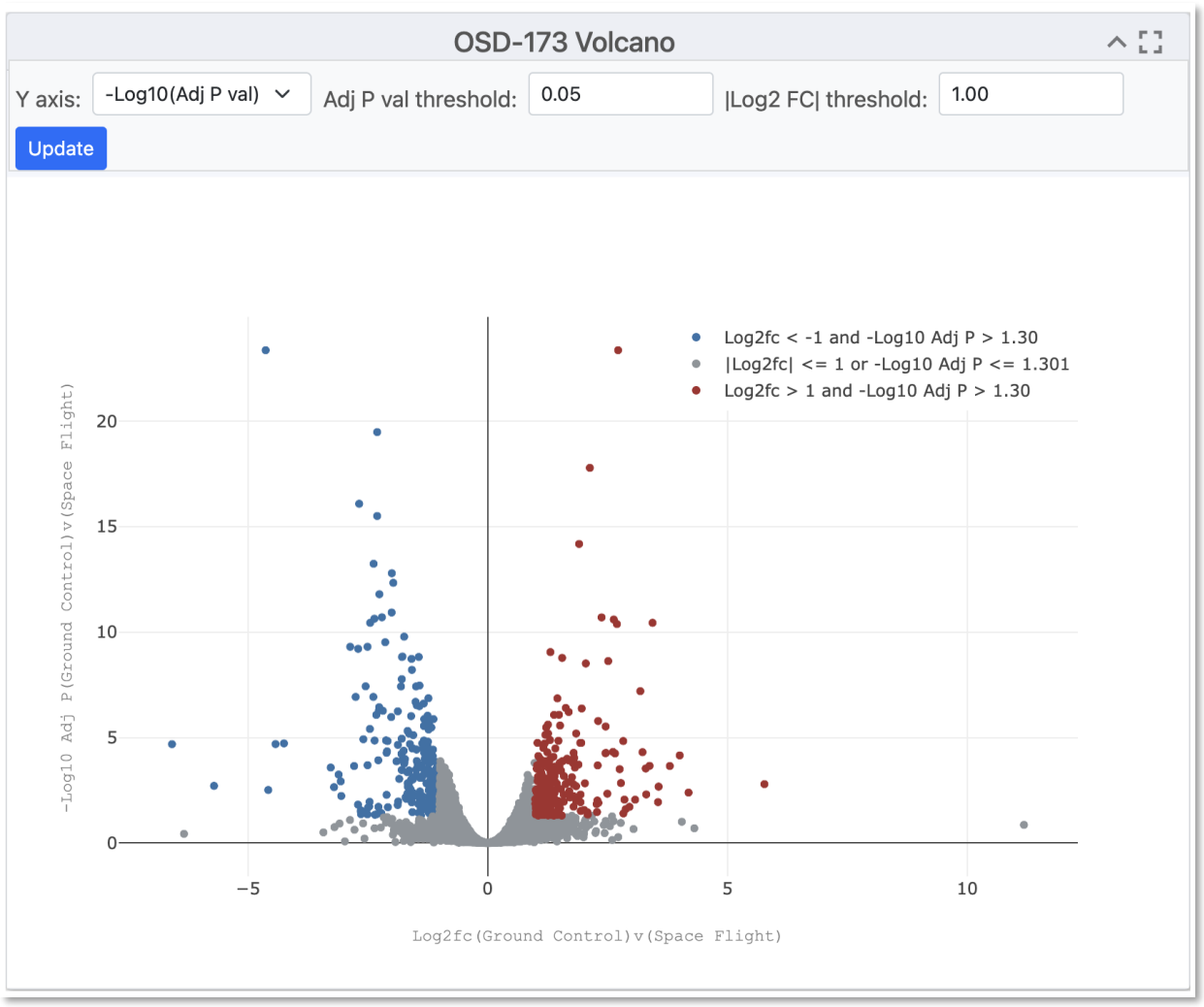


Volcano Plots

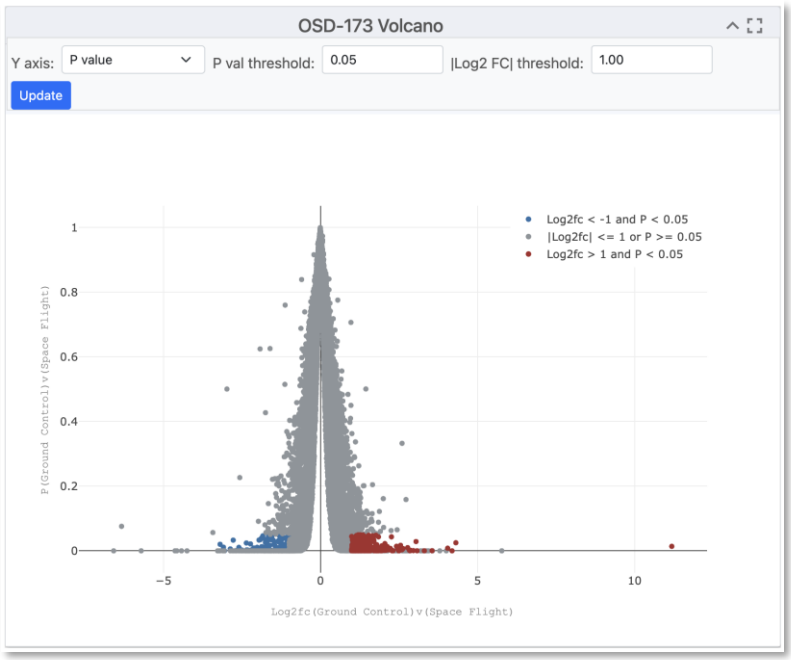
[Link to read about Volcano Plots]

A volcano plot is useful for identifying events that differ significantly between two groups of experimental subjects. The name volcano plot comes from its resemblance to a volcanic eruption with the most significant points at the top, like spewed pieces of molten lava. Each point on the graph represents a gene. The log2-fold differences between the groups are plotted on the x-axis and the -log10 P value differences are plotted on the y-axis. The horizontal dashed line represents the significance threshold specified in the analysis, usually derived using a multiple testing correction.

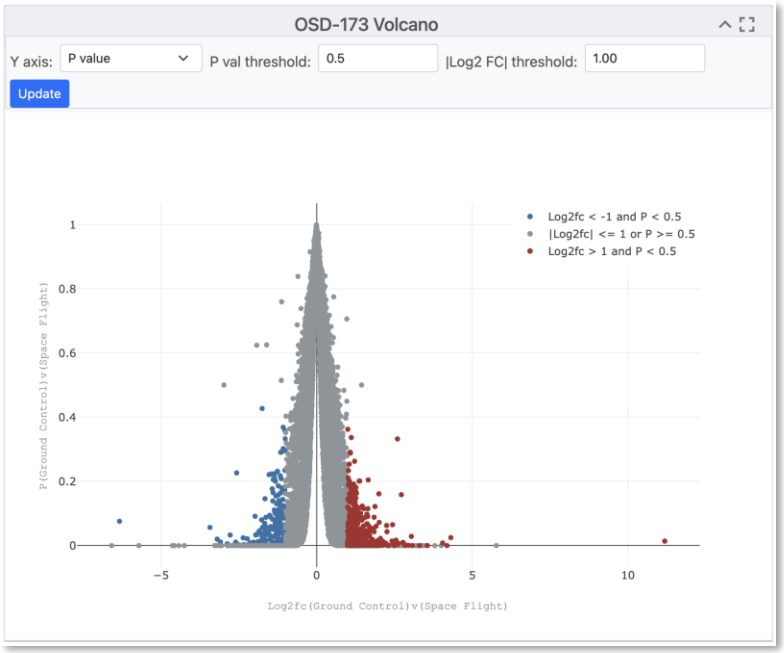
The default display for Volcano Plots will have the -Log10(Adj P Value) with and Adj P Value threshold of 0.05 and a Log2 FC threshold of 1.00 as shown below.



Users have the ability to change the type of data displayed on the Y axis, and the options from the dropdown menu include "P Value, Adjusted P Value, and -Log10(P Value)". Below is an example of the "P value" display for a volcano plot.



