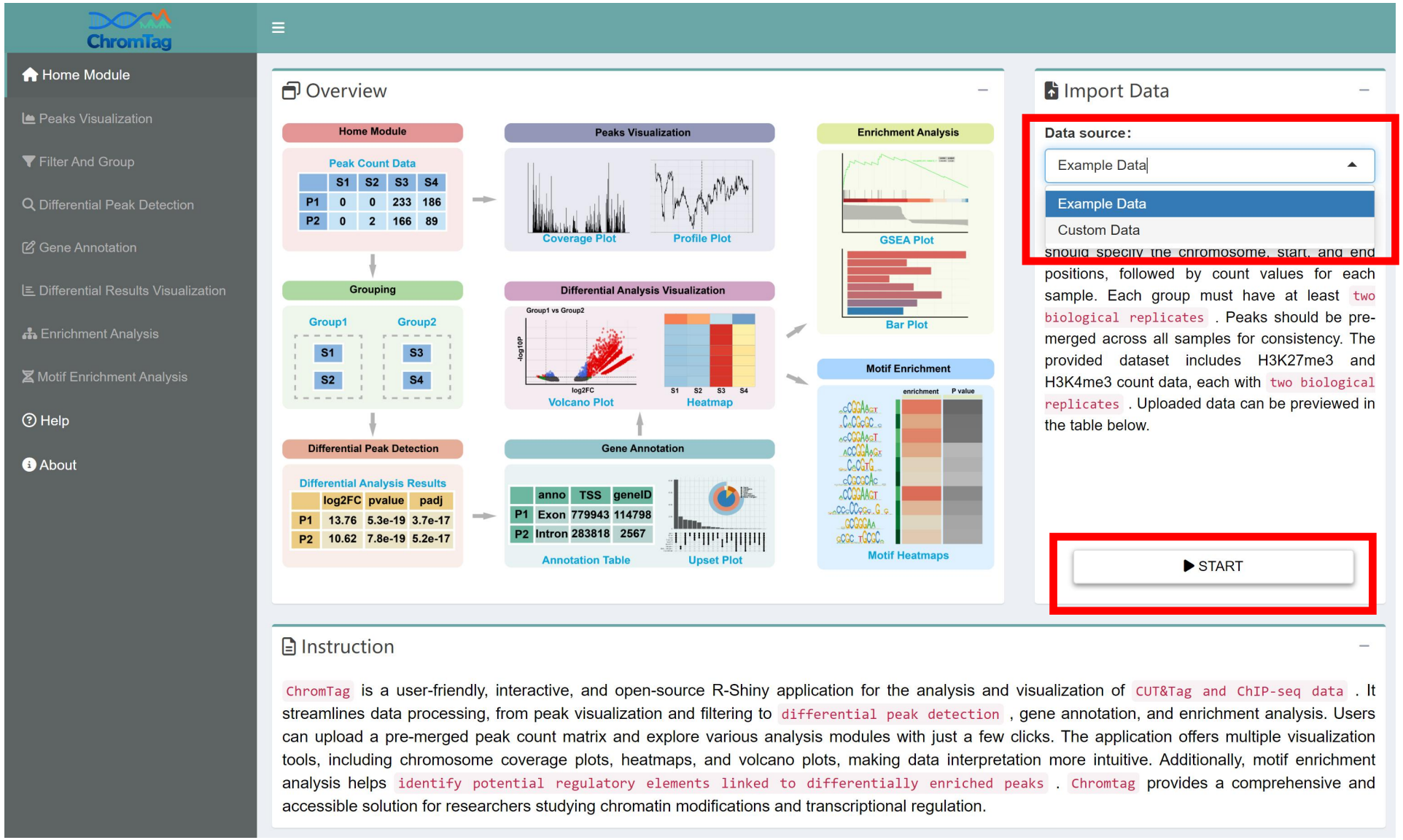
ChromTag

Basic Tutorial

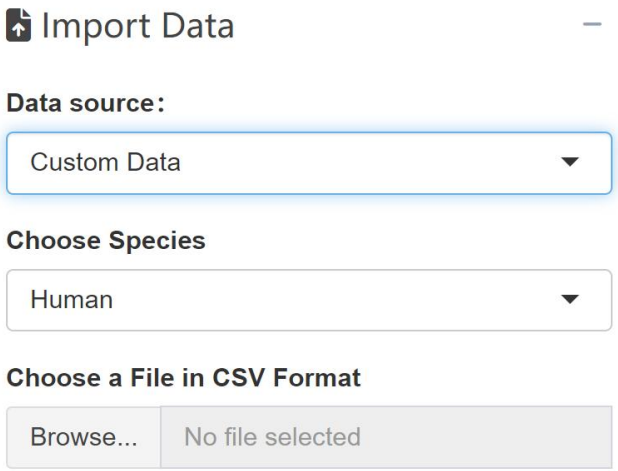
Home Module



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



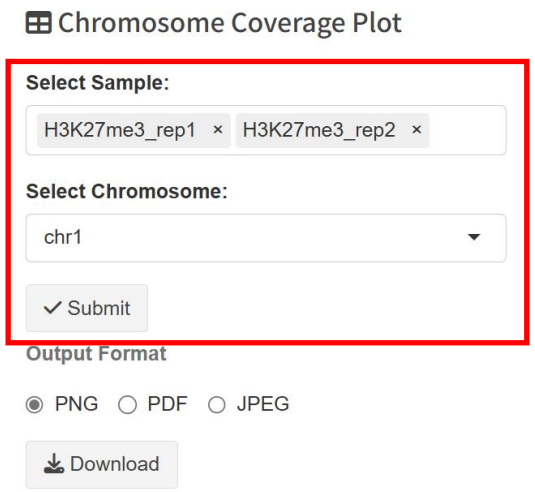
Users can choose between "Example Data" or "Custom Data". If you select "Example Data", the application will automatically load a pre-merged peak count matrix provided by the system, which includes both H3K27me3 and H3K4me3 count data for two biological replicates per group.

