



Basic Tutorial

# Table of Contents

<u>Topic</u>	<u>Page</u>
Getting Started	2
Data Input	5
Data Formats	6
Tutorial – Getting Data from GEO	7
ROGUE – Differential Expression Analysis	21
ROGUE – Create Groups	31
ROGUE – Groups Differential Expression Analysis	40
ROGUE – Single Sample Gene Expression Analysis	41
ROGUE – Groups Gene Expression Analysis	43
ROGUE – Gene Set Enrichment Analysis	44
ROGUE – Gene Ontology Analysis	46
ROGUE – Statistical Group Comparison (Biomarker Discovery)	49
ROGUE – tSNE	52
ROGUE – Gene Ontology Expression Comparison	54
ROGUE – Differentially Expressed Gene Ontologies	57

# Getting Started

The screenshot displays the ROGUE software interface, which is a RNAseq & Ontology Graphic User Environment. The left sidebar contains a navigation menu with the following items:

- View ROGUE Manual
- ?
- ROGUE
- Load Count Data:(EdgeR)
- Load Expression Data
- Create Groups
- EdgeR Group Comparison
- Gene Comparison (Samples)
- Gene Comparison (Groups)
- Advanced Analysis
- Download Report

The main area of the interface shows several analysis workflows and their corresponding visualizations:

- edgeR Analysis (DEG)**: Perform basic DEG analysis. Visualization: A scatter plot with red and blue points.
- Grouped edgeR (DEG)**: Perform basic DEG analysis with grouped samples. Visualization: A scatter plot with multiple colored groups of points.
- Gene Expression Comparison**: Basic gene expression comparisons. Visualization: A bar chart with multiple colored bars.
- Grouped Gene Expression Comparison**: Gene expression comparisons of grouped samples. Visualization: A bar chart with multiple colored bars.
- Gene Ontology**: GO analysis on a list of genes. Visualization: A bar chart with various colored bars.
- Group Statistical Analysis**: Identify consistently differentially expressed genes between groups (potential biomarkers). Visualization: A 3D t-SNE plot.
- Group Stats Summary Plots**: Summary plots for group statistics.
- Group Stats Gene List**: Gene lists for group statistics.
- t-SNE**: t-SNE analysis on Samples and Genes. Visualization: A 2D t-SNE plot.
- Group Expressed Gene Ontology**: Differential expression analysis of ontology specific genes. Visualization: A density plot.
- Differentially Expressed Gene Ontology**: Differentially regulated pathways/biological processes. Visualization: A bar chart with red and blue bars.

A large central logo for "ROGUE" is positioned between the analysis sections.

# Getting Started

**In-app help**  
Click the '?' for tips on using the app

**edgeR Analysis (DEG)**  
Perform basic DEG analysis

**Grouped edgeR (DEG)**  
Perform basic DEG analysis with grouped samples

**Gene Expression Comparison**  
Basic gene expression comparisons

**Grouped Gene Expression Comparison**  
Gene expression comparisons of grouped samples

**Gene Ontology**  
GO analysis on a list of genes

**Group Statistical Analysis**

**Group Stats Summary Plots**

**Group Stats Gene List**

**t-SNE**  
t-SNE analysis on Samples and Genes

**Group Expressed Gene Ontology**  
Differential expression analysis of ontology specific genes

**Differentially Expressed Gene Ontology**  
Differentially regulated pathways/biological processes

**ROGUE**

# Getting Started

The screenshot shows the ROGUE interface with a sidebar on the left containing navigation links like 'View ROGUE Manual', 'ROGUE', 'Load Count Data:(EdgeR)', 'Load Expression Data', 'Create Groups', 'EdgeR Group Comparison', 'Gene Comparison (Samples)', 'Gene Comparison (Groups)', 'Advanced Analysis', and 'Download Report'. A red arrow points from a 'View ROGUE Manual' link to a white cloud-shaped callout box.

**In-app help:**  
Describes the tools in the app. Click 'Next' to see more tips.

The main area displays several analysis tools:

- edgeR Analysis (DEG)**: Perform basic DEG analysis.
- Grouped edgeR (DEG)**: Perform basic DEG analysis with grouped samples.
- Gene Expression Comparison**: Basic gene expression comparisons.
- Grouped Gene Expression Comparison**: Gene expression comparisons of grouped samples.
- Gene Ontology**: GO analysis on a list of genes.
- Group Statistical Analysis**: Identify consistently differentially expressed genes between groups (potential biomarkers).
- Group Stats Summary Plots**: t-SNE
- Group Stats Gene List**: I-SNE analysis on Samples and Genes
- Group Expressed Gene Ontology**: Differential expression analysis of ontology specific genes.
- Differentially Expressed Gene Ontology**: Differentially regulated pathways/biological processes.

A red circle with the number '1' is positioned over the 'Group Statistical Analysis' tool. A callout box with the number '1' contains the text: "If your RNAseq data is in raw reads or counts, click the 'Load Count Data:(EdgeR)' in the left menu bar to convert to RPKM and perform differential gene expression analysis."

At the bottom of the callout box are buttons for 'Skip', '← Back', and 'Next →'.

# Data Input

ROGUE can use count/reads data (raw or normalized) and expression data (FPKM/RPKM) as input.

## Counts/Reads Data

The screenshot shows the ROGUE application interface. On the left, a dark sidebar lists various analysis options. The 'Load Count Data:(EdgeR)' option is highlighted with a red arrow. The main panel is titled 'Choose Reads File' and contains fields for 'Browse...' and 'No file selected'. Below this are sections for 'Select Input Source' (with 'Database' selected), 'Select Data Library' (set to 'Grace'), 'Select Dataset' (empty dropdown), 'Select Reads Status' (with 'Raw' selected), and a 'Load Data' button at the bottom.

## Expression Data

The screenshot shows the ROGUE application interface. On the left, a dark sidebar lists various analysis options. The 'Load Expression Data' option is highlighted with a red arrow. The main panel is titled 'Choose Expression File' and contains fields for 'Browse...' and 'No file selected'. Below this are sections for 'Select Input Source' (with 'Database' selected), 'Select Data Library' (set to 'GEO\_Pediatric\_Cancer'), 'Select Dataset' (empty dropdown), and a 'Load File' button at the bottom.

# Data Formats

## Raw Counts

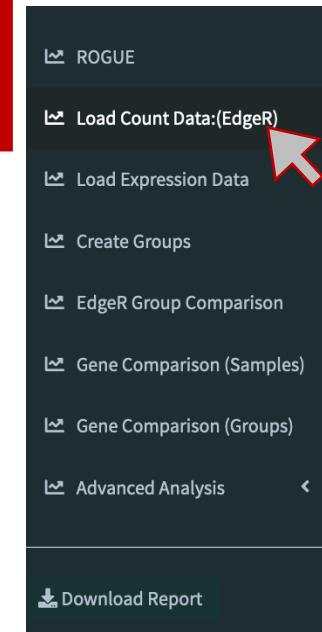
gene_name	symbol	len	Sample_1	Sample_2	Sample_3
NM_177327	Wwp1	5070	687	579	761
NM_177326	Pak2	4099	5194	4724	5962
NM_177325	Tsr1	3385	1204	1421	2334

**Note:** Tab delimited files:  
Raw counts can be submitted genenames, symbols, len ('len' column are the gene lengths in bases and needs to be labeled 'len'), followed by sample columns with raw count values

## Raw Counts

symbol	Sample_1	Sample_2	Sample_3
Wwp1	687	579	761
Pak2	5194	4724	5962
Tsr1	1204	1421	2334

**Note:** Tab delimited files: Raw counts can be submitted with the genename/symbol column followed by the Sample columns with the raw counts. The tool will use pre-determined gen lengths to normalize counts.



## Normalized Counts (by length)

Symbol	Sample_1	Sample_2	Sample_3
ENSG00000000419	687.43	579.63	761.19
ENSG00000000457	5194.64	4724.91	5962.03
ENSG00000000460	1204.01	1421.63	2334.32

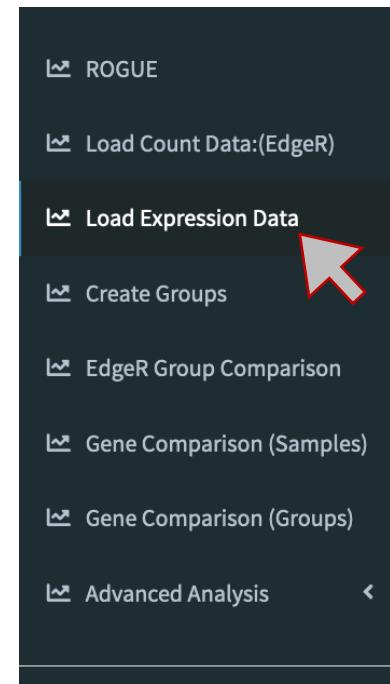
**Note:** Tab delimited files:  
Normalized counts can be submitted with the genename/symbol column followed by the Sample columns with the normalized counts.

# Data Formats

**Note:** Tab delimited files: Expression values can be submitted with the genename/symbol column followed by the Sample columns with the expression values (eg. FPKM/RPKM).

## Expression Values

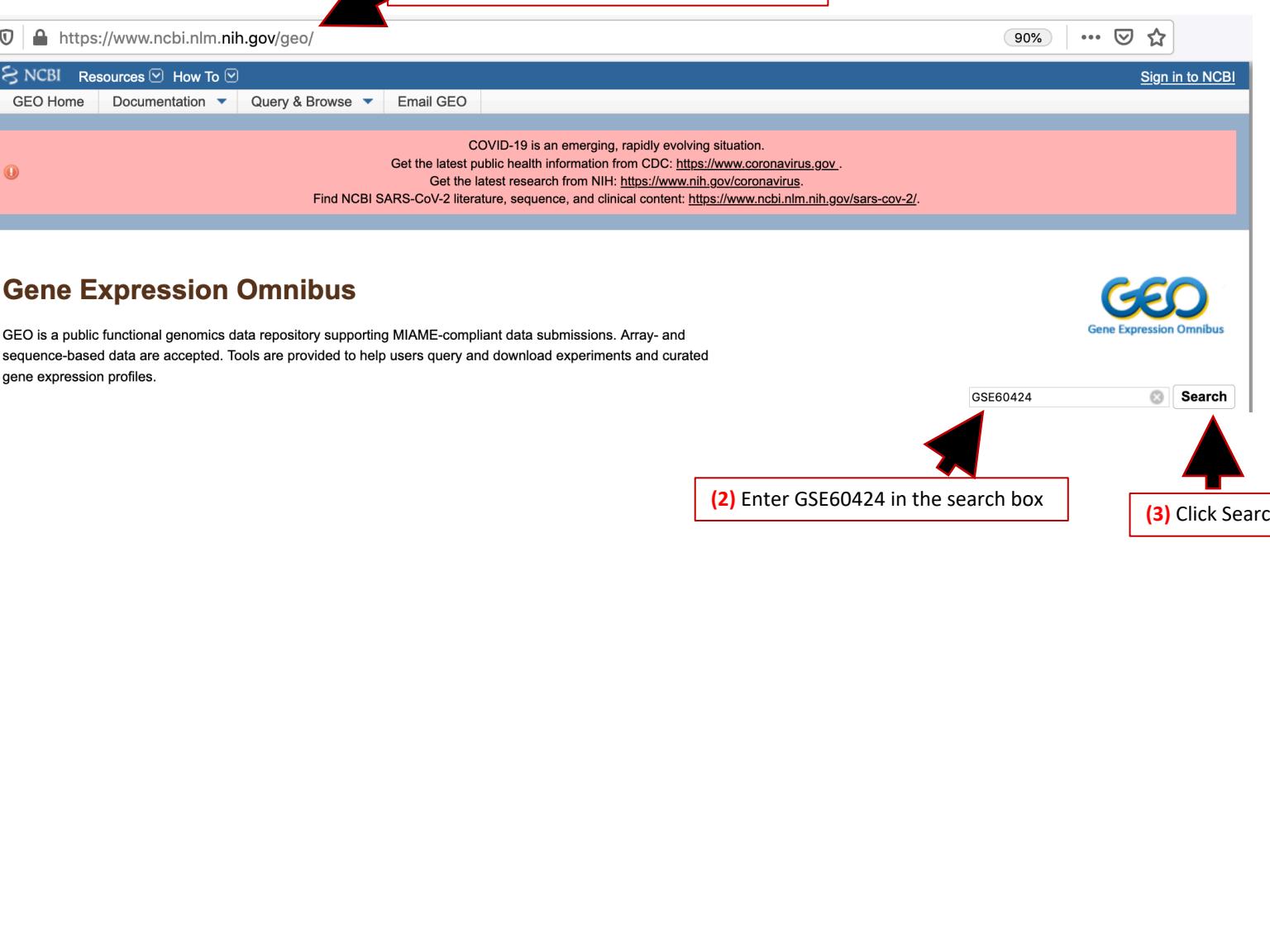
genenames	Sample_1	Sample_2	Sample_3
TSPAN6	17.84	16.37	18.45
TNMD	0	0.18	0
DPM1	23.05	21.24	15.8



The screenshot shows a 'Choose Expression File' dialog box. It includes fields for 'Choose Expression File' (with 'Browse...' and 'No file selected' buttons), 'Select Input Source' (with 'Database' and 'Upload File' radio buttons, where 'Database' is selected), 'Select Data Library' (with a dropdown menu showing 'GEO\_Pediatric\_Cancer'), 'Select Dataset' (with a dropdown menu), and a 'Load File' button at the bottom.

# Tutorial-Getting Data from GEO

(1) Go to <https://www.ncbi.nlm.nih.gov/geo/>



(2) Enter GSE60424 in the search box

(3) Click Search

(4) You will be redirected to this page



<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE60424>

https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE60424

NCBI Gene Expression Omnibus

COVID-19 is an emerging, rapidly evolving situation. Get the latest public health information from CDC: <https://www.coronavirus.gov>. Get the latest research from NIH: <https://www.nih.gov/coronavirus>. Find NCBI SARS-CoV-2 literature, sequence, and clinical content: <https://www.ncbi.nlm.nih.gov/sars-cov-2/>.

GEO Home Documentation Query & Browse Email GEO Sign in to NCBI

GEO Gene Expression Omnibus

GEO accession: GSE60424

Scope: Self Format: HTML Amount: Quick GEO accession: GSE60424

**Series GSE60424**

Status	Public on Jan 06, 2015
Title	Next generation sequencing of human immune cell subsets across diseases
Organism	<i>Homo sapiens</i>
Experiment type	Expression profiling by high throughput sequencing
Summary	This study compared whole transcriptome signatures of 6 immune cell subsets and whole blood from patients with an array of immune-associated diseases. Fresh blood samples were collected from healthy subjects and subjects diagnosed type 1 diabetes, amyotrophic lateral sclerosis, and sepsis, as well as multiple sclerosis patients before and 24 hours after the first treatment with IFN-beta. At the time of blood draw, an aliquot of whole blood was collected into a Tempus tube (Invitrogen), while the remainder of the primary fresh blood sample was processed to highly pure populations of neutrophils, monocytes, B cells, CD4 T cells, CD8 T cells, and natural killer cells. RNA was extracted from each of these cell subsets, as well as the whole blood samples, and processed into RNA sequencing (RNAseq) libraries (Illumina TruSeq). Sequencing libraries were analyzed on an Illumina HiScan, with a target read depth of ~20M reads. Reads were demultiplexed, mapped to human gene models (ENSEMBL), and tabulated using HTSeq. Read count data were normalized by the TMM procedure (edgeR package).
Overall design	We performed whole genome RNAseq profiling of immune cell subsets and whole blood from subjects with an array of immune-associated diseases.
Contributor(s)	Speake C, Linsley PS, Whalen E, Chaussabel D, Presnell SR, Mason MJ, Gersuk VH, O'Brien KK, Nguyen Q, Greenbaum CJ, Buckner JH, Malhotra U
Citation(s)	Linsley PS, Speake C, Whalen E, Chaussabel D. Copy number loss of the interferon gene cluster in melanomas is linked to reduced T cell infiltrate and poor patient prognosis. <i>PLoS One</i> 2014;9(10):e109760. PMID: 25314013
Submission date	Aug 14, 2014
Last update date	May 15, 2019
Contact name	Scott Presnell
E-mail(s)	<a href="mailto:SPresnell@benaroya-research.org">SPresnell@benaroya-research.org</a>
Organization name	Benaroya Research Institute
Department	Systems Immunology
Street address	1201 Ninth Ave.
City	Seattle
State/province	WA
ZIP/Postal code	98101
Country	USA

**Status** Public on Jan 06, 2015

**Title** Next generation sequencing of human immune cell subsets across diseases

**Organism** *Homo sapiens*

**Experiment type** Expression profiling by high throughput sequencing

**Summary** This study compared whole transcriptome signatures of 6 immune cell subsets and whole blood from patients with an array of immune-associated diseases. Fresh blood samples were collected from healthy subjects and subjects diagnosed type 1 diabetes, amyotrophic lateral sclerosis, and sepsis, as well as multiple sclerosis patients before and 24 hours after the first treatment with IFN-beta. At the time of blood draw, an aliquot of whole blood was collected into a Tempus tube (Invitrogen), while the remainder of the primary fresh blood sample was processed to highly pure populations of neutrophils, monocytes, B cells, CD4 T cells, CD8 T cells, and natural killer cells. RNA was extracted from each of these cell subsets, as well as the whole blood samples, and processed into RNA sequencing (RNAseq) libraries (Illumina TruSeq). Sequencing libraries were analyzed on an Illumina HiScan, with a target read depth of ~20M reads. Reads were demultiplexed, mapped to human gene models (ENSEMBL), and tabulated using HTSeq. Read count data were normalized by the TMM procedure (edgeR package).

**Overall design** We performed whole genome RNAseq profiling of immune cell subsets and whole blood from subjects with an array of immune-associated diseases.

