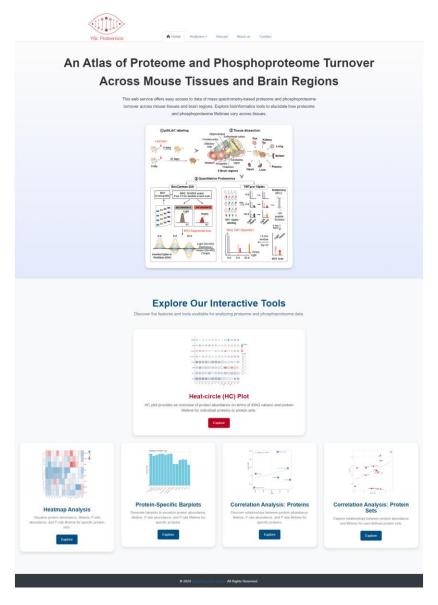
# User Manual for Tissue Proteome and Phosphoproteome Turnover Web Application

# **Table of Contents**

- 1. Data Preparation
- 2. Analysis modules
- 3. How to run this tool locally?

Tissue-PPT is a comprehensive resource that maps protein abundance (PA) and protein lifetime (PT) across eight mouse tissues and nine brain regions, providing in-depth insights into tissue-specific proteostasis. Using advanced proteomic techniques of data-independent acquisition (DIA) and tandem mass tagging (TMTpro), Tissue-PPT analyzed 11,000 proteins and 40,000 phosphosites with their PA and PT profiles, offering unmatched coverage and precision. Tissue-PPT uncovers e.g., tissue-specific short- and long-lived proteins, the role of phosphorylation in regulating protein stability, and how protein-protein interactions and organellar localization influence proteostasis and protein lifetime. By integrating multi-omics datasets, Tissue-PPT provides a detailed atlas of protein dynamics and tissue-specific regulation, offering new therapeutic insights. Accessible through an interactive web portal, Tissue-PPT serves as a valuable tool for studying proteome and phosphoproteome turnover in health and disease.

Users are welcome to visit this web site: https://yslproteomics.shinyapps.io/tissuePPT. We provide the option for the users to install this tool locally (Please check "3. How to run this tool locally?" part below). Then the website homepage can be shown like this:



# 1. Data Preparation

Users could either paste/upload a list of proteins of interest (POIs) or type in a gene name/UniProt ID.

#### 1.1. A list of POIs

POIs should be protein names, like below:

1	Names
2	Psmb2
3	Psmb4
4	Psma6
5	Psmb6
6	Psmb1
7	Psmb3
8	Psma4
9	Psma1
10	Psma3
11	Psma5

Or UniProt IDs, as below:

1	UniProtIDs
2	A8MPP1
3	B7ZAQ6
4	C9JRZ8
5	O00193
6	O00743
7	O14791
8	O15504
9	O43731
10	O60293
11	O60487
12	O60675
13	O60927
14	O75177

All of these names or IDs can be saved in a .csv/.txt file. There is only one column in this file and each row is a protein name or UniProt ID. In addition, Tissue-PPT also supports users to paste these POIs as described in the "2. Import Data" part.

# 1.2. A gene name/UniProt ID

Herein, users could directly type in a gene name (e.g. Snca), or a UniProt ID (e.g. O55042).

# 2. Analysis modules

There are five main modules currently supported in this application:

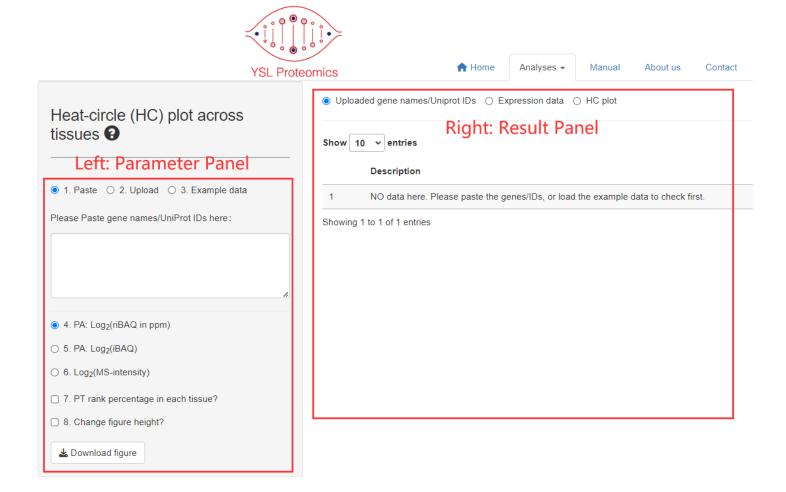
- 2.1. Heat-circle (HC) Plot: Visualize protein abundance and lifetime across tissues.
- 2.2. Heatmap Analysis: Generate heatmaps for protein or phosphosite data.
- 2.3. Protein-Specific Barplots: Create barplots for abundance and lifetime data of specific proteins.
- 2.4. Correlation Analysis: Individual Proteins: Explore relationships between protein and phosphorylation datasets of abundance and lifetime for a single protein.
- 2.5. Correlation Analysis: Protein Sets: Analyze correlations between protein sets or between protein sets and or the proteome.