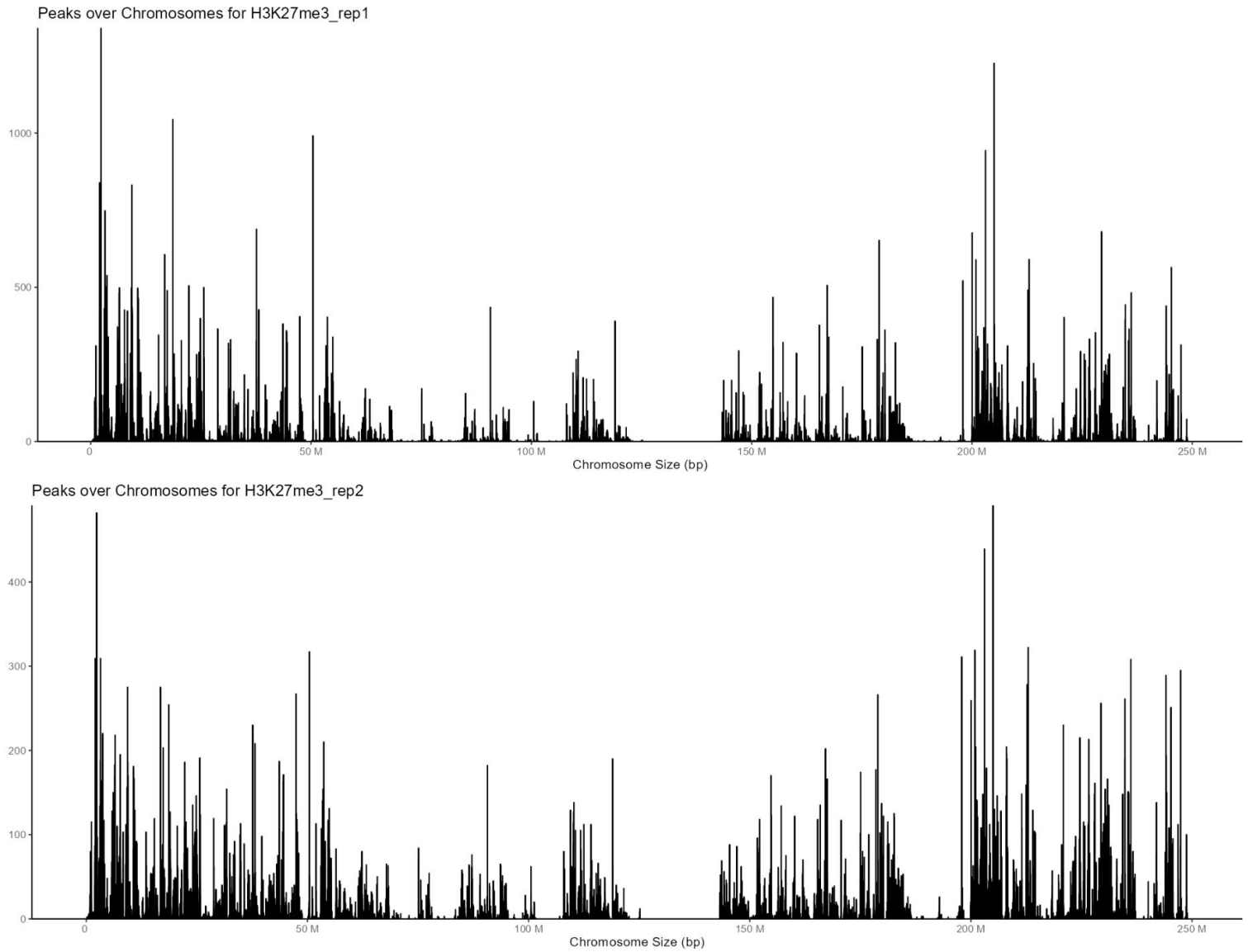
	Chromosome	Start	End	H3K27me3_rep1	H3K27me3_rep2	H3K4me3_rep1	H3K4me3_rep2
2	chr1	27805	30656	0	0	233	186
3	chr1	135727	140083	0	2	166	89
4	chr1	198335	200968	0	0	219	169
5	chr1	391014	392146	1	3	0	1
6	chr1	392617	393317	2	1	0	0
7	chr1	393808	394451	2	1	0	0
8	chr1	440860	441530	4	2	0	1
9	chr1	442100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

