



**Basic Tutorial**

# Home Module

ChromTag

Home Module

Peaks Visualization

Filter And Group

Differential Peak Detection

Gene Annotation

Differential Results Visualization

Enrichment Analysis

Motif Enrichment Analysis

Help

About

Overview

Home Module

Peaks Visualization

Enrichment Analysis

Grouping

Differential Analysis Visualization

Differential Peak Detection

Gene Annotation

Motif Enrichment

Peak Count Data

	S1	S2	S3	S4
P1	0	0	233	186
P2	0	2	166	89

Grouping

Group1Group2

S1S2S3S4

Differential Analysis Results

	log2FC	pvalue	padj
P1	13.76	5.3e-19	3.7e-17
P2	10.62	7.8e-19	5.2e-17

Annotation Table

	anno	TSS	geneID
P1	Exon	779943	114798
P2	Intron	283818	2567

Upset Plot

Coverage Plot

Profile Plot

GSEA Plot

Bar Plot

Volcano Plot

Heatmap

Motif Heatmaps

Import Data

Data source:

Example Data

Example Data

Custom Data

should specify the chromosome, start, and end positions, followed by count values for each sample. Each group must have at least two biological replicates. Peaks should be pre-merged across all samples for consistency. The provided dataset includes H3K27me3 and H3K4me3 count data, each with two biological replicates. Uploaded data can be previewed in the table below.

START

Instruction

ChromTag is a user-friendly, interactive, and open-source R-Shiny application for the analysis and visualization of CUT&Tag and ChIP-seq data. It streamlines data processing, from peak visualization and filtering to differential peak detection, gene annotation, and enrichment analysis. Users can upload a pre-merged peak count matrix and explore various analysis modules with just a few clicks. The application offers multiple visualization tools, including chromosome coverage plots, heatmaps, and volcano plots, making data interpretation more intuitive. Additionally, motif enrichment analysis helps identify potential regulatory elements linked to differentially enriched peaks. Chromtag provides a comprehensive and accessible solution for researchers studying chromatin modifications and transcriptional regulation.

Users can choose between "Example Data" or "Custom Data". If you select "Example Data", the application will automatically load a pre-merged peak count matrix provided by the system, which includes both H3K27me3 and H3K4me3 count data for two biological replicates per group.

Once the data source is selected, users can click the "Start" button to load the selected data and proceed to the next module. After the data is successfully loaded, the system will display the pre-processed data.

# Home Module

Import Data

Data source:

Custom Data

Choose Species

Human

Choose a File in CSV Format

Browse...

No file selected

Users can choose between "Example Data" or "Custom Data". If you select "Example Data", the application will automatically load a pre-merged peak count matrix provided by the system, which includes both H3K27me3 and H3K4me3 count data for two biological replicates per group.

1	Chromosome	Start	End	H3K27me3_rep1	H3K27me3_rep2	H3K4me3_rep1	H3K4me3_rep2
2	chr1	27805	30656	0	0	233	186
3	chr1	135727	140083	0	2	166	89
4	chr1	198335	200968	0	0	219	169
5	chr1	391014	392146	1	3	0	1
6	chr1	392617	393317	2	1	0	0
7	chr1	393808	394451	2	1	0	0
8	chr1	440860	441530	4	2	0	1
9	chr1	442100	442830	2	0	0	0
10	chr1	492154	494761	0	1	73	36

If you choose to upload custom data, you must also select the corresponding species. The uploaded custom data should be pre-merged across samples and must follow a specific format: the data should be in CSV format with a header row. The first column should list the chromosome (starting with "chr"), the second and third columns should contain the start and end positions, and the subsequent columns should represent the count values for each sample. Additionally, each group must have at least two biological replicates for proper analysis.

# Peaks Visualization

This step is optional and can be skipped.

