

# Step-by-step Guide to Perform Differential Expression Analysis using FlexStat Pipeline

## - Upload experimental results

Scribe 



This feature facilitates pairwise differential expression analysis with integrated multiple-testing corrections. Users have the option to filter results by configuring cutoffs for log fold change and p-values.

This functionality includes visual representations of differential expression results, including boxplots, volcano plots, and heat maps.

Protein type-specific principal component analysis is a prominent aspect of this feature.

This tutorial is based on uploading an experimentally generated protein expression profile into the application.

1

Navigate to <https://jglab.shinyapps.io/flexstatv1-pipeline-only/>

2

Go to "Differential Expression" tab.

- 3** Prepare data having rows as the samples and protein names in the columns and having a column for experimental condition/class as shown in the right-side panel.

## Sample Data

Condition	O76070	P01344	P01579	P00709	P41159	P00918	P01112	Q15843	P10636
A	28.41	27.36	27.40	27.14	28.23	28.04	26.95	25.87	29.26
A	28.46	27.40	27.37	27.05	28.14	28.07	27.15	25.68	29.32
A	28.41	27.47	27.37	27.08	28.23	28.03	27.15	25.49	29.24
B	24.28	24.63	23.63	22.84	24.47	24.17	23.46	22.01	25.93
B	24.28	24.73	23.44	23.03	24.72	24.52	23.76	21.78	25.56
B	24.20	24.66	23.68	22.76	24.66	24.47	23.70	21.95	25.55

- 4** Click "Browse..." and upload your file

## Limma Analysis

Select CSV File to Import i

**Browse...**

ccRCC\_patients\_norm.csv

Show head

Upload complete

Use Sample Data

Data

Results

## Original Data

protein

1	kidneyTisue1
2	kidneyTisue2
3	kidneyTisue3
4	kidneyTisue4
5	kidneyTisue5
6	kidneyTisue6

- 5 Select columns need to be removed.

## Limma Analysis

Select CSV File to Import 

ccRCC\_patients\_norm.csv

Show head

Upload complete

Use Sample Data

Data

Results

Top 50

### Select columns to remove

patient\_id  histological\_type

V1

protein

class

P09110

P05166

Not Selected

### Contrast variable

Not Selected

Contrast other classes 

## Original Data

	protein	class
1	kidneyTisue1	Normal
2	kidneyTisue2	Tumor
3	kidneyTisue3	Normal
4	kidneyTisue4	Tumor
5	kidneyTisue5	Normal
6	kidneyTisue6	Tumor

- 6 Select Class variable from the dropdown

Transpose data

Log2 Transform

Log10 Transform

### Select columns to remove

protein  patient\_id  histological\_type

### Class Variable

Not Selected

Not Selected

protein

class

patient\_id

histological\_type

Not Selected

Contrast other classes 

Log fold-change variable 

P-value variable 

	protein	class
1	kidneyTisue1	Normal
2	kidneyTisue2	Tumor
3	kidneyTisue3	Normal
4	kidneyTisue4	Tumor
5	kidneyTisue5	Normal
6	kidneyTisue6	Tumor



When a “wrong” column is selected for differential expression analysis in the “Class Variable”. E.g. “Sample” instead of “Condition”.

- 7 Users will not be able to proceed, until the correct class column is selected.

#### Class Variable

patient\_id

Chosen variable is not applicable for limma analysis

#### Class of Interest

Not Selected

#### Contrast variable

Not Selected

Contrast other classes

**8** Select Contrast variable from the dropdown

protein patient\_id histological\_type

Class Variable

Class of Interest

Contrast variable

Normal 

Contrast other classes 

Log fold-change variable 

P-value variable 

Adjust P-values for Multiple Comparisons

3	kidneyTisue3	Normal
4	kidneyTisue4	Tumor
5	kidneyTisue5	Normal
6	kidneyTisue6	Tumor