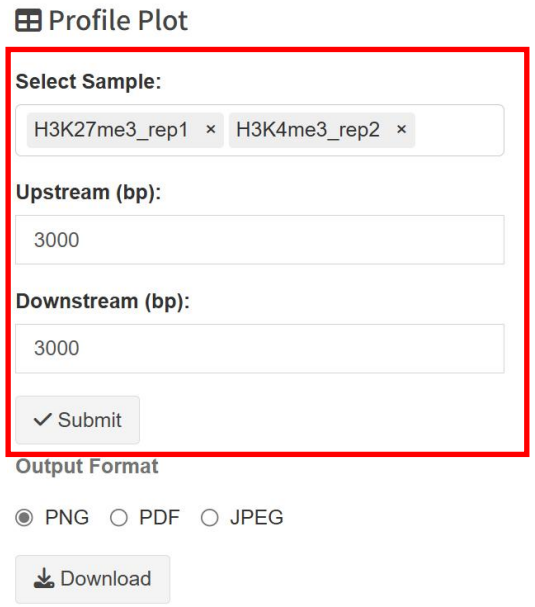


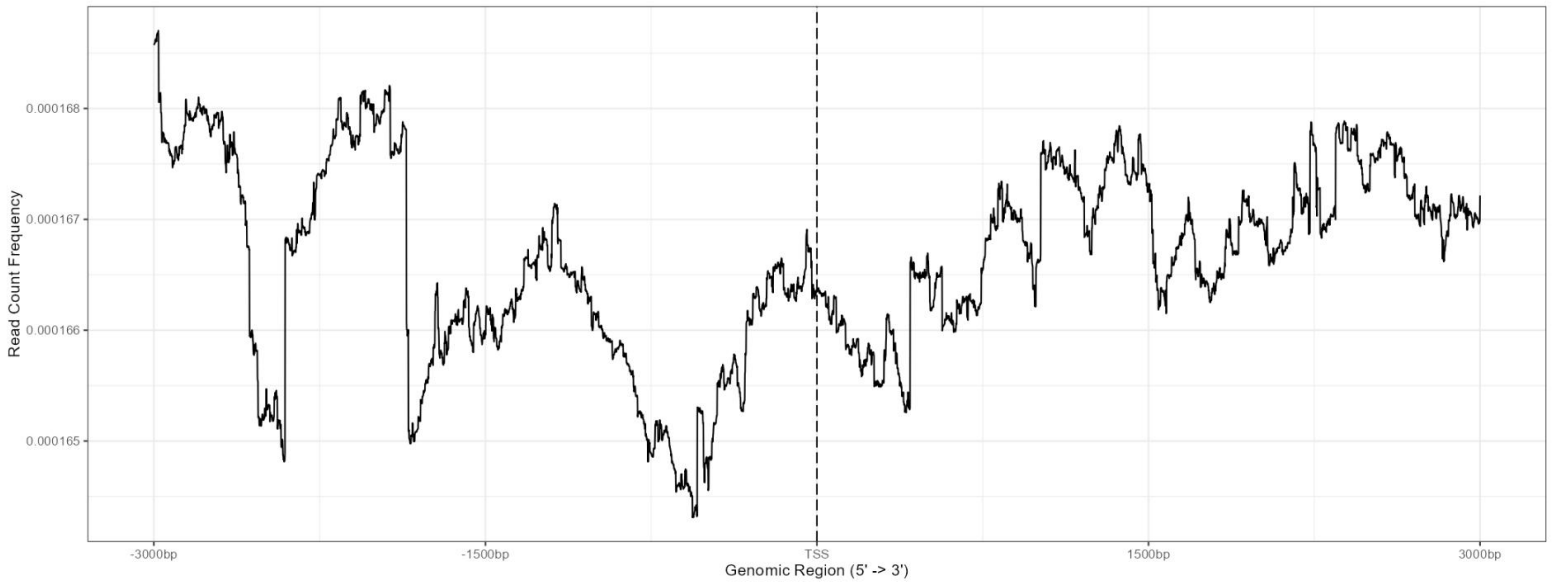


Users can select one or more samples of a specific chromosome to view simultaneously. After selecting the desired samples and chromosome, click the "Submit" button and wait for the image to generate.