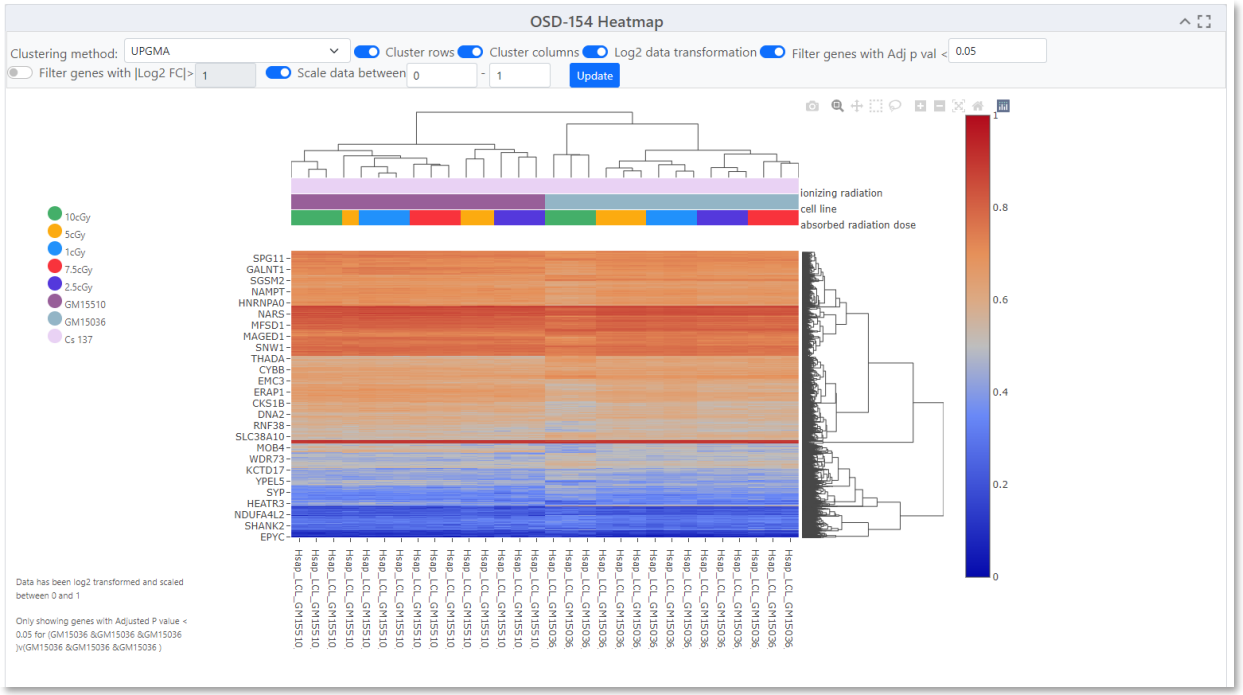
The ability to change the P value threshold is available and the image below shows a P value threshold increase to 0.5.



Heatmap

[Link to read about Heatmaps]

Heatmaps allow researchers to quickly and easily identify patterns of gene expression that are associated with specific conditions or treatments and uses color coding to indicate the magnitude of values. By measuring the number of RNA molecules produced by genes in a particular sample, researchers can determine the level of gene expression. The default settings for the heat map are shown in the image below. The heatmap links genes depending on how alike they are based on the conditions set in the experiment.

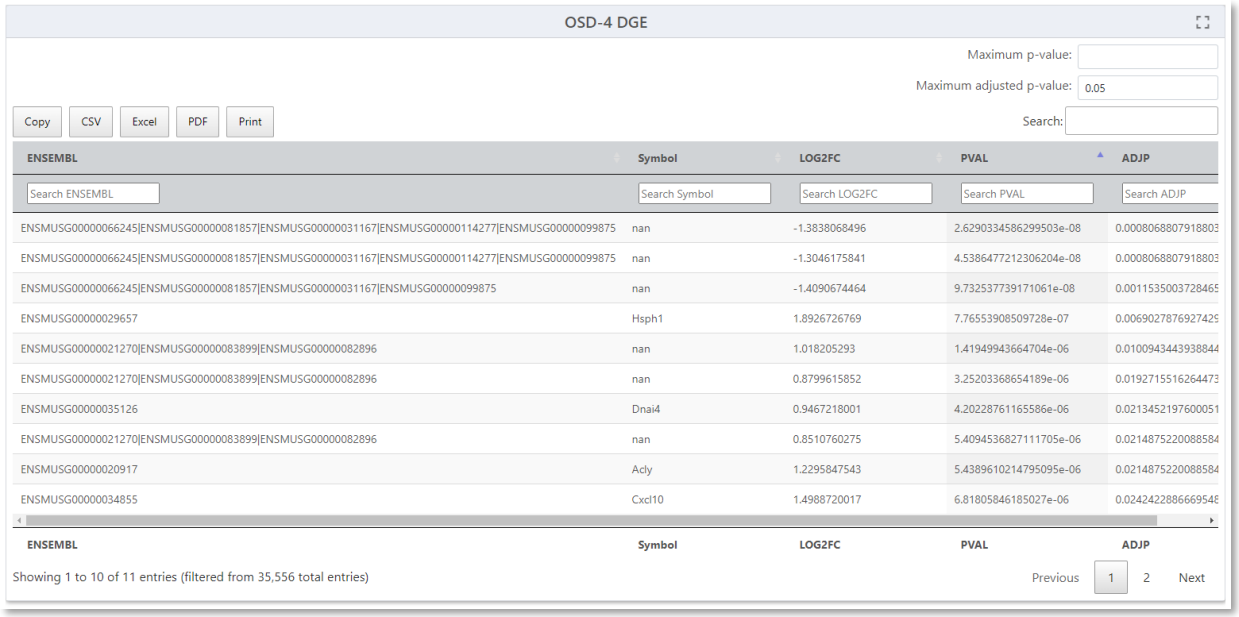


- **Choosing Clustering Method:** Users can select a clustering method to display results. The default is set to UPGMA (Unweighted Pair Group Method with Arithmetic Mean). Clustering helps group genes with similar expression profiles, making patterns more apparent.
- **Toggling Rows and Columns:** Users can toggle off rows or columns of genes. This affects how the heatmap links genes based on similarity. Toggling off rows or columns can help focus on specific groups of genes and their expression patterns. The following example shows the heatmap with clustered columns but not rows.

- **Log2 Transformation:** The Log2 transformation is available to enhance the display of genes with more pronounced differences. Applying Log2 transformation can reveal subtler variations in gene expression between conditions.
- **Filtering by Significance:** Users can filter the genes displayed on the heatmap based on their significance. By default the heatmap shows only the genes with an Adjusted P Value < 0.05 for the selected group. This filtering helps highlight genes with statistically significant expression changes. If the study has no genes with an Adjusted P Value smaller than the value entered, an error message will be displayed saying “No genes passed through the filtering condition”. The user can modify the value and Update the plot.

DGE Table

Each study will have an associated Differential Gene Expression (DGE) table available that includes information on each individual sample from the study.



To export a Differential Gene Expression (DGE) table from the study visualization page, follow these steps:

- **Locate the DGE Table:** Scroll down to the bottom of the study visualization page. There, you'll find the Differential Gene Expression table containing valuable data.

- **Copy the Table to Clipboard:** Identify the "Copy" button within the DGE table. It should be prominently displayed. Click the "Copy" button. This action will copy the entire table, including all data, headers, and values, to your device's clipboard.

- **Save as CSV, Excel, PDF, or Print:** To save the data in various file formats, look for the corresponding buttons.

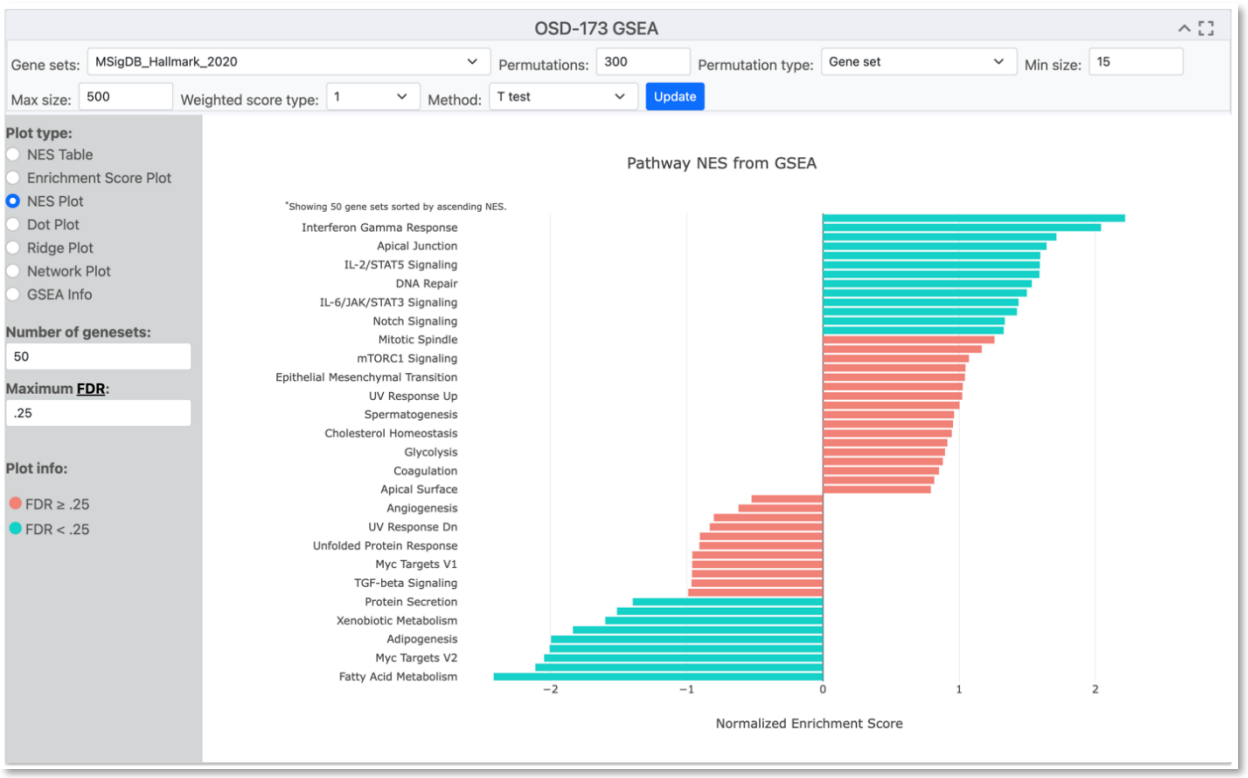
    - *For a CSV file:* Locate and click the "CSV" button. This will prompt a download of the DGE table data in CSV format to your device.
    - *For an Excel file:* Look for the "Excel" button. Click it to initiate the download of the DGE table data in Excel format (XLSX) to your device.
    - *For a PDF file:* Find and select the "PDF" button. This action will convert the DGE table into a PDF file that you can save to your device.
    - *For Printing:* Spot the "Print" button. Clicking this will open a new window displaying a printer-friendly version of the DGE table. You can then use your browser's print functionality to print the table directly.