**Contributor(s)** Speake C, Linsley PS, Whalen E, Chaussabel D, Presnell SR, Mason MJ, Gersuk VH, O'Brien KK, Nguyen Q, Greenbaum CJ, Buckner JH, Malhotra U

**Citation(s)** Linsley PS, Speake C, Whalen E, Chaussabel D. Copy number loss of the interferon gene cluster in melanomas is linked to reduced T cell infiltrate and poor patient prognosis. *PLoS One* 2014;9(10):e109760. PMID: [25314013](#)

**Submission date** Aug 14, 2014

**Last update date** May 15, 2019

**Contact name** Scott Presnell

**E-mail(s)** [SPresnell@benaroyaresearch.org](mailto:SPresnell@benaroyaresearch.org)

**Organization name** Benaroya Research Institute

**Department** Systems Immunology

**Street address** 1201 Ninth Ave.

**City** Seattle

**State/province** WA

**ZIP/Postal code** 98101

**Country** USA

**Platforms (1)** [GPL15456](#) Illumina HiScanSQ (*Homo sapiens*)

**Samples (134)** [GSM1479433](#) lib221  
More... [GSM1479434](#) lib222  
[GSM1479435](#) lib223

**Relations**

BioProject [PRJNA258216](#)  
SRA [SRP045500](#)

**Download family**

Supplementary file	Size	Download	File type/resource
GSE60424_GEOSubmit_FC1to11_normalized_counts.txt.gz	2.3 Mb	(ftp)(http)	TXT

# Tutorial-Getting Data from GEO

<https://ftp.ncbi.nlm.nih.gov/geo/series/GSE60nnn/GSE60424/matrix/>

## Index of /geo/series/GSE60nnn/GSE60424/matrix

Name	Last modified	Size
<a href="#">Parent Directory</a> <a href="#">GSE60424_series_matrix.txt.gz</a>	2020-08-13 16:12	8.5K

**(2) Download 'GSE60424\_series\_matrix.txt.gz'**



**(1) Click 'Series Matrix File(s)'**



**(3) Click (ftp) to download the normalized counts file.**



# Tutorial-Getting Data from GEO

**Note:**

GSE60424\_GEOSubmit\_FC1to11\_normalized\_counts.txt is a normalized counts matrix. The file should look like this:

GSE60424\_GEOSubmit\_FC1to11\_normalized\_counts.txt

genenames	lib221	lib222	lib223	lib224	.....	lib355
ENSG00000000003	1	0	1	1		
ENSG00000000005	0	0	0	0		
ENSG00000000419	23	22	25	28		
ENSG00000000457	11	11	14	17		
ENSG00000000460	3	3	3	3		
ENSG00000000938	1141	872	1068	629		
ENSG00000000971	3	1	1	3		
ENSG00000001036	17	21	27	17		
ENSG00000001084	19	14	19	19		
ENSG00000001167	35	33	33	44		
...	...	...	...	...		
...	...	...	...	...		
...	...	...	...	...		
...	...	...	...	...		

# Tutorial-Getting Data from GEO

## GSE60424 series\_matrix.txt

```
!Series_title "Next generation sequencing of human immune cell subsets across diseases"
!Series_geo_accession "GSE60424"
!Series_status "Public on Jan 06 2015"
!Series_submission_date "Aug 14 2014"
!Series_last_update_date "May 15 2019"
!Series_pubmed_id "25314013"
!Series_summary "This study compared whole transcriptome signatures of 6 immune cell subsets and whole blood from patients with an array of immune-associated diseases. Fresh blood samples were collected from healthy subjects and subjects diagnosed type 1 diabetes, amyotrophic lateral sclerosis, and sepsis, as well as multiple sclerosis patients before and 24 hours after the first treatment with IFN-beta. At the time of blood draw, an aliquot of whole blood was collected into a Tempus tube (Invitrogen), while the remainder of the primary fresh blood sample was processed to highly pure populations of neutrophils, monocytes, B cells, CD4 T cells, CD8 T cells, and natural killer cells. RNA was extracted from each of these cell subsets, as well as the whole blood samples, and processed into RNA sequencing (RNAseq) libraries (Illumina TruSeq). Sequencing libraries were analyzed on an Illumina HiScan, with a target read depth of ~20M reads. Reads were demultiplexed, mapped to human gene models (ENSEMBL), and tabulated using HTSeq. Read count data were normalized by the TMM procedure (edgeR package)."
!Series_overall_design "We performed whole genome RNAseq profiling of immune cell subsets and whole blood from subjects with an array of immune-associated diseases."
!Series_type "Expression profiling by high throughput sequencing"
!Series_contributor "Cate,,Speake"
!Series_contributor "Peter,S,Linsley"
!Series_contributor "Elizabeth,,Whalen"
!Series_contributor "Damien,,Chaussabel"
!Series_contributor "Scott,R,Presnell"
!Series_contributor "Michael,J,Mason"
!Series_contributor "Vivian,H,Gersuk"
!Series_contributor "Kimm,K,O'Brien"
!Series_contributor "Quynh-Anh,,Nguyen"
!Series_contributor "Carla,J,Greenbaum"
!Series_contributor "Jane,H,Buckner"
!Series_contributor "Uma,,Malhotra"
!Series_sample_id "GSM1479433 GSM1479434 GSM1479435 GSM1479436 GSM1479437 GSM1479438 GSM1479439 GSM1479440 GSM1479441 GSM1479442 GSM1479443 GSM1479444 GSM1479445 GSM1479446 GSM1479447 GSM1479448
GSM1479449 GSM1479450 GSM1479451 GSM1479452 GSM1479453 GSM1479454 GSM1479455 GSM1479456 GSM1479457 GSM1479458 GSM1479459 GSM1479460 GSM1479461 GSM1479462 GSM1479463 GSM1479464 GSM1479465 GSM1479466 GSM1479467
GSM1479468 GSM1479469 GSM1479470 GSM1479471 GSM1479472 GSM1479473 GSM1479474 GSM1479475 GSM1479476 GSM1479477 GSM1479478 GSM1479479 GSM1479480 GSM1479481 GSM1479482 GSM1479483 GSM1479484 GSM1479485 GSM1479486
GSM1479487 GSM1479488 GSM1479489 GSM1479490 GSM1479491 GSM1479492 GSM1479493 GSM1479494 GSM1479495 GSM1479496 GSM1479497 GSM1479498 GSM1479499 GSM1479500 GSM1479501 GSM1479502 GSM1479503 GSM1479504 GSM1479505
GSM1479506 GSM1479507 GSM1479508 GSM1479509 GSM1479510 GSM1479511 GSM1479512 GSM1479513 GSM1479514 GSM1479515 GSM1479516 GSM1479517 GSM1479518 GSM1479519 GSM1479520 GSM1479521 GSM1479522 GSM1479523 GSM1479524
GSM1479525 GSM1479526 GSM1479527 GSM1479528 GSM1479529 GSM1479530 GSM1479531 GSM1479532 GSM1479533 GSM1479534 GSM1479535 GSM1479536 GSM1479537 GSM1479538 GSM1479539 GSM1479540 GSM1479541 GSM1479542 GSM1479543
GSM1479544 GSM1479545 GSM1479546 GSM1479547 GSM1479548 GSM1479549 GSM1479550 GSM1479551 GSM1479552 GSM1479553 GSM1479554 GSM1479555 GSM1479556 GSM1479557 GSM1479558 GSM1479559 GSM1479560 GSM1479561 GSM1479562
GSM1479563 GSM1479564 GSM1479565 GSM1479566 "
!Series_contact_name "Scott,,Presnell"
!Series_contact_email "SPresnell@benaroyaresearch.org"
!Series_contact_department "Systems Immunology"
!Series_contact_institute "Benaroya Research Institute"
!Series_contact_address "1201 Ninth Ave."
!Series_contact_city "Seattle"
!Series_contact_state "WA"
!Series_contact_zip/postal_code "98101"
!Series_contact_country "USA"
!Series_supplementary_file "ftp://ftp.ncbi.nlm.nih.gov/geo/series/GSE60nnn/GSE60424/suppl/GSE60424_GEOSubmit_FC1to11_normalized_counts.txt.gz"
!Series_platform_id "GPL15456"
!Series_platform_taxid "9606"
!Series_sample_taxid "9606"
!Series_relation "BioProject: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA258216"
!Series_relation "SRA: https://www.ncbi.nlm.nih.gov/sra?term=SRP045500"

!
Sample_title "lib221" "lib222" "lib223" "lib224" "lib225" "lib226" "lib227" "lib228" "lib229" "lib230" "lib231" "lib232"
"lib233" "lib234" "lib235" "lib236" "lib237" "lib238" "lib239" "lib240" "lib241" "lib242" "lib243" "lib244" "lib245"
"lib246" "lib247" "lib248" "lib249" "lib250" "lib251" "lib252" "lib253" "lib254" "lib255" "lib256" "lib257" "lib258"
"lib259" "lib260" "lib261" "lib262" "lib263" "lib264" "lib265" "lib266" "lib267" "lib268" "lib269" "lib270" "lib271"
"lib272" "lib273" "lib274" "lib275" "lib276" "lib277" "lib278" "lib279" "lib280" "lib281" "lib282" "lib283" "lib284"
"lib285" "lib286" "lib288" "lib289" "lib290" "lib291" "lib292" "lib293" "lib294" "lib295" "lib296" "lib297" "lib298"
"lib299" "lib300" "lib301" "lib302" "lib303" "lib304" "lib305" "lib306" "lib307" "lib308" "lib309" "lib310" "lib311"
"lib312" "lib313" "lib314" "lib315" "lib316" "lib317" "lib318" "lib319" "lib320" "lib321" "lib322" "lib323" "lib324"
"lib325" "lib326" "lib327" "lib328" "lib329" "lib330" "lib331" "lib332" "lib333" "lib334" "lib335" "lib336" "lib337"
"lib338" "lib339" "lib340" "lib341" "lib342" "lib343" "lib344" "lib345" "lib346" "lib347" "lib348" "lib349" "lib350"
"lib351" "lib352" "lib353" "lib354" "lib355"
```

Note: This is a standard series matrix file that is uploaded with data to GEO. It describes the data using standardized fields defining filenames, sample labels, and sample characteristics, and the related publication.

# Tutorial-Getting Data from GEO

Go to 'ROGUE\_Companion' to prepare files.

[reslnmaris01.research.chop.edu/ROGUE\\_Companion](https://reslnmaris01.research.chop.edu/ROGUE_Companion)

**Note:** The GEO ROGUE Companion tool was created to help prepare a data matrix and group file from data downloaded from GEO for analysis using ROGUE. Both a data matrix and Series matrix file are required.

Follow this tutorial as an example.

# Tutorial-Getting Data from GEO

The screenshot shows a web browser window for the GEO ROGUE Companion tool. The URL is https://reslnmaris01.research.chop.edu/ROGUE.Companion/. The page title is "GEO ROGUE Companion". On the left, there's a section for selecting whether data is a single matrix or multiple files, with "Single Matrix" selected. Below it is a "Upload GSE Series Matrix File" section with a "Browse..." button and a message "No file selected". A red box labeled "(1) Upload 'Series Matrix File'" surrounds this area. To the right is a file selection dialog titled "Select Series Matrix field to Match data columns/files". It shows a list of files in the "Downloads" folder:

- GSE60424\_GEOSubmit\_FC1to11\_normalized\_counts.txt
- GSE60424\_GEOSubmit\_FC1to11\_normalized\_counts.txt.gz
- GSE60424\_series\_matrix.txt** (highlighted with a blue selection bar)
- GSE60424\_series\_matrix.txt.gz

A red box labeled "(2) Select GSE60424\_series\_matrix.txt" surrounds the selected file. At the bottom of the dialog are "Cancel" and "Open" buttons, with a red box labeled "(3) Click 'Open'" surrounding the "Open" button.

**Note:** Upload the files downloaded from GSE60424.

# Tutorial-Getting Data from GEO

The screenshot shows the 'GEO ROGUE Companion' application interface. On the left, there's a sidebar for selecting data type and uploading files. The main area has sections for 'Select Series Matrix field to Match data columns/files' (set to 'Sample\_title'), 'Details' (listing sample names like lib221, lib222, lib223, lib224), 'Data Matrix Sample Names/' (listing files: GSE60424\_GEOSubmit\_FC1to11\_normalized\_counts.txt, GSE60424\_GEOSubmit\_FC1to11\_normalized\_counts.txt.gz, GSE60424\_series\_matrix.txt, GSE60424\_series\_matrix.txt.gz), and 'Select Common Columns'.

**(1) Upload 'Counts Matrix File'**: A red box highlights the 'Browse...' button for the 'Upload Data matrix' section. A black arrow points to this button.

**(2) Select GSE60424\_GEOSubmit\_FC1to11\_normalized\_counts.txt**: A red box highlights the file 'GSE60424\_GEOSubmit\_FC1to11\_normalized\_counts.txt' in the 'Details' list. A black arrow points to this file.

**(3) Click 'Open'**: A red box highlights the 'Open' button at the bottom right of the file selection dialog. A black arrow points to this button.

**Note:** Upload the files downloaded from GSE60424.  
When the series\_matrix file is uploaded the field headers will populate the 'Series Matrix field' dropdown box and the details/members of the selected field will fill the 'Details' text box.

# Tutorial-Getting Data from GEO

## GEO ROGUE Companion

Select whether data is a single matrix or multiple files

Single Matrix  
 Multiple Files

Upload GSE Series Matrix File

Browse... GSE60424\_series\_matrix.txt  
Upload complete

Upload Data matrix

Browse... GSE60424\_GEOSubmit\_FC1to11\_normalized\_.txt  
Upload complete

Select Series Matrix field to Match data columns/files

Sample\_title

Sample\_title  
Series\_title  
Series\_pubmed\_id  
Series\_sample\_id  
Series\_platform\_id  
Series\_platform\_taxid  
Series\_sample\_taxid  
Sororita\_covid\_oh1  
lib221  
lib222  
lib223  
... -- .

(1) This Drop down menu is filled with ID and title Fields from the Series Matrix file. When an ID or Title field is selected the 'Details' text box is populated with the members of that field as one line per member.

## GEO ROGUE Companion

Select whether data is a single matrix or multiple files

Single Matrix  
 Multiple Files

Upload GSE Series Matrix File

Browse... GSE60424\_series\_matrix.txt  
Upload complete

Upload Data matrix

Browse... GSE60424\_GEOSubmit\_FC1to11\_normalized\_.txt  
Upload complete

Select Series Matrix field to Match data columns/files

Sample\_title

Details

lib221  
lib222  
lib223  
lib224  
... -- .

Data Matrix Sample Names/headers

genenames  
lib221  
lib222  
lib223  
... -- .

(2) When an ID or Title field is selected the 'Details' text box is populated with the members of that field as one line per member.

It is important that the user chooses the ID or Title field that has details with the same names as the column headers listed in the 'Data Matrix Samples Names/headers' text box.

(3) When a Data matrix is loaded, the column headers fill the 'Data Matrix Samples Names/headers' text box.

# Tutorial-Getting Data from GEO

## Create Groups

### Group Name

Type Group Name here

(1) Enter a group name.



### Select Features

age

32

(3) Choose the feature details of the members of the group.



### Add Group Feature

### Group Features

Selected group features will appear here

(2) Choose a feature field that is extracted from the “!characteristics” fields from the series matrix file.



### Create\_Group

### Groups

Groups appear here

(4) The added features will appear in this box. Multiple features can be added.  
When a group is created, this tool will include the samples that have all the selected characteristics as members of the group.



(5) When ‘Create Group’ Button is clicked and a group is successfully created, it will appear in the format:  
**‘Group\_Name:Member1;Member2;Member3’**  
This is the same group format required by the main ROGUE tool.

# Tutorial-Getting Data from GEO

**Note:** Follow these steps to create groups that will be used in the ROGUE tutorial.

## Create Groups

Group Name

(1) Type 'Healthy\_CD4' as the Group Name

Select Features

celtype

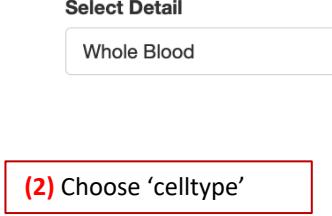
- age
- cellcount
- celtype
- collectiondate
- diseasestatus
- donorid
- gender
- index
- Create\_Group

Select Detail

Whole Blood

## Create Groups

Group Name

(2) Choose 'celtype'

Select Features

celtype

Add Group Feature

Group Features

Selected group features will appear here

Create\_Group

Select Detail

Whole Blood

- Whole Blood
- Neutrophils
- Monocytes
- B-cells
- CD4
- CD8
- NK

## Create Groups

Group Name

(3) Choose 'CD4'

Select Features

celtype

Select Detail

CD4

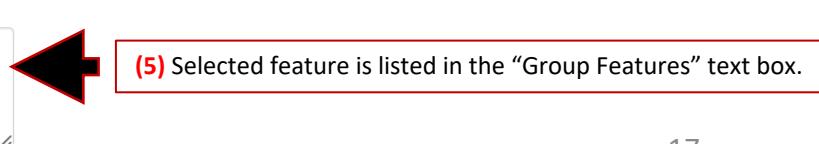
Add Group Feature

Group Features

Selected group features will appear here

## Create Groups

Group Name

(4) Click 'Add Group Feature'

Select Features

celtype

Select Detail

CD4

Add Group Feature

Group Features

celtype: CD4

# Tutorial-Getting Data from GEO

Create Groups

Group Name

Select Features

diseasestatus

- age
- cellcount
- celltype
- collectiondate
- diseasestatus
- donorid
- gender
- index\_Group

Select Detail

CD4

(6) Choose 'diseasestatus'

Create Groups

Group Name

Select Features

diseasestatus

Add Group Feature

Group Features

celltype: CD4

Create\_Group

(7) Choose 'Healthy Control'

Create Groups

Group Name

Select Features

diseasestatus

Select Detail

Healthy Control

Add Group Feature

Group Features

celltype: CD4

(8) Click 'Add Group Feature'

Create Groups

Group Name

Select Features

diseasestatus

Select Detail

Healthy Control

Add Group Feature

Group Features

celltype: CD4  
diseasestatus: Healthy Control

(9) Selected feature is added to the "Group Features" text box.

