

User Manual for scCompare

scCompare is a web-based platform for scRNA-seq comparisons in autoimmune diseases.

To begin with the scCompare tool, the user can select the cell type (orange box in Figure 1). Users can either type the cell type of interest in the search box or click on triangle to expand the drop-down list that include 121 cell types. Users are allowed to select multiple cell types. Then, the users are required to select disease of interest (blue box in Figure 1). Users can select one or multiple diseases of interest by typing in the disease name in input box or by clicking on the disease from the drop-down menu. This menu is updated based on the selection made in the cell type section.

After selecting cell type (fibroblast) and disease (AD, CD and PSO), users can select other optional attributes (e.g. tissue or treatment of interest) highlighted in yellow box in Figure 1. Keep in mind that dropdown options change based on the selection made in the previous step.

After confirming selections (green box in Figure 1), related datasets will be displayed in the main console, users can click on checkboxes to select datasets. Users can also click on the dataset_acc (black box in Figure 1) to view the summary of a particular dataset (shown in Figure 2).

scCompare

HomeDataset SelectionResults Panel

Developed by Immunology Computational Biology - GRC

REQUIRED

Select cell types of interest

Search

cell

abnormal cell

ciliated cell

connective tissue cell

fibroblast

stromal cell

contractile cell

embryonic cell

epithelial cell

hematopoietic cell

melanocyte

mural cell

neural cell

perivascular cell

precursor cell

secretory cell

transit amplifying cell

Select diseases of interest

AD CD PSO

OPTIONAL

Select tissues of interest

Select treatment of interest

confirm selections

Compare Datasets

The table below allows you to explicitly select datasets that you would like to compare. By default, all datasets are selected. If you are only interested in a few datasets, you can click the **deselect all** button and then make your selections. Click the **Compare Datasets** button to display results across selected data. Results will be displayed in the Results Panel.

geneQueryDEGfinderDEGCompare

Enter a comma-separated list of gene names (max 10 genes)

CTHRC1, COL1A1, KIF26B, POSTN, ADAM12, COL3A1, MMP13, SPARC

Toggle Selection

dataset_acc	title	disease	source	treatment
<input type="checkbox"/> GSE134809	Single-Cell Analysis of Crohn's Disea...	CD	Ileum	NA
<input type="checkbox"/> GSE147424	Single-cell transcriptome analysis of...	AD;Control	Skin	NA
<input type="checkbox"/> GSE150672	Highly Efficient, Massively-Parallel ...	PSO;Control	Skin	NA
<input type="checkbox"/> GSE162183	Single cell transcriptional zonation ...	PSO;Control	Skin	NA
<input type="checkbox"/> GSE180885	Single-cell analysis reveals innate L...	AD;Control	Skin	NA
<input type="checkbox"/> GSE198805	Single-cell transcriptome characteris...	PSO	Skin	NA
<input type="checkbox"/> GSE214695	IBD single cell data from colonic mucosa	CD;UC;Control	Colon	NA
<input type="checkbox"/> GSE220116	Single-cell transcriptomics suggest d...	HS;PSO;Control	Skin	Secukinumab
<input type="checkbox"/> PMID33290721_CTRL_CD	Single-Cell Sequencing of Developing ...	CD;Control	Intestine	NA
<input type="checkbox"/> PMID33479125_10X	Developmental cell programs are co-op...	AD;PSO;Control	Skin	NA