By following these steps, you can export the Differential Gene Expression table data from the study visualization page in a variety of formats. Choose the method that best suits your needs to access and analyze the DGE data efficiently.

GSEA

[Link to read about GSEA]

GSEA stands for gene set enrichment analysis, a method to identify gene groups that are overrepresented in a large gene set. It uses statistics to pinpoint significantly enriched or depleted gene classes.



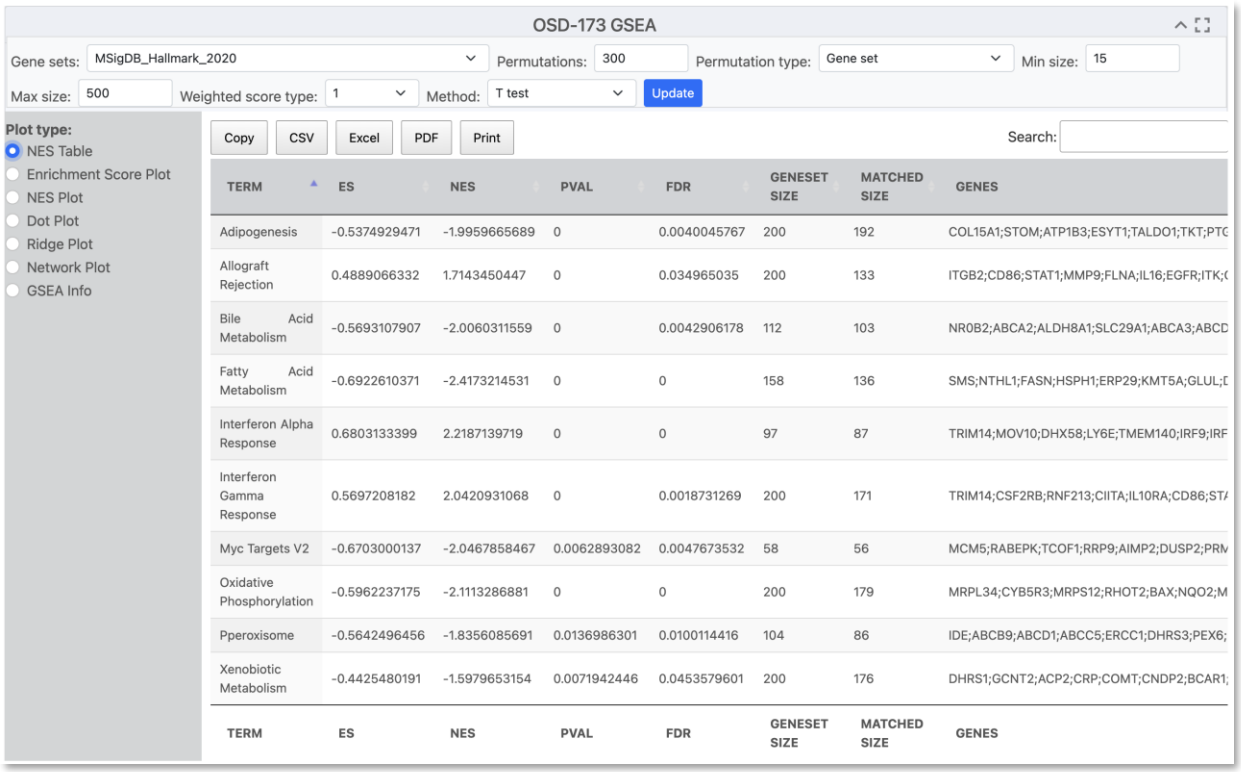
On the Gene Lab Visualization Portal, you'll find a dedicated GSEA section for each study. Within this GSEA section, there are various parameters you can customize:

- **Choose Gene Sets:** Opt for the gene sets to filter from. The default is "KEGG 2019".
- **Permutations:** Decide on the number of permutations you desire and whether they're based on phenotypes or gene sets. The default is 300.
- **Gene Number Range**: Adjust the minimum and maximum gene sizes. Increasing the minimum size omits genes with fewer than 15 data points, same for the maximum size.

- **Weighted Score Type:** Defaults to one, representing the t-test. Alternatively, choose signal-to-noise, fold change, or log2 fold change.
- **Statistical Method:** Select your preferred statistical method. The default is the t-test.

To update the plot with your changes, simply click "Update." A range of plot types is available:

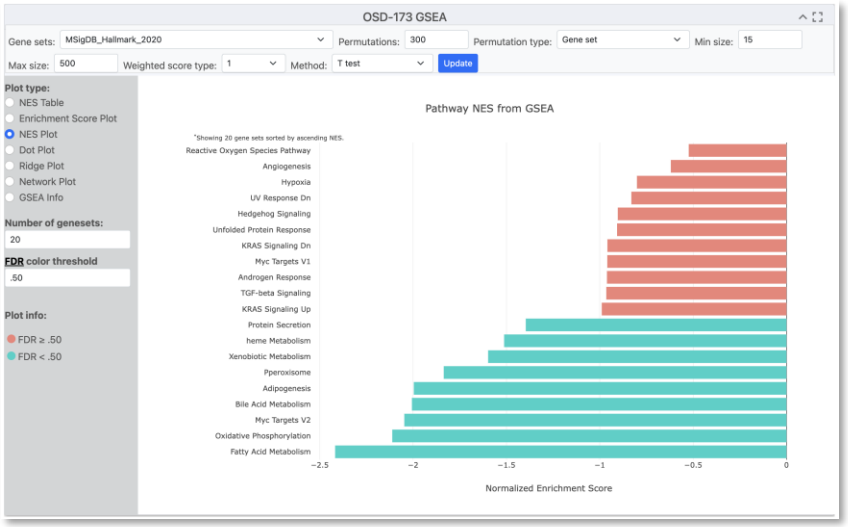
- **NES Table:** View different gene sets in a table format. Export this table using the options at the top. The table includes the Enrichment Score (ES), Normalized Enrichment Score (NES), P value (PVAL), False Discovery Rate (FDR), Gene set size, Matched size and genes for each resulting gene set.



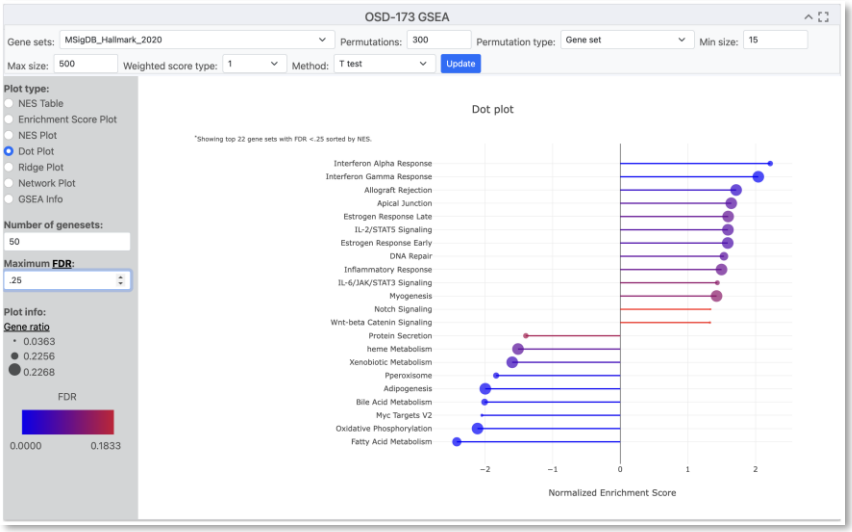
TERM	ES	NES	PVAL	FDR	GENESET SIZE	MATCHED SIZE	GENES
Adipogenesis	-0.5374929471	-1.9959665689	0	0.0040045767	200	192	COL15A1;STOM;ATP1B3;ESYT1;TALDO1;TKT;PTC
Allograft Rejection	0.4889066332	1.7143450447	0	0.034965035	200	133	ITGB2;CD86;STAT1;MMP9;FLNA;IL16;EGFR;ITK;(
Bile Acid Metabolism	-0.5693107907	-2.0060311559	0	0.0042906178	112	103	NR0B2;ABCA2;ALDH8A1;SLC29A1;ABCA3;ABCD
Fatty Acid Metabolism	-0.6922610371	-2.4173214531	0	0	158	136	SMS;NTHL1;FASN;HSPH1;ERP29;KMT5A;GLUL;[
Interferon Alpha Response	0.6803133399	2.2187139719	0	0	97	87	TRIM14;MOV10;DHX58;LY6E;TMEM140;IRF9;IRF
Interferon Gamma Response	0.5697208182	2.0420931068	0	0.0018731269	200	171	TRIM14;CSF2RB;RNF213;CIITA;IL10RA;CD86;STA
Myc Targets V2	-0.6703000137	-2.0467858467	0.0062893082	0.0047673532	58	56	MCM5;RABEPK;TCOF1;RRP9;AIMP2;DUSP2;PRM
Oxidative Phosphorylation	-0.5962237175	-2.1113286881	0	0	200	179	MRPL34;CYB5R3;MRPS12;RHOT2;BAX;NQO2;M
Pperoxisome	-0.5642496456	-1.8356085691	0.0136986301	0.0100114416	104	86	IDE;ABCB9;ABCD1;ABCC5;ERCC1;DHRS3;PEX6;
Xenobiotic Metabolism	-0.4425480191	-1.5979653154	0.0071942446	0.0453579601	200	176	DHRS1;GCNT2;ACP2;CRP;COMT;CNDP2;BCAR1,
TERM	ES	NES	PVAL	FDR	GENESET SIZE	MATCHED SIZE	GENES

- **NES Plot:** The default plot displays normalized enrichment scores based on gene sets. The number of gene sets displayed can be modified by adjusting the number under “Number of gene sets” under the plot options. By default, the gene sets with an FDR bigger or equal to 0.25 are displayed in red while the gene sets with an FDR smaller than 0.25 are displayed in blue. FDR indicates the likelihood that a result is valid, e.g., FDR of 0.25 means a 25% chance of validity. This threshold can also be modified.