# 2.1. Heat-circle (HC) Plot

#### 2.1.1. What is HC-Plot?

HC-plot provides a synchronized overview of protein abundance (PA, in terms of iBAQ values, relative iBAQ values or riBAQ, and MS-intensities) and protein lifetime (PT or protein T50 determined by pulsed SILAC labeling) for individual proteins or protein sets. This allows researchers to explore relationships between protein abundance and lifetime across tissues.

#### 2.1.2. How to Use HC-Plot?

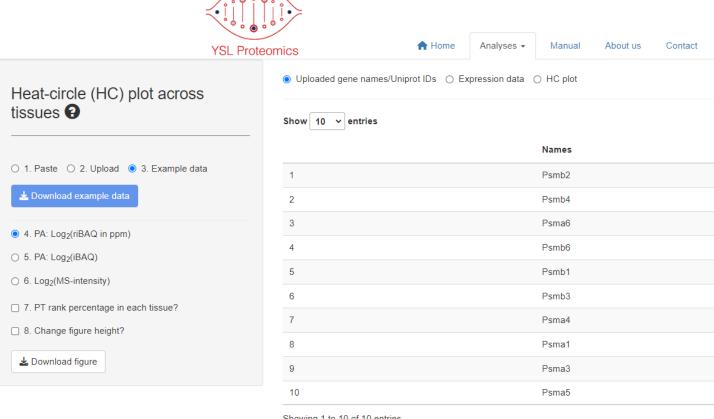


#### A. Parameters:

- A.1. Paste. Users can paste the gene names/UniProt IDs directly here. We found this option very useful for direct inspection of PA and PT values for given genes/proteins.
- A.2. Upload. This means users can upload the gene names/UniProt IDs in a .csv file.
- A.3. Example data. Here shows an example data for users.
- A.4. PA: Log2(riBAQ in ppm). riBAQ mean relative iBAQ. ppm means the relative contribution of iBAQ values to the total protein amount. If users select this parameter, this tool will transform the iBAQ values (log2) into ppm (log2). The values where Log2(riBAQ in ppm) < -1 or Log2(riBAQ in ppm) > 9 are defined as outliers.
- A.5. PA: Log2(iBAQ). If users select this parameter, this tool will match the iBAQ values (log2) for the input gene names/UniProt IDs.
- A.6. Log2(MS-intensity). If users select this parameter, this tool will match the protein intensity (log2) for the input gene names/UniProt IDs.
- A.7. PT rank percentage in each tissue? If true, this tool will calculate the rank of the protein lifetime value and normalize every rank to the maximum rank in terms of rank percentage (0-100%) for each tissue.
- A.7.1. Row scaled across tissues? If true in A.7. above, users will see this parameter, which means this tool will additionally scale the rank percentage value across all tissues for the current gene/protein list (Z score across tissues).
- A.8. Change figure height? If enabled, users can manually adjust the figure height using the parameter A.8.1 (described below). If disabled, the tool will automatically adjust the figure height based on the number of input gene names or UniProt IDs.
- A.8.1. Figure Height. This parameter adjusts the height of the plot.
- B. Results:
- B.1. Uploaded gene names/Uniprot IDs. Here shows the input gene names/Uniprot IDs. If users input nothing, it shows 'NO data here. Please paste the genes/IDs, or load the example data to check first.'
- B.2. Expression data. Here shows the matched results of the input gene names/Uniprot IDs, including protein iBAQ values and protein lifetime values. Please note that if users select parameters 4, 5, and/or 6 mentioned above, the corresponding results will be displayed based on the selected parameters.
- B.3. HC plot. Here shows the HC plot based on the matched results.

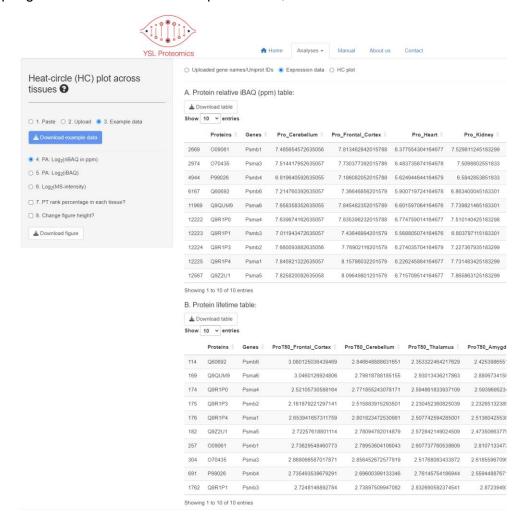
# 2.1.3. Example Output

Users can choose the '3. Example data' in the left parameter panel. Shown as below:



Showing 1 to 10 of 10 entries

Then users click 'Expression data', this tool will match relative iBAQ values in ppm (log2) and lifetime values (log2) for the input genes based on the default parameters, shown as below:

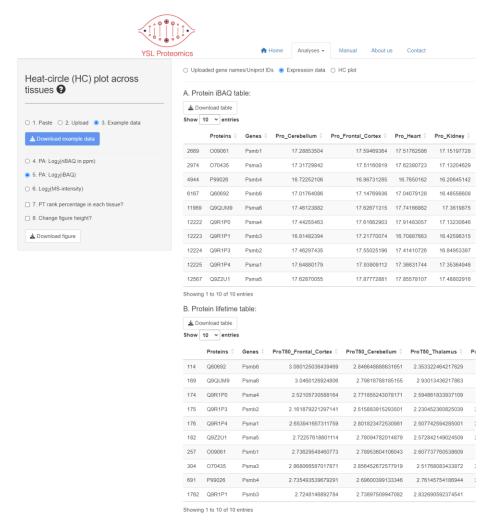


If users click the "Download table" button, the corresponding table will be downloaded to their local device.

The HC-plot based on these values is shown as below:



If users choose '5. PA: Log2 (iBAQ)', this tool will match iBAQ values (log2) and lifetime values (log2) for the input genes, shown as below:

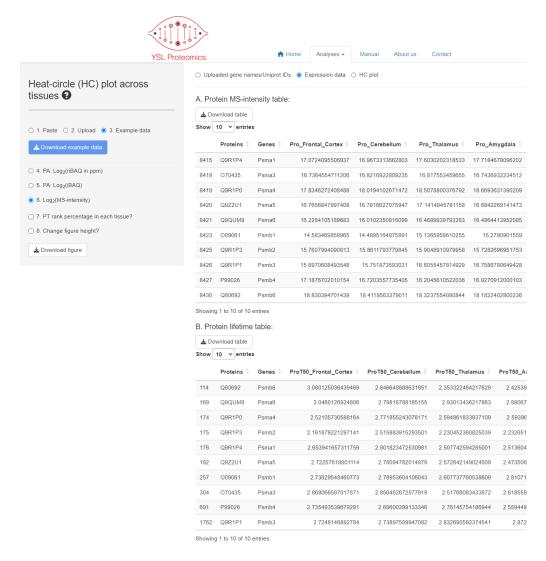


If users click the "Download table" button, the corresponding table will be downloaded to their local device.

The HC-plot based on these values is shown as below:



If users choose '6. Log2(MS-intensity)', this tool will match protein intensities (log2) and lifetime values (log2) for the input genes, shown as below:

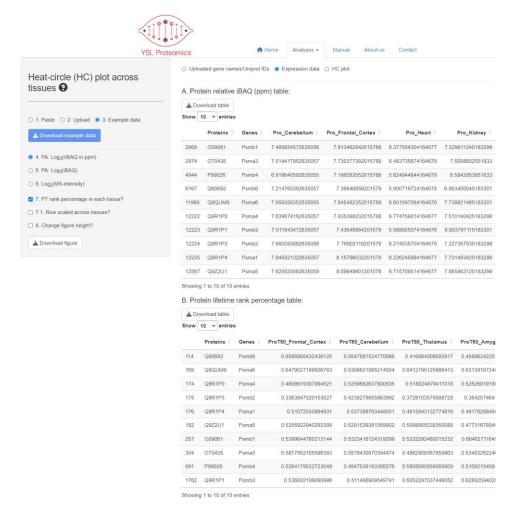


If users click the "Download table" button, the corresponding table will be downloaded to their local device.

The HC-plot based on these values is shown as below:

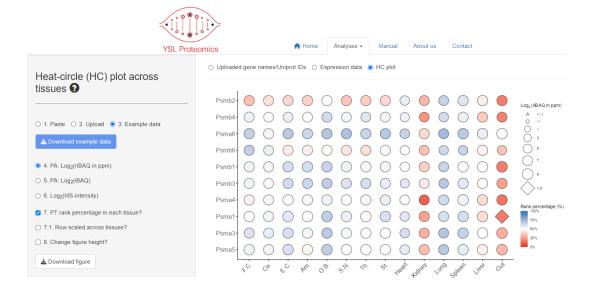


If users choose '4. PA: Log2(riBAQ in ppm)' and '7. PT rank percentage in each tissue?', this tool will match relative iBAQ values in ppm (log2) and then calculate rank percentage of each protein lifetime value for each tissue, shown as below:

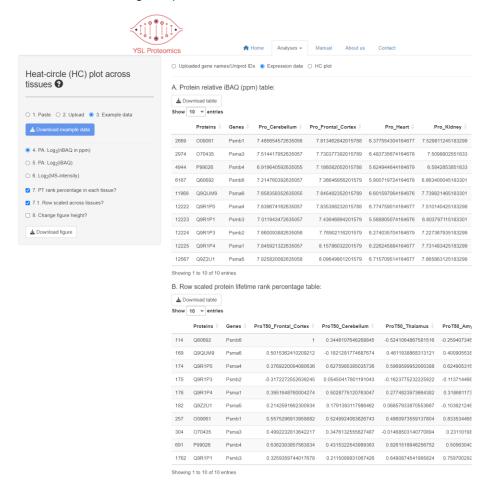


If users click the "Download table" button, the corresponding table will be downloaded to their local device.