**9** Click "Perform Limma"

Log fold-change variable 

P-value variable 

Adjust P-values for Multiple Comparisons 

Benjamini-Hochberg

 Perform Limma

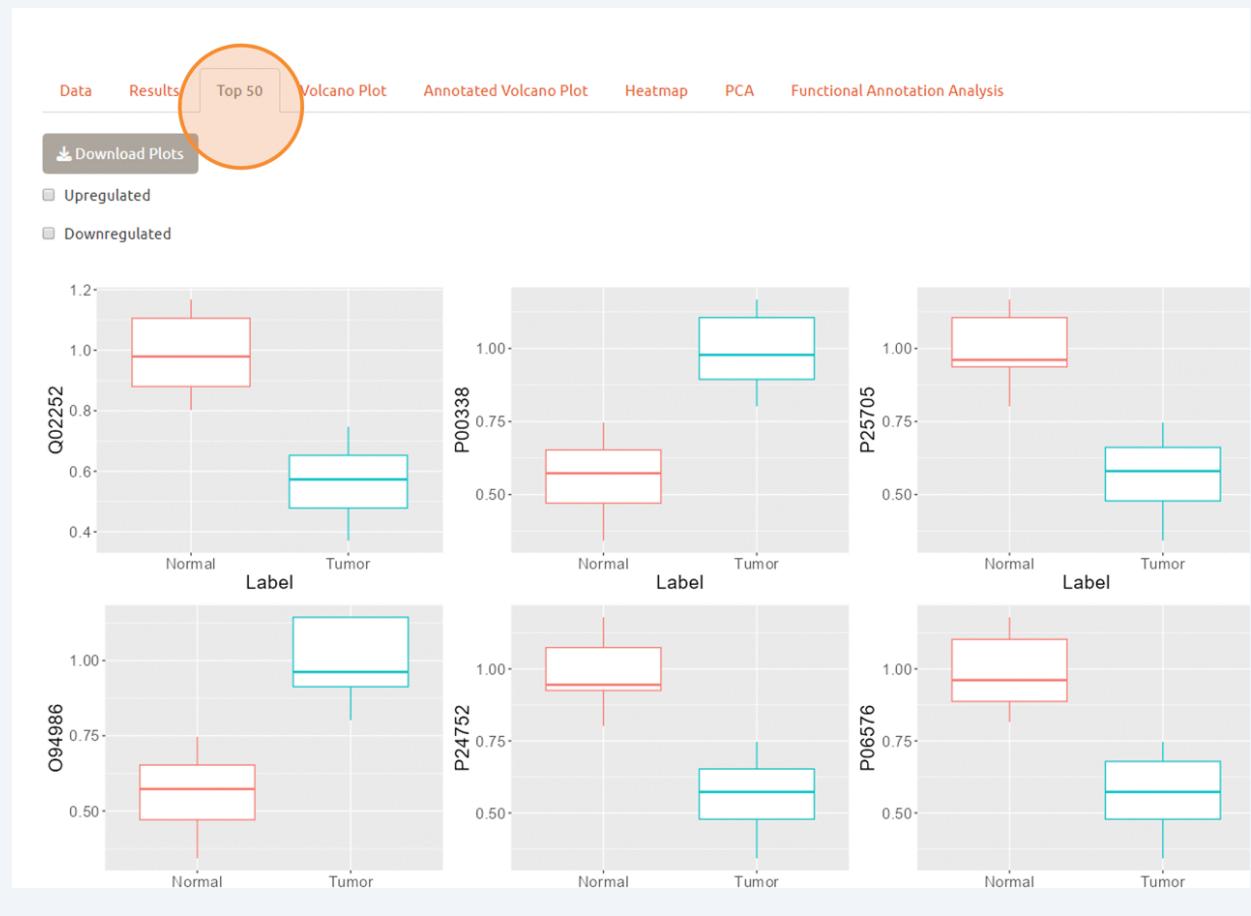
Check the Result Tab

- 10 Click "Results" to check the differential expression results.

The screenshot shows a web-based application for differential expression analysis. At the top, there is a navigation bar with links: "Automated Differential Expression", "Consensus Clustering", and "About". Below the navigation bar, there is a toolbar with several tabs: "Data" (disabled), "Results" (highlighted with a red oval), "Top 50", "Volcano Plot", and "Annotated Volcano Plot". On the left side, there is a form area with a "Complete" button, a "Use Sample Data" checkbox, and a "Log10 Transform" checkbox. There are also two large input fields and a dropdown menu. On the right side, under the "Original Data" section, there is a table with the following data:

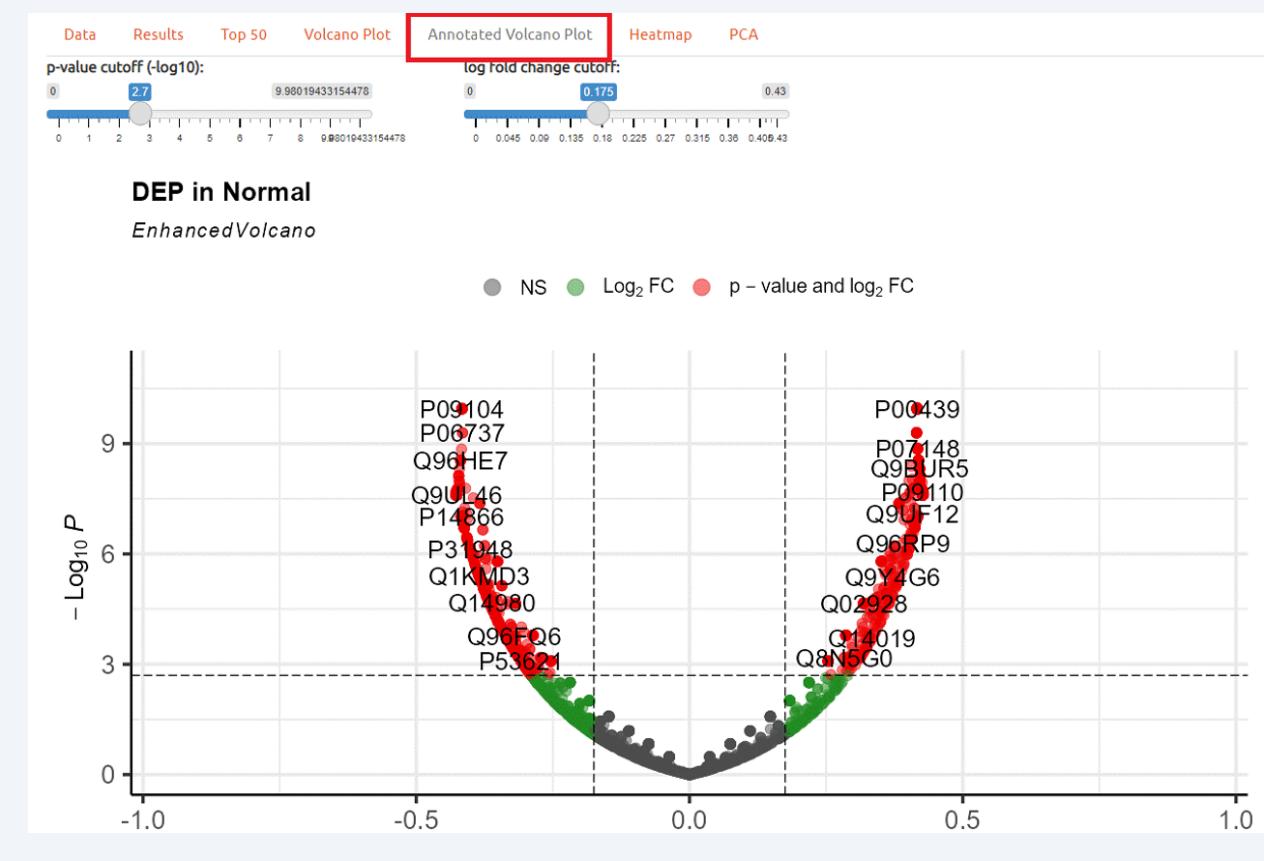
	protein	class	patient_id	histological_type	P09110	P05
1	kidneyTissue1	Normal	patient_1	non-tumorous	1.12	
2	kidneyTissue2	Tumor	patient_1	clear cell RCC	0.72	
3	kidneyTissue3	Normal	patient_2	non-tumorous	0.96	
4	kidneyTissue4	Tumor	patient_2	clear cell RCC	0.56	
5	kidneyTissue5	Normal	patient_3	non-tumorous	1.17	
6	kidneyTissue6	Tumor	patient_3	clear cell RCC	0.62	