# Tutorial-Getting Data from GEO

**Create Groups**

**Group Name**  
Healthy\_CD4

**Select Features**  
diseasestatus

**Select Detail**  
Healthy Control

**Add Group Feature**

**Group Features**

celltype: CD4  
diseasestatus: Healthy Control

**Create\_Group**

**Groups**  
Groups appear here

**(10) Click 'Create\_Group'**

**Create Groups**

**Group Name**  
Healthy\_CD4

**Select Features**  
diseasestatus

**Select Detail**  
Healthy Control

**Add Group Feature**

**Group Features**

Selected group features will appear here

**Create\_Group**

**Groups**  
Healthy\_CD4:lib229;lib291;lib298;lib312

**(11) Groups are created in format  
'Group\_Name:Member1;Member2;Member3'**

# Tutorial-Getting Data from GEO

## Create Groups

Group Name  
Healthy\_NK

Select Features  
diseasestatus

Select Detail  
Healthy Control

Add Group Feature

### Group Features

Selected group features will appear here

Create\_Group

### Groups

Healthy\_CD4:lib229;lib291;lib298;lib312  
Healthy\_CD8:lib230;lib292;lib299;lib313  
Healthy\_NK:lib231;lib293;lib300;lib314



(12) Create Healthy CD8 Groups and Healthy NK Groups (This Should be the result).

Select whether to create table with original gene ID or human gene symbol

- Original IDs
- Human Gene Symbols

[Download Data Table](#)

[Download Groups](#)

Select whether to create table with original gene ID or human gene symbol

Original IDs

Human Gene Symbols

(13) Select 'Human Gene Symbols' if you want to download a table with gene symbols instead of Ensembl IDs

[Download Data Table](#)

[Download Groups](#)

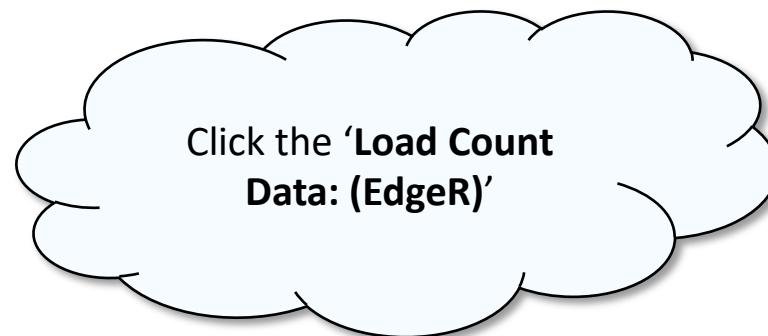
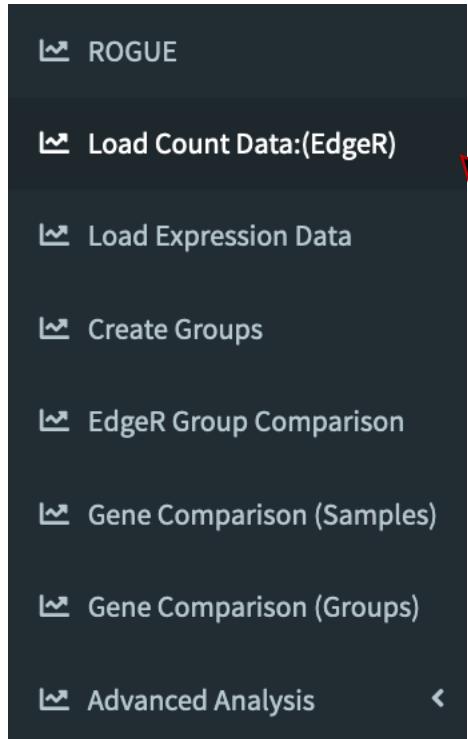


(14) Download Data Tables with Gene Symbols and Group Files

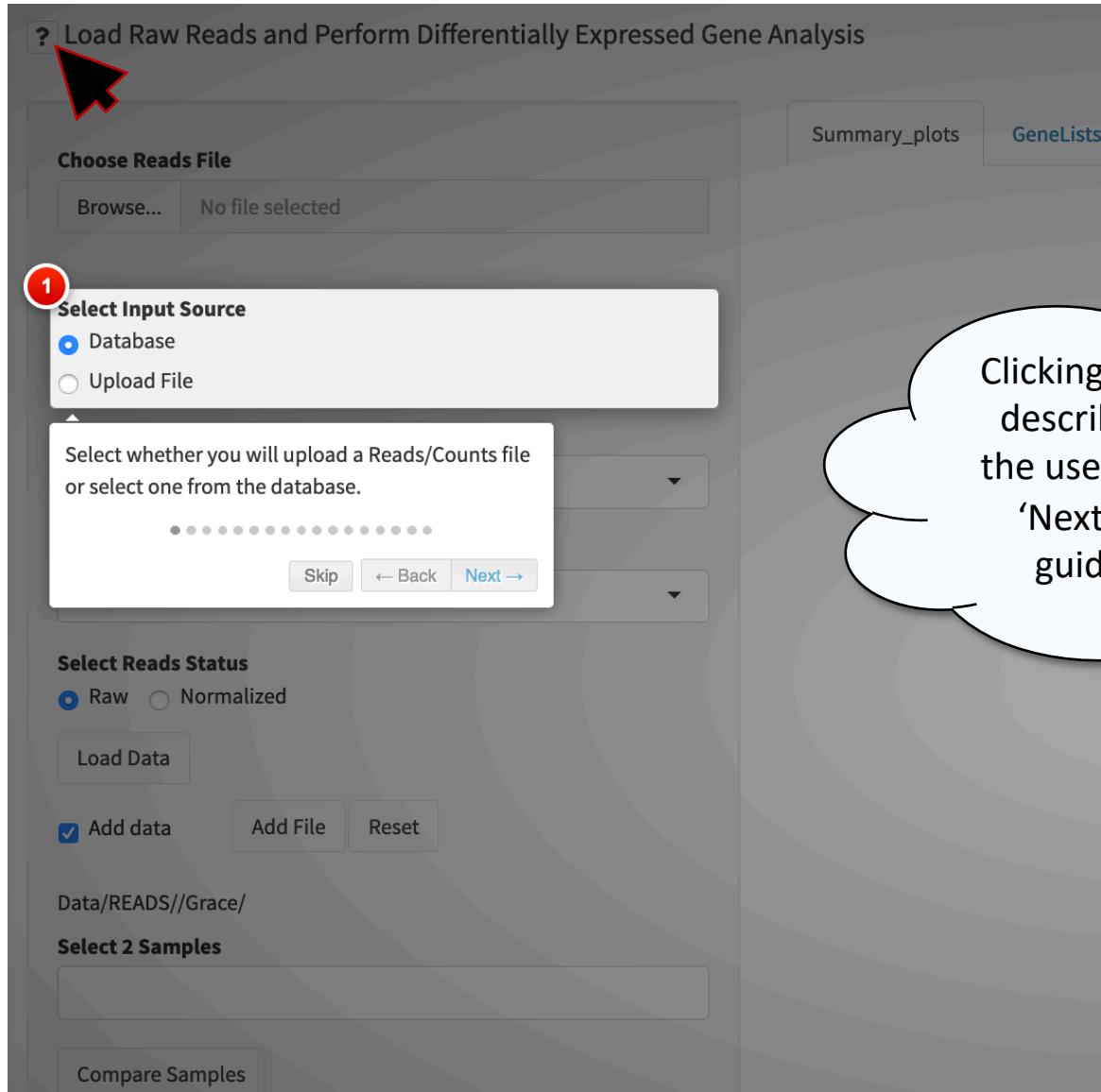
Note: Download the Data Table as 'Data\_Matrix.txt' and the groups file as 'Groups\_File.txt'

# ROGUE – Differential Expression Analysis

<https://reslnmaris01.research.chop.edu/ROGUE/>



# ROGUE – Differential Expression Analysis



Clicking the in-app help button will describe each options and guide the user through the process. Click 'Next' to see more tips and be guided through the options.

# ROGUE – Differential Expression Analysis

? Load Raw Reads and Perform Differentially Expressed Gene Analysis

**Choose Reads File**

No file selected

**Select Input Source**

Database  
 Upload File

**Select Data Library**

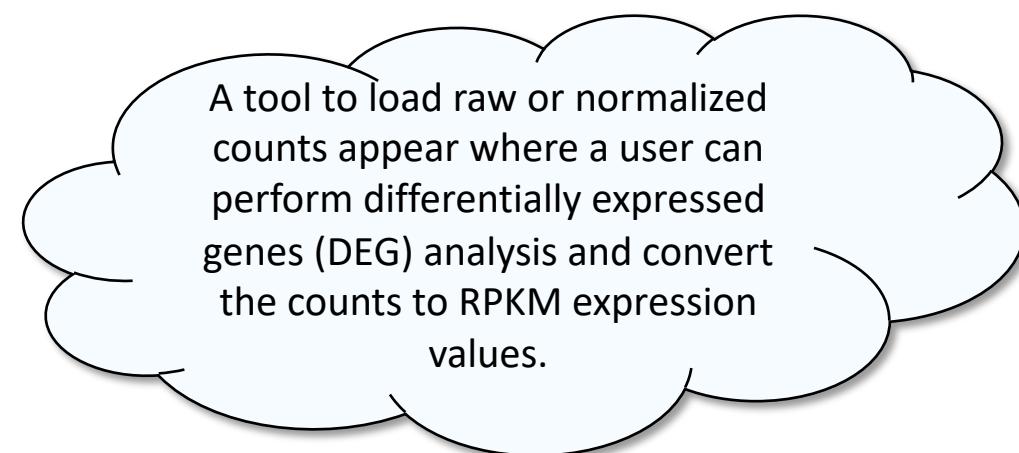
**Select Dataset**

**Select Reads Status**

Raw  Normalized

Add data

**Select 2 Samples**



A tool to load raw or normalized counts appear where a user can perform differentially expressed genes (DEG) analysis and convert the counts to RPKM expression values.

# ROGUE – Differential Expression Analysis

**Choose Reads File**

No file selected



**Select Input Source**

Database  
 Upload File



**Select Data Library**

GSE60424

**Select Dataset**

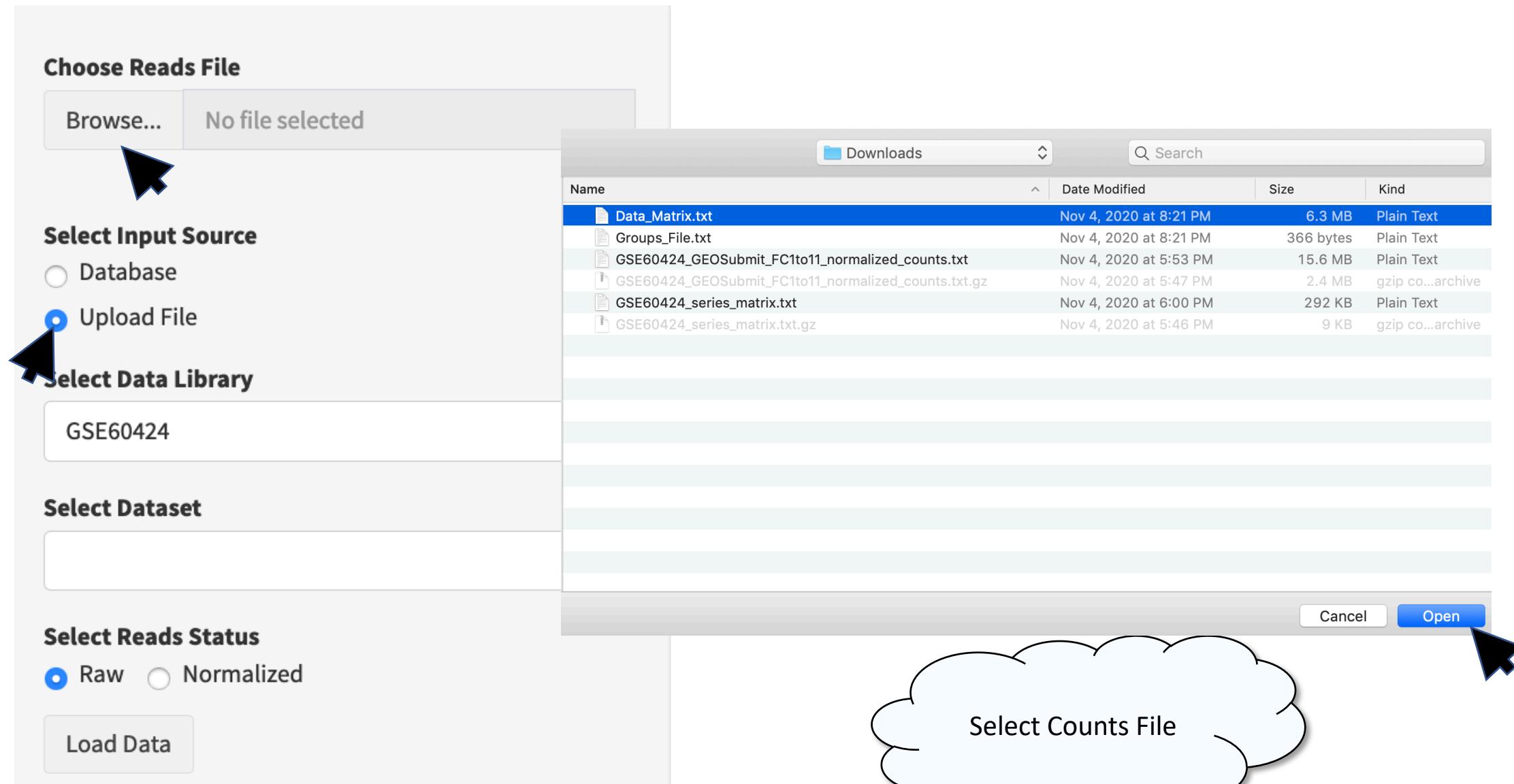
**Select Reads Status**

Raw  Normalized

**Load Data**

Check the upload radio button  
and click 'Browse' to  
select a Counts file

# ROGUE – Differential Expression Analysis



# ROGUE – Differential Expression Analysis

**Choose Reads File**

Data\_Matrix.txt

Upload complete

**Select Input Source**

Database  
 Upload File

**Select Data Library**

GSE60424

**Select Dataset**

▼

**Select Reads Status**

Raw  Normalized

This file has normalized counts (by gene length). Select the '**Normalized**' radio button.