1-10 of 12 rows

Previous12Next

Figure1: Data Selection in scCompare

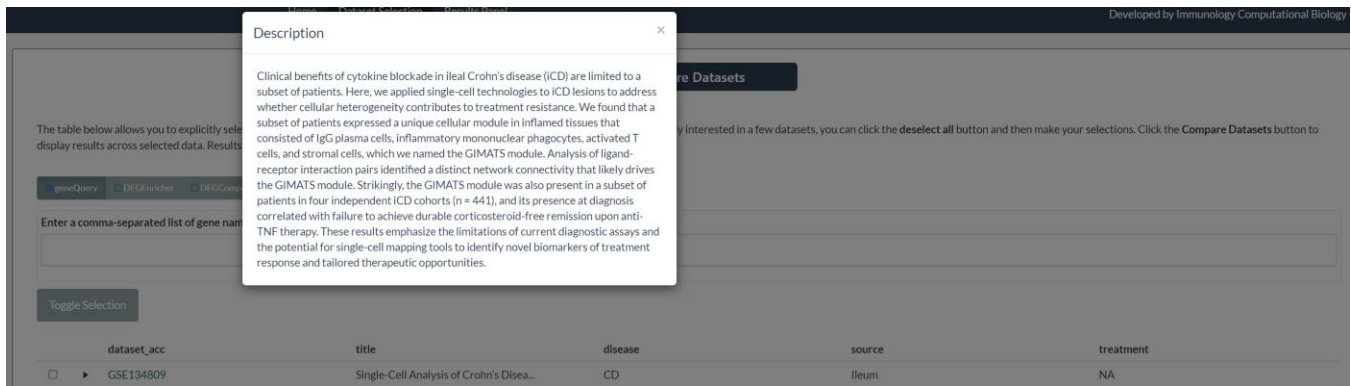


Figure 2: Data Description after clicking on dataset_acc

Now, there are three analysis type options available – **geneQuery**, **DEGEnricher**, **DEGCompare** (brown box in Figure 3).

1.1 geneQuery

Users need to enter a list of comma-separated genes, a maximum of 10 genes are allowed.

Compare Datasets

The table below allows you to explicitly select datasets that you would like to compare. By default, all datasets are selected. If you are only interested in a few datasets, you can click the **deselect all** button and then make your selections. Click the **Compare Datasets** button to display results across selected data. Results will be displayed in the Results Panel.

geneQuery DEGEnricher DEGCompare

Enter a comma-separated list of gene names (max 10 genes)

CTHRC1, COL1A1, KIF26B, POSTN, ADAM12, COL3A1, MMP13, SPARC

Toggle Selection

dataset_acc	title	disease	source	treatment
<input checked="" type="checkbox"/> GSE134809	Single-Cell Analysis of Crohn's Disea...	CD	Ileum	NA
<input checked="" type="checkbox"/> GSE147424	Single-cell transcriptome analysis of...	AD;Control	Skin	NA
<input type="checkbox"/> GSE150672	Highly Efficient, Massively-Parallel ...	PSO;Control	Skin	NA
<input checked="" type="checkbox"/> GSE162183	Single cell transcriptional zonation ...	PSO;Control	Skin	NA
<input type="checkbox"/> GSE180885	Single-cell analysis reveals innate I...	AD;Control	Skin	NA
<input checked="" type="checkbox"/> GSE198805	Single-cell transcriptome characteris...	PSO	Skin	NA
<input checked="" type="checkbox"/> GSE214695	IBD single cell data from colonic mucosa	CD;UC;Control	Colon	NA
<input type="checkbox"/> GSE220116	Single-cell transcriptomics suggest d...	HS;PSO;Control	Skin	Secukinumab
<input checked="" type="checkbox"/> PMID33290721_CTRL_CD	Single-Cell Sequencing of Developing ...	CD;Control	Intestine	NA
<input type="checkbox"/> PMID33479125_10X	Developmental cell programs are co-op...	AD;PSO;Control	Skin	NA

Figure 3: Matching Datasets based on user provided parameters

Next step is to select datasets, users can do so by clicking on the checkboxes or using Toggle Selection button (grey box in Figure 3) to select/deselect all the datasets.

Click on “Compare Datasets” button (orange box in Figure 3).

1.2 DEGENricher

DEGENricher performs Enrichment Analysis, users are required to enter their gene list comprising of at least 10 genes (brown box Figure 4). Users can specify Geneset Threshold or Pathway Threshold parameters (blue box in Figure 4) and select the datasets for enrichment analysis. Click on “Compare Datasets” button (orange box in Figure 4) to view results.