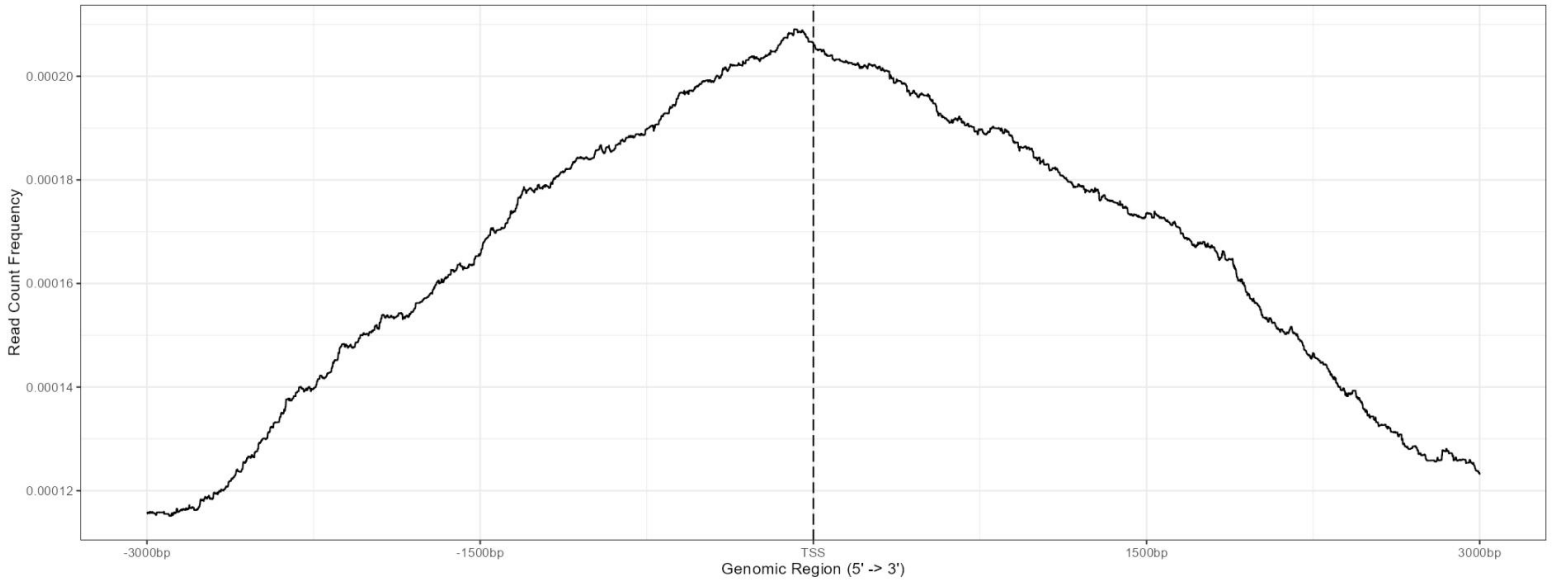
Peaks Visualization

This step is optional and can be skipped.

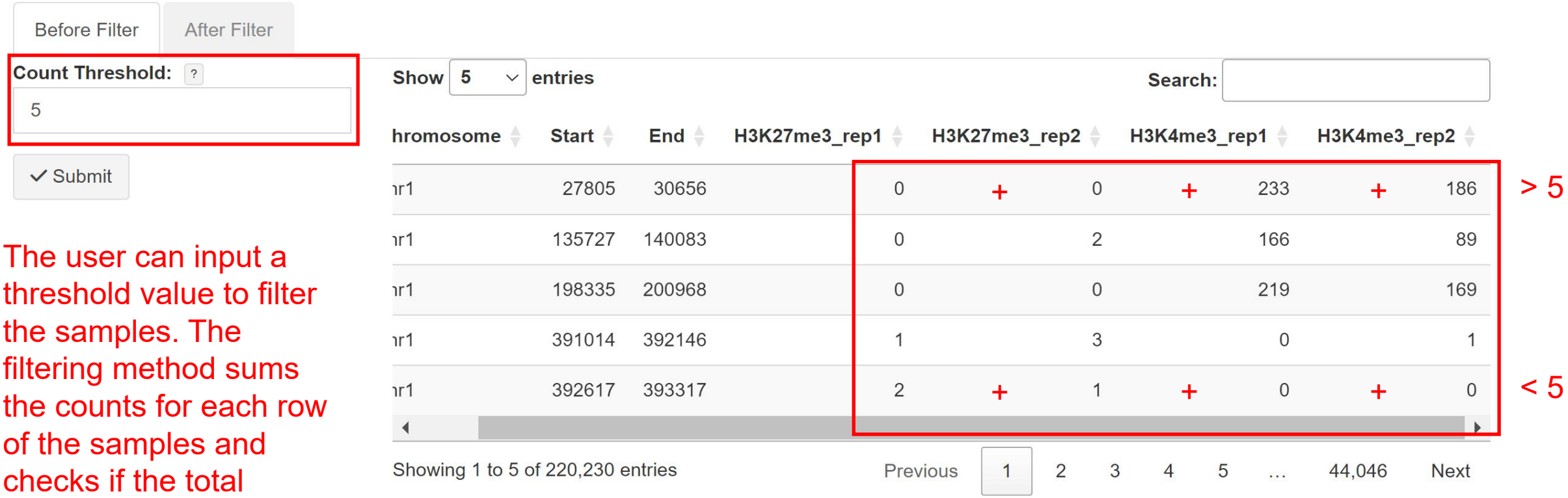




Users can select one or more samples and customize the genomic region for observation. After making the selections, click the "Submit" button and wait for the image to be generated.