In the example below the plot has been modified to show only the first 10 gene sets ordered by ascending NES and gene sets with an FDR smaller than 0.50 are displayed in blue.

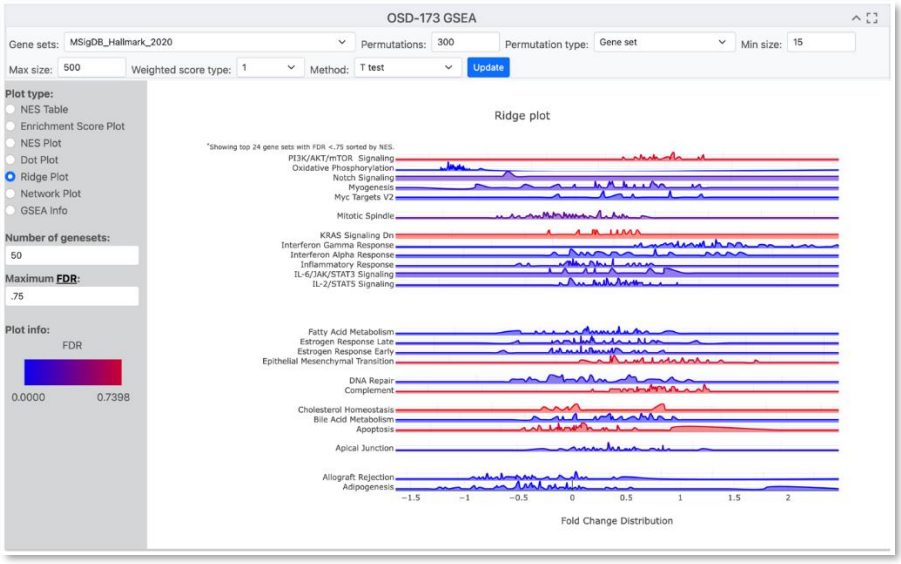


- **Dot Plot:** Similar to NES Plot, it showcases the top gene sets sorted by NES and filtered by FDR. The FDR threshold for gene sets to be displayed can be modified by adjusting the number under “Maximum FDR”. The size of each dot represents the gene ratio. The color of each dot represents its FDR, blue being the lowest value and red being the highest.

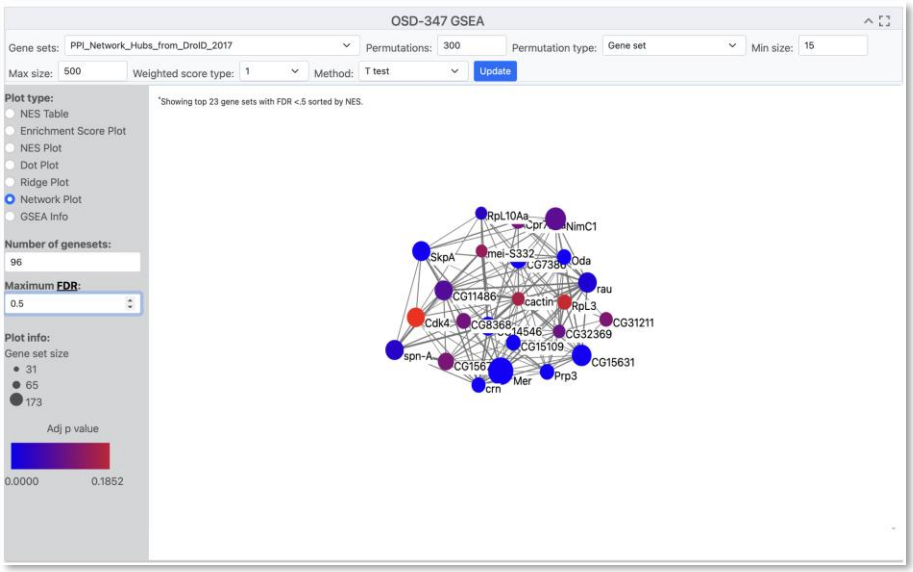


- **Enrichment Plot:** This reveals the fold change distribution of the top gene sets with an FDR of under 0.25. As with previous plots, the number of gene sets displayed, and FDR threshold can be modified. The color of each plot represents

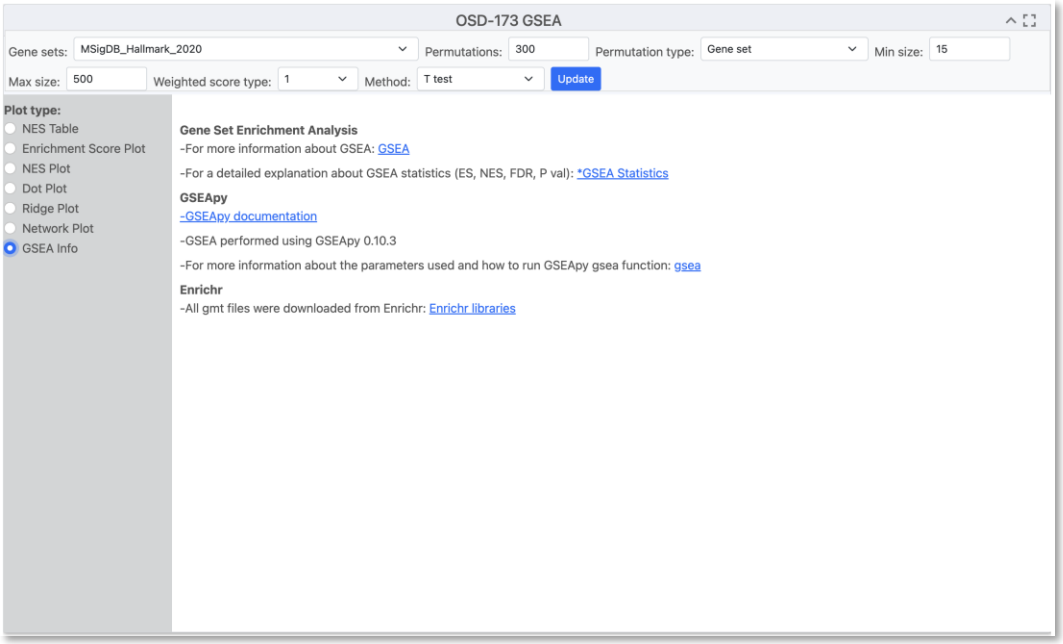
the FDR value for the gene set, blue being the lowest value and red being the highest value.



- **Network Plot:** Allows you to visualize relationships between gene sets using a network plot. The size of each dot represents the gene set size, while the color of each dot represents the adjusted p value for the gene set (blue being the lowest value and red being the highest). As with previous plots, the number of gene sets displayed, and FDR threshold can be modified. Hover overing one dot will highlight the gene sets it is connected to.

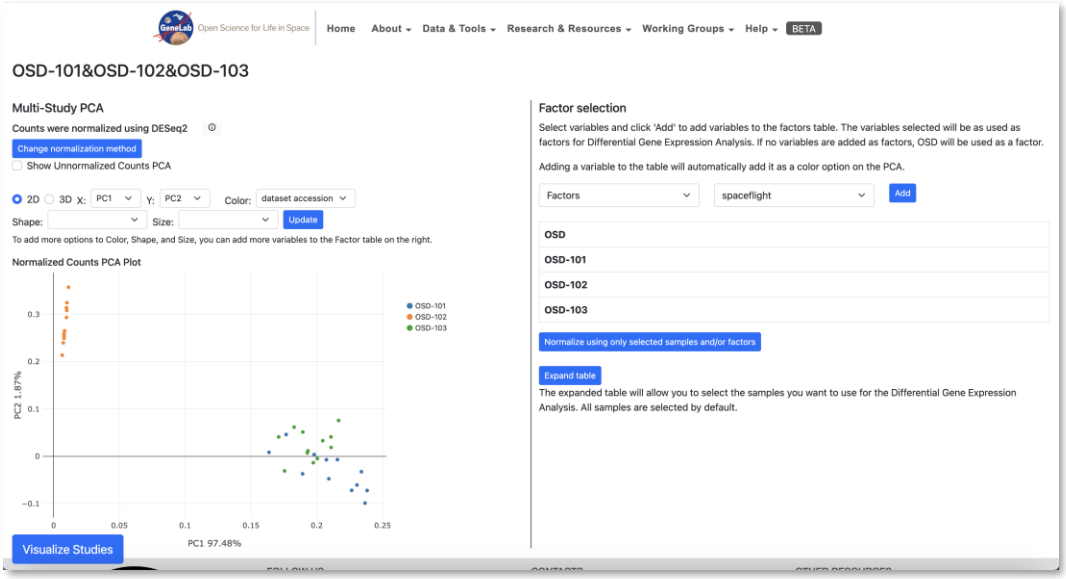


- **GSEA Info:** For in-depth details about GSEA creation, statistics, and plot documentation.