Compare Datasets

The table below allows you to explicitly select datasets that you would like to compare. By default, all datasets are selected. If you are only interested in a few datasets, you can click the **deselect all** button and then make your selections. Click the **Compare Datasets** button to display results across selected data. Results will be displayed in the Results Panel.

geneQueryDEGENricherDEGCompare

Please enter your gene list (comprising at least 10 genes) and select comparisons to perform Enrichment Analysis on the genes

SKAP2
PPP5C
CD6
RUNX3
AKAP11
TAB2
PRDM1
HDAC7
TNFRSF1A
FOSL2
FAP
PTPRC
IL12RB2
CYLD
GCKR
SIRPG
CREM

Geneset Threshold Selection

☐ p-value☒ adjusted p-value

Set a threshold for the p-value

Set a threshold for the logFC

0.05

0.25

☒ up☐ down

Enrichment Significance Level

☒ FDR☐ TOP

0.05

Toggle Selection

dataset_acc	title	disease	source	treatment
<input checked="" type="checkbox"/> GSE134809	Single-Cell Analysis of Crohn's Disea...	CD	Ileum	NA
<input checked="" type="checkbox"/> GSE147424	Single-cell transcriptome analysis of...	AD;Control	Skin	NA
<input type="checkbox"/> GSE150672	Highly Efficient, Massively-Parallel ...	PSO;Control	Skin	NA
<input checked="" type="checkbox"/> GSE162183	Single cell transcriptional zonation ...	PSO;Control	Skin	NA
<input type="checkbox"/> GSE180885	Single-cell analysis reveals innate I...	AD;Control	Skin	NA
<input checked="" type="checkbox"/> GSE198805	Single-cell transcriptome characteris...	PSO	Skin	NA

Figure 4: DEGENricher Tab in scCompare

1.3 DEGCompare

“DEGCompare” allows to compare DEGs and perform pathway enrichment analysis. Users need to specify the threshold for p-value or log fold change (green box in Figure 5) to filter the DEGs and then select datasets of interest. Click on “Compare Datasets” button (orange box in Figure 5) to view results.

Compare Datasets

The table below allows you to explicitly select datasets that you would like to compare. By default, all datasets are selected. If you are only interested in a few datasets, you can click the **deselect all** button and then make your selections. Click the **Compare Datasets** button to display results across selected data. Results will be displayed in the Results Panel.

☐ geneQuery ☐ DEGENricher ☒ DEGCompare

☐ p-value ☒ adjusted p-value

Set a threshold for the p-value:
 Set a threshold for the logFC:

up

down

Toggle Selection

	dataset_acc	title	disease	source	treatment
<input checked="" type="checkbox"/>	GSE134809	Single-Cell Analysis of Crohn's Disea...	CD	Ileum	NA
<input checked="" type="checkbox"/>	GSE147424	Single-cell transcriptome analysis of...	AD;Control	Skin	NA
<input type="checkbox"/>	GSE150672	Highly Efficient, Massively-Parallel ...	PSO;Control	Skin	NA
<input checked="" type="checkbox"/>	GSE162183	Single cell transcriptional zonation ...	PSO;Control	Skin	NA
<input type="checkbox"/>	GSE180885	Single-cell analysis reveals Innate I...	AD;Control	Skin	NA
<input checked="" type="checkbox"/>	GSE198805	Single-cell transcriptome characteris...	PSO	Skin	NA
<input checked="" type="checkbox"/>	GSE214695	IBD single cell data from colonic mucosa	CD;UC;Control	Colon	NA

Figure 5: DEGCompare tab in scCompare

2. Results

The comparison results are shown in the “Results Panel” tab.

2.1 geneQuery Result

For geneQuery, there are two sub tabs: Comparisons Table and Expression Visuals. In the “Comparison Tables (Results)” tab, “Summary of Query and Matching Datasets” panel includes query parameters and number of datasets related to each query gene based on each disease (red box in Figure 6). Results panel includes the comparison results with sub-tabs for each disease. The table includes the description of each dataset and comparison and whether the genes of interest are significant in the comparison.