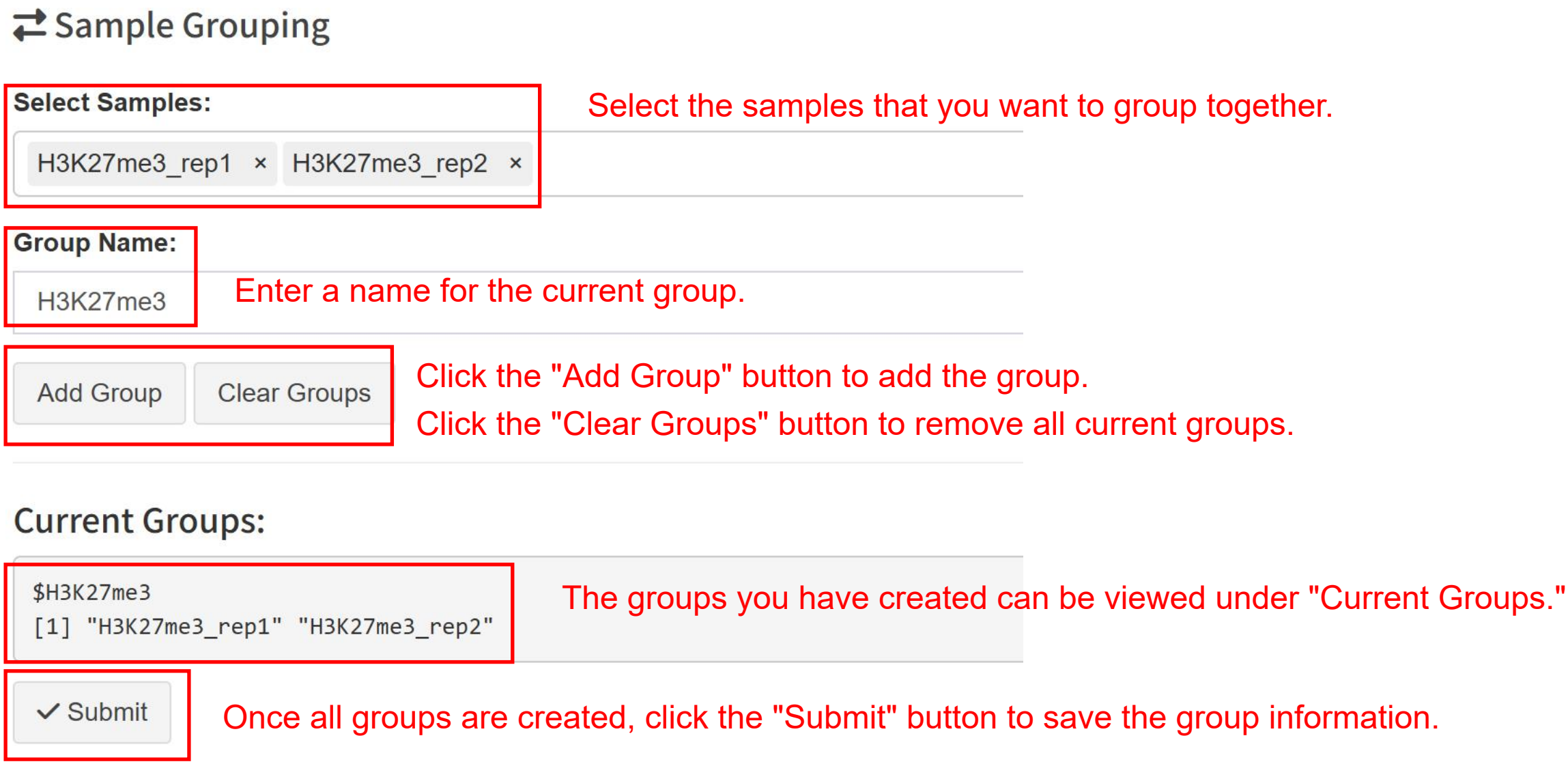


Filter and Group

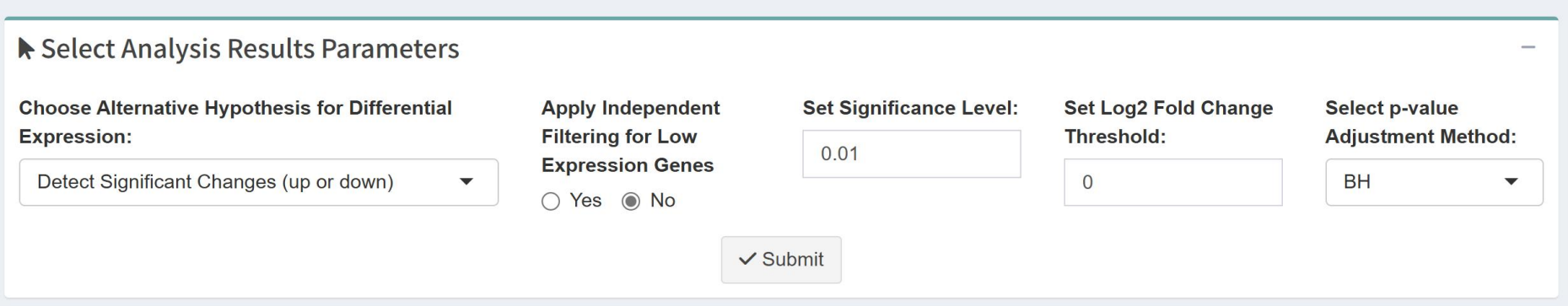


The user can input a threshold value to filter the samples. The filtering method sums the counts for each row of the samples and checks if the total exceeds the threshold value entered by the user.

Filter and Group



Differential Peak Detection



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.

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

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.

Log2 Fold Change Threshold

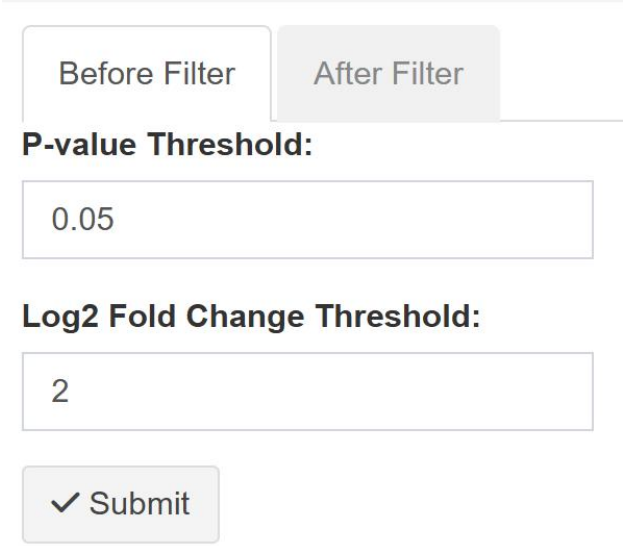
This threshold sets the minimum difference between groups for a peak to be considered significant. A log2 fold change of 1 represents a 2-fold difference between groups.

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.