The HC-plot based on these values is shown as below:



If users choose '7.1. Row scaled across tissues?', this tool will scale the rank percentage of each protein lifetime value across all tissues for the current gene/protein list, shown as below:



If users click the "Download table" button, the corresponding table will be downloaded to their local device.

The HC-plot based on these values is shown as below:

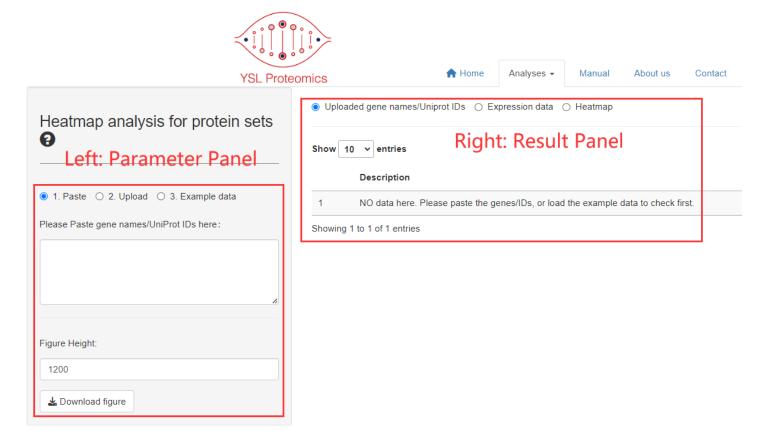


#### 2.2. Heatmap Analysis

# 2.2.1. What is Heatmap Analysis?

Visualize protein abundance, lifetime, phospho abundance, and phospho lifetime across tissues with heatmaps. This module provides a rapid and comprehensive view of data trends for protein sets.

# 2.2.2. How to Use Heatmap Analysis?

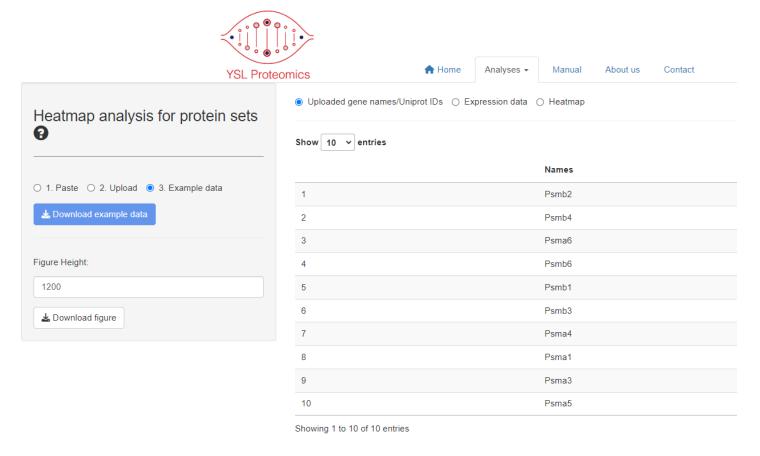


#### A. Parameters:

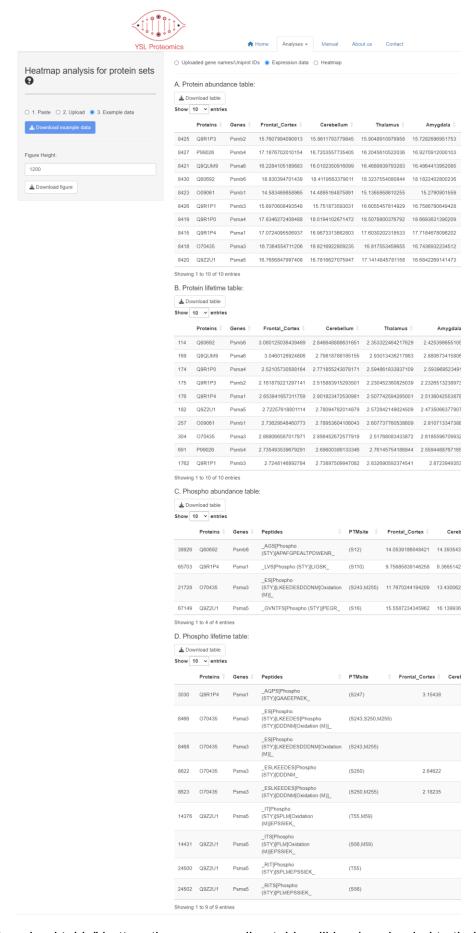
- A.1. Paste. This means users can paste the gene names/UniProt IDs directly here.
- A.2. Upload. This means users can upload the gene names/UniProt IDs in a .csv file.
- A.3. Example data. Here shows an example data for users.
- A.4. Figure Height. This parameter adjust the height of the plot.
- B. Results:
- B.1. Uploaded gene names/Uniprot IDs. Here shows the input gene names/Uniprot IDs. If users input nothing, it shows 'NO data here. Please paste the genes/IDs, or load the example data to check first.'
- B.2. Expression data. Here shows the matched results of the input gene names/Uniprot IDs, including protein abundance (MS-intensities derived from BoxCarmax-DIA data), protein lifetime, phospho abundance, and phospho lifetime.
- B.3. Heatmap. Here shows the Heatmap based on the matched results.