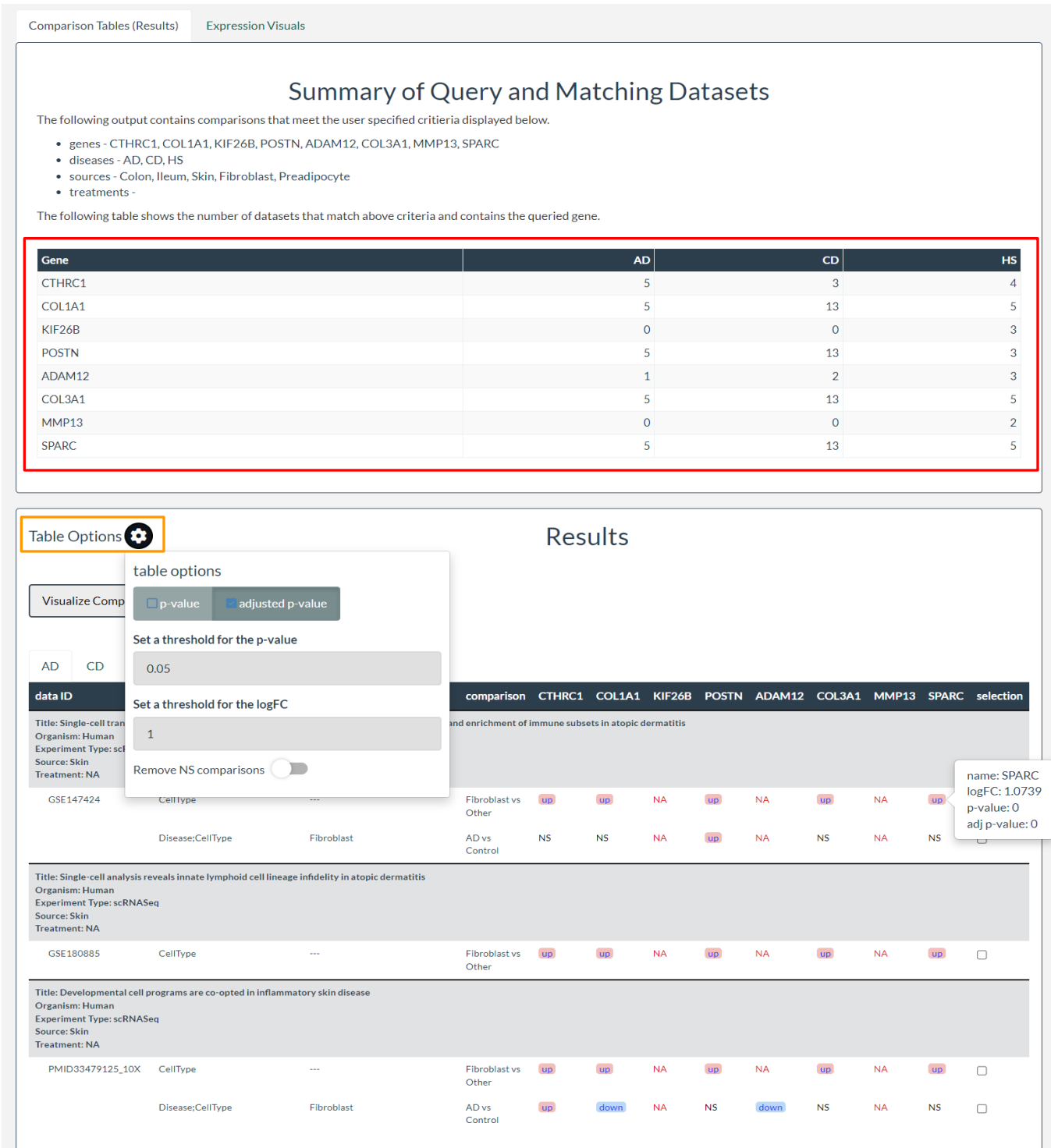
“NS” means gene of interest is not significant in the comparison while red and blue “info” button represent the gene of interest is significantly up- or down-regulated in the comparison. Hovering over the “info” button can get the detailed statistic result. The directionality of the comparison can be found in the “comparison” column. For example, “Connective Tissue vs Others” means up-regulated gene has higher expression in Connective Tissue compared with the other cells.

User can modify the threshold to change the status of each gene in each comparison by clicking the “Table Options” button (orange box in Figure 6).

Users can select one or multiple comparisons for each disease and click on “Visualize Comparison” button to plot the heatmap for each gene in each comparison in the “Expression Visuals” panel.