11 Click "Top 50" to check the activity differences in the top 50 protein candidates.

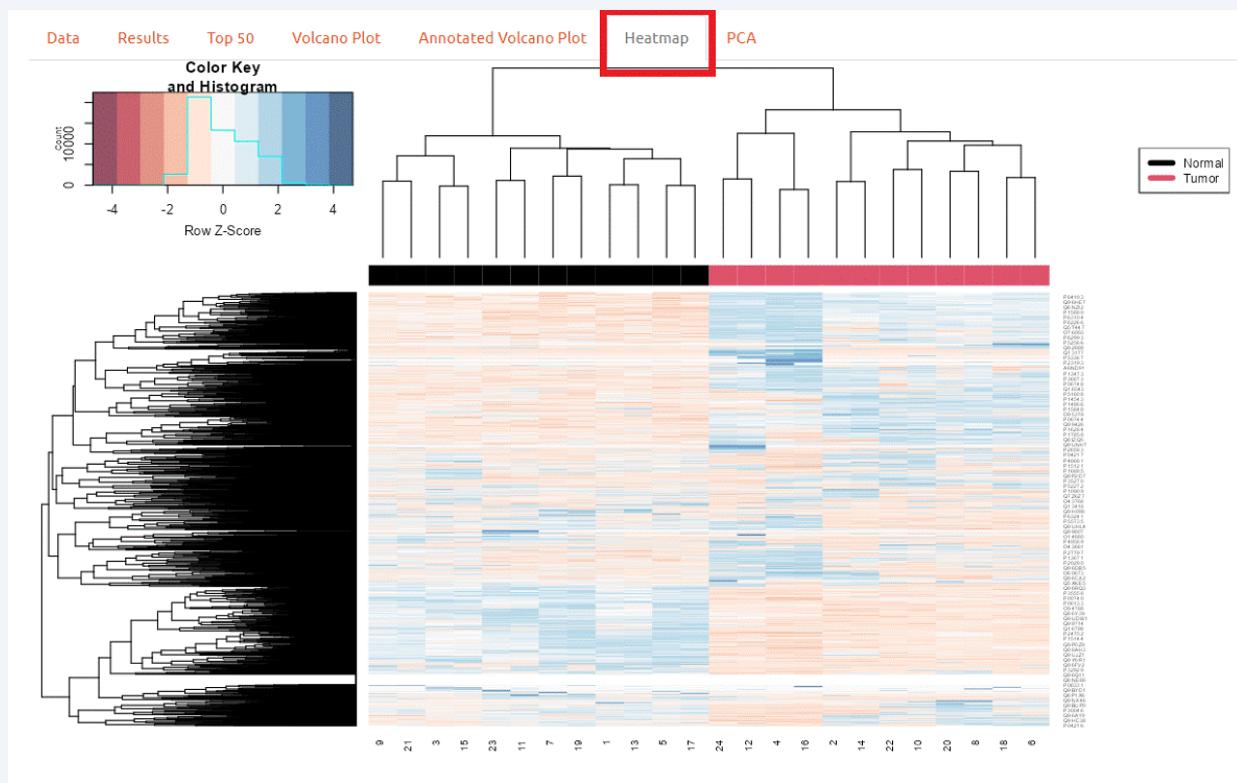


12

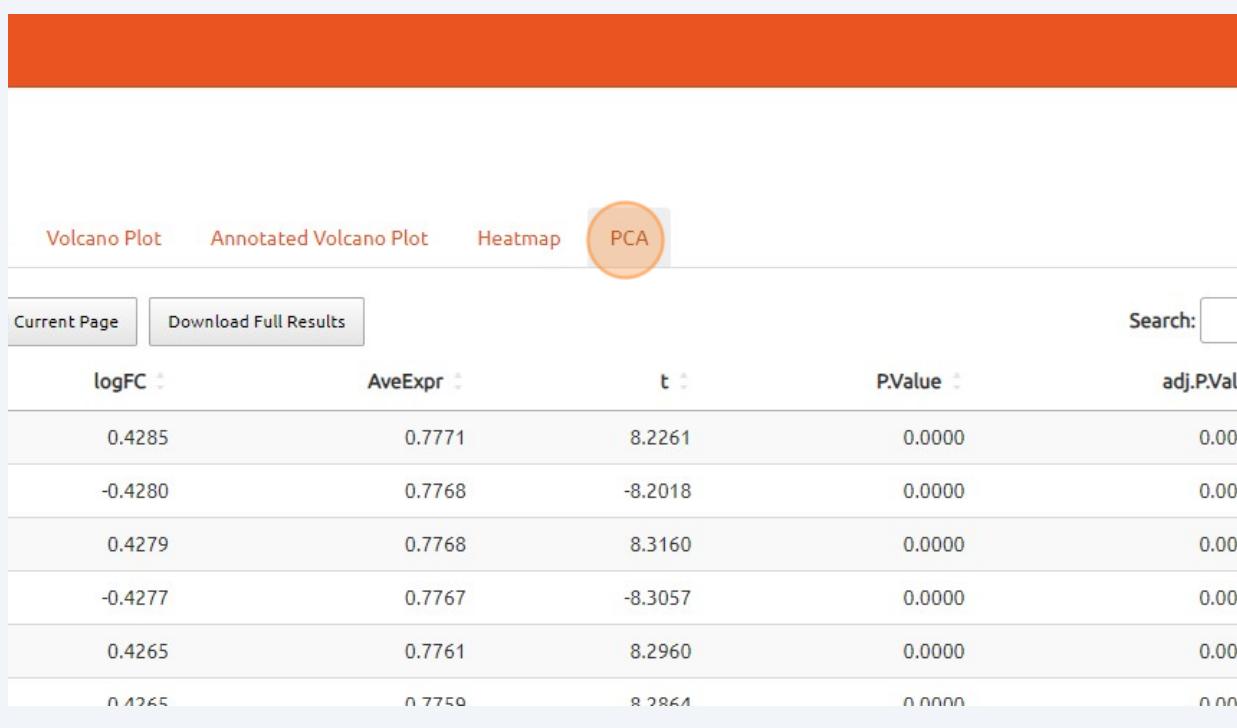
Obtain volcano plots from the "Volcano plot" and "Annotated Volcano plot" tabs