Click the '**Load Data**' Button

Processing Reads Progress: 80%

# ROGUE – Differential Expression Analysis

? Load Raw Reads and Perform Differentially Expressed Gene Analysis

**Choose Reads File**

Data\_Matrix.txt  
**Upload complete**

**Select Input Source**

Database  
 Upload File

**Select Data Library**

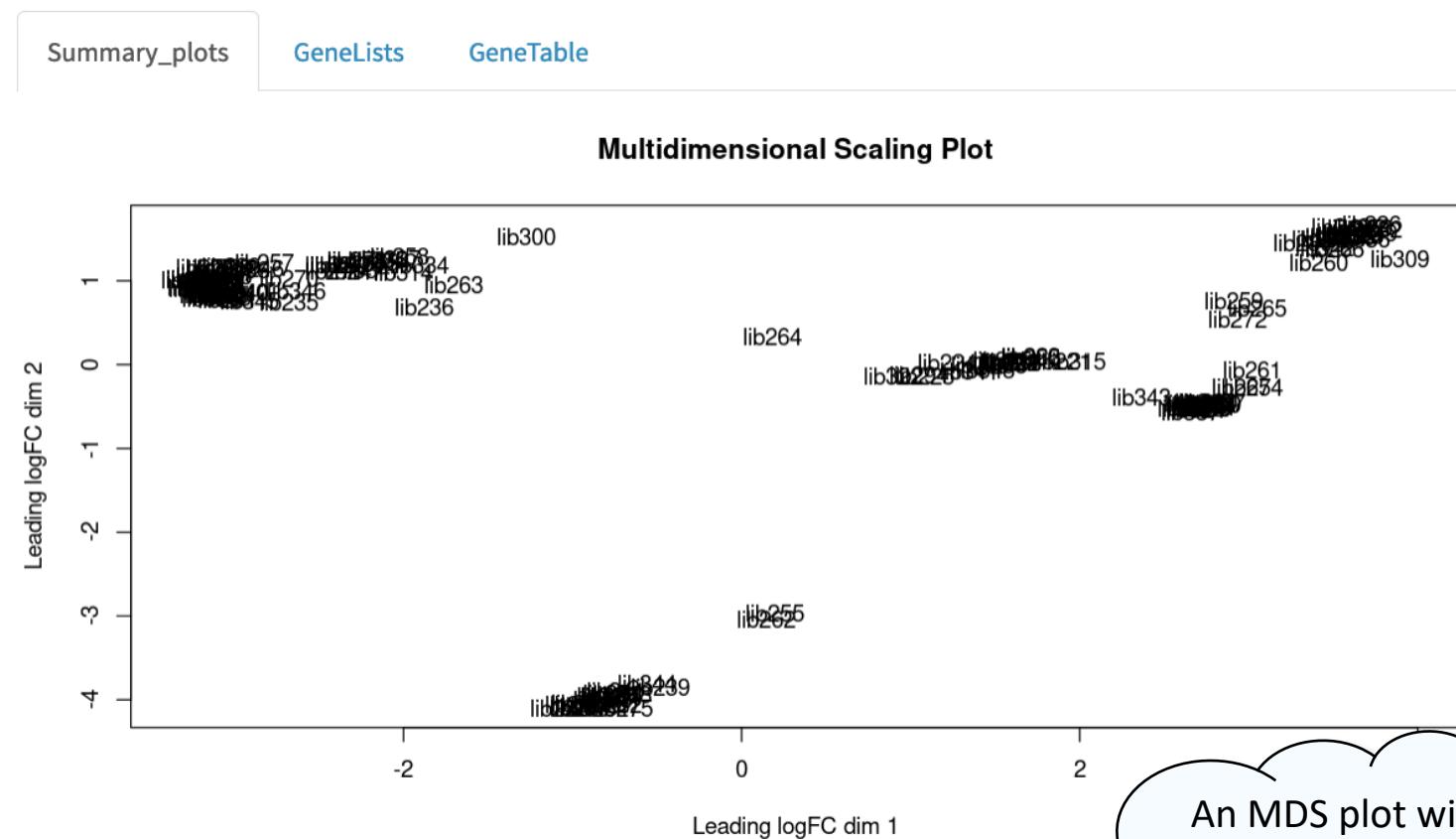
GSE60424

**Select Dataset**

▼

**Select Reads Status**

Raw  Normalized



An MDS plot with sample names appears when the data is loaded

# ROGUE – Differential Expression Analysis

? Load Raw Reads and Perform Differentially Expressed Gene Analysis

**Choose Reads File**

Browse... Data\_Matrix.txt  
**Upload complete**

**Select Input Source**  
 Database  
 Upload File

**Select Data Library**  
GSE60424

**Select Dataset**  
Select

**Select Reads Status**  
 Raw  Normalized

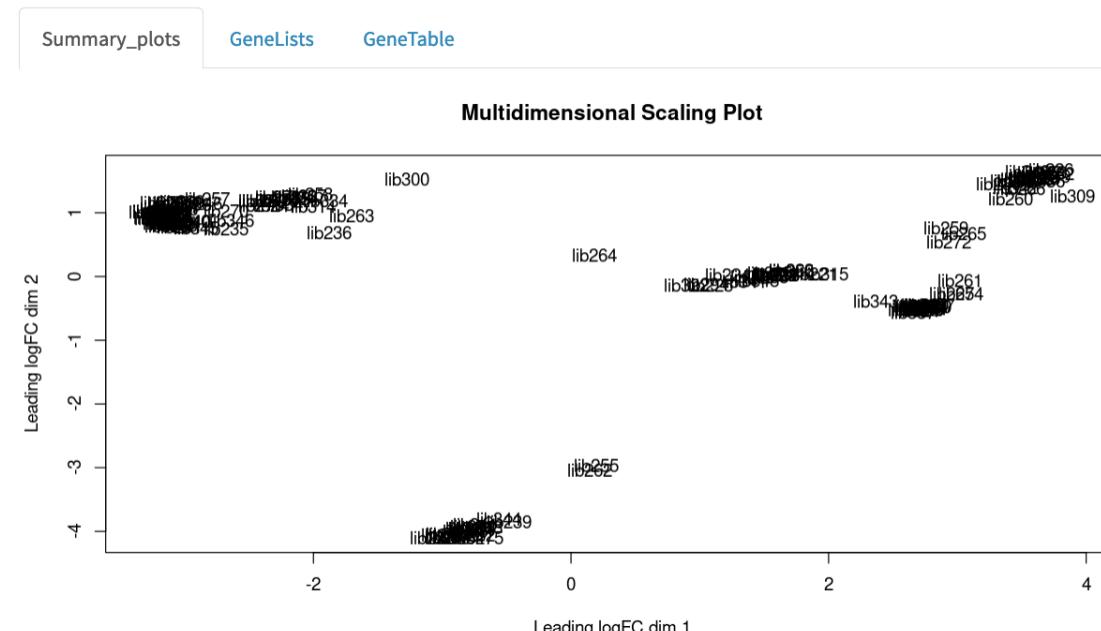
Load Data

Add data

Data/READS//GSE60424/

**Select 2 Samples**  
lib221 lib222 lib223 lib224 lib225 lib226 lib227 lib228

**P-value correction method**  
fdr



Select two Libraries (Sample Names) to perform DEG analysis from the '**Select Library**' dropdown menu.

# ROGUE – Differential Expression Analysis

**Choose Reads File**

Browse... Data\_Matrix.txt  
Upload complete

**Select Input Source**

Database  
 Upload File

**Select Data Library**

GSE60424

**Select Dataset**

▼

**Select Reads Status**

Raw  Normalized

**Load Data**

Add data

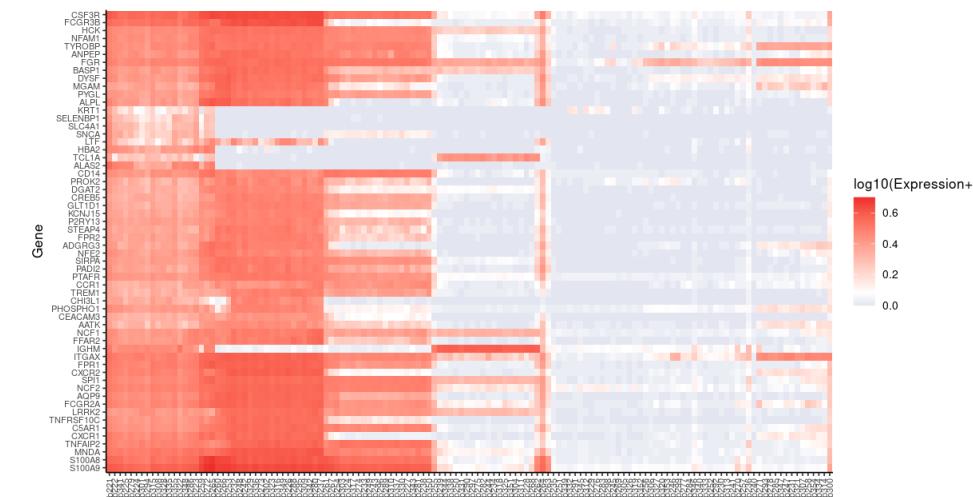
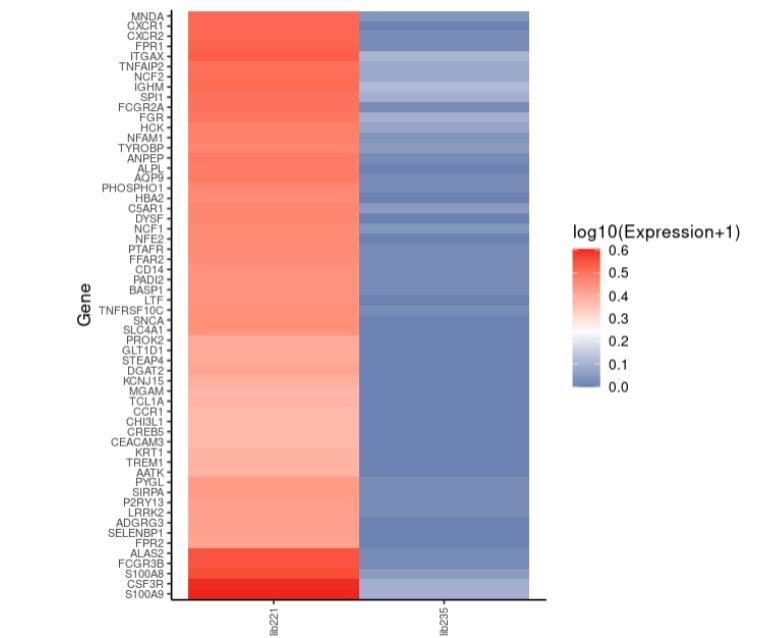
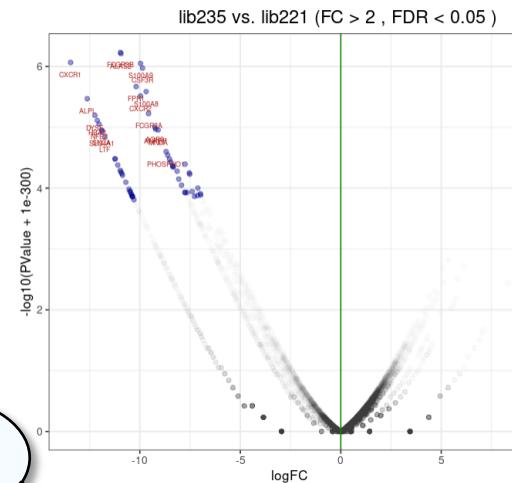
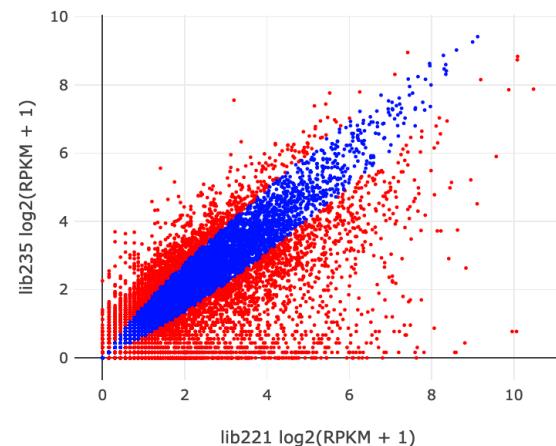
Data/READS//GSE60424/

**Select 2 Samples**

lib221 lib235

**Compare Samples**

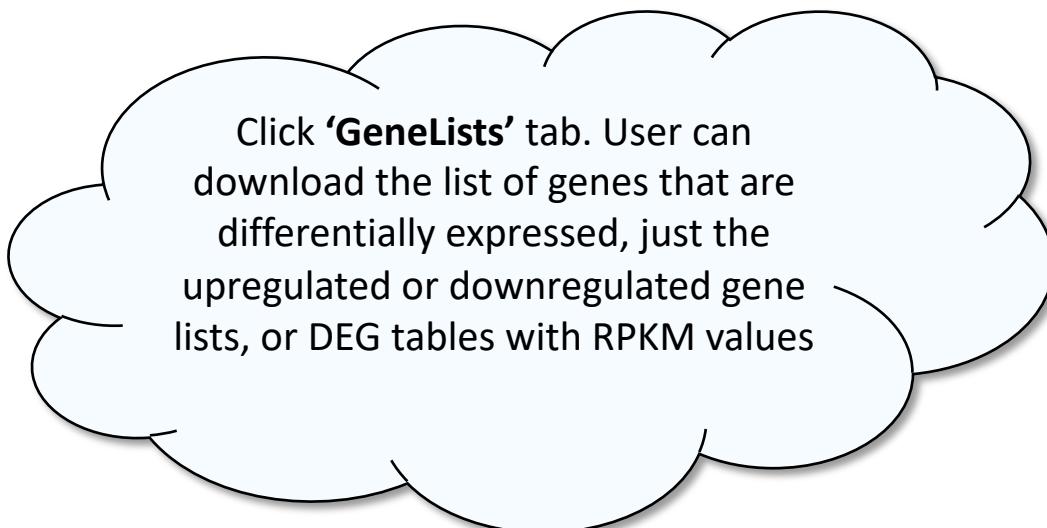
Click 'Compare Libraries' button to perform DEG analysis and display plots.



# ROGUE – Differential Expression Analysis

The screenshot shows the ROGUE software interface. On the left, there's a panel titled 'Choose Reads File' with a 'Browse...' button, a file path 'GSE60424\_GEOSubmit\_FC1to11\_normalized\_cou', and a progress bar indicating 'Upload complete'. Below it is a 'Select Input Source' section with 'Database' and 'Upload File' options, where 'Upload File' is selected. On the right, there are two tabs: 'Summary\_plots' (selected) and 'GeneLists' (highlighted with a black arrow). Under each tab are download links:

- Under 'Summary\_plots': Download FC Gene List, Download RPKM Table
- Under 'GeneLists': Download Upregulated FC Gene List, Download Upregulated RPKM Data, Download Downregulated Gene List, Download Downregulated RPKM Data



# ROGUE – Create Groups

The screenshot shows the ROGUE interface with the 'Create Groups' tab selected. The left sidebar lists various analysis options, and the main panel provides a form for creating groups. A red arrow points to the 'Create Groups' tab in the sidebar, and another red arrow points to the 'Create Groups' button in the main panel. Two callout bubbles provide instructions: one pointing to the 'Create Groups' tab and another pointing to the help button.

ROGUE RNAseq & Ontology Graphic User Environment

View ROGUE Manual ?

ROGUE

Load Count Data:(EdgeR)

Load Expression Data

Create Groups

EdgeR Group Comparison

Gene Comparison (Samples)

Gene Comparison (Groups)

Advanced Analysis <

Download Report

? Create Groups

Group Name Enter Group Name

Select Group Members

Create Group

Choose 'Groups' File

Browse... No file selected

Upload Groups Confirm Groups

Groups

Download Groups

Click the 'Create Groups' tab.

Click '?' button for step by step tips on using this tool

# ROGUE – Create Groups

The screenshot shows the ROGUE software interface with a dark theme. On the left, a vertical sidebar contains links: View ROGUE Manual, ROGUE, Load Count Data:(EdgeR), Load Expression Data, Create Groups (selected), EdgeR Group Comparison, Gene Comparison (Samples), Gene Comparison (Groups), Advanced Analysis, and Download Report. The main area is titled 'Create Groups'. It has sections for 'Group Name' (with a text input field 'Enter Group Name' and a 'Create Group' button), 'Choose 'Groups' File' (with 'Browse...' and 'No file selected' buttons), 'Upload Groups' (with a button), and 'Groups' (an empty table). A dropdown menu titled 'Select Group Members' is open, listing sample identifiers: lib221, lib222, lib223, lib224, lib225, lib226, lib227, and lib228. A red arrow points to the 'lib221' option. At the bottom, there is a 'Confirm Groups' button followed by a 'lib220' placeholder.

Create groups by choosing samples/libraries from the **'Select Group Members'** dropdown list.

# ROGUE – Create Groups

 Create Groups

**Group Name**

Group1

**Select Group Members**

lib221 lib222 lib223 lib224

Create Group



Assign a name to the the group in the '**Group Name**' textbox then click the '**Create Group**' button.

# ROGUE – Create Groups

? Create Groups

**Group Name**  
Enter Group Name

**Select Group Members**

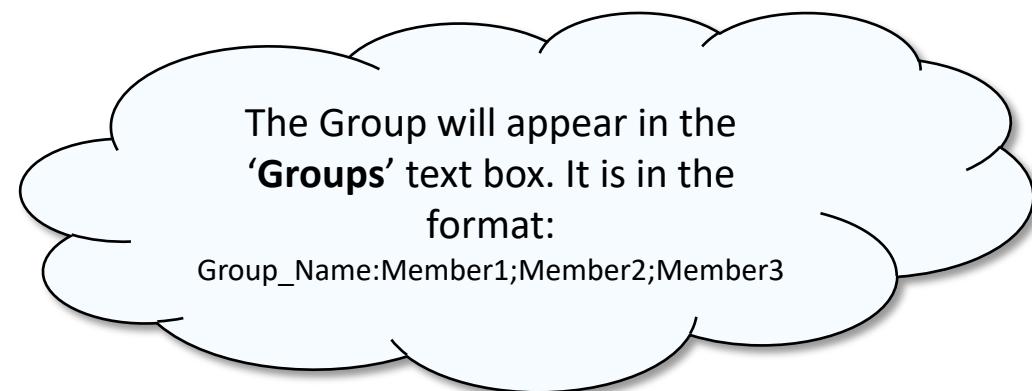
**Create Group**

**Choose 'Groups' File**  
Browse... No file selected

**Upload Groups**      **Confirm Groups**

**Groups**  
Group1:lib221;lib222;lib223;lib224

Download Groups



# ROGUE – Create Groups

? Create Groups

**Group Name**  **Select Group Members**

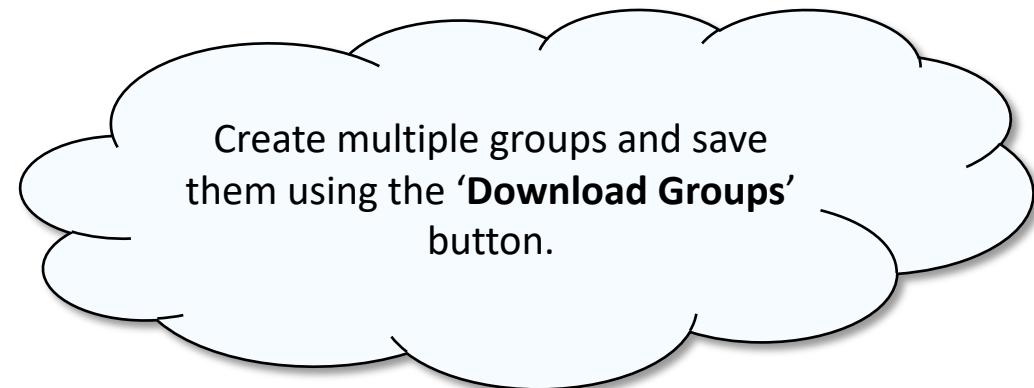
**Create Group**

**Choose 'Groups' File**

**Upload Groups** **Confirm Groups**

**Groups**  
Group1:lib221;lib222;lib223;lib224

**Download Groups** 



# ROGUE – Create Groups

**Create Groups**

**Group Name**

**Select Group Members**

**Create Group**

**Choose 'Groups' File**

**Browse...** **No file selected** 

**Upload Groups** **Confirm Groups**

**Groups**  
Group1:lib221;lib222;lib223;lib224

**Download Groups**

Name

Name	Date Modified	Size	Kind
Data_Matrix.txt	Nov 4, 2020 at 8:21 PM	6.3 MB	Plain Text
<b>Groups_File.txt</b>	<b>Nov 4, 2020 at 8:21 PM</b>	<b>366 bytes</b>	<b>Plain Text</b>
GSE60424_GEOSubmit_FC1to11_normalized_counts.txt	Nov 4, 2020 at 5:53 PM	15.6 MB	Plain Text
GSE60424_GEOSubmit_FC1to11_normalized_counts.txt.gz	Nov 4, 2020 at 5:47 PM	2.4 MB	gzip co...archive
GSE60424_series_matrix.txt	Nov 4, 2020 at 6:00 PM	292 KB	Plain Text
GSE60424_series_matrix.txt.gz	Nov 4, 2020 at 5:46 PM	9 KB	gzip co...archive

**Enter groups manually by typing in the format:  
Group\_Name:Member1;Member2;Member3  
Or upload saved groups by clicking 'Browse' button**

# ROGUE – Create Groups

**Create Groups**

**Group Name**

Enter Group Name

**Select Group Members**

**Create Group**

**Choose 'Groups' File**

Browse... No file selected

**Upload Groups**      **Confirm Groups**

**Groups**

Group1:lib221;lib222;lib223;lib224

**Download Groups**

A screenshot of the ROGUE software interface for creating groups. On the left, there's a form with fields for 'Group Name' (containing 'Enter Group Name'), 'Select Group Members' (with a placeholder box), and a 'Create Group' button. Below these are sections for choosing a 'Groups' file (with 'Browse...' and 'No file selected' buttons) and managing groups ('Upload Groups', 'Confirm Groups', 'Groups' list, and 'Download Groups'). A central part of the interface is a file selection dialog box. This dialog has a title bar 'Groups File Selection' and a message 'Select the Groups file' inside a cloud-shaped callout. It contains a table with columns 'Name', 'Date Modified', 'Size', and 'Kind'. The table lists several files: 'Data\_Matrix.txt', 'Groups\_File.txt' (which is highlighted in blue), 'GSE60424\_GEOSubmit\_FC1to11\_normalized\_counts.txt', 'GSE60424\_GEOSubmit\_FC1to11\_normalized\_counts.txt.gz', 'GSE60424\_series\_matrix.txt', and 'GSE60424\_series\_matrix.txt.gz'. The 'Groups\_File.txt' row shows details: Nov 4, 2020 at 8:21 PM, 366 bytes, Plain Text. At the bottom of the dialog are 'Cancel' and 'Open' buttons, with a red arrow pointing to the 'Open' button.