Please keep in mind that if a dataset has multiple diseases, some comparisons are common across the diseases. For example, in PMID33479125_10X, Connective Tissue Cell vs Other will be displayed under

both the tabs- AD and PSO. Users can select this comparison under either one (AD or PSO) or both tabs to generate the heatmap.



Under Expression Visuals Tab, users can plot the heatmap based on log (Fold change) or signed -log(p-value). Users can also modify the scale for the heatmap plot using the Select ceiling/floor option.



Figure 7: Heatmap under Expression Visuals Tab in scCompare

2.2 DEGENricher

After entering the gene list and clicking on “Compare Datasets”, user is redirected to “Enrichment Analysis” Tab under Results Panel. It displays the enrichment analysis report identifying enriched differential gene lists based on user provided gene list input.

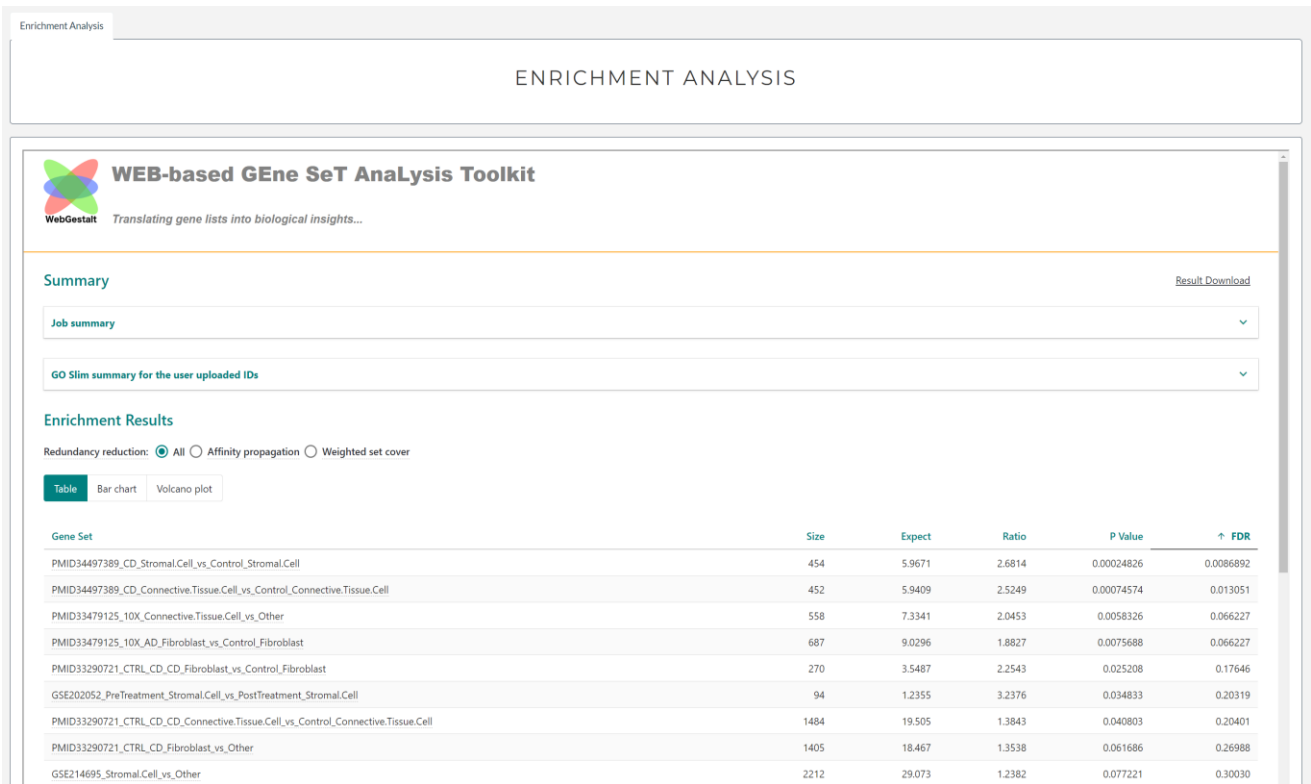


Figure 8: Enrichment results using WebgestaltR

2.3 DEGCompare

DEGCompare allows interactive comparisons of multiple differential gene lists and performs pathway enrichment analyses. Users can select the gene lists from the left-hand side menu (blue box in Figure 9), maximum 7 gene lists can be selected at once.