## Chromosome Coverage Plot

Select Sample:

H3K27me3\_rep1 × H3K27me3\_rep2 ×

Select Chromosome:

chr1

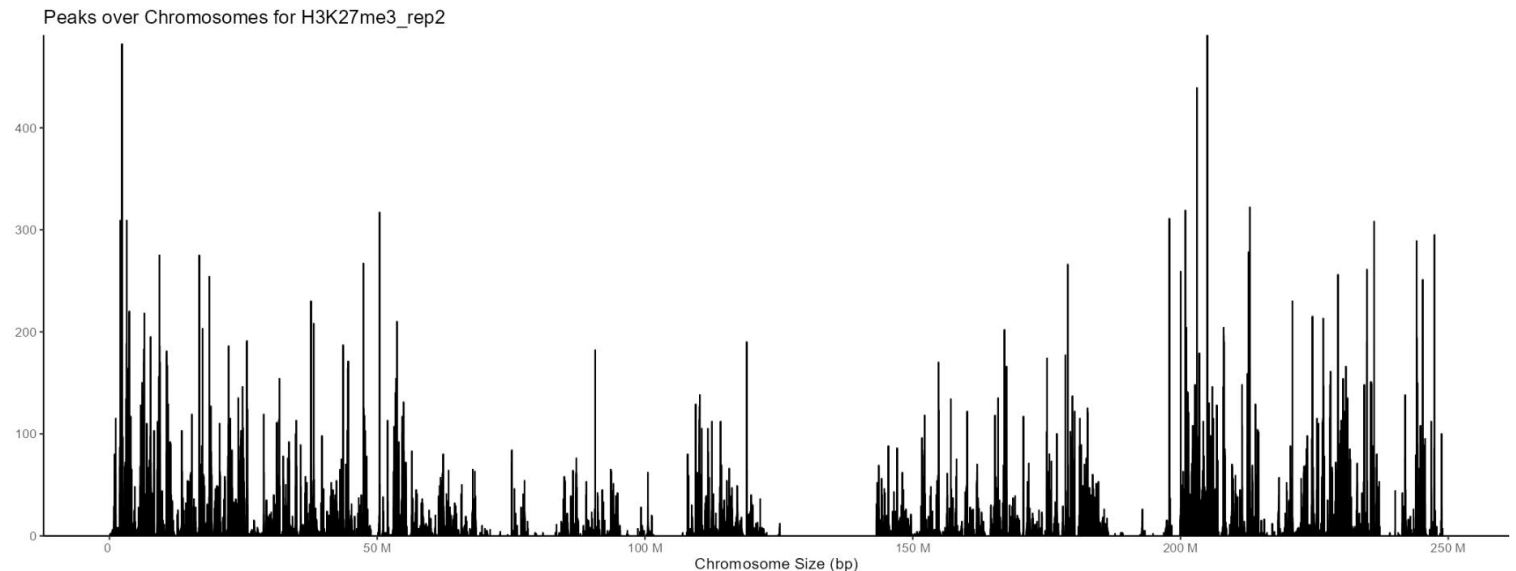
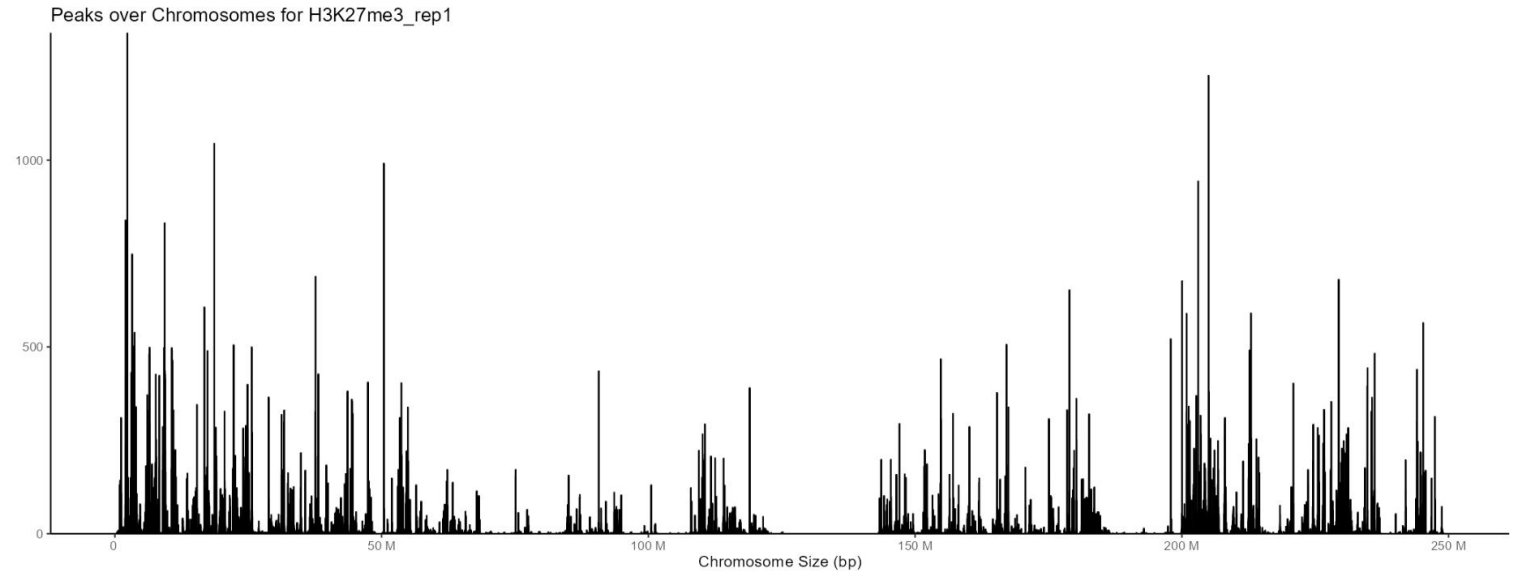
✓ Submit

Output Format

☒ PNG ☐ PDF ☐ JPEG

Download

Users can select one or more samples of a specific chromosome to view simultaneously. After selecting the desired samples and chromosome, click the "Submit" button and wait for the image to generate.



# Peaks Visualization

This step is optional and can be skipped.

## Profile Plot

Select Sample:

H3K27me3\_rep1 × H3K4me3\_rep2 ×

Upstream (bp):

3000

Downstream (bp):

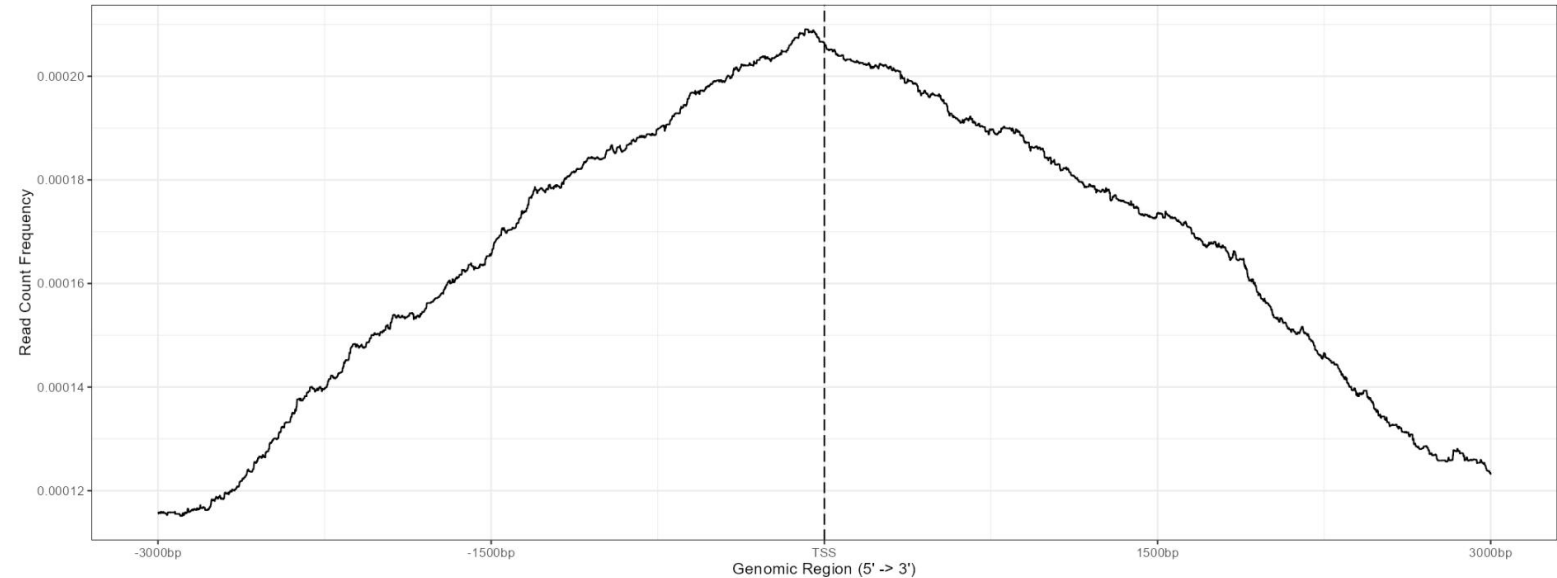
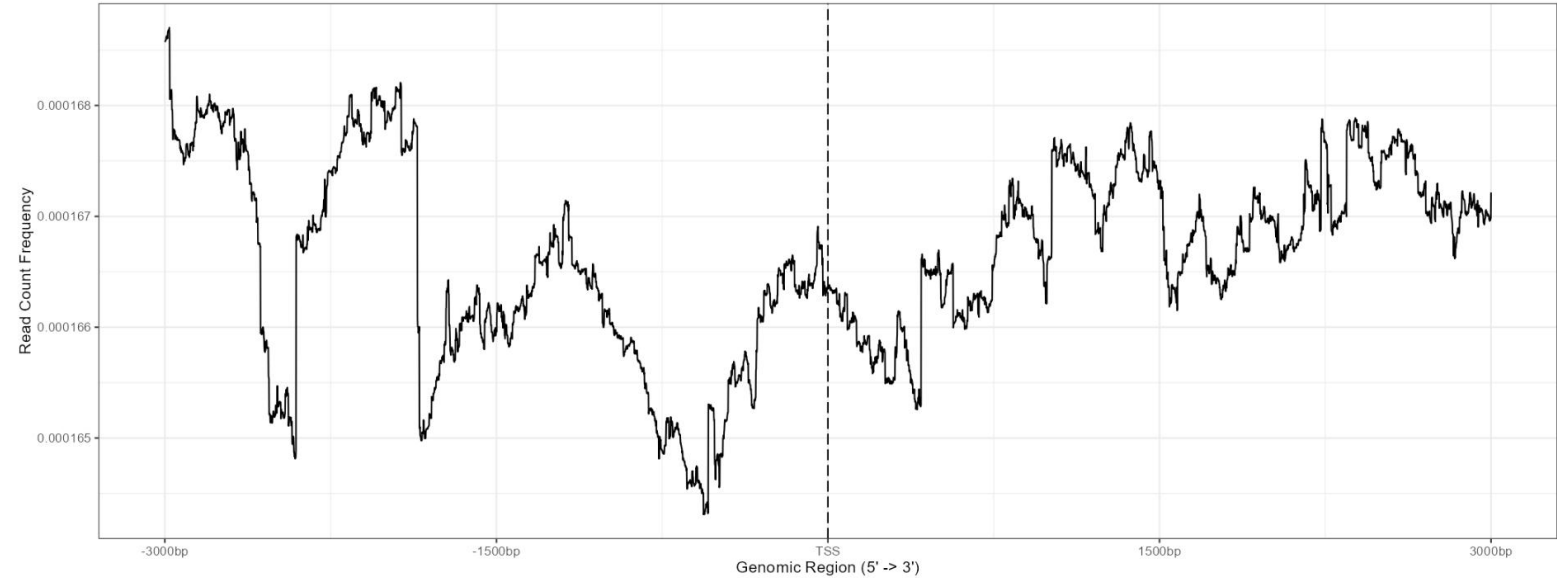
3000