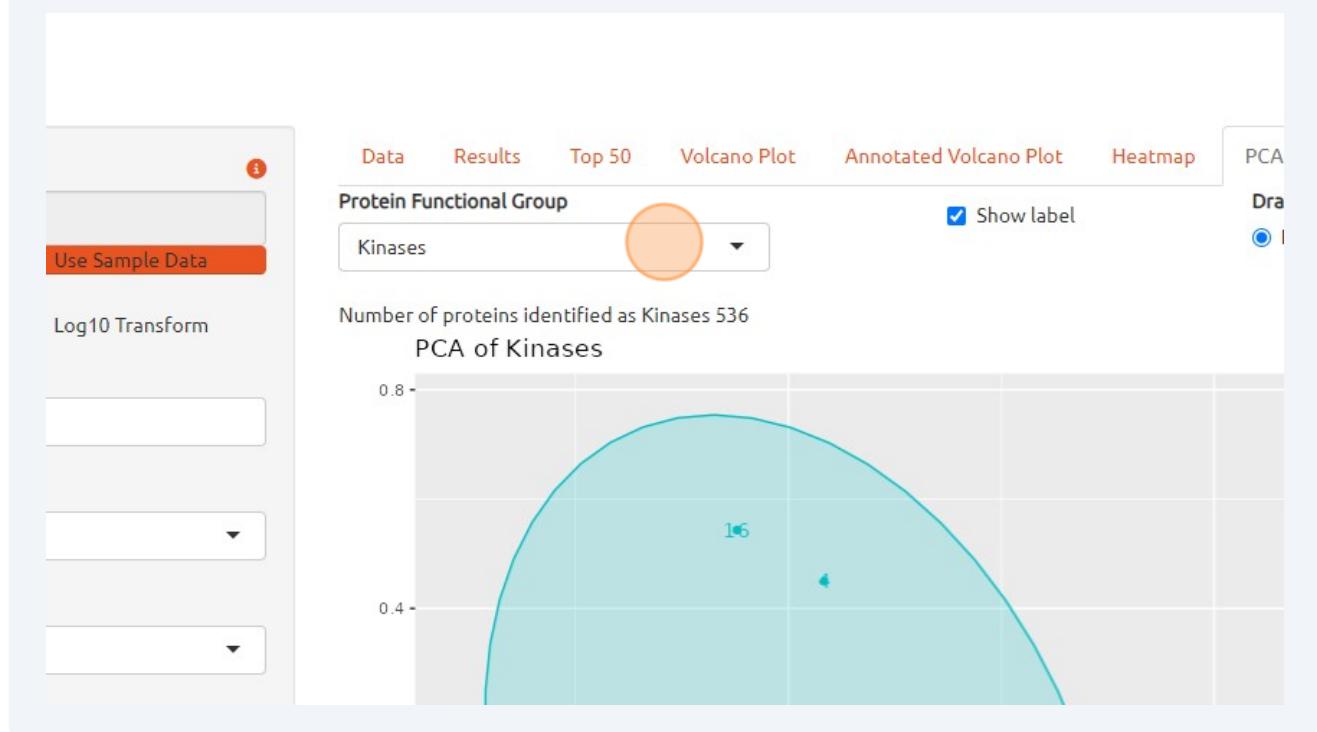
### 13 Obtain heatmap from the "Heatmap" tab.