While comparing the genes among these selected gene lists, users can click on the checkboxes to include or exclude a particular gene list from the visualization (orange box in Figure 9). Users can also click on Enrichment Analysis button (green box in figure 9) to perform pathway enrichment analysis on the selected genes.

“Download Full Results” button (black box in figure 9) allows users to download complete list of genes based on selections made in include/exclude option.

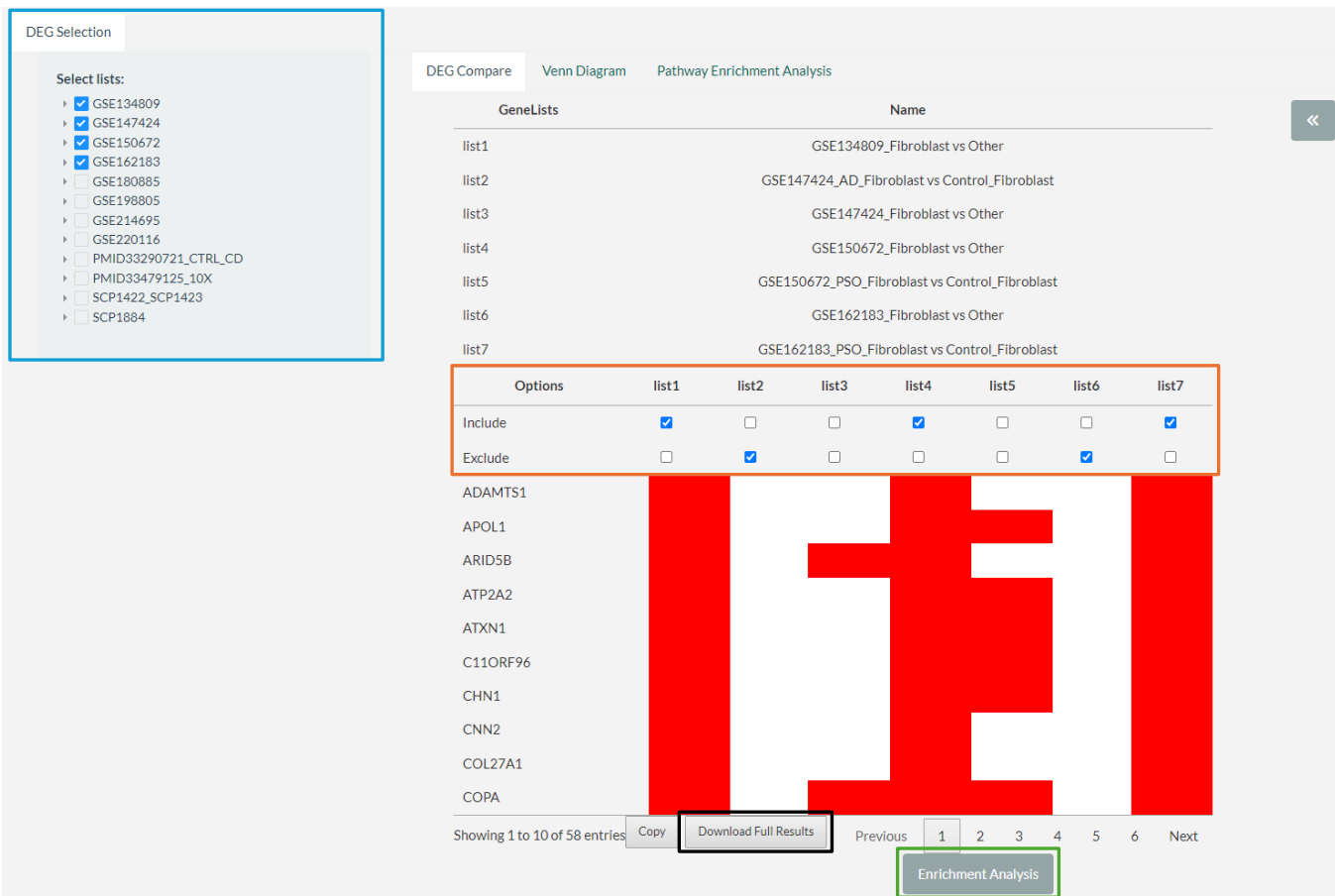


Figure 9: DEGCompare tab under Results Panel

Under “Venn Diagram” tab, by default 4 gene lists are selected. A maximum of 4 gene lists can be selected at once. It provides