With these steps, you can effectively navigate and utilize the GSEA section, gaining insights into gene set enrichment analysis for your study.

Multi-Study



Multi-Study analysis is an advanced tool designed for analyzing and visualizing multiple RNA sequencing studies concurrently. The multi-study page is used to initialize the parameters for data visualization of the multiple studies. Researchers can uncover intricate patterns of gene expression associated with specific conditions or treatments across a variety of experiments.

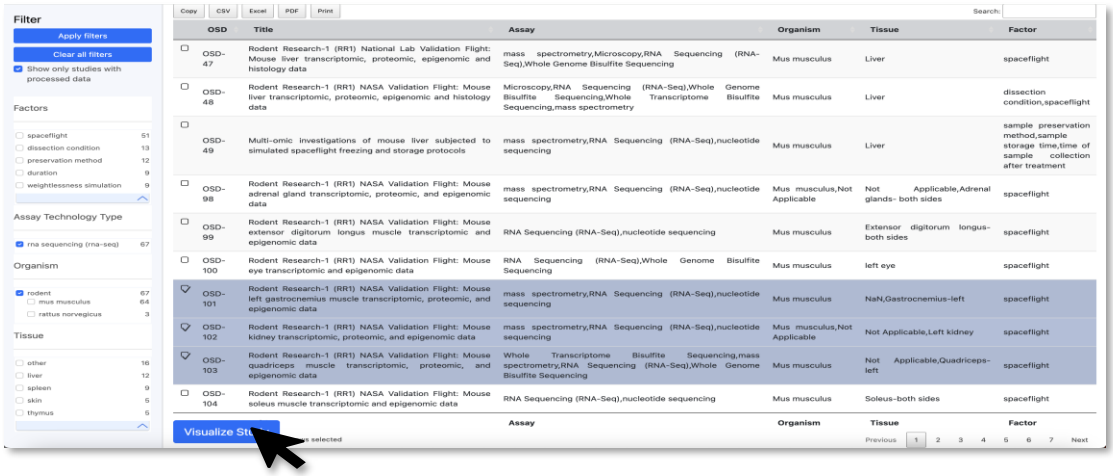
To combine studies, there are two main steps: selecting the samples and factors you wish to use and running the Differential Gene Expression (DGE) analysis. Factors are used as the groups between which differential expression is calculated.

To help you select the samples and variables to be used as factors, you will be re-directed to a page containing the PCA plot for the combined studies, a table to add variables to be used as factors and a table to select samples. This page has the following components:

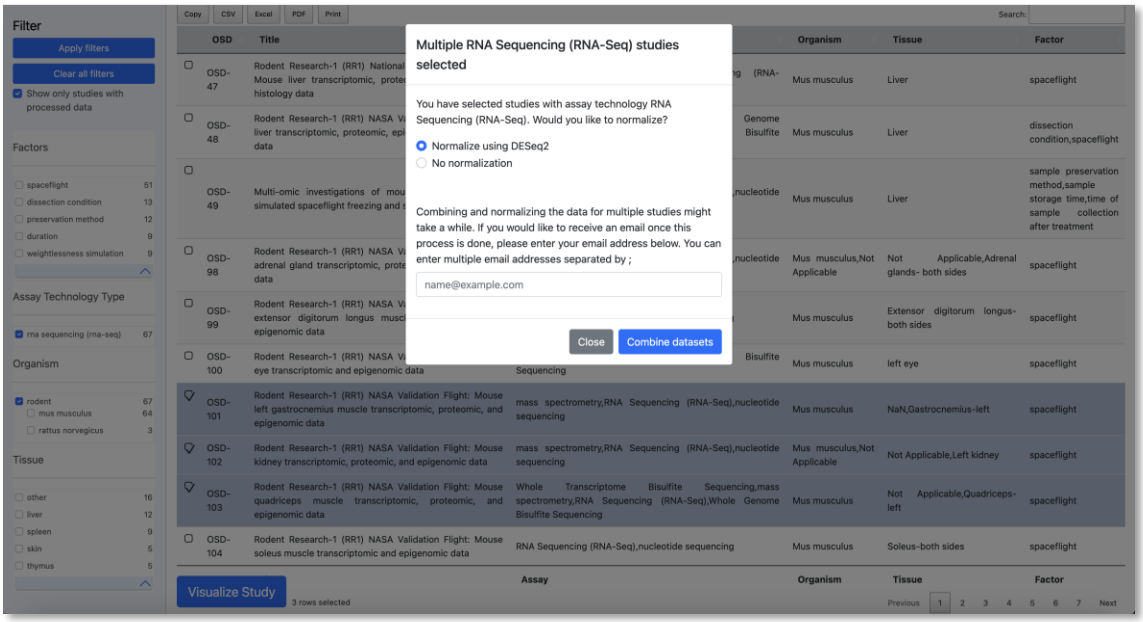
- Information about how studies were combined and normalized.
- PCA plot of normalized counts.
- PCA plot of unnormalized counts.
- Factor selection table: Displays each studies' values for the variables added to the table.
- Sample selection table
- Option to normalize data using a subset of samples and specific variables as factors.

Below are detailed instructions on how to effectively navigate and utilize the Multi-Study Page:

- Selecting studies for Multi-Study analysis:
    - For your initial test, let's use rodent studies as an example.
    - Start by selecting "rodent" as the organism of interest.
        - *Since combining DNA microarray assays is not yet supported, ensure to filter by both "rodent" and "RNA sequencing" in the assay technology type.*
    - Choose multiple different rodent studies that encompass various tissue types. For this example, OSD-101, OSD-102 and OSD-103 were selected.
    - Mark the checkboxes beside the selected studies in the studies table.
    - Click the "Visualize Study" button to proceed.