### 14 Click "PCA"



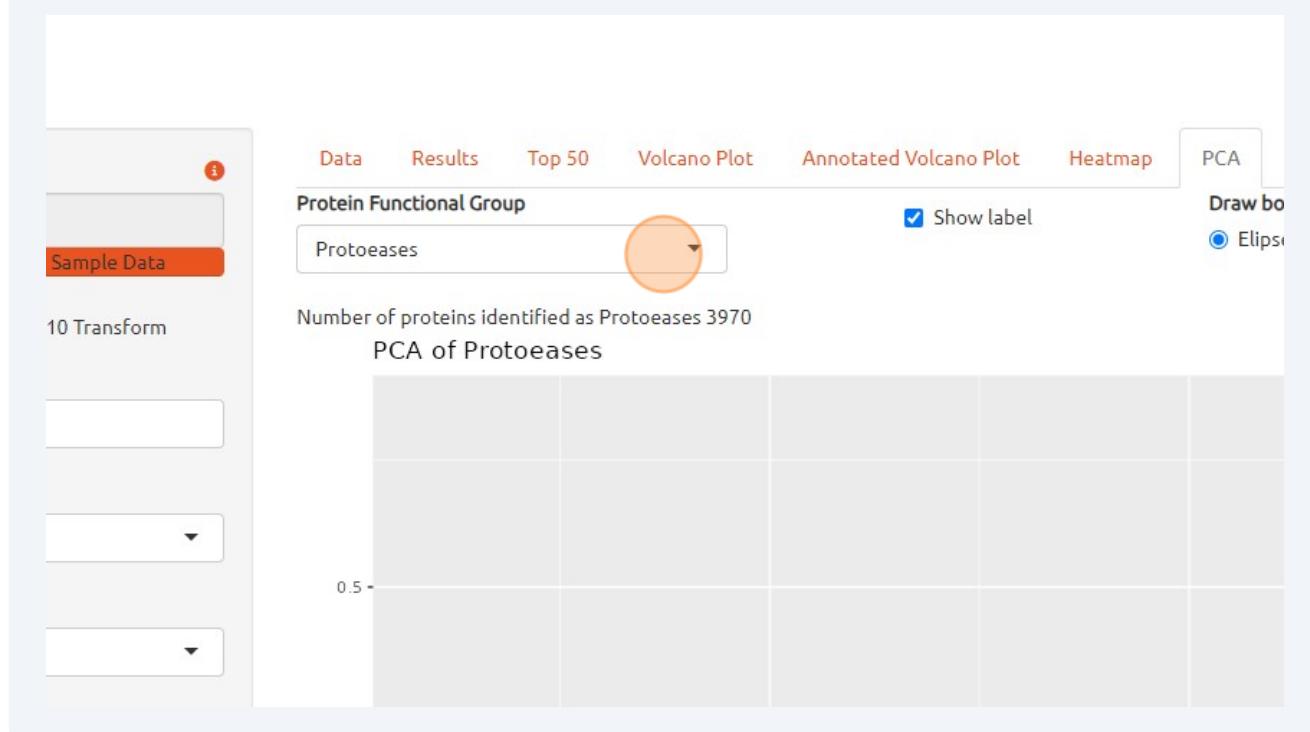
- 15 Click "Kinases" to filter PCA with kinases.



## 16 Filtered by kinases



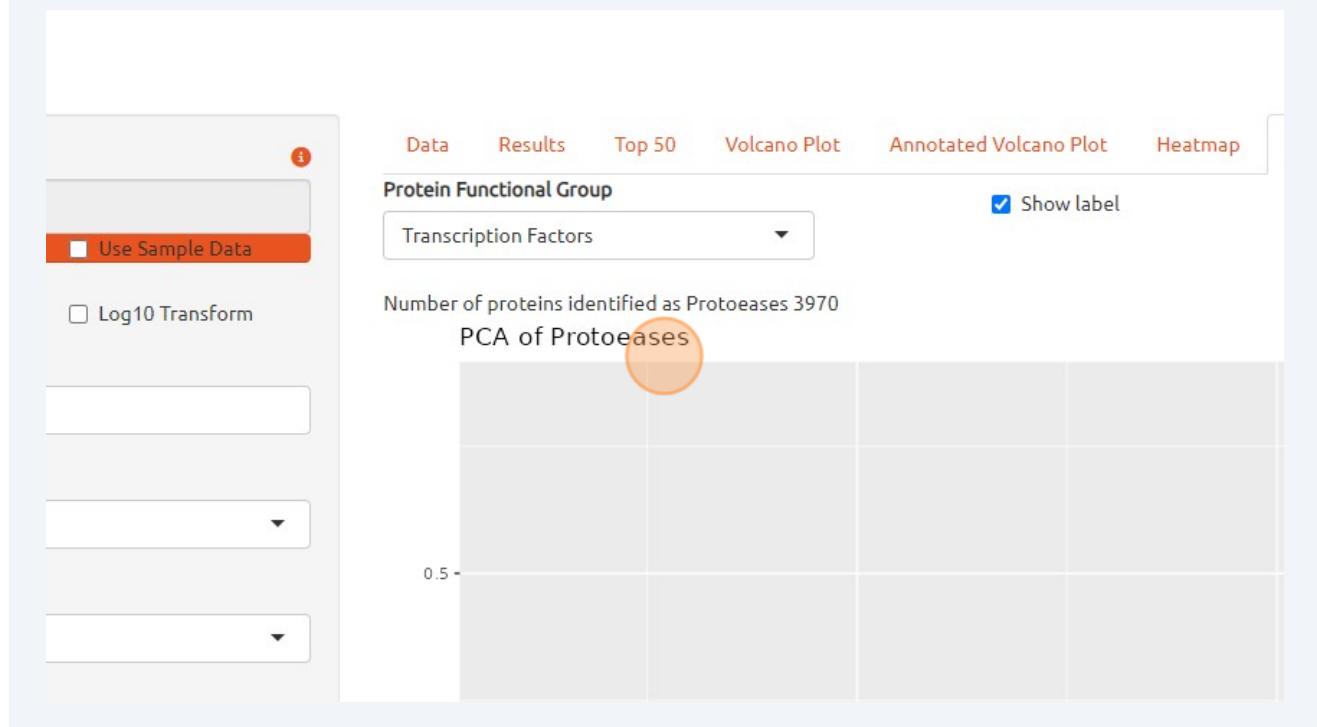
17 Click "Proteases" to filter PCA with proteases.