# 2.2.3. Example Output

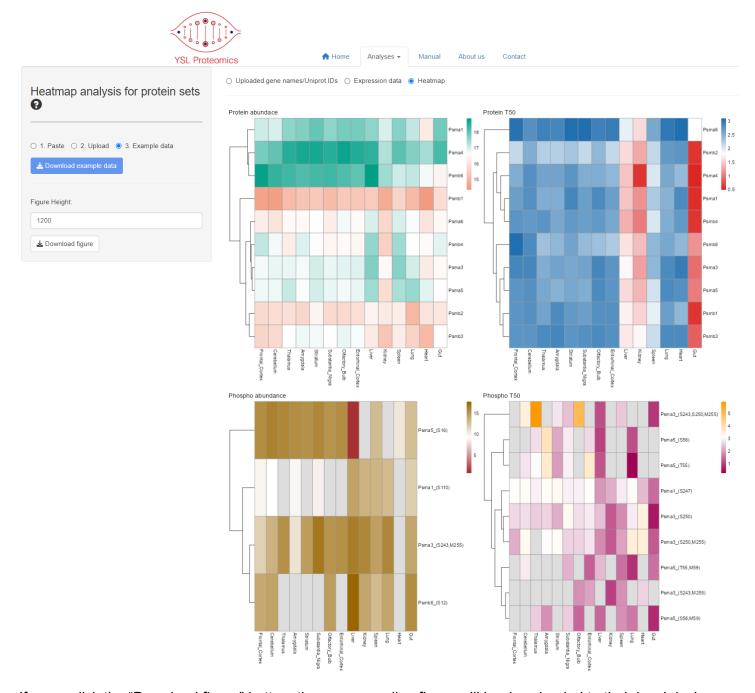
Users can choose the '3. Example data' in the left parameter panel. Shown as below:



When users click on 'Expression data', the tool retrieves and matches four types of log2-transformed values for the input genes: Protein abundance, Protein lifetime, Phospho abundance, and Phospho lifetime. The matched data is displayed as shown below:



The Heatmap based on these values is shown as below:

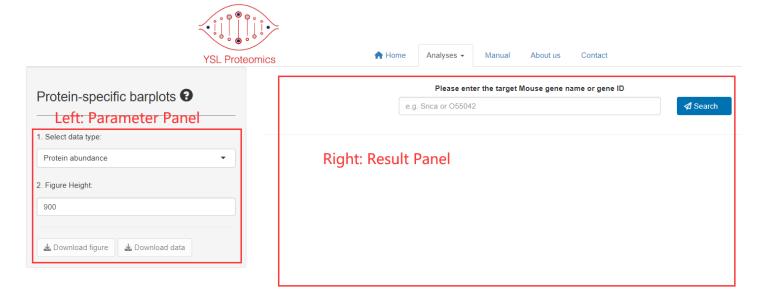


#### 2.3. Protein-Specific Barplots

# 2.3.1. What is Protein-Specific Barplots?

Protein-Specific Barplots are visual tools designed to display the distribution of protein-related data, such as protein abundance, lifetime, phosphosite abundance, or phosphosite lifetime, across different tissues for a specific protein. These barplots allow researchers to analyze and compare how a particular protein's characteristics vary across multiple tissues. By entering a protein's name or ID, users can generate detailed plots that summarize the protein's behavior, making it easier to identify patterns, trends, or anomalies in the data.

#### 2.3.2. How to Use Protein-Specific Barplots?



#### A. Parameters:

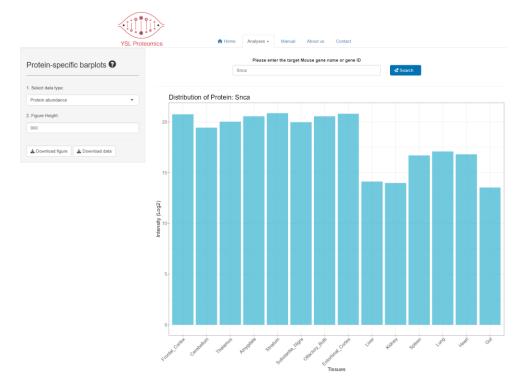
- A.1. Select data type. There are four types of datasets, including Protein abundance (MS-intensities derived from BoxCarmax-DIA data), Protein lifetime, Phosphosite abundance, and Phosphosite lifetime. Users should choose one of them.
- A.2. Figure Height. This parameter adjusts the height of the plot.

#### B. Results:

Please enter the target Mouse gene name or gene ID. Here users should type in a mouse gene name or UniProt ID that they want to check, for example, Snca or O55042. Case does not matter.