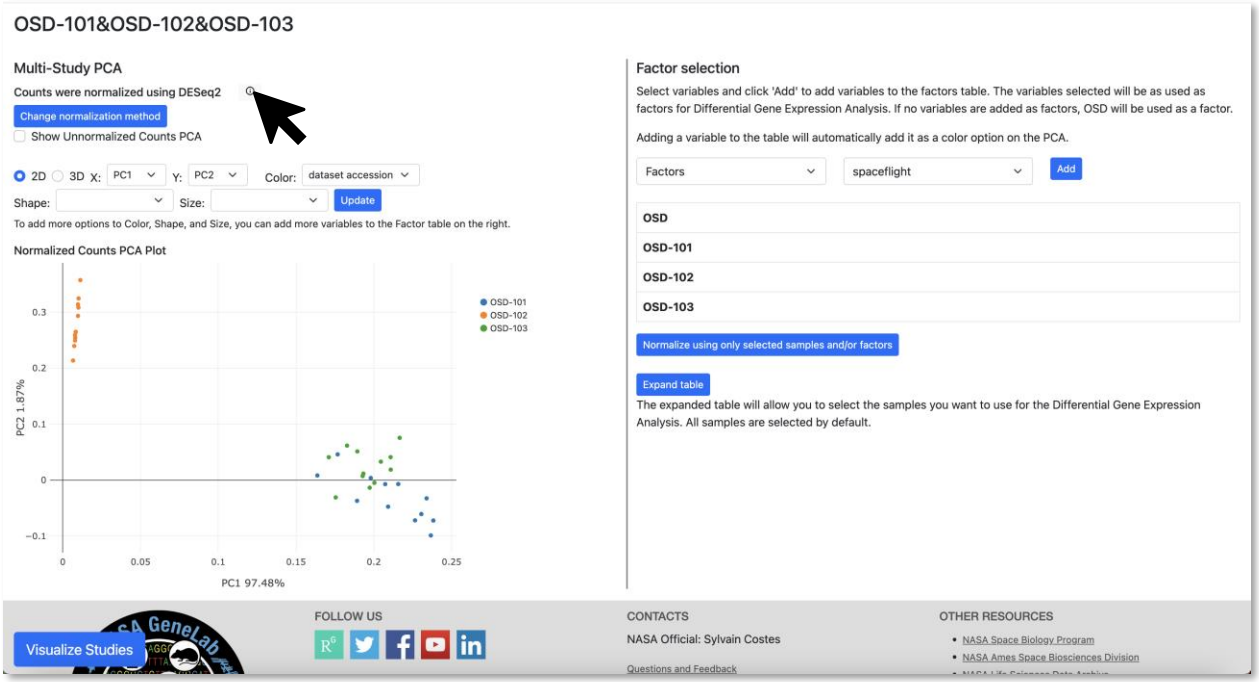


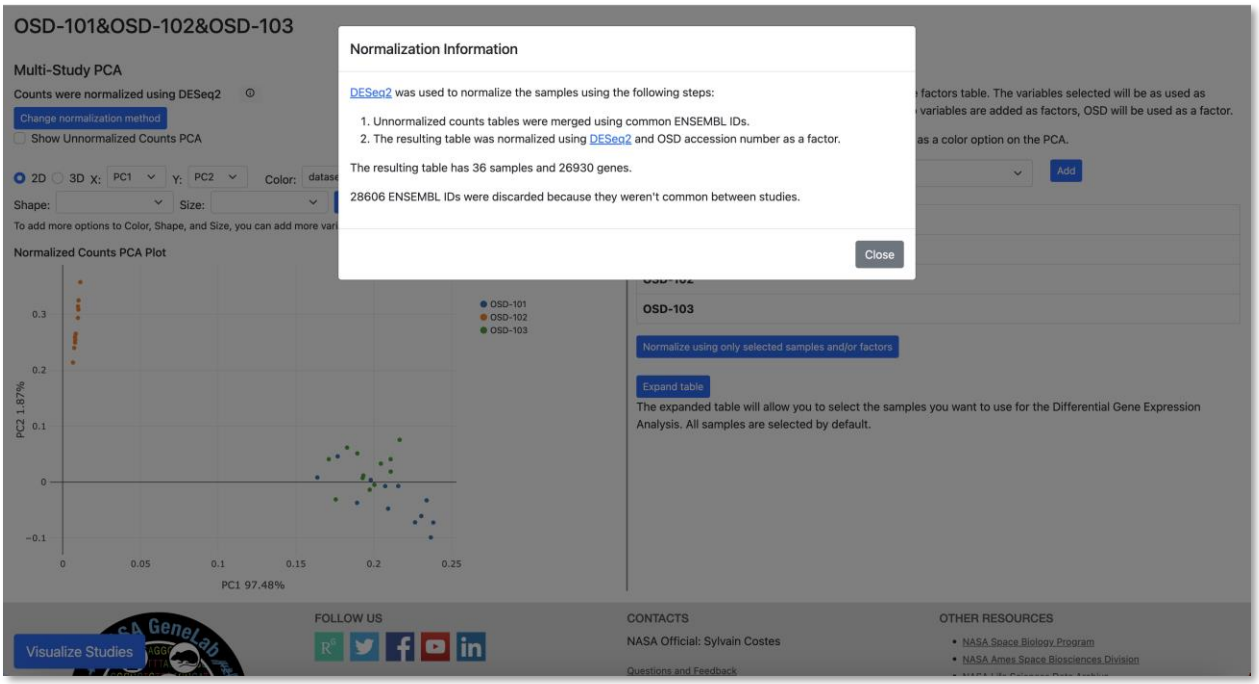
- Data Normalization:
  - A dialog box will appear to prompt you for data normalization options.
  - The default selection is "DESeq2" for normalization, but you can also choose "No Normalization". If normalization is chosen, the studies will be normalized by their accession number.
  - If desired, you can enter your email address to receive a notification when the studies have been combined. The email will contain a URL you can use to access the combined studies.
    - Alternatively, proceed without entering an email address.



Exploring the Multi-Study Page

Information about how the studies were combined and normalized and the resulting number of samples, genes and discarded genes can be displayed by clicking on the information icon next to the label “Counts were normalized using DESeq2”.

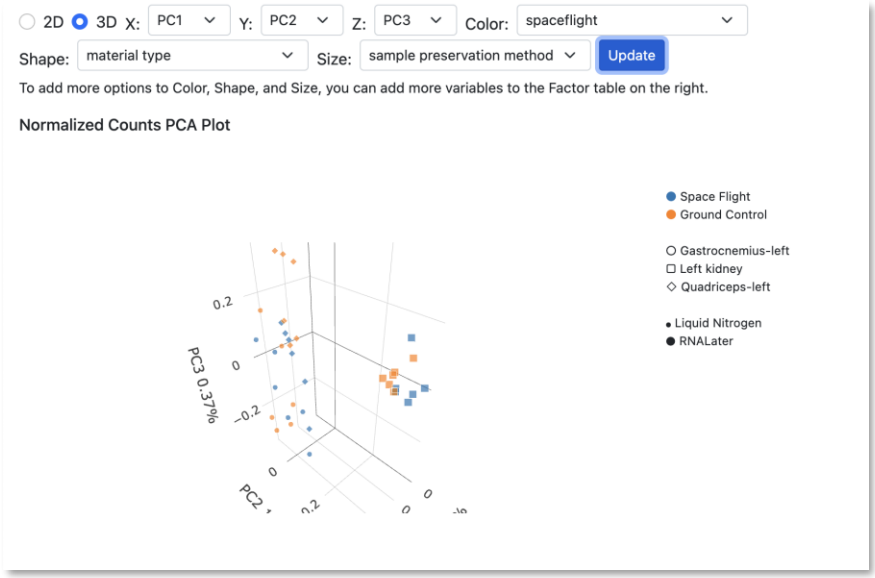




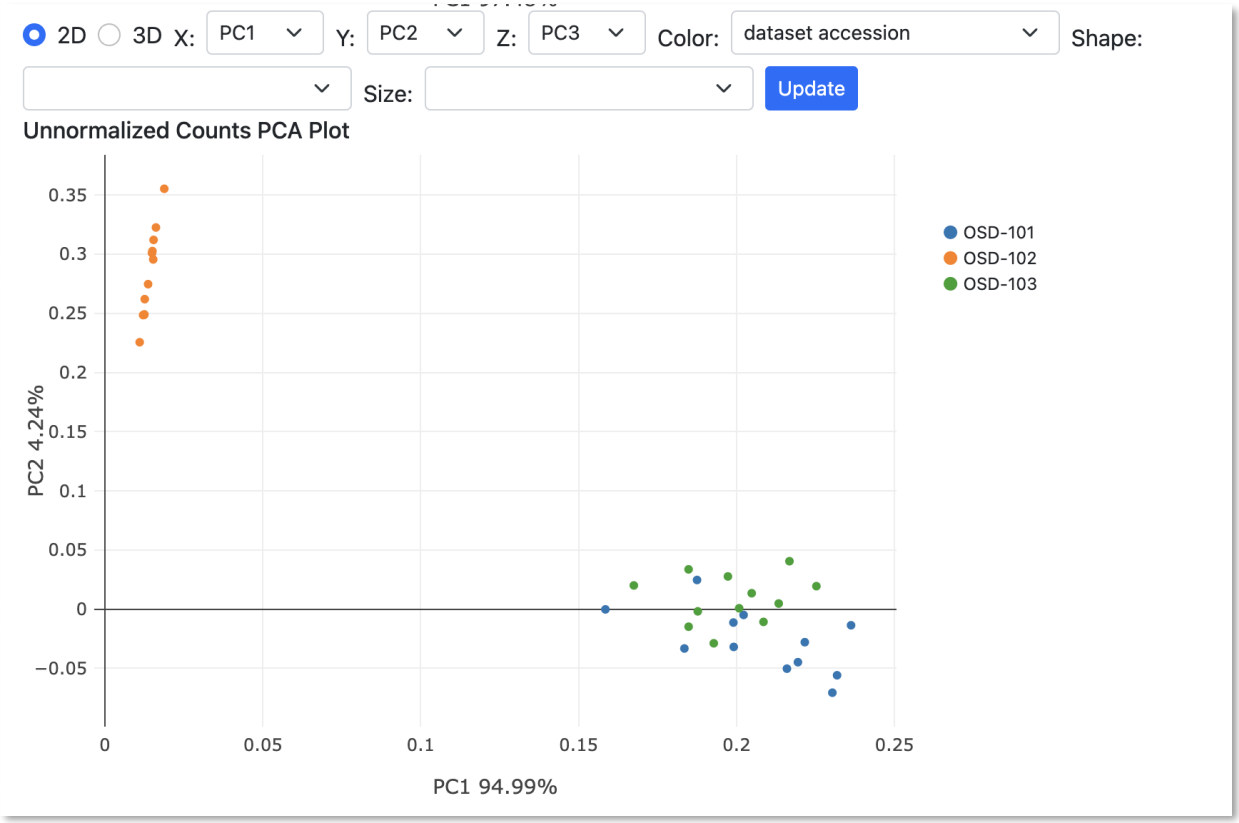
Multi-Study PCA Plot

A PCA plot for data visualization will be included on the multi-study page. If the data was normalized, the PCA plot for normalized counts will be displayed at the top. You can use the drop-down menus to change the color, shape and size on each dot based on each sample's condition. The variables available on each drop-down menu are dynamically added when you add variables to the Factors table.

The PCA plot can be displayed as a 2D or 3D plot



You can display the PCA plot of the Unnormalized counts by clicking on the checkbox next to “Show Unnormalized Counts PCA”. The unnormalized counts PCA plot has the same features as the normalized counts PCA, it can be displayed as a 2D or 3D plot and the color, shape and size of the dots can be customized.



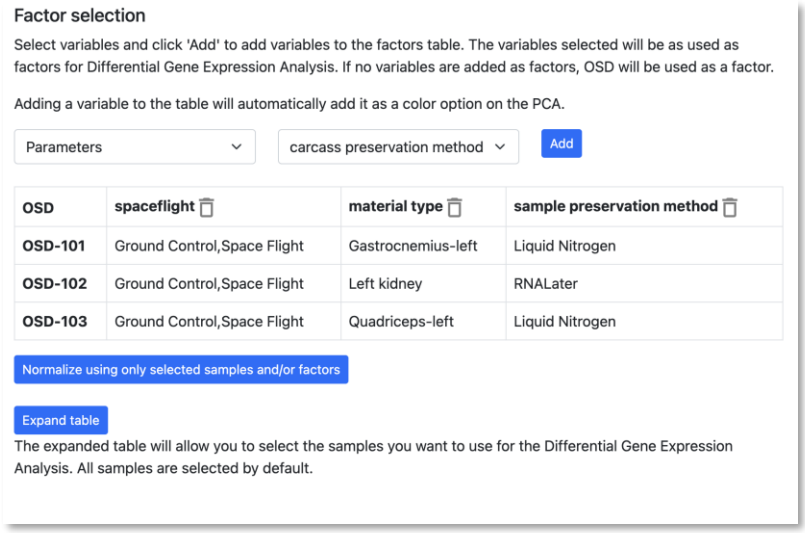
Factor Selection for Gene Expression Analysis

The variables selected as factors will be used to group samples and calculate differential expression. The counts table will also be normalized based on these groups. For example, on the image below the user has added material type to the factors table. If the user choses to use material type as a factor, all samples with material type “Gastrocnemius-left” will be treated as one group, samples with material type “Left kidney” will be treated as a second group and samples with material type “Quadriceps-left” will be treated as a third group. DGE Analysis will then show comparisons (Fold Change, P value, Adjusted P value) between Gastrocnemius-left & Left kidney, Gastrocnemius-left & Quadriceps-left, Left kidney & Quadriceps-left.