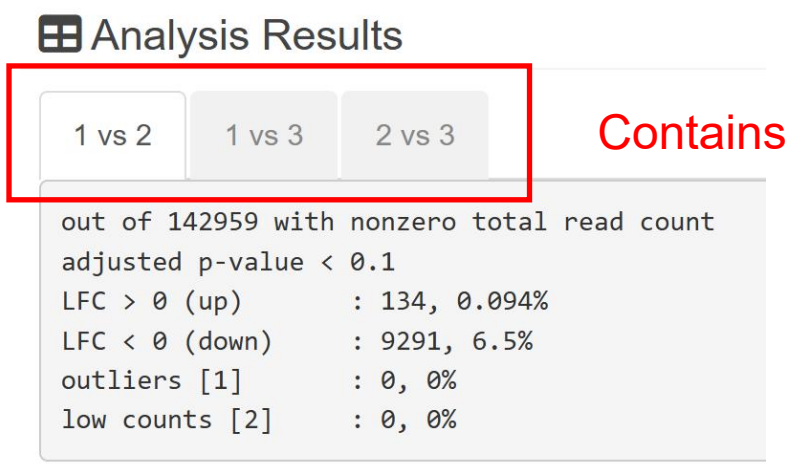
Differential Peak Detection

·Comparison between two groups

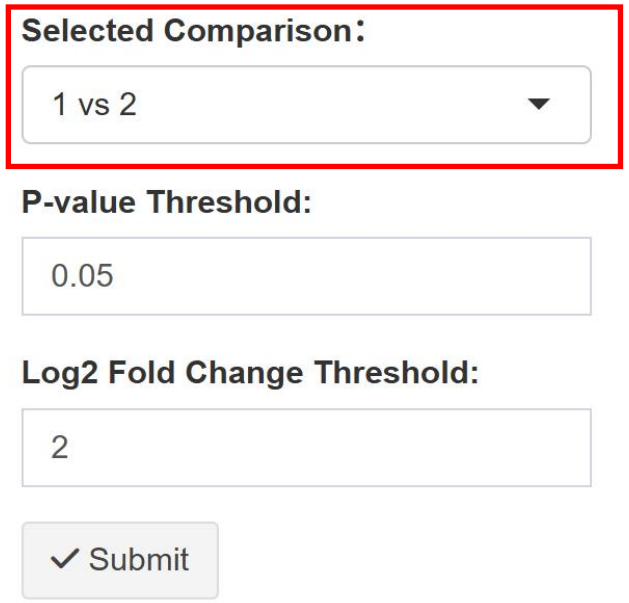


P-value and Log2 Fold Change are used as criteria to filter the results.