# 2.3.3. Example Output

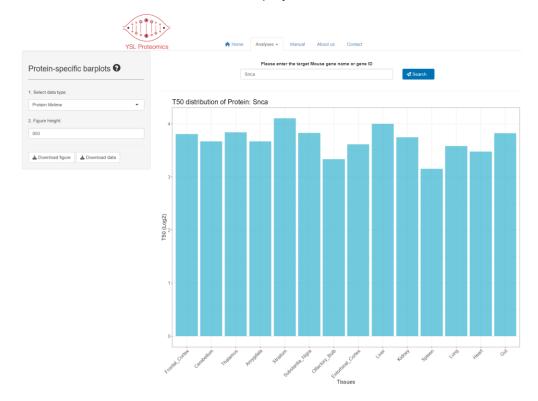
Users can type in a mouse gene name or UniProt ID that they want to check, for example, Snca. Then click the 'Search' button and the results are shown as below:



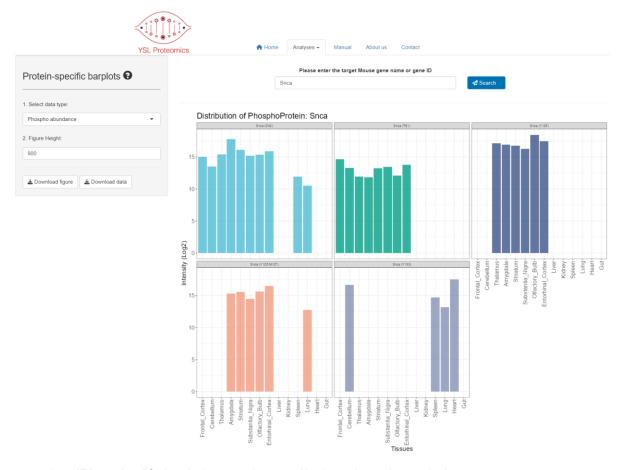
If users click the "Download figure" button, the corresponding figure will be downloaded to their local device.

If users click the "Download data" button, the corresponding data used for the barplot will be downloaded to their local device.

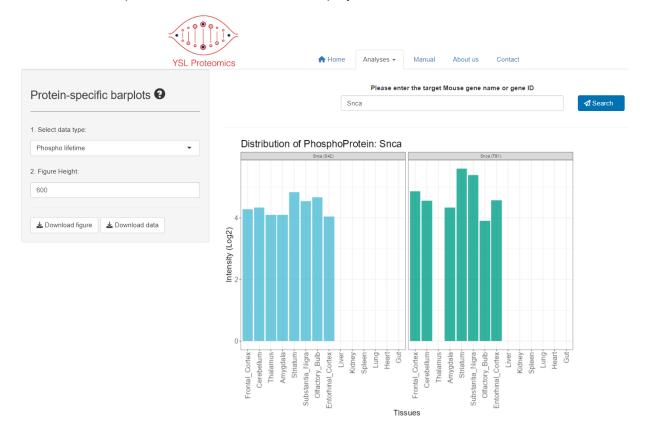
When users select 'Protein lifetime', the results are displayed as shown below:



When users select 'Phospho abundance', the results are displayed as shown below:



When users select 'Phospho lifetime', the results are displayed as shown below:

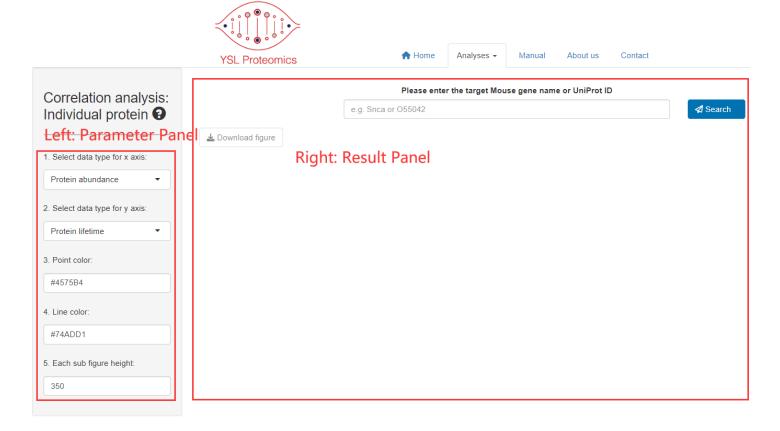


# 2.4. Correlation Analysis: Individual Proteins

# 2.4.1. What is Correlation Analysis: Individual Proteins?

Correlation Analysis: Individual Protein refers to a method for examining the relationship between a specific protein's abundance, lifetime, phospho abundance, or phospho lifetime across different tissues. By default, both Pearson and Spearman correlations are calculated for users. This analysis helps researchers understand how the behavior of an individual protein aligns with other variables or datasets. It provides valuable insights into tissue-specific dynamics, functional roles, or potential regulatory mechanisms associated with the protein, enabling a deeper understanding of its biological significance.