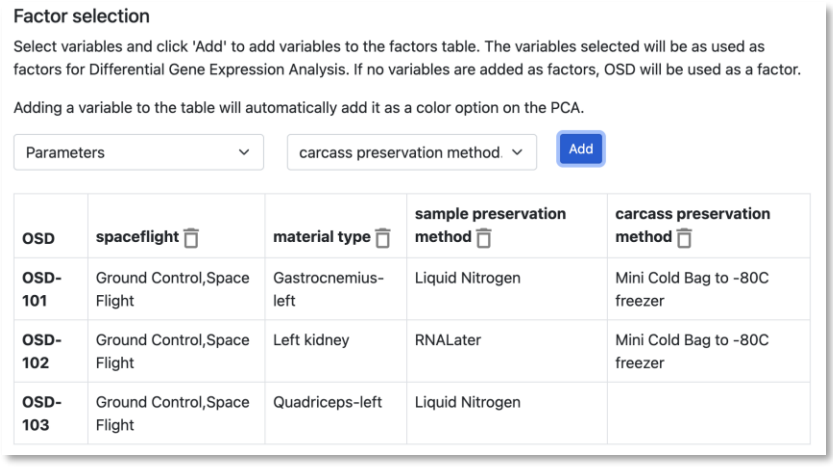
Any characteristic, parameter or factor on the studies’ metadata can be added to the table. Adding a variable to the table will also add it as an option to change the color,

size, and shape of the PCA plots. Under "Factor Selection," choose variables to generate a factors table for differential gene analysis.

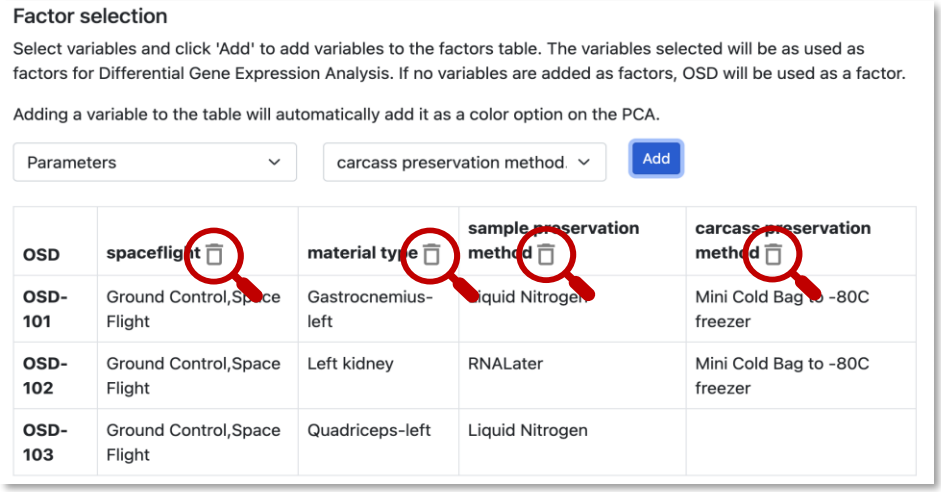
Select parameters, characteristics, or factors from the dropdown list to add to the table. The factors, characteristics and parameters are retrieved from the studies metadata ISA files. More information about the ISA model and what defines a parameter, characteristic or factor can be found here https://isa-specs.readthedocs.io/en/latest/isamodel.html .



The table will display the values each study contains for the added variables. If a variable is not common between studies and the information is not available on one of the studies' metadata, the table will display an empty value and it will be filled with NA for the next steps.

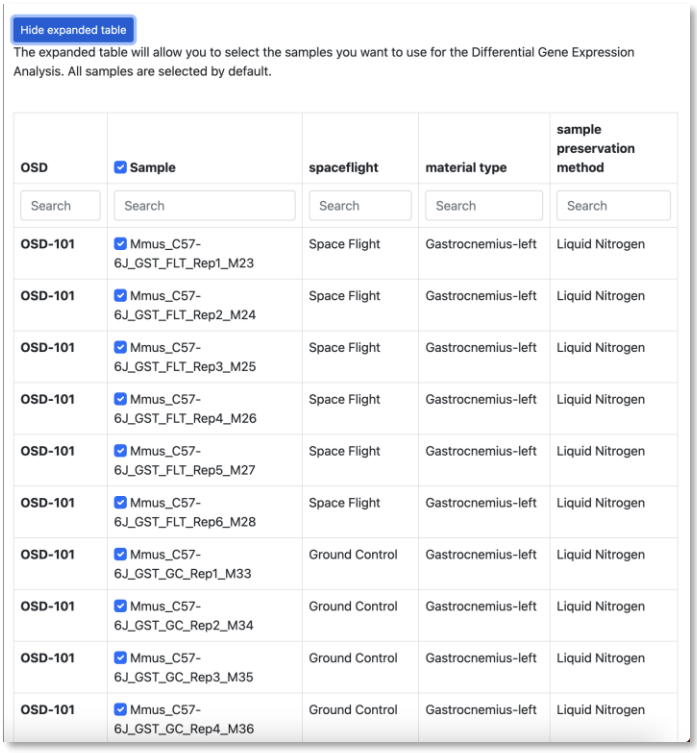


Variables can be deleted from the table by clicking on the trash icon next to each column header.

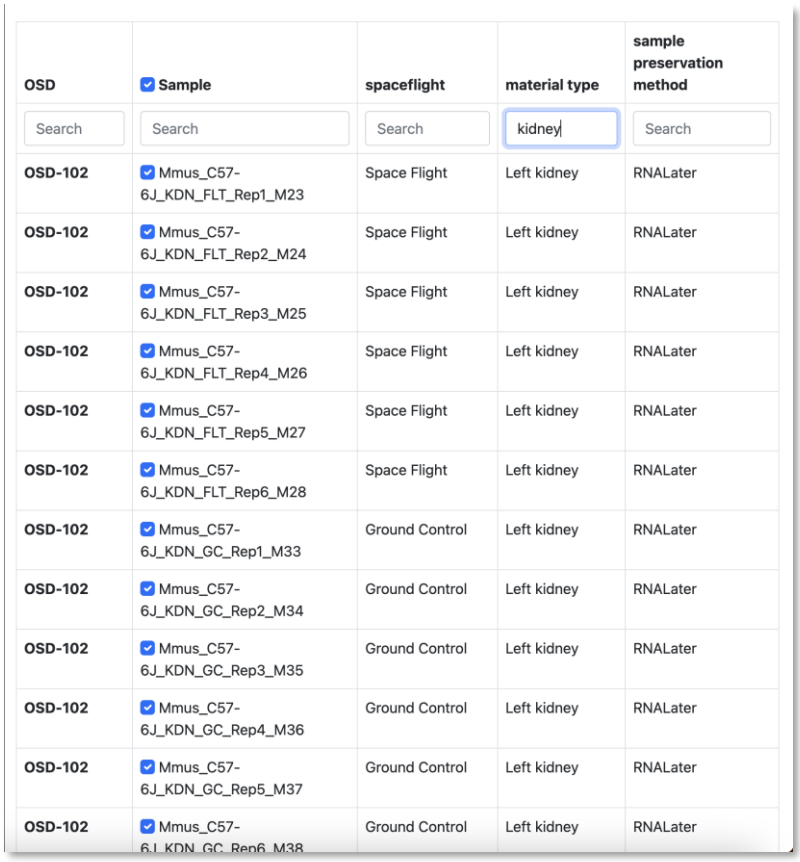


Sample Selection for Gene Expression Analysis

Select specific samples by clicking the "Expand table" button. This will display a table with all the samples and their values for the variables added to the factors table.



You can use the checkbox next the Samples table header to select all samples. The search boxes underneath each column header can be helpful to filter samples with a specific value.