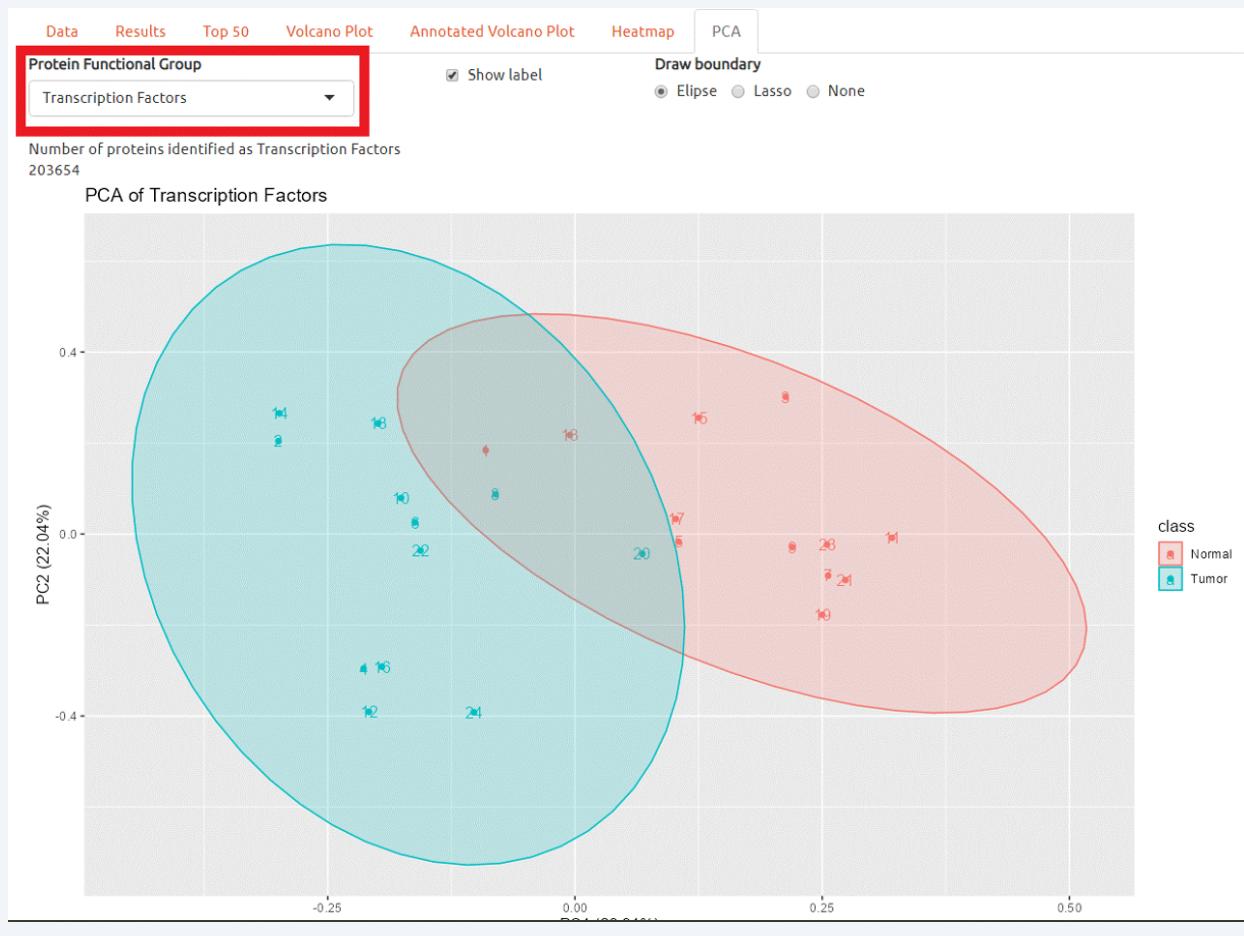
## 18 Filtered by protoeases



**19** Click "Transcription Factors" to filter PCA with Transcription Factors.



## 20 Filtered by transcription factors



21

Click "Functional Annotation Analysis". This will prompt a list of online available tools.