·Comparison between multiple groups

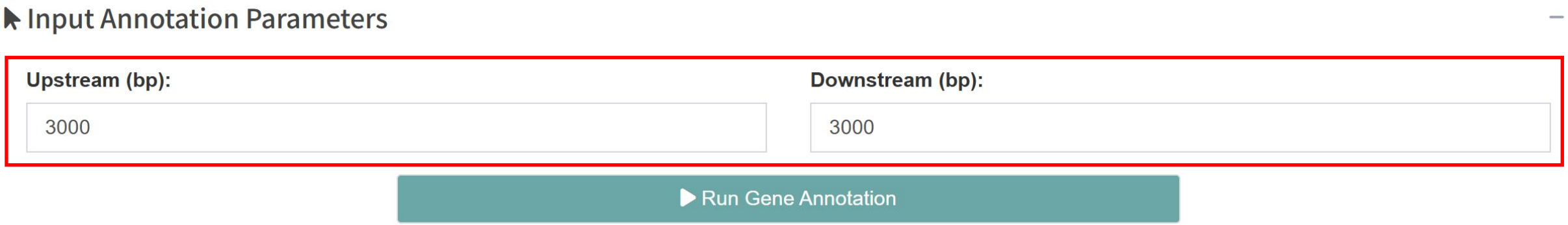


Contains multiple analysis results



Select one group for subsequent analysis

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

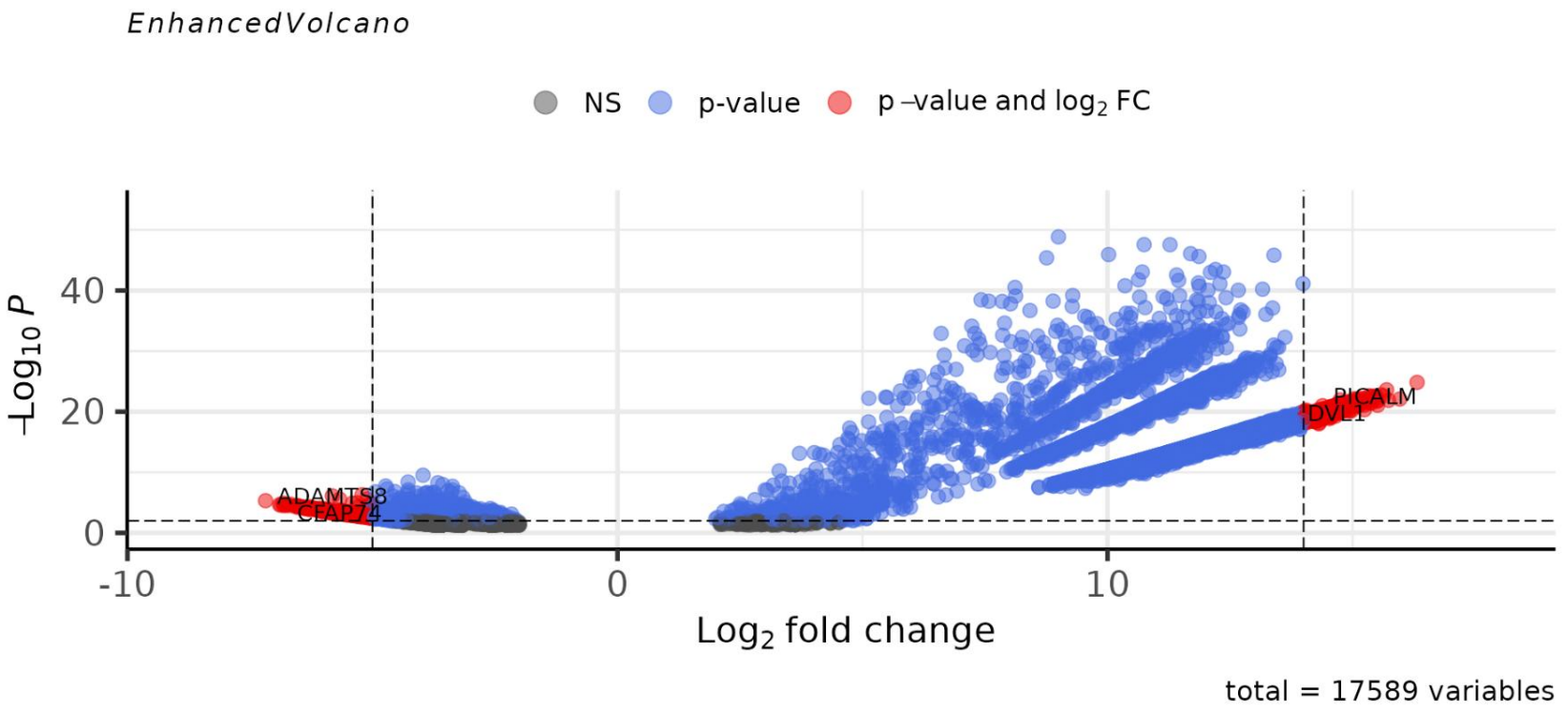
3

Label Size

4

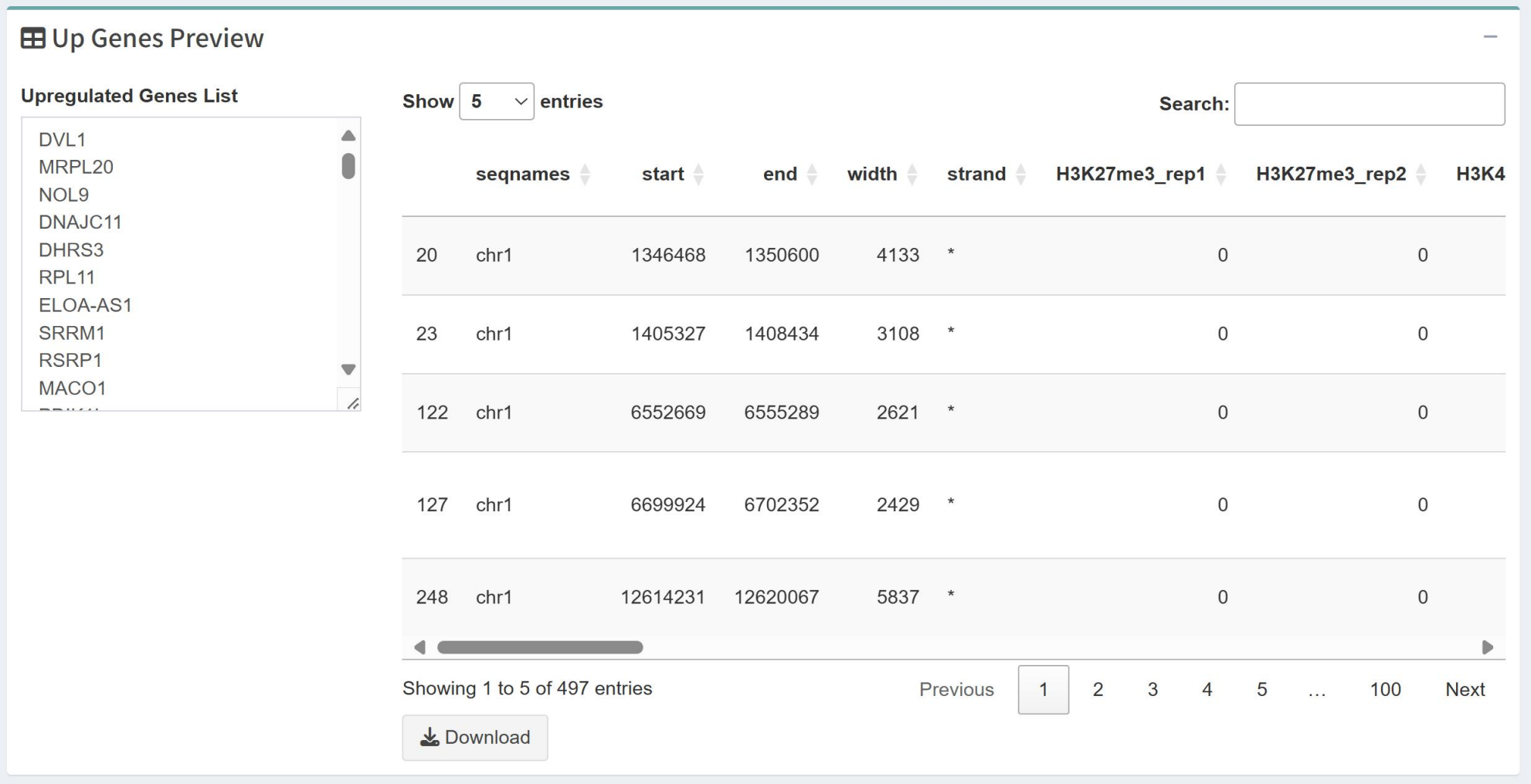
✓ Submit

Significant peaks are selected by adjusting these three thresholds.