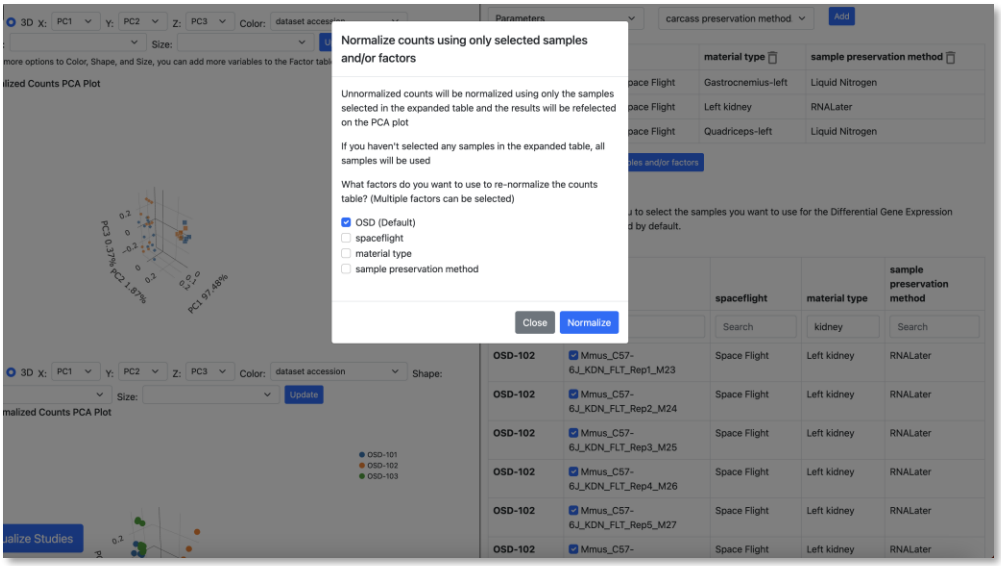
OSD	Sample	spaceflight	material type	sample preservation method
Search	Search	Search	kidney	Search
OSD-102	Mmus_C57-6J_KDN_FLT_Rep1_M23	Space Flight	Left kidney	RNALater
OSD-102	Mmus_C57-6J_KDN_FLT_Rep2_M24	Space Flight	Left kidney	RNALater
OSD-102	Mmus_C57-6J_KDN_FLT_Rep3_M25	Space Flight	Left kidney	RNALater
OSD-102	Mmus_C57-6J_KDN_FLT_Rep4_M26	Space Flight	Left kidney	RNALater
OSD-102	Mmus_C57-6J_KDN_FLT_Rep5_M27	Space Flight	Left kidney	RNALater
OSD-102	Mmus_C57-6J_KDN_FLT_Rep6_M28	Space Flight	Left kidney	RNALater
OSD-102	Mmus_C57-6J_KDN_GC_Rep1_M33	Ground Control	Left kidney	RNALater
OSD-102	Mmus_C57-6J_KDN_GC_Rep2_M34	Ground Control	Left kidney	RNALater
OSD-102	Mmus_C57-6J_KDN_GC_Rep3_M35	Ground Control	Left kidney	RNALater
OSD-102	Mmus_C57-6J_KDN_GC_Rep4_M36	Ground Control	Left kidney	RNALater
OSD-102	Mmus_C57-6J_KDN_GC_Rep5_M37	Ground Control	Left kidney	RNALater
OSD-102	Mmus_C57-6J_KDN_GC_Rep6_M38	Ground Control	Left kidney	RNALater

Use the checkbox next to each sample name to select the samples you wish to use on the next steps.

Normalize using only selected samples and/or factors:

You can choose to normalize the combined counts table using only a subset of samples and any factor (or factors) added to the factors table. The PCA plot will be updated, and this can help you make a better decision on which variables to use as factors for DGE analysis.

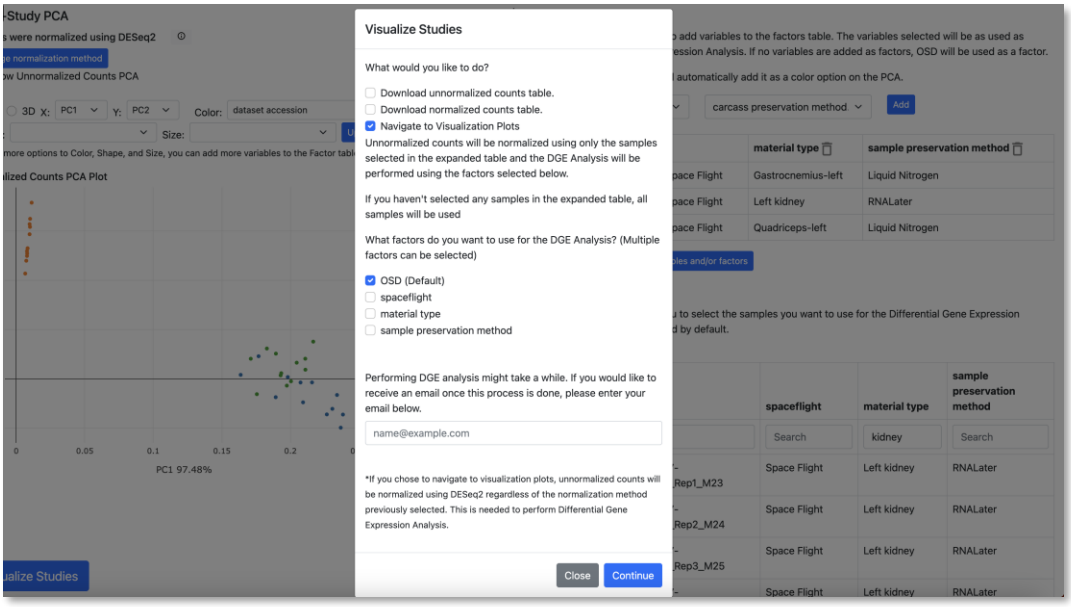
- Click on the Normalize using only selected samples and/or factors button. This will display a modal with the different options for normalization.



- Use the checkboxes to select the factor or factors you want to use to normalize. Only the samples selected on the samples table will be used.

Visualizing and Downloading Results:

Click "Visualize Studies" to proceed to visualization plots or download the counts table. A modal will be displayed with the options to continue. You can choose to download the unnormalized counts table, download the normalized counts table, or navigate to visualization plots.

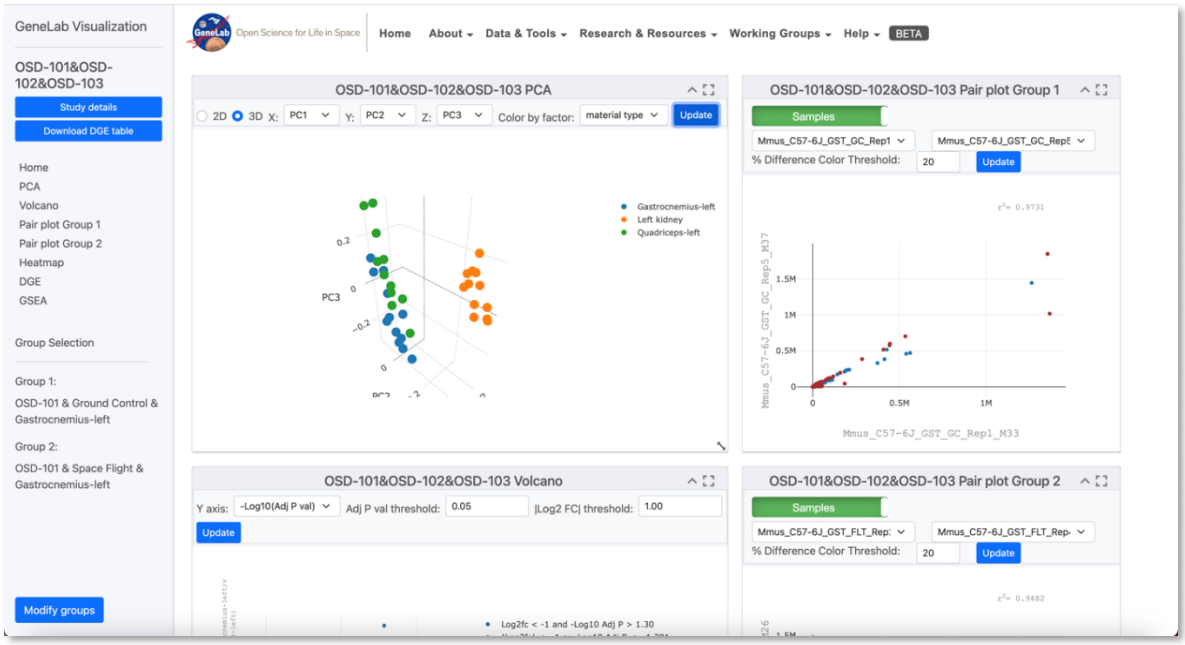


If you wish to navigate to visualization plots you can choose which variables you would like to use as factors. Only samples selected on the samples table will be used. You will have the option to enter your email to be notified when the DGE analysis is ready. The email sent will contain a URL you can use to access the visualization plots resulting of the DGE analysis

- Click the continue button to perform the selected action.

Exploring Visualization Plots:

Upon completion, the page will direct you to a range of visualization plots and graphs for your data analysis.



The visualization portal will have the same plots and capabilities as the single study visualization portal.


GeneLab Visualization
OSD-101&OSD-102&OSD-103
Study details
Download DGE table
Home
PCA
Volcano
Pair plot Group 1
Pair plot Group 2
Heatmap
DGE
GSEA
Group Selection
Group 1:
OSD-101 & Ground Control & Gastrocnemius-left
Group 2:
OSD-101 & Space Flight & Gastrocnemius-left
Modify groups
Home
About
Data & Tools
Research & Resources
Working Groups
Help
BETA
OSD-101&OSD-102&OSD-103 PCA
OSD-101&OSD-102&OSD-103 Pair plot Group 1
OSD-101&OSD-102&OSD-103 Volcano
OSD-101&OSD-102&OSD-103 Pair plot Group 2
OSD-101&OSD-102&OSD-103 Heatmap
OSD-101&OSD-102&OSD-103 DGE
OSD-101&OSD-102&OSD-103 GSEA
Pathway NES from GSEA
Normalized Enrichment Score

With these comprehensive instructions, you are well-equipped to navigate the Multi-Study Page efficiently. This tool empowers you to analyze multiple RNA sequencing studies simultaneously, uncovering complex gene expression patterns and gaining valuable insights into your experimental data.