- To perform functional enrichment: gene set, pathways - [DAVID](#)
- To perform functional enrichment: gene set, pathways (with visualizations)- [ShinyGO](#)
- protein-protein interaction network analysis - [STRING](#)

The screenshot shows a horizontal navigation bar with several items: Data, Results, Top 50, Volcano Plot, Annotated Volcano Plot, Heatmap, PCA, and Functional Annotation Analysis. The 'Functional Annotation Analysis' item is circled in red, indicating it is the selected or next step. Below the bar, there is a brief text description and a bulleted list of three analysis options, each with a corresponding link.

erform functional annotation analysis on DEP proteins.

- DAVID Functional Annotation Analysis:  
[navigate to DAVID platform](#)
- ShinyGO Functional Enrichment Analysis:  
[navigate to ShinyGO platform](#)
- STRING Protein-Protein Interaction Networks Functional Enrichment Analysis:  
[navigate to STRING platform](#)

## 22

## DAVID Functional Annotation Analysis

The link will automatically input the differentially expressed protein list into the web server and the results will be prompted.

- The documentation to the DAVID web server is available [here\\*](#).
- The tutorials on how to use DAVID is available [here\\*](#). (\*) Adapted from DAVID website

The screenshot shows the DAVID Functional Annotation Analysis interface. At the top, there are tabs for Data, Results, Top 50, Volcano Plot, Annotated Volcano Plot, Heatmap, PCA, and Functional Annotation Analysis (which is selected). Below the tabs, a banner says 'Perform functional annotation analysis on DEP proteins.' A blue arrow points from the 'DAVID Functional Annotation Analysis' section in the text below to the 'Functional Annotation Analysis' tab in the header. The main content area has three sections: 'DAVID Functional Annotation Analysis' (with a circled 'DAVID' and a blue arrow pointing to it), 'ShinyGO Functional Enrichment Analysis' (with a 'navigate to ShinyGO platform' link), and 'STRING Protein-Protein Interaction Network' (with a 'navigate to STRING platform' link). On the right, the 'Gene List Manager' panel shows a list of species: Use All Species - Saccharomyces cerevisiae, Homo sapiens(46), Bos taurus(1), and a 'Select Species' dropdown. It also includes a 'List Manager' section with 'List\_1' and buttons for Use, Rename, Remove, Combine, Show Gene List, and View Unmapped Ids. To the right of the manager is the 'Annotation Summary Results' panel, which lists 276 DAVID IDs and various annotation categories: Disease (2 selected), Functional\_Annotations (5 selected), Gene\_Ontology (3 selected), Gene\_Report\_Categories (0 selected), General\_Annotations (0 selected), Interactions (1 selected), Literature (0 selected), Pathways (3 selected), Protein\_Domains (4 selected), and Tissue\_Expression (0 selected). Below this is a 'Combined View for Selected Annotation' section with buttons for Functional Annotation Clustering, Functional Annotation Chart, and Functional Annotation Table.

## 23 ShinyGO Geneset Annotation Analysis

The user instructions are available [here](#).

The screenshot shows the ShinyGO 0.80 interface. At the top, there's a navigation bar with tabs: Results, Top 50, Volcano Plot, Annotated Volcano Plot, Heatmap, PCA, and Functional Annotation Analysis. Below the navigation bar, a message says "Functional annotation analysis on DEP proteins." There are two main sections: "Functional Annotation Analysis" and "Functional Enrichment Analysis". A blue arrow points from the "Functional Enrichment Analysis" section down to the main analysis form. The form includes fields for "Select a species (Required)" (set to Human), "Background (recommended)", "FDR cutoff" (0.05), "# pathways to show" (20), and "Pathway size: Min." and "Max.". To the right of the form is a sidebar with links to DAVID, ShinyGO, STRING, and ShinyGO 0.77. The sidebar also contains a brief history of the tool, citation information, and a link to the GitHub repository.

## 24 STRING Protein-protein interaction network analysis

The user guide is available [here](#).

The screenshot shows the STRING interface. At the top, there's a navigation bar with tabs: Results, Top 50, Volcano Plot, Annotated Volcano Plot, Heatmap, PCA, and Functional Annotation Analysis. Below the navigation bar, a message says "Perform functional annotation analysis on DEP proteins." There are three main sections: "DAVID Functional Annotation Analysis", "ShinyGO Functional Enrichment Analysis", and "STRING Protein-Protein Interaction Networks Function". A blue arrow points from the "STRING Protein-Protein Interaction Networks Function" section down to the main analysis form. The form is titled "SEARCH" and includes a "List of Names:" input field with a list of protein identifiers (Q09666, Q60610, E7EX90, Q8N3D0, Q15075, Q92814) and a "Organisms:" dropdown set to "Homo sapiens". Below these are "Advanced Settings" and a "SEARCH" button. The left sidebar lists various search options: Protein by name, Multiple proteins, Proteins by sequences, Proteins with Values/Ranks, Protein families ("COGs"), Pathway / Process / Disease, Add organism, Organisms, Examples, and Random entry.