### 2.4.2. How to Use Correlation Analysis: Individual Proteins?



#### A. Parameters:

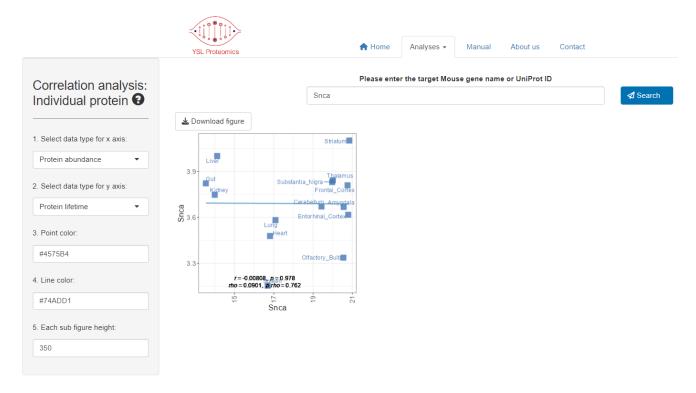
- A.1. Select data type for x axis. There are four types of datasets, including Protein abundance, Protein lifetime, Phosphosite abundance, and Phosphosite lifetime. Users should choose one of them for x axis.
- A.2. Select data type for y axis. Similar to above, but for y axis.
- A.3. Point color. Users can type in a color name to change the point color in the correlation plot.
- A.4. Line color. Users can type in a color name to change the line color in the correlation plot.
- A.5. Each sub figure height. This parameter adjust the height of each subgraph.

#### B. Results:

Please enter the target Mouse gene name or gene ID. Here users should type in a mouse gene name or UniProt ID that they want to check, for example, Snca or O55042. Case does not matter.

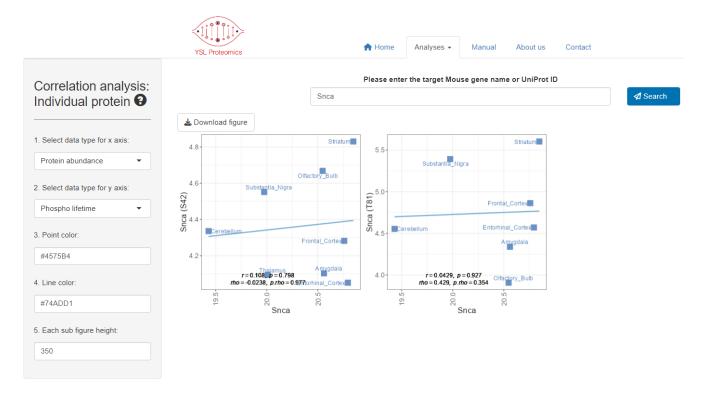
# 2.4.3. Example Output

Users can type in a mouse gene name or UniProt ID that they want to check, for example, Snca. Then click the 'Search' button and the results are shown as below:



*r* means Pearson correlation coefficient, *p* means the p value of the correlation test with a pearson method. *rho* means Spearman correlation coefficient, *p.rho* means the p value of the correlation test with a spearman method.

When users select 'Protein abundance' for x axis and 'Phospho lifetime' for y axis, the results are displayed as shown below:



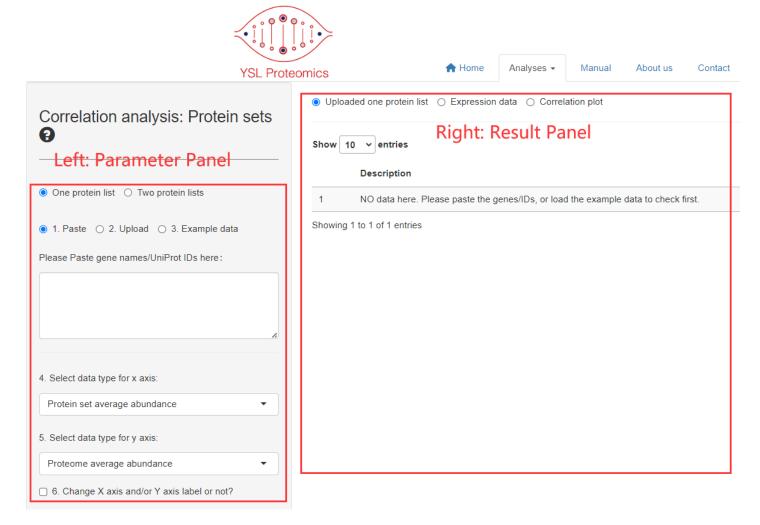
#### 2.5. Correlation Analysis: Protein Sets

# 2.5.1. What is Correlation Analysis: Protein Sets?

Correlation Analysis: Protein Set supports discovery of relationship and dependency between protein abundance and protein lifetime for a given protein set or two protein sets or the averaged protein levels. By default, both Pearson and Spearman correlations are calculated for users.

# 2.5.2. How to Use Correlation Analysis: Individual Proteins?

I. One protein list. If users choose 'One protein list' here, they can input a single list of proteins, and this tool will perform correlation analysis between the proteins in the provided list and the entire proteome dataset integrated within the tool.