a Venn plot to identify the number of common genes between the lists. Users can select which four lists they would like to compare from the dropdown menu (black box in Figure 10). Users can also relabel these gene lists in the text box (purple box in Figure 10) below each gene list name.



Figure 10: Venn Diagram tab under Results Panel

After clicking on “Enrichment Analysis”, User is redirected to Pathway Enrichment Analysis tab, where enrichment report is generated using WebGestaltR. Reactome Pathway enrichment analysis is performed for the filtered genes from DEGCompare.

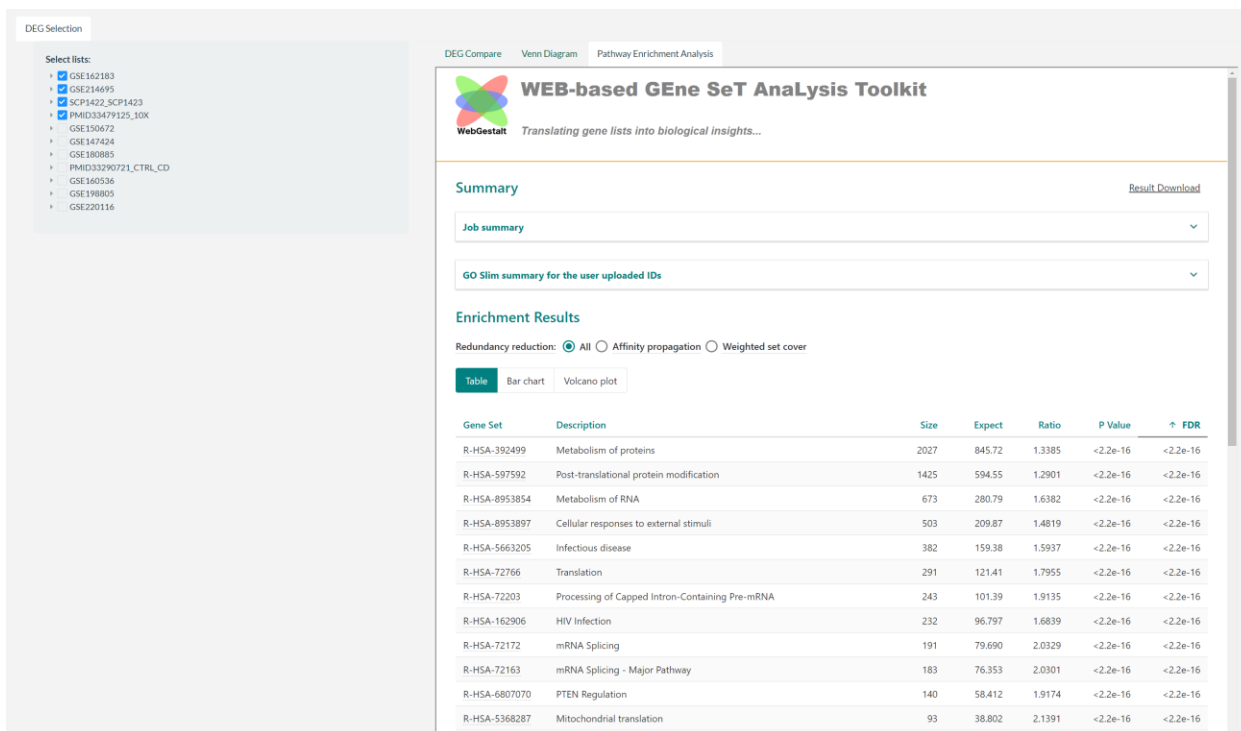


Figure 11: Pathway enrichment analysis Tab under Results Panel