Name	Date Modified	Size	Kind
Data_Matrix.txt	Nov 4, 2020 at 8:21 PM	6.3 MB	Plain Text
Groups_File.txt	Nov 4, 2020 at 8:21 PM	366 bytes	Plain Text
GSE60424_GEOSubmit_FC1to11_normalized_counts.txt	Nov 4, 2020 at 5:53 PM	15.6 MB	Plain Text
GSE60424_GEOSubmit_FC1to11_normalized_counts.txt.gz	Nov 4, 2020 at 5:47 PM	2.4 MB	gzip co...archive
GSE60424_series_matrix.txt	Nov 4, 2020 at 6:00 PM	292 KB	Plain Text
GSE60424_series_matrix.txt.gz	Nov 4, 2020 at 5:46 PM	9 KB	gzip co...archive

37

# ROGUE – Create Groups

? Create Groups

**Group Name**  
Enter Group Name

**Select Group Members**

**Create Group**

**Choose 'Groups' File**

Browse... Groups\_File  
**Upload complete**

**Upload Groups**  Confirm Groups

**Groups**  
Group1:lib221;lib222;lib223;lib224

**Download Groups**

When the Browse box says, 'Upload complete' click the '**Upload Groups**' button.

# ROGUE – Create Groups

? Create Groups

Group Name

Enter Group Name

Select Group Members

Create Group

Choose 'Groups' File

Browse... Groups\_File

Upload complete

Upload Groups

Confirm Groups

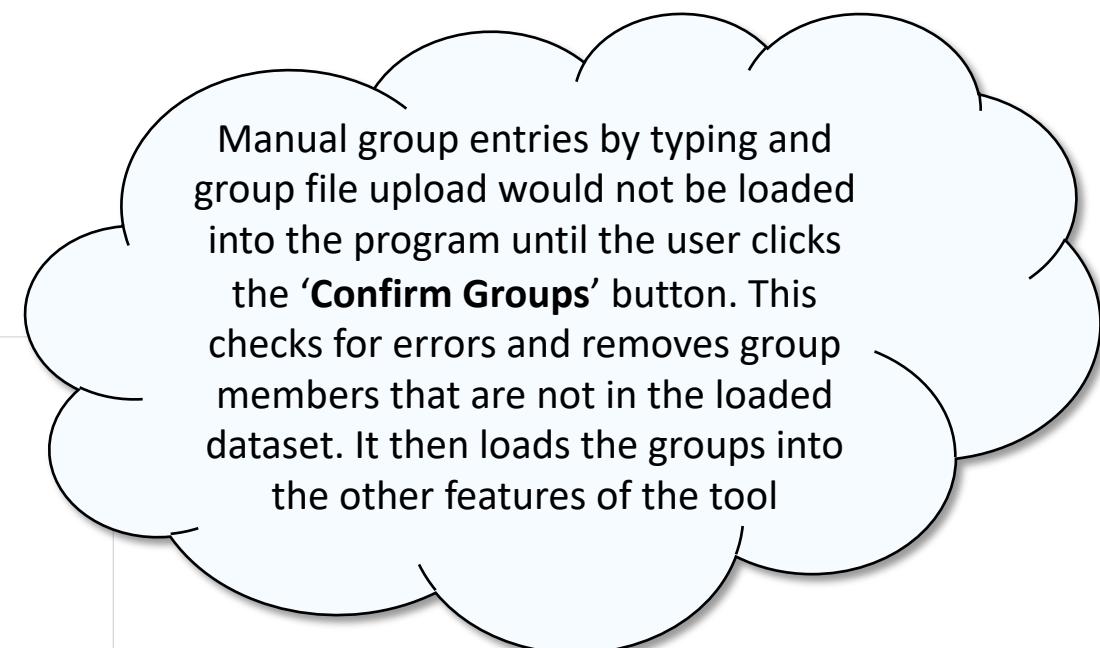
Groups

Healthy\_CD4:lib229;lib291;lib298;lib312

Healthy\_CD8:lib230;lib292;lib299;lib313

Healthy\_NK:lib231;lib293;lib300;lib314

Download Groups



Manual group entries by typing and group file upload would not be loaded into the program until the user clicks the '**Confirm Groups**' button. This checks for errors and removes group members that are not in the loaded dataset. It then loads the groups into the other features of the tool

# ROGUE – Groups Differential Expression Analysis

## ? Groups' Differentially Expressed Genes Analysis

Select Group 1

Healthy\_CD4

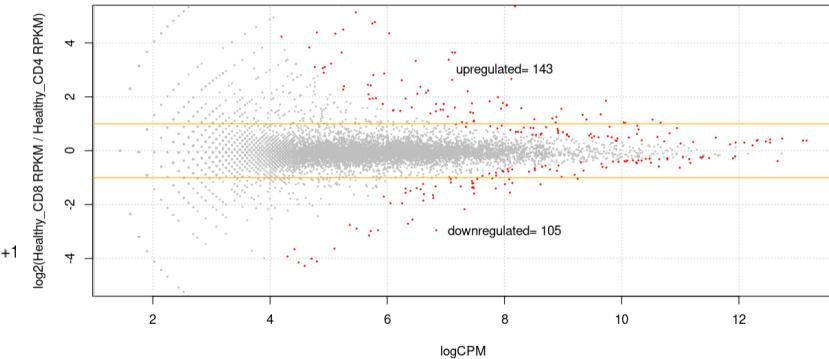
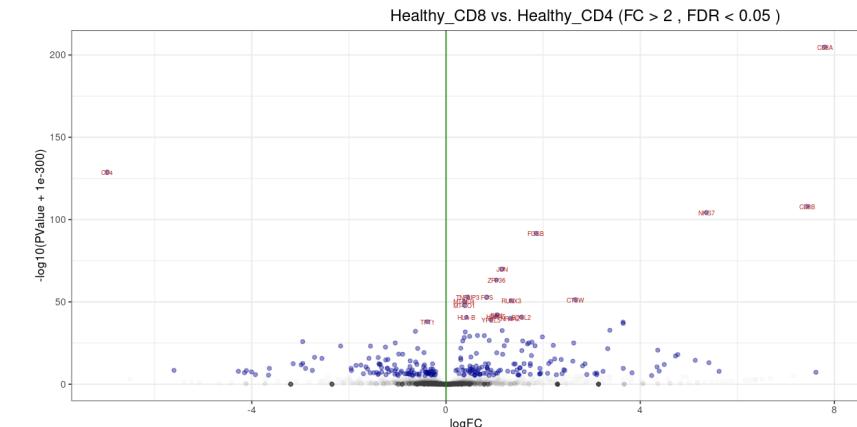
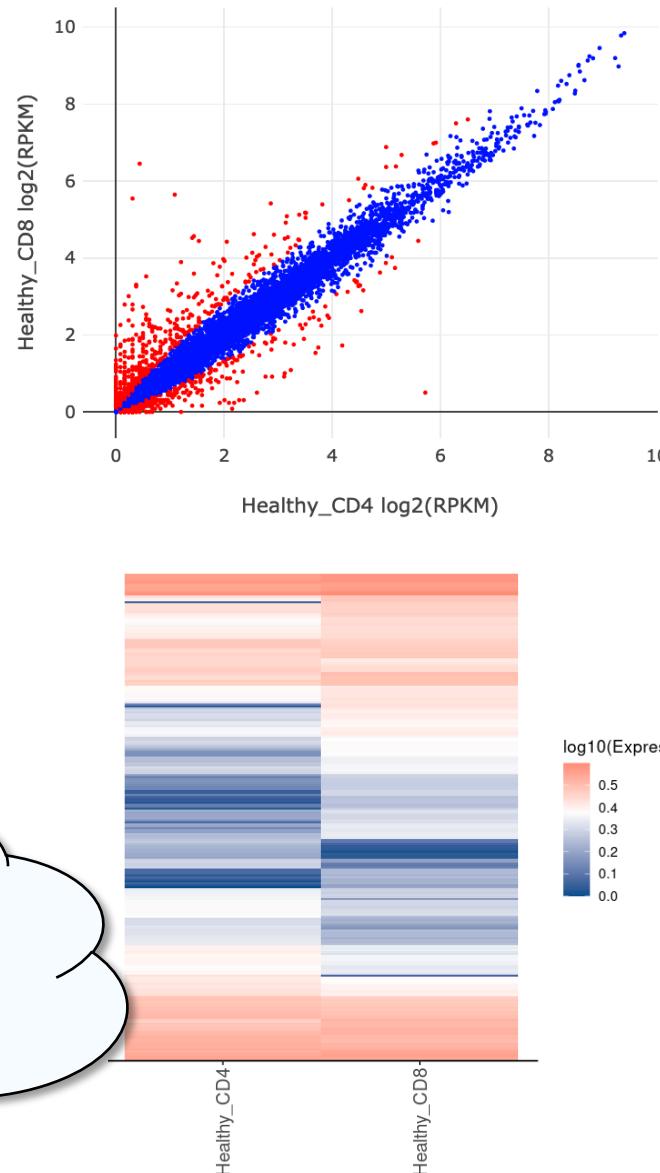
Select Group 2

Healthy\_CD8

Compare Groups

Select 'Healthy\_CD8' from the 'Select Group 2' dropdown box, and click 'Compare Groups'

Plots will be generated on the summary plots page



# ROGUE – Single Sample Gene Expression Analysis

The screenshot shows the ROGUE software interface with the 'Gene Comparison (Samples)' tab selected. The left sidebar contains navigation links: ROGUE, Load Count Data:(EdgeR), Load Expression Data, Create Groups, EdgeR Group Comparison, Gene Comparison (Samples) (highlighted with a red arrow), Gene Comparison (Groups) (highlighted with a red arrow), Advanced Analysis, and Download Report.

The main panel has the following sections:

- Select Genes (max:12):** A dropdown menu containing "IL24", "PPIL2", "FADS2", "stat5a", and "STAT5A". A red arrow points to the "STAT5A" entry.
- Paste Gene List (max:12):** A text input field containing "IL21R", "MYC", and "IL15". A red arrow points to the "IL15" entry.
- Select Samples:** A dropdown menu containing "lib229", "lib230", and "lib231". A red arrow points to the "lib231" entry.

Below these sections are various configuration controls:

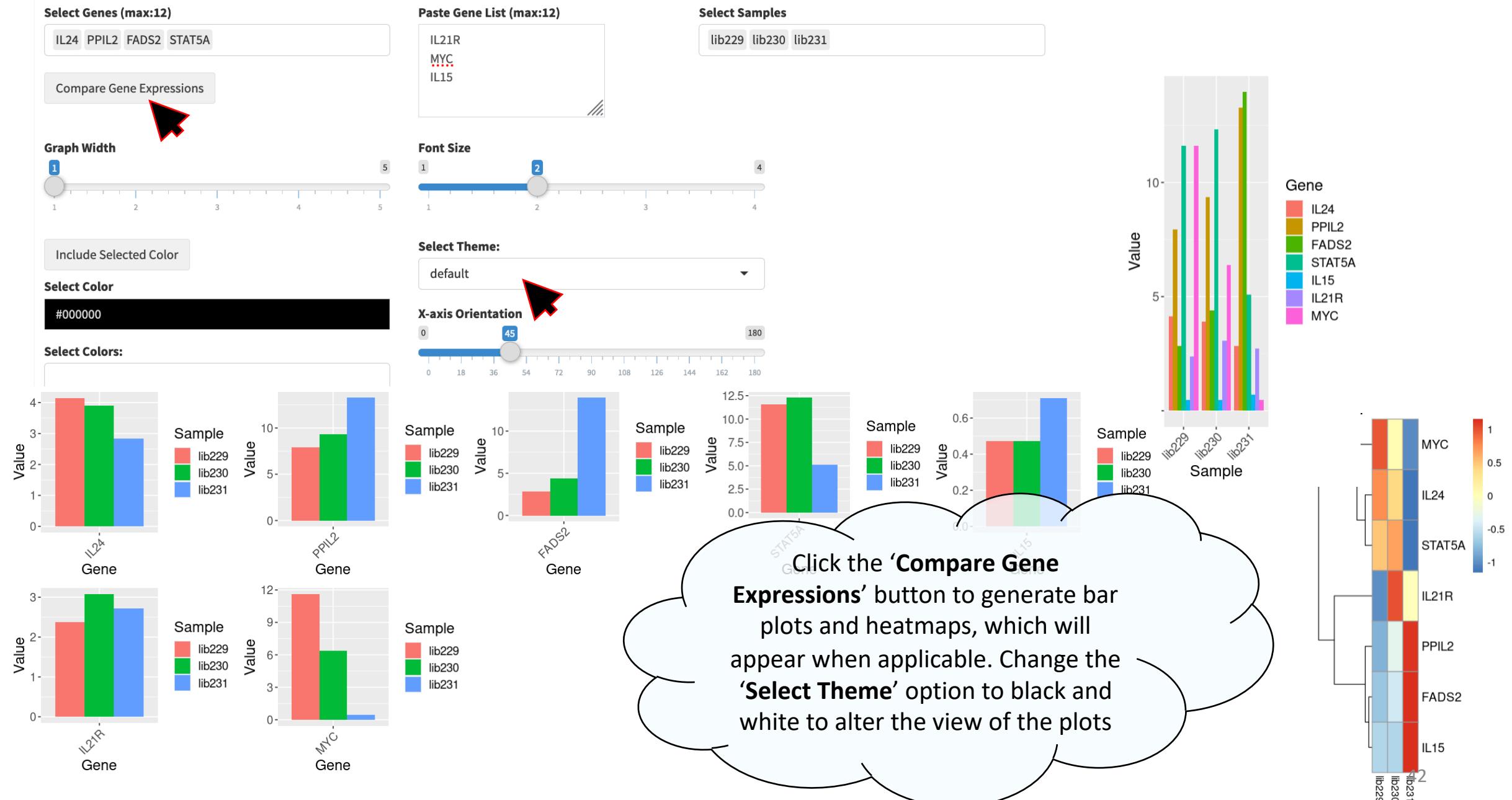
- Graph Width:** A slider set to 2.
- Font Size:** A slider set to 2.
- Select Theme:** A dropdown menu set to "default".
- X-axis Orientation:** A slider set to 45.

A large callout bubble on the right side provides instructions:

Click '**Gene Comparison (Samples)**' tab. User can select genes from the '**Select Genes**' dropdown list or type/paste genes in the '**Paste Gene List**' text box. Select Libraries or samples to compare from the '**Select Samples**' dropdown list.