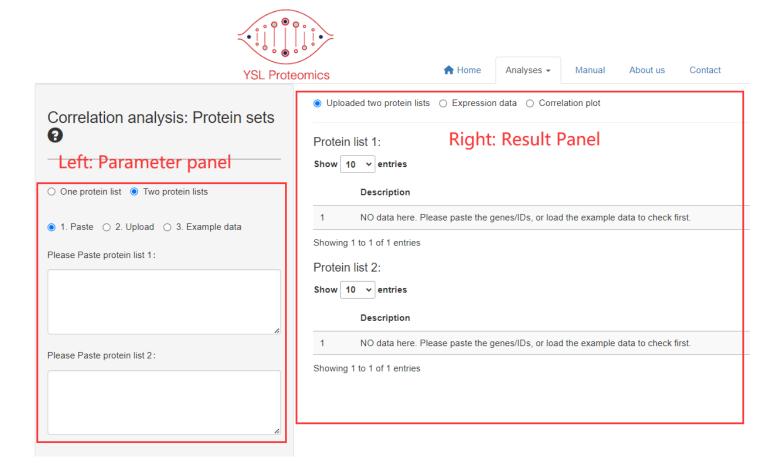


#### A. Parameters:

- A.1. Paste. This means users can paste the gene names/UniProt IDs directly here.
- A.2. Upload. This means users can upload the gene names/UniProt IDs in a .csv file.
- A.3. Example data. Here shows an example data for users.
- A.4. Select data type for x axis. There are four types of datasets, including Protein set average abundance (the average abundance of all proteins within the input protein set across different tissues), Protein set average lifetime (the average lifetime of all proteins within the input protein set across different tissues), Proteome

average abundance (the average abundance of all proteins within the whole integrated proteome dataset across different tissues), Proteome average lifetime (the average lifetime of all proteins within the whole integrated proteome dataset across different tissues). Users should choose one of them for x axis.

- A.5. Select data type for y axis. Similar to above, but for y axis.
- A.6. Change X axis and/or Y axis label or not? If ture, users can define the X/Y axis label by themselves.
- A.7. Point color. Users can type in a color name to change the point color in the correlation plot.
- A.8. Line color. Users can type in a color name to change the line color in the correlation plot.
- A.9. Each sub figure height. This parameter adjust the height of each subgraph.
- B. Results:
- B.1. Uploaded one protein list. Here shows the input one list of proteins. If users input nothing, it shows 'NO data here. Please paste the genes/IDs, or load the example data to check first.'
- B.2. Expression data. Here shows the matched results of the input gene names/Uniprot IDs.
- B.3. Correlation plot. Here shows the Correlation plot based on the matched results.
- II. Two protein lists. Users can input two lists of proteins, and this tool will perform correlation analysis between the two lists of proteins.

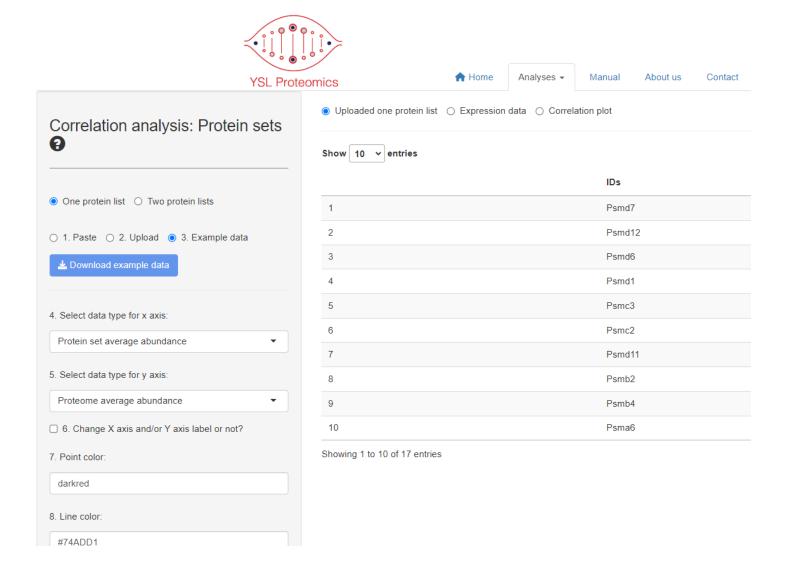


A. Parameters:

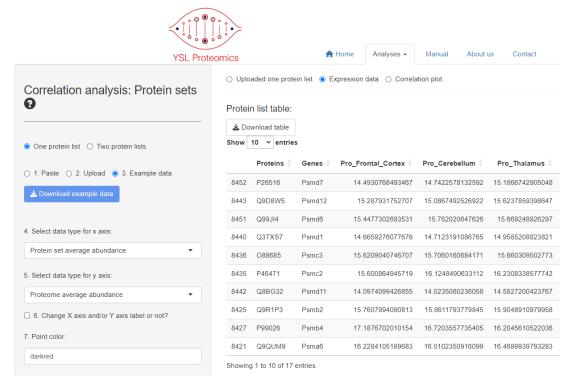
- A.1. Paste. This means users can paste the gene names/UniProt IDs directly here. Please note, users should input two lists of proteins here.
- A.2. Upload. This means users can upload the gene names/UniProt IDs in a .csv file. Please note, users should upload two lists of proteins here.
- A.3. Example data. Here shows two example data for users.
- A.4. Other parameters are similar to above.
- B. Results:
- B.1. Uploaded two protein lists . Here shows the input two lists