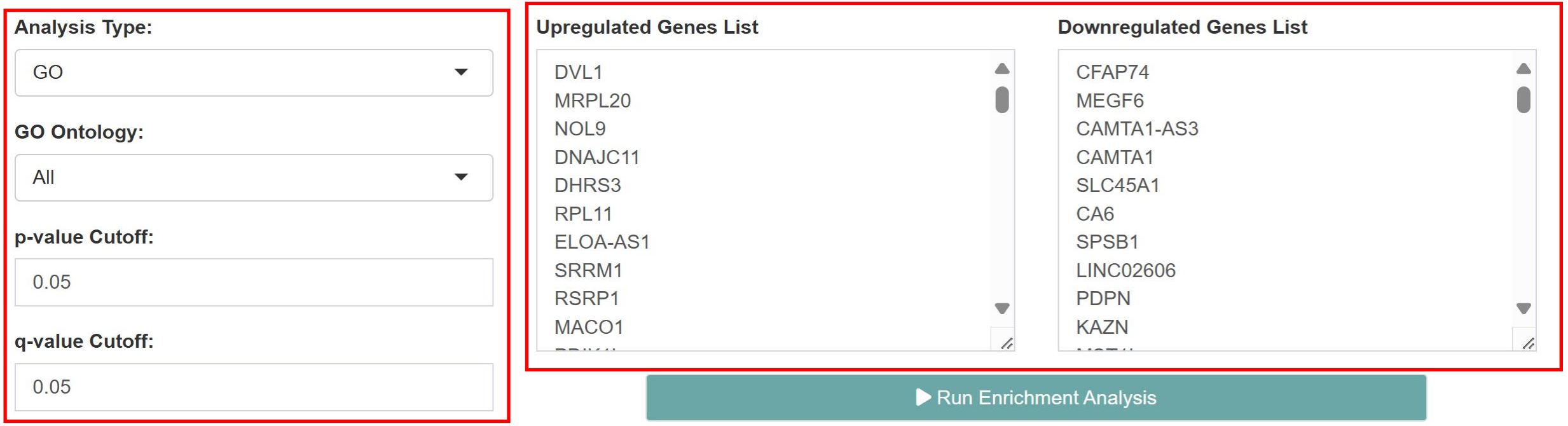
Click the submit button to refresh the image and submit the selected significant peaks.

Differential Results Visualization



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

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

Result Table

Show 5 entries

Search:

	ID	Description	setSize	enrichmentScore	NES	pvalue	p.adjust	qvalue	rank	leadin
1	HALLMARK_MYC_TARGETS_V1	HALLMARK_MYC_TARGETS_V1	21	0.5878462624601615	1.761019444635097	0.0003722606194540109	0.01302912168089038	0.01136374522543823	309	tags=7 list=31 signal=
2	HALLMARK_MTORC1_SIGNALING	HALLMARK_MTORC1_SIGNALING	10	0.6536144578313253	1.693874467460625	0.003560572490301716	0.06231001858028002	0.05434558011513145	355	tags=1 list=35 signal=
3	HALLMARK_G2M_CHECKPOINT	HALLMARK_G2M_CHECKPOINT	11	0.6135845562677349	1.625287973804026	0.007167248219702037	0.08361789589652376	0.07292989416538914	296	tags=8 list=29 signal=

Showing 1 to 3 of 3 entries

Previous 1 Next

Download

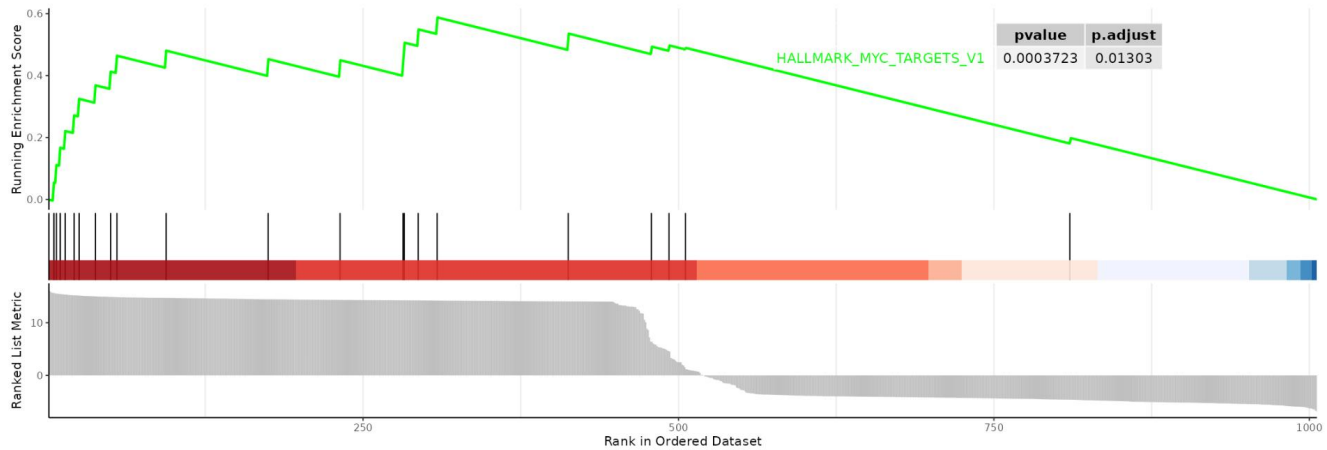
Enrichment Result Visualization

GSEA Plot | Dot Plot | Ridge Plot

Output Format

PNG PDF JPEG

Download



If GSEA is selected, you need to choose an entry you wish to plot before viewing the GSEA plot.

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

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

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.