41

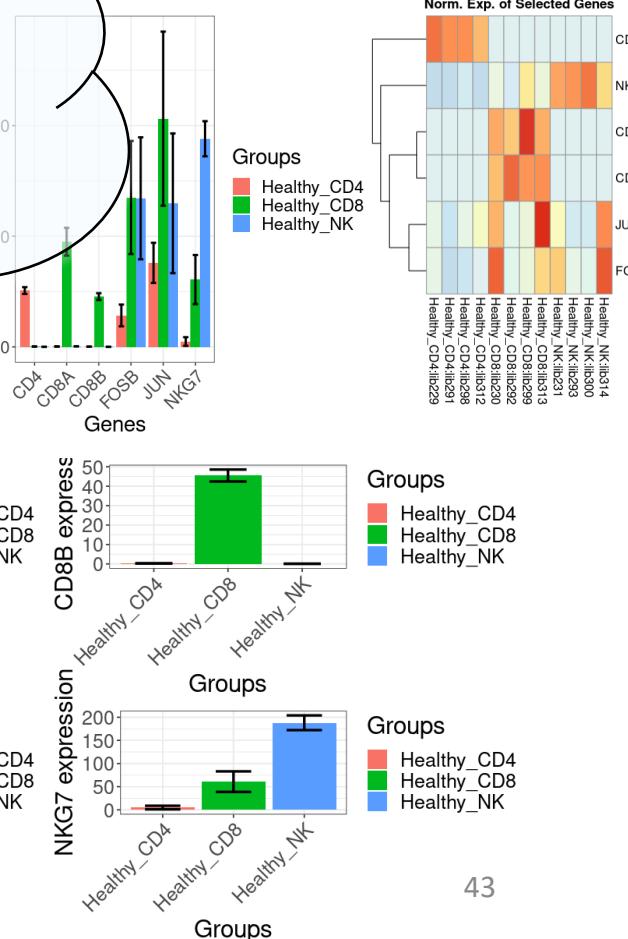
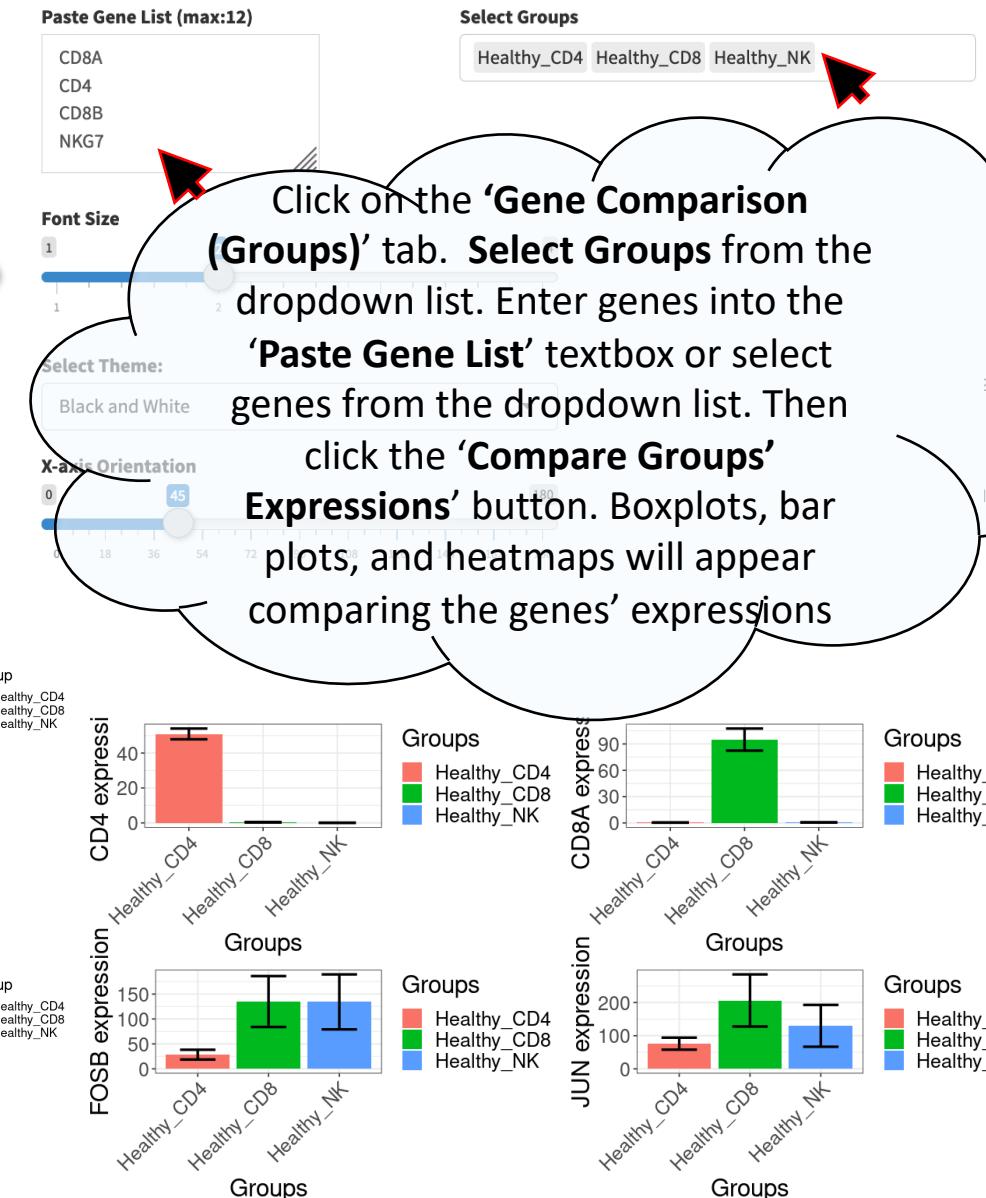
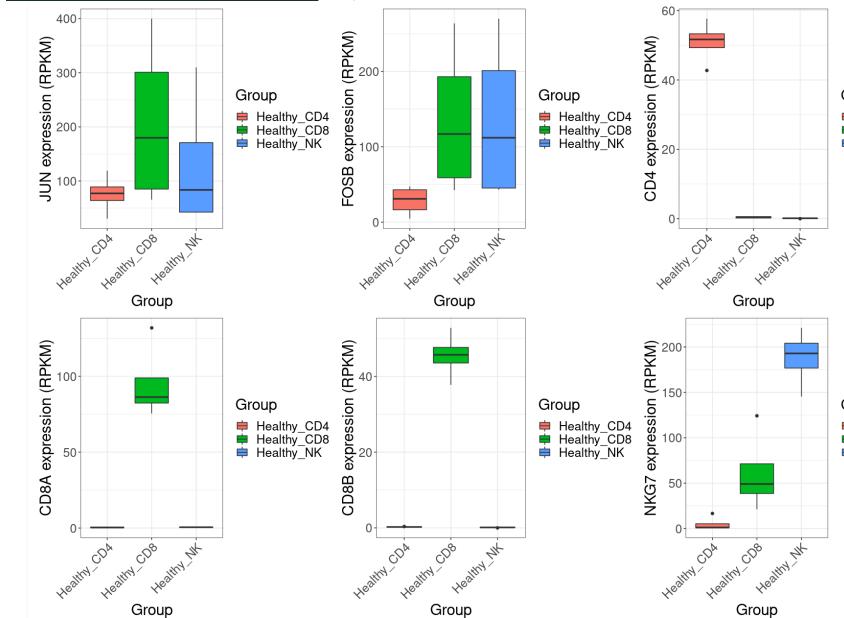
## ROGUE – Single Sample Gene Expression Analysis



View ROGUE Manual ?

- ROGUE
- Load Count Data:(EdgeR)
- Load Expression Data
- Create Groups
- EdgeR Group Comparison
- Gene Comparison (Samples)
- Gene Comparison (Groups)**
- Advanced Analysis <

Download Report



# ROGUE – Gene Set Enrichment Analysis

The screenshot shows the ROGUE software interface. On the left is a dark sidebar with a list of menu items: View ROGUE Manual, ROGUE, Load Count Data:(EdgeR), Load Expression Data, Create Groups, EdgeR Group Comparison, Gene Comparison (Samples), Gene Comparison (Groups), Advanced Analysis, Gene Set Enrichment Analysis (highlighted with a red arrow), Gene Ontology, Group Statistical Comparison, Group Expr. Ontologies, Differentially Expressed Ontolog, and Download Report. The main area displays the 'Perform Gene Set Enrichment Analysis' dialog box. It includes fields for Analyze Groups or Samples? (Groups is selected), Select Control (lib221), Select Subjects (lib355), Upregulated Enriched GeneSets (10), Downregulated Enriched GeneSets (10), Select GSEA Collection (hallmark gene sets is selected), Find Enriched Gene Signatures, Select Gene Set, and Select Gene Limit (set to 0). A question mark icon in the top right corner of the dialog box is also highlighted with a red arrow.

Click the 'Advanced Analysis' tab.  
Then Click on 'Gene Set Enrichment Analysis'

Click '?' button for step  
by step tips on using this  
tool

# ROGUE – Gene Set Enrichment Analysis

Select ‘Groups’ radio button.

Then select groups for ‘Select Control’ and ‘Select Subjects’ drop down box. Select GSEA sets from the ‘Select GSEA Collection’ drop down box.

Click the ‘Find Enrichment Gene Signatures’ button.

? Perform Gene Set Enrichment Analysis

Analyze Groups or Samples?

Groups    Sample

**Select Control**

Healthy\_CD4

**Select Subjects**

Healthy\_CD8

**Upregulated Enriched GeneSets** 10

**Downregulated Enriched GeneSets** 10

**Select GSEA Collection**

hallmark gene sets   GO gene sets

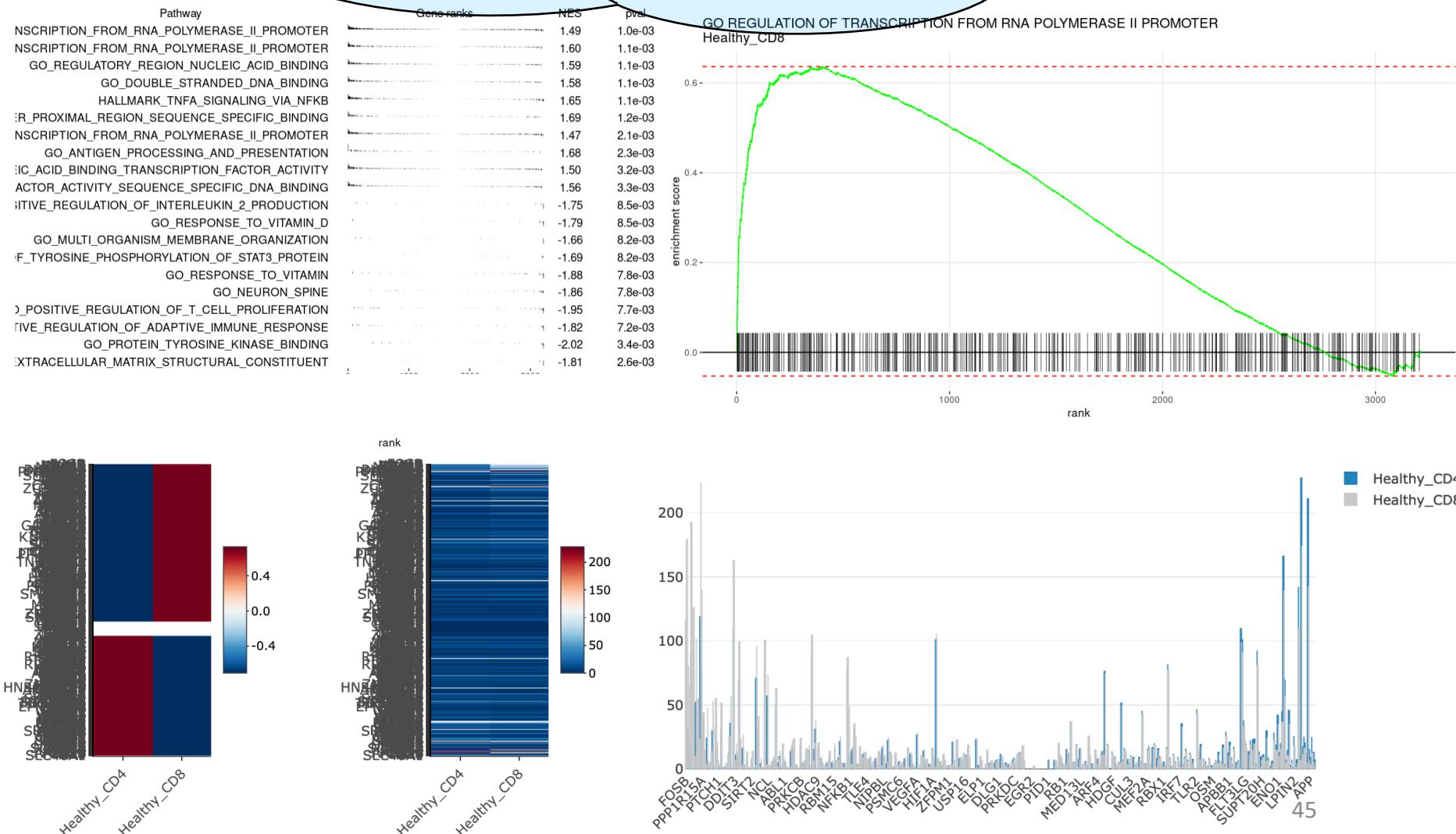
**Find Enriched Gene Signatures**

**Select Gene Set**

GO\_REGULATION\_OF\_TRANSCRIPTION\_FROM\_RNA\_P

**Select Gene Limit**

2 1,784



# ROGUE – Gene Ontology Analysis

The screenshot shows the ROGUE RNAseq+ Ontology Graphic User Environment. The main window displays the 'Gene Ontology Analysis' page. Key features include:

- Left Sidebar:** Includes links for ROGUE, Load Count Data, Load Expression Data, Create Groups, EdgeR Group Comparison, Gene Comparison (Samples), Gene Comparison (Groups), Advanced Analysis, Gene Set Enrichment Analysis, Gene Ontology (selected), Group Statistical Comparison, Group Expr. Ontologies, Differentially Expressed Ontologies, and Download Report.
- Top Bar:** Shows the ROGUE logo and 'RNAseq+ Ontology Graphic User Environment'.
- Header:** 'Gene Ontology Analysis' with a help icon.
- Form Fields:**
  - Select Organism: Human (checked), Mouse (checked).
  - Enter Gene List: GZMB, ITGAX (highlighted with a red arrow).
  - Select Ontology Type: biological\_process (selected).
  - P-value: 0.05.
  - Correction: none.
  - Select X-axis Labels: GO ID (selected).
  - Get Ontologies button (highlighted with a red arrow).
  - Select Ontology Confidence Code:
    - EXP - Inferred from Experiment (checked)
    - IDA - inferred from direct assay (checked)
    - IPI - inferred from physical interaction (checked)
    - IMP - inferred from mutant phenotype (checked)
    - IGI - inferred from genetic interaction (checked)
    - IEP - inferred from expression pattern (checked)
    - TAS - traceable author statement (checked)
- Right Panel:** 'Select Ontologies' tab selected. It includes 'Select GO' and 'Genes' buttons.
- Annotations:**
  - A callout bubble points to the 'Get Ontologies' button with the text: 'Click 'Advanced Analysis' tab. Click on the 'Gene Ontology tab'. Users can type or paste a list of genes in the 'Enter Gene List' textbox. Click 'Get Ontologies' button. The initial processing of this function may take a few minutes.'
  - A callout bubble points to the 'Get Ontologies' button with the text: 'Don't forget to use the '?' button for tips.'
  - A callout bubble points to the 'Select Ontology Confidence Code' section with the text: 'If the gene list is long, it is **HIGHLY RECOMMENDED** to use an external gene ontology/pathway database.'
- Status Bar:** 'Determining Gene Ontology' with 'Progress: 20%'.
- Page Number:** 46.

# ROGUE – Gene Ontology Analysis

Select Organism

Human  Mouse

Enter Gene List

GZMB  
ITGAX

Select Ontology Type

molecular\_function  
 biological\_process  
 cellular\_component

P-value P-val adj

0.05 none▼

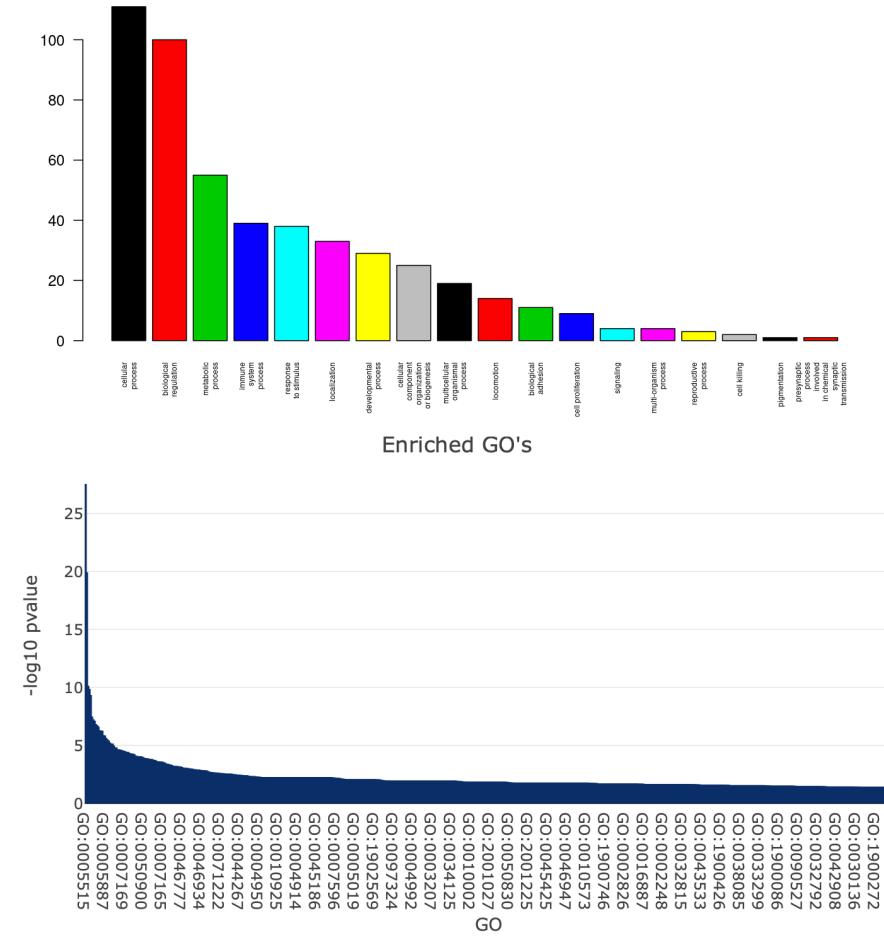
Select X-axis Labels

GO ID  GO Name

Get Ontologies

Select Ontology Confidence Code

EXP - Inferred from Experiment  
 IDA - inferred from direct assay  
 IPI - inferred from physical interaction  
 IMP - inferred from mutant phenotype  
 IGI - inferred from genetic interaction  
 IEP - inferred from expression pattern  
 TAS - traceable author statement  
 ISS - inferred from sequence similarity  
 IEA - inferred from electronic annotation  
 NAS - non-traceable author statement  
 ND - no biological data available  
 IC - inferred by curator  
 RCA - inferred from reviewed computational analysis



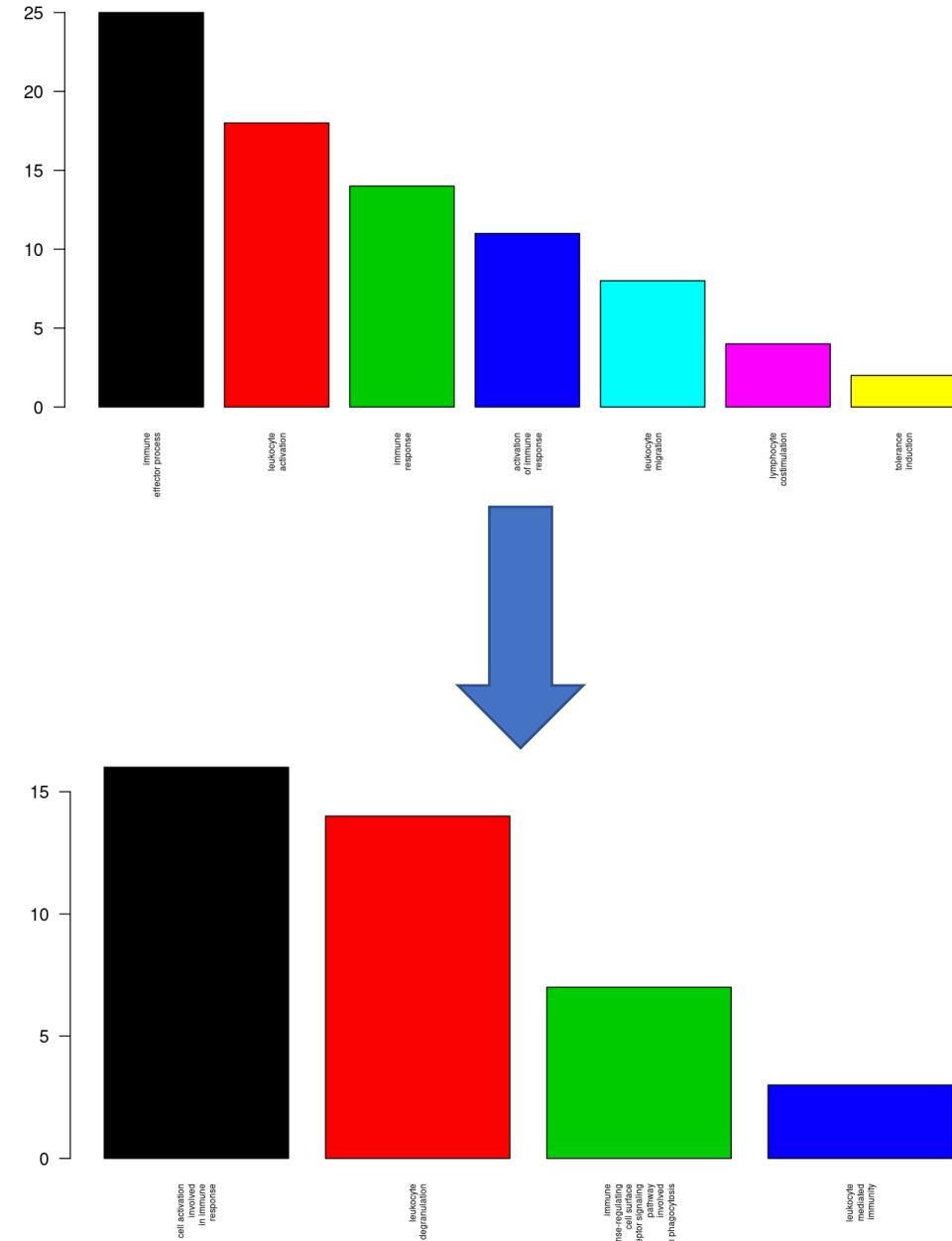
- Select Ontologies
- cellular process
  - biological regulation
  - metabolic process
  - immune system process
  - response to stimulus
  - localization
  - developmental process
  - cellular component organization or biogenesis
  - multicellular organismal process
  - locomotion
  - biological adhesion
  - cell proliferation
  - signaling
  - multi-organism process
  - reproductive process
  - cell killing
  - pigmentation
  - presynaptic process involved in chemical synaptic transmission

Select GO  
Represented Genes

DENND3  
GSAP

Gene Ontology bar plots appear. Choose ontology subcategories and click 'Select GO' button to explore even deeper subcategories within the selected ontology.

# ROGUE – Gene Ontology Analysis



Go Back

## Select Ontologies

- immune effector process
- leukocyte activation
- immune response
- activation of immune response
- leukocyte migration
- lymphocyte costimulation
- tolerance induction

Select GO

### immune system process:Represented Genes

MGAM  
TYROBP



**Note:** This text box lists the genes represented in the bar plot.

Go Back

## Select Ontologies

- cell activation involved in immune response
- leukocyte degranulation
- immune response-regulating cell surface receptor signaling pathway involved in phagocytosis
- leukocyte mediated immunity

Select GO

### immune effector process:Represented Genes

MGAM  
TYROBP

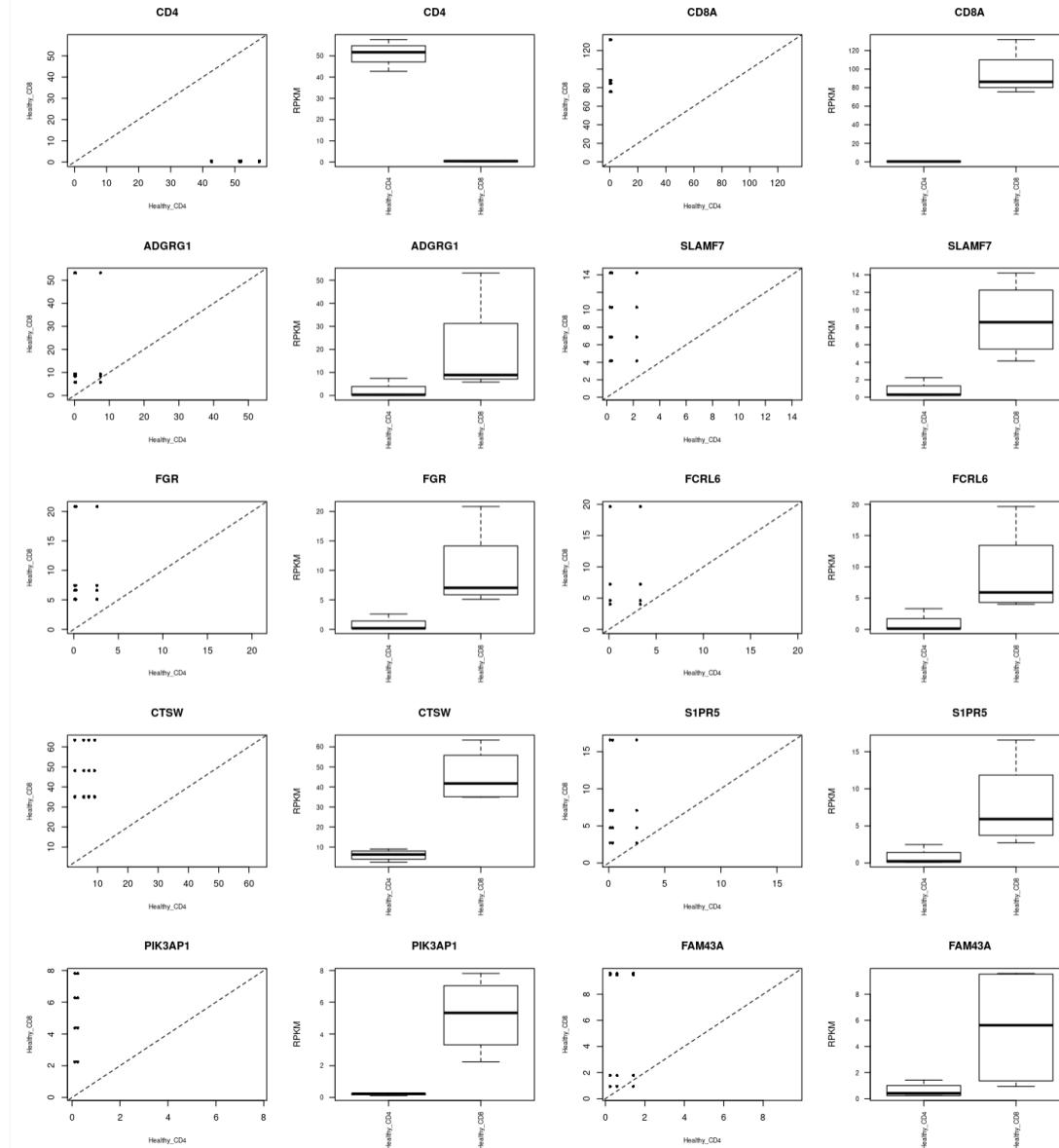
Gene Ontology bar plots appear. Choose ontology subcategories and click 'Select GO' button to explore even deeper subcategories within the selected ontology.

# ROGUE – Statistical Group Comparison (Biomarker Discovery)

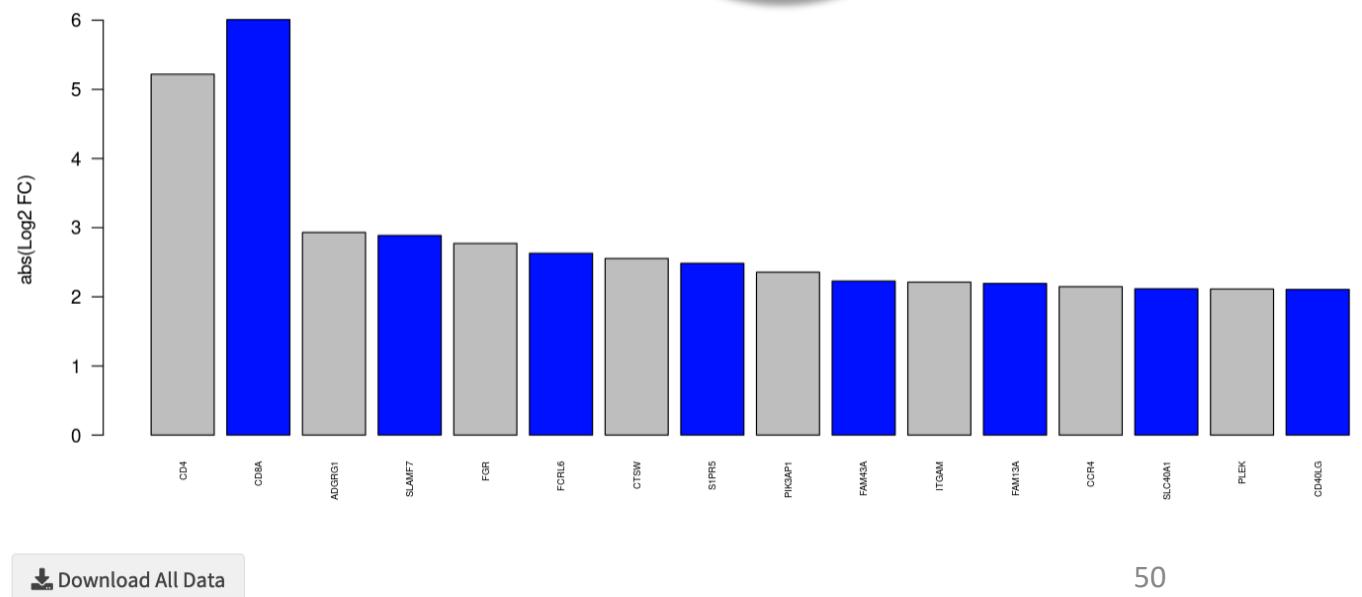
The screenshot shows the ROGUE software interface. On the left, a sidebar lists various analysis options. A red arrow points to the 'Group Statistical Comparison' option. The main panel displays the 'Statistical Biomarker Prediction Tool' under the 'Group\_Stats\_GeneList' tab. It includes dropdown menus for 'Select Group1' (Healthy\_CD4) and 'Select Group2' (Healthy\_CD8), a checkbox for 'Fold Change' (checked), another for 'Mean/SD' (checked), and others for 'Wilcox' and 'Ttest'. A 'Find Genes' button is highlighted with a red arrow. Below it is a slider for 'Select Pvalue' set at 0.05. There are also input fields for 'Min RPKM' (5) and 'Min log2FC' (2). A checkbox 'Show log2FC in gene list?' is checked. At the bottom, a 'Display Top Genes' field is set to 10.

Click '**Group Statistical Comparison**' tab. This tool attempts to identify potential biomarkers by comparing gene expression across all members of each group. Select groups to be compared in the '**Select Groups**' dropdown list. check '**Fold Change**' and '**Mean/SD**' checkboxes. Click '**Find Genes**' button.

# ROGUE – Statistical Group Comparison (Biomarker Discovery)



Plots of expression values and most consistently differentially expressed genes across groups will be displayed.  
Only the top 10 (default) genes will be displayed in the dot and box plots.  
This value can be changed. All data and plots can be downloaded to a pdf file.



[Download All Data](#)

# ROGUE – Statistical Group Comparison (Biomarker Discovery)

**Select Group1**

Healthy\_CD4

**Select Group2**

Healthy\_CD8

**Select Tests**

Fold Change

Mean/SD

Wilcox

Ttest

**Find Genes**

**Select Pvalue**

0  1

**Min RPKM**

5

**Min log2FC**

2

**Display Top Genes**

10

Show log2FC in gene list?

[Group\\_Stats\\_Summary\\_plots](#)

[Group\\_Stats\\_GeneList](#)

[t\\_SNE](#)

**Top scoring genes with corresponding fold change**

CD4	-5.21903072483948
CD8A	6.00962199376225
ADGRG1	2.93043913372405
SLAMF7	2.88655602890384
FGR	2.77220028537918
FCRL6	2.62932885057784
CTSW	2.55439201369152
S1PR5	2.48415368235614
PIK3AP1	2.35513136913033
FAM43A	2.22750735888574
ITGAM	2.21170495335981
FAM13A	-2.19363429579282
CCR4	-2.14631226734269
SLC40A1	-2.1157391470026

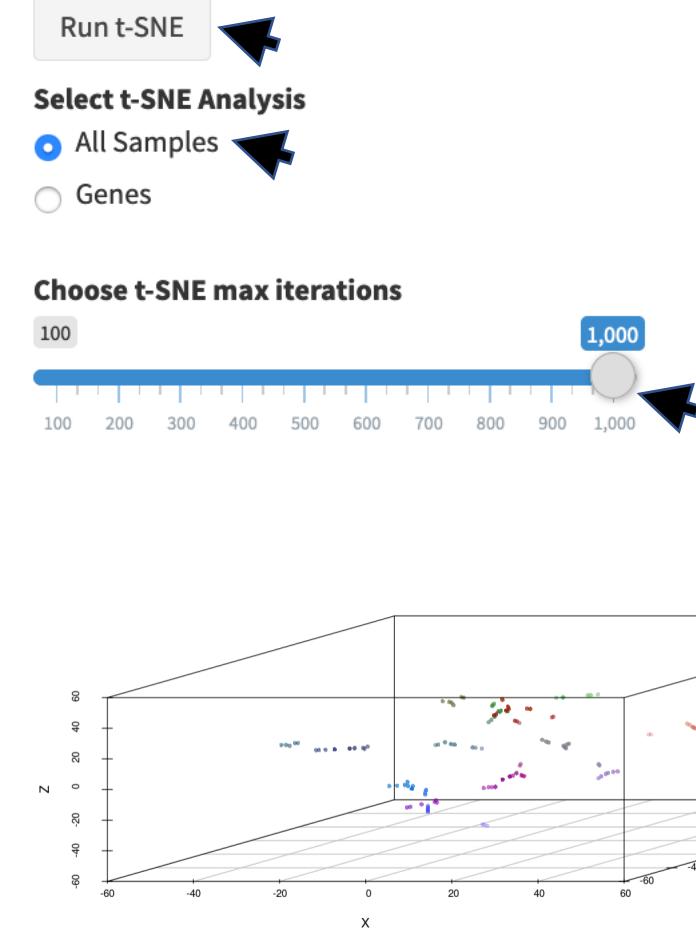
Click on  
**'Group\_Stats\_GeneList'** tab.  
A list of genes and the log2 fold change will appear in a textbox ranked from highest to lowest fitting the parameters set in the left panel.

# ROGUE – tSNE

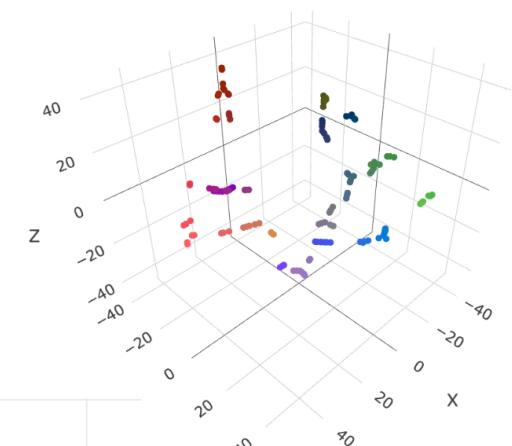
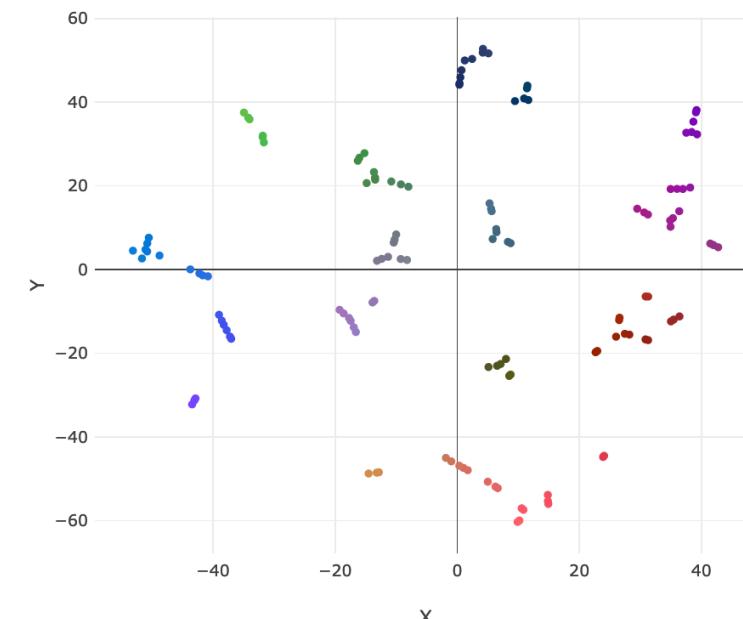
The screenshot shows the ROGUE interface with the 't-SNE' tab selected. The left sidebar contains various analysis options like ROGUE, Load Count Data, Create Groups, etc. The main panel has sections for 'Select Group1' (Healthy\_CD4), 'Select Group2' (Healthy\_CD8), 'Select Tests' (Fold Change, Mean/SD checked; Wilcox, Ttest unchecked), 'Find Genes', 'Select Pvalue' (0.05), 'Min RPKM' (5), 'Min log2FC' (2), 'Display Top Genes' (10), and 'Show log2FC in gene list?' (checked). The 't-SNE' tab is highlighted with a red arrow. Below it are settings for 't-SNE tool': 'Run t-SNE' button, 'Select t-SNE Analysis' (All Samples selected, Genes unselected), 'Choose t-SNE max iterations' slider (set to 1,000), 'Choose point size' slider (set to 3), 'Set Seed' input field (4000), and a checkbox for 'Generate distance p-value lists? (May take some time)'. There are also optional fields for 'Enter Gene List' and 'Highlight Gene (Optional)'.