✓ Submit

Output Format

☒ PNG ☐ PDF ☐ JPEG

Download



Users can select one or more samples and customize the genomic region for observation. After making the selections, click the "Submit" button and wait for the image to be generated.

# Filter and Group

Before Filter

After Filter

Count Threshold: ?

5

✓ Submit

The user can input a threshold value to filter the samples. The filtering method sums the counts for each row of the samples and checks if the total exceeds the threshold value entered by the user.

Show 5 entries

Search:

hromosome	Start	End	H3K27me3_rep1	H3K27me3_rep2	H3K4me3_rep1	H3K4me3_rep2
1r1	27805	30656	0	0	233	186
1r1	135727	140083	0	2	166	89
1r1	198335	200968	0	0	219	169
1r1	391014	392146	1	3	0	1
1r1	392617	393317	2	1	0	0

Showing 1 to 5 of 220,230 entries

Previous

1

2

3

4

5

...

44,046

Next

> 5

< 5



# Filter and Group

## ↔ Sample Grouping

Select Samples:

H3K27me3\_rep1 × H3K27me3\_rep2 ×

Select the samples that you want to group together.

Group Name:

H3K27me3

Enter a name for the current group.

Add Group Clear Groups

Click the "Add Group" button to add the group.  
Click the "Clear Groups" button to remove all current groups.

## Current Groups:

\$H3K27me3  
[1] "H3K27me3\_rep1" "H3K27me3\_rep2"

The groups you have created can be viewed under "Current Groups."

✓ Submit

Once all groups are created, click the "Submit" button to save the group information.

# Differential Peak Detection

▶ Select Analysis Results Parameters

Choose Alternative Hypothesis for Differential Expression:

Detect Significant Changes (up or down) ▼

Apply Independent Filtering for Low Expression Genes

☐ Yes ☒ No

Set Significance Level:

0.01

Set Log2 Fold Change Threshold:

0

Select p-value Adjustment Method:

BH ▼

✓ Submit

## Alternative Hypothesis

This parameter allows users to select the type of changes to detect, either "upregulated" or "downregulated" peaks. It helps focus the analysis on the specific direction of change.

## Log2 Fold Change Threshold

This threshold sets the minimum difference between groups for a peak to be considered significant. A log2 fold change of 1 represents a 2-fold difference between groups.

## Significance Level

The significance level defines the threshold for statistical significance. For example, a p-value threshold of 0.01 means peaks with p-values lower than 0.01 are considered significant.

## p-value Adjustment Method

Choose a method to adjust p-values for multiple testing. The Benjamini-Hochberg (BH) method is commonly used to control the false discovery rate (FDR) in genomics.

## p-value Adjustment Method

This option lets you choose a method for adjusting p-values to handle multiple comparisons. The Benjamini-Hochberg (BH) method controls the false discovery rate (FDR) and is more powerful with many tests. The Holm method controls the family-wise error rate (FWER) by adjusting p-values stepwise, reducing false positives but potentially increasing the risk of missing true positives.



# Differential Peak Detection

## Comparison between two groups

Before Filter

After Filter

P-value Threshold:

0.05

Log2 Fold Change Threshold:

2

✓ Submit

P-value and Log2 Fold Change are used as criteria to filter the results.

## Comparison between multiple groups

Analysis Results

1 vs 2

1 vs 3

2 vs 3

Contains multiple analysis results

out of 142959 with nonzero total read count  
adjusted p-value < 0.1  
LFC > 0 (up) : 134, 0.094%  
LFC < 0 (down) : 9291, 6.5%  
outliers [1] : 0, 0%  
low counts [2] : 0, 0%

Selected Comparison:

1 vs 2

▼

Select one group for subsequent analysis

P-value Threshold:

0.05

Log2 Fold Change Threshold:

2

✓ Submit

# Gene Annotation

## Input Annotation Parameters

Upstream (bp):

Downstream (bp):

▶ Run Gene Annotation

Customize the annotated genomic region and click the 'Run Gene Annotation' button to run it.

# Differential Results Visualization

-log10(p) Cutoff

2

Positive Log2 FC Threshold

14

Negative Log2 FC Threshold

5

Point Size

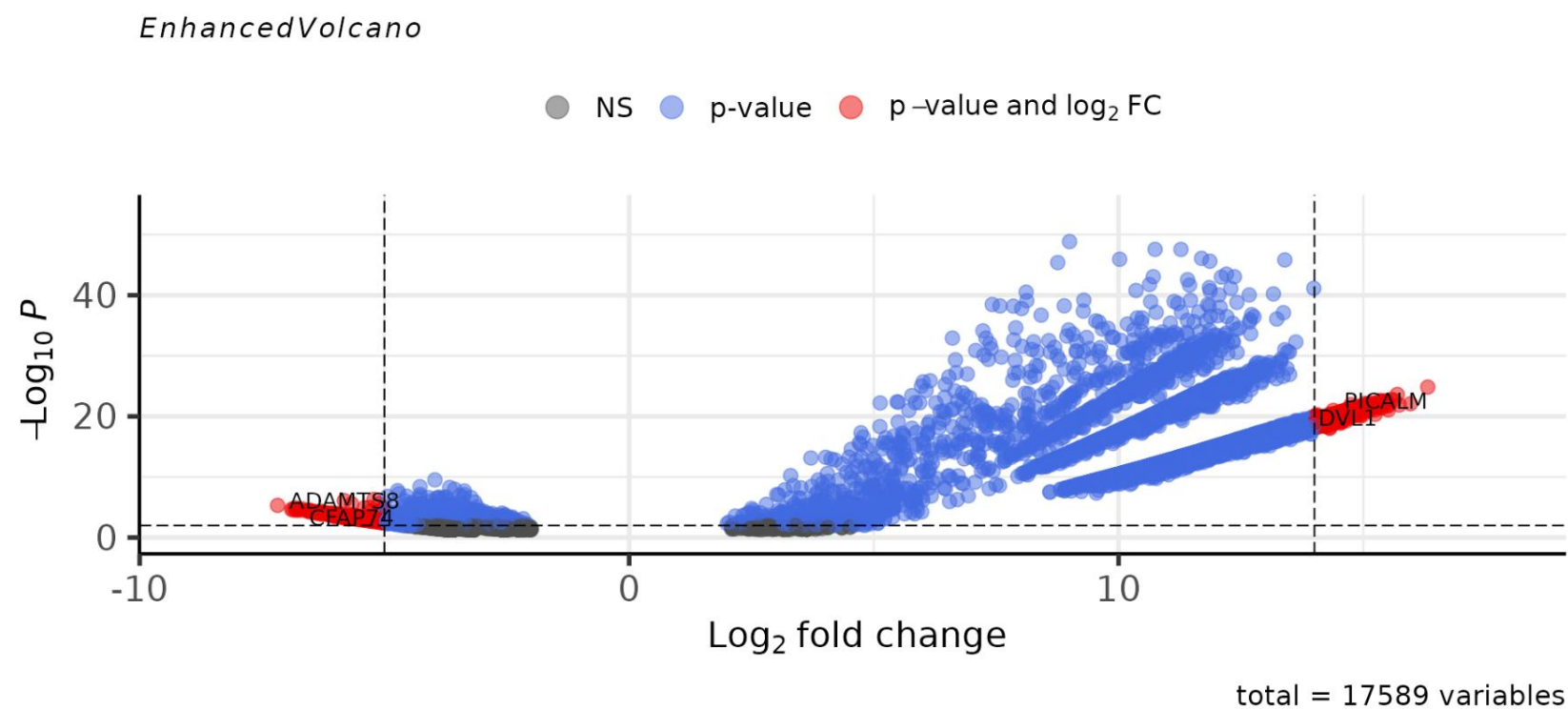
3

Label Size

4

✓ Submit

Significant peaks are selected by adjusting these three thresholds.



Click the submit button to refresh the image and submit the selected significant peaks.

# Differential Results Visualization

Up Genes Preview

Upregulated Genes List

DVL1

MRPL20

NOL9

DNAJC11

DHRS3

RPL11

ELOA-AS1

SRRM1

RSRP1

MACO1

Show 5 entries

Search:

	seqnames	start	end	width	strand	H3K27me3_rep1	H3K27me3_rep2	H3K4
20	chr1	1346468	1350600	4133	*	0	0	
23	chr1	1405327	1408434	3108	*	0	0	
122	chr1	6552669	6555289	2621	*	0	0	
127	chr1	6699924	6702352	2429	*	0	0	
248	chr1	12614231	12620067	5837	*	0	0	

Showing 1 to 5 of 497 entries

Download

This section allows you to view the up/down gene list or peak table obtained from the filtering process.

# Enrichment Analysis

Analysis Type:

GO

GO Ontology:

All

p-value Cutoff:

0.05

q-value Cutoff:

0.05

Upregulated Genes List

DVL1  
MRPL20  
NOL9  
DNAJC11  
DHRS3  
RPL11  
ELOA-AS1  
SRRM1  
RSRP1  
MACO1

Downregulated Genes List

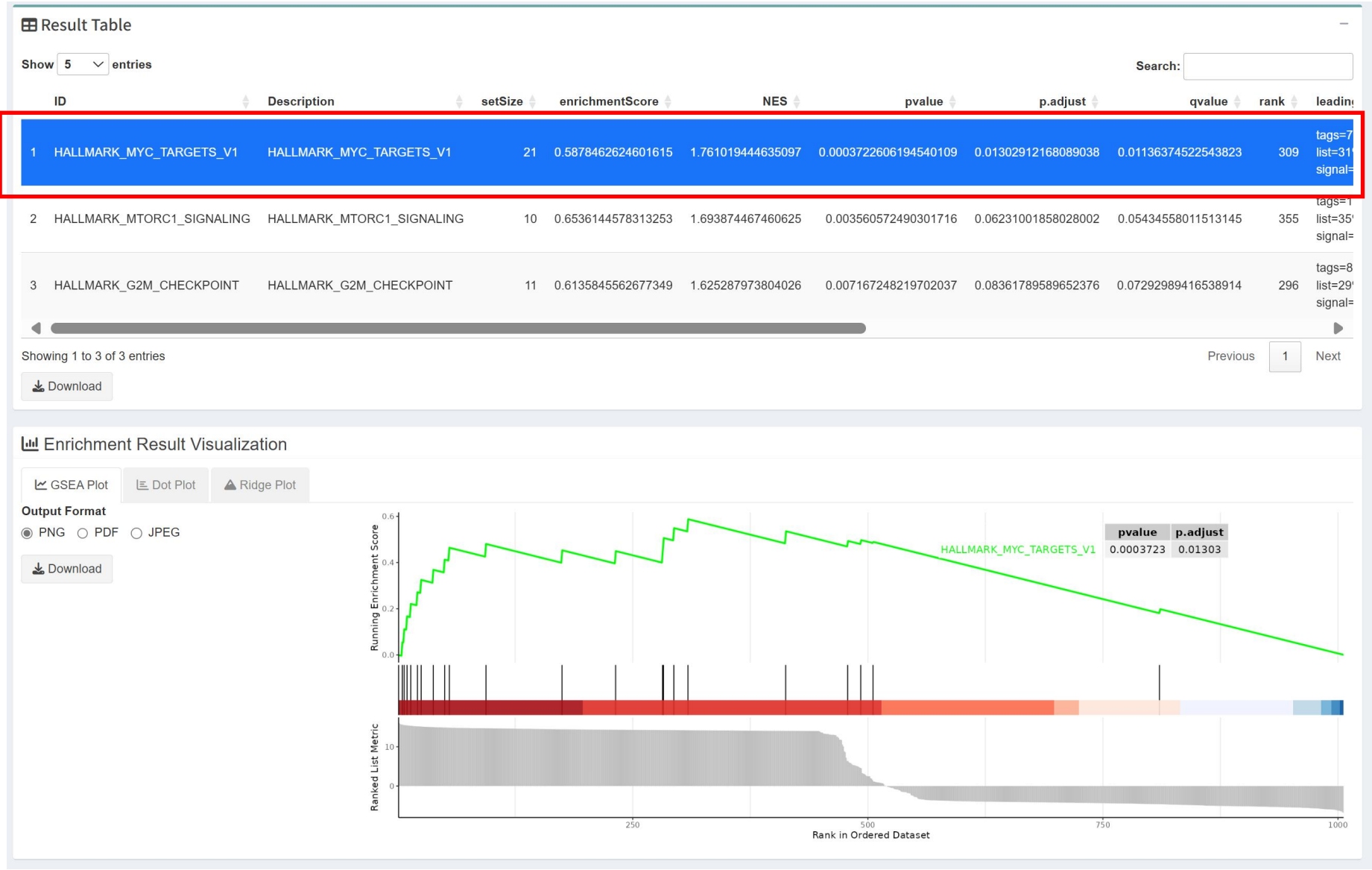
CFAP74  
MEGF6  
CAMTA1-AS3  
CAMTA1  
SLC45A1  
CA6  
SPSB1  
LINC02606  
PDPN  
KAZN

▶ Run Enrichment Analysis

Adjust the types and parameters you need to analyze

For GO or KEGG analysis, you can enter or modify the gene list using SYMBOLs to specify the genes for enrichment analysis. If GSEA is selected, the input is restricted to all genes annotated in the Gene Annotation module.

# Enrichment Analysis





# Motif Enrichment Analysis

## Upregulated Peaks List

chr3-150407290-150415877	▲
chr16-2150741-2157314	●
chr16-88972494-88978084	
chr6-119343897-119351533	
chr12-56682168-56690455	
chr8-38173859-38178882	
chr9-128688198-128692391	
chr11-65420641-65428391	
chr10-92688412-92694886	
chr6-27889944-27897392	▼
chr11-22222222-22222222	✎

## Downregulated Peaks List

chr7-154746354-154754198	▲
chr11-20155235-20165969	●
chr7-27100472-27112603	
chr1-208072367-208087339	
chr1-8323134-8331215	
chr11-118150179-118161808	
chr20-47665285-47675999	
chr14-64734713-64750256	
chr1-28810214-28820062	
chr4-182814697-182830407	▼
chr12-118887818-11889817	✎

This section displays the top 200 upregulated and 200 downregulated peaks, which are selected from the differential peaks identified through the volcano plot filtering step. You can also input and modify your own lists of upregulated and downregulated peaks for further analysis.

## Number of Top Transcription Factors to Plot

10

## Display Motif GC Content

- ☒ Yes
- ☐ No

## Enable Clustering

- ☒ Yes
- ☐ No

- Number of Top Transcription Factors to Plot:** Set how many top transcription factors you want to visualize (default is 10).
- Display Motif GC Content:** Choose whether to show the GC content of motifs.
- Enable Clustering:** Decide whether to enable clustering for grouping similar motifs in the plot.