Click on 't-SNE' tab

# ROGUE – tSNE



Select '**All Samples**' radio button and change '**Choose t-SNE max iterations**' sliding bar to 1000. 2D and 3D t-SNE plots will be generated.



# ROGUE – Gene Ontology Expression Comparison

The screenshot shows the ROGUE software interface with the 'Expressed Ontologies Tool' tab selected. On the left, a sidebar lists various analysis options: ROGUE, Load Count Data:(EdgeR), Load Expression Data, Create Groups, EdgeR Group Comparison, Gene Comparison (Samples), Gene Comparison (Groups), Advanced Analysis, Gene Set Enrichment Analysis, Gene Ontology, Group Statistical Comparison, Group Expr. Ontologies (highlighted with a red arrow), and Differentially Expressed Ontolog. At the bottom, there is a 'Download Report' button.

**Expressed Ontologies Tool**

Analyze Groups or Samples?  Groups  Samples

Select Control: Healthy\_CD4

Select Subjects: Healthy\_CD8, Healthy\_NK

Select GO Class: immune system process

Select GO Class: molecular\_function

Combine GOs with - Optional: (Push GO Class Button)

Get GO Gene's Fold Change



Click on 'Group Expr. Ontologies' tab.  
Select 'Groups' radio button  
Select one **Control** and either one or multiple **Subjects** from the dropdown lists.  
Select 'immune system process' from '**Select GO Class**' drop down menu.

# ROGUE – Gene Ontology Expression Comparison

Analyze Groups or Samples?

Groups  Samples

Select Control

Healthy\_CD4

Select Subjects

Healthy\_CD8 Healthy\_NK

Select GO Class

immune system process

Select GO (Push GO Class Button Above)

innate

innate immune response

innate immunity memory response

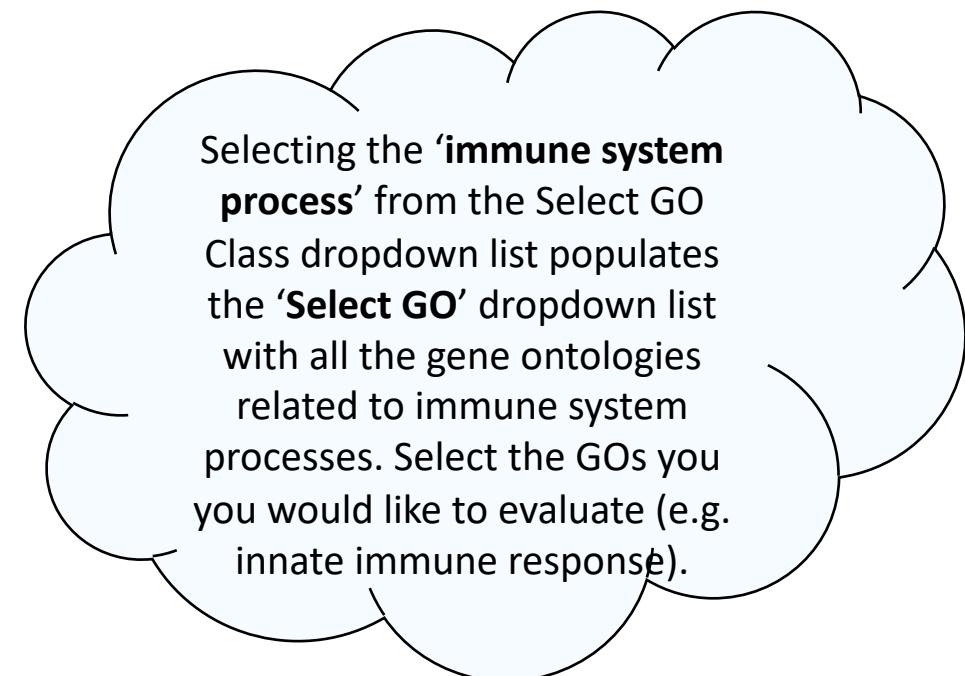
innate immune response in mucosa

innate immune response-activating signal transduction

innate immune response activating cell surface receptor signaling pathway

antifungal innate immune response

activation of innate immune response



# ROGUE – Gene Ontology Expression Comparison

? Expressed Ontologies Tool

Analyze Groups or Samples?

Groups    Samples

Select Control

Healthy\_CD4

Select Subjects

Healthy\_CD8   Healthy\_NK

Select GO Class

immune system process

Select GO (Push GO Class Button Above)

innate immune response

Select GO Class

molecular\_function

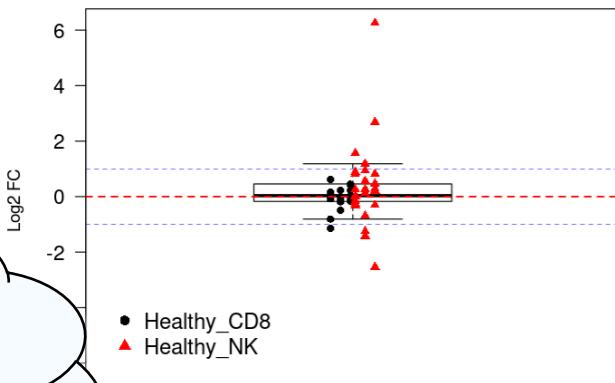
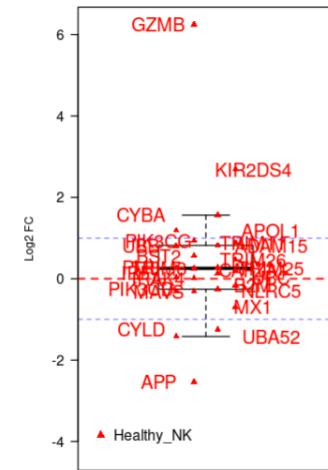
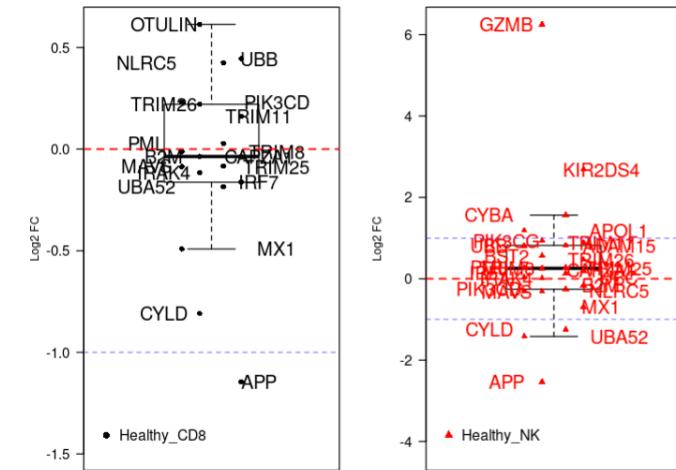
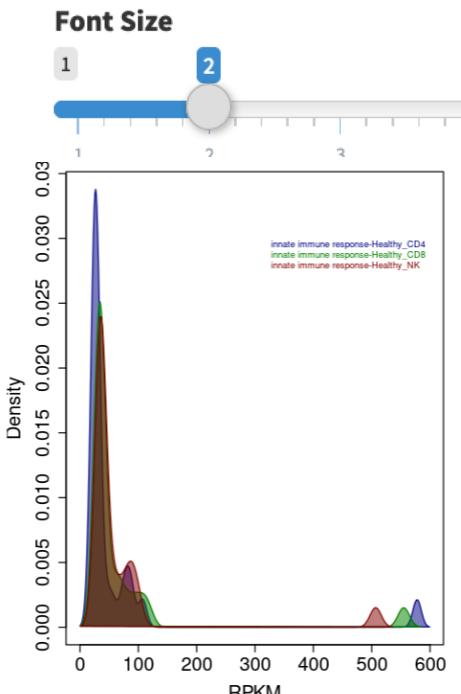
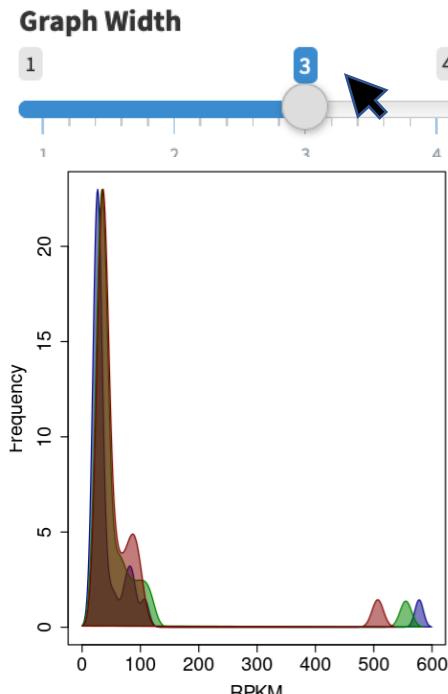
Combine GOs with - Optional:(Push GO Class Button)

Get GO Gene's Fold Change

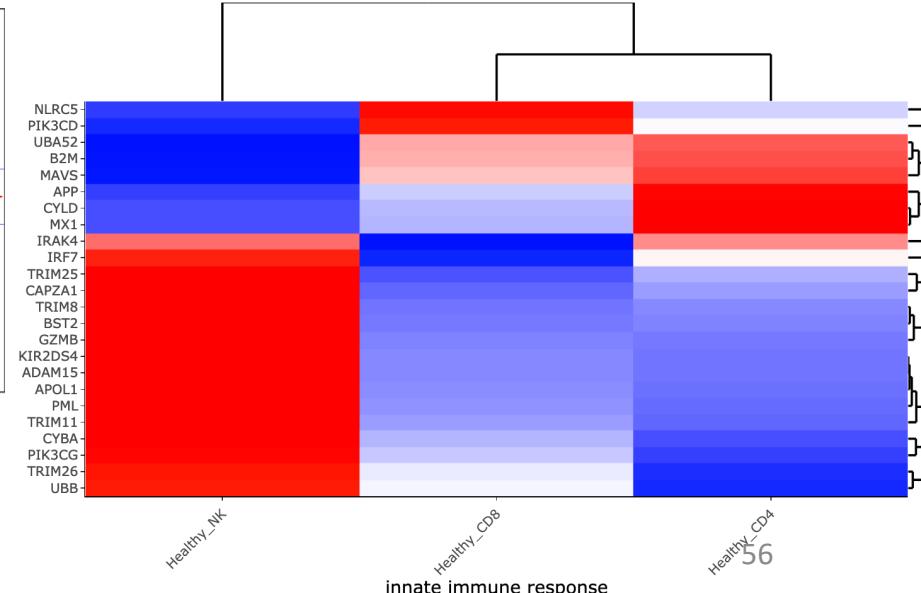
Select Display Style

Overlay    Beside

Label points?



Click the 'Beside' radio button and the 'Label points' check box then click 'Get GO Gene's Fold Change' button. Adjust the 'Graph Width' parameter.



# ROGUE – Differentially Expressed Gene Ontologies

**ROGUE**  
RNAseq + Ontology Graphic User Environment

[View ROGUE Manual](#) [?](#)

- ROGUE
- Load Count Data:(EdgeR)
- Load Expression Data
- Create Groups
- EdgeR Group Comparison
- Gene Comparison (Samples)
- Gene Comparison (Groups)
- Advanced Analysis
- Gene Set Enrichment Analysis
- Gene Ontology
- Group Statistical Comparison
- Group Expr. Ontologies
- Differentially Expressed Ontologs
- [Download Report](#)

**Differentially Expressed Ontologies Tool**

Analyze Groups or Samples?  
 Groups  Samples

Select Control  
Healthy\_CD4

Select Subjects  
Healthy\_NK

Select GO Class  
immune system process

Choose Keywords

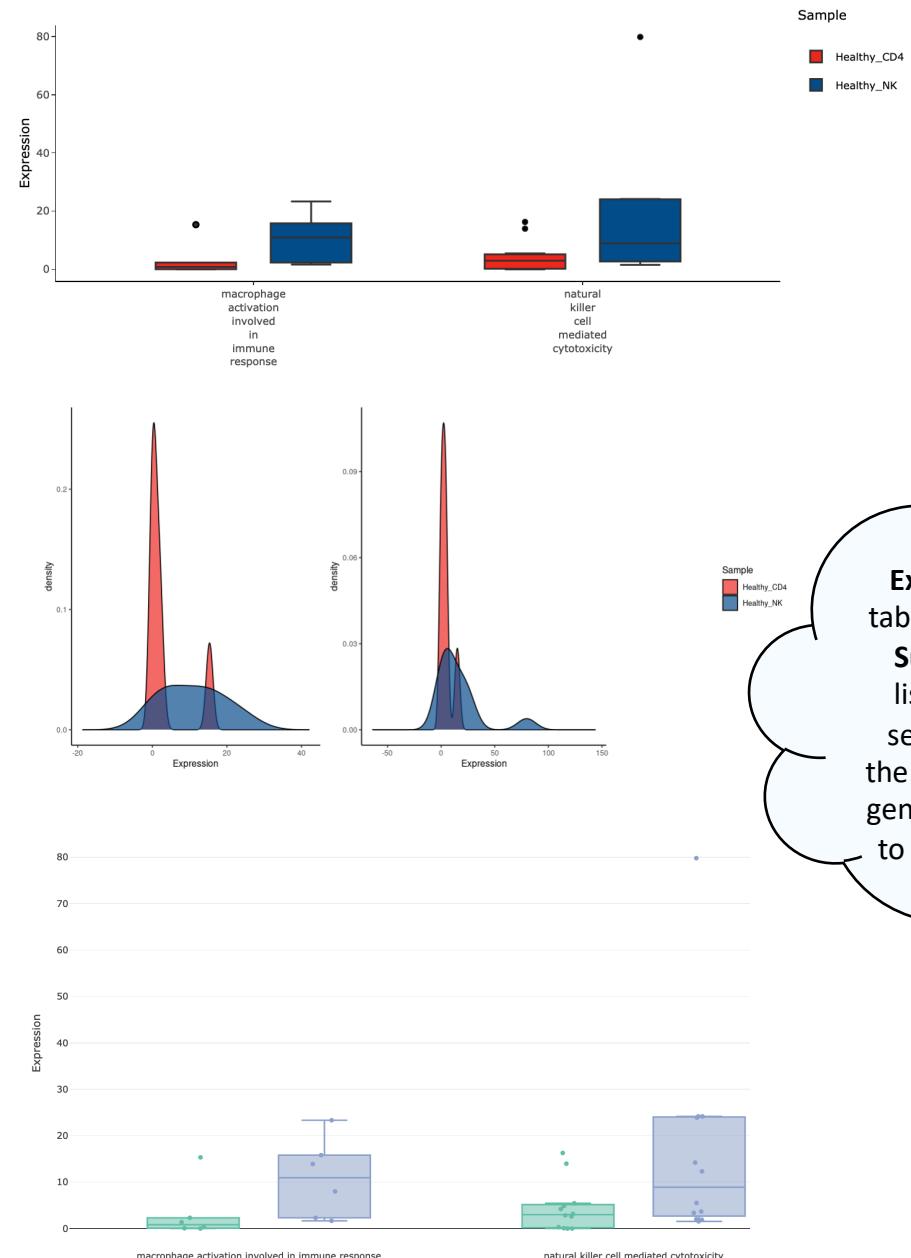
Find Differentially Expressed GOs

Label absFC Threshold  
0

Min RPKM base  
3

Min log2 Fold Change  
0.7

Max p-value  
0.1



Click on 'Differentially Expressed Gene Ontologies' tab. Select one **Control** and one **Subject** from the dropdown lists. Select GO Class and/or select keywords to include in the search. The tool will look for gene ontologies that are related to the GO class or that include the keywords.

# ROGUE – Differentially Expressed Gene Ontologies

Sample

- Healthy\_CD4
- Healthy\_NK

**ROGUE**  
RNAseq & Ontology Graphic User Environment

[View ROGUE Manual](#) [?](#)

[ROGUE](#)

[Load Count Data:\(EdgeR\)](#)

[Load Expression Data](#)

[Create Groups](#)

[EdgeR Group Comparison](#)

[Gene Comparison \(Samples\)](#)

[Gene Comparison \(Groups\)](#)

[Advanced Analysis](#)

[Gene Set Enrichment Analysis](#)

[Gene Ontology](#)

[Group Statistical Comparison](#)

[Group Expr. Ontologies](#)

[Differentially Expressed Ontolog](#)

**Differentially Expressed Ontologies Tool**

**Analyze Groups or Samples?**  
 Groups  Samples

**Select Control**  
Healthy\_CD4

**Select Subjects**  
Healthy\_NK

**Select GO Class**  
biological regulation

**Choose Keywords**

**Find Differentially Expressed GOs**

**Label absFC Threshold**

Perform analysis  
searching for  
differences in  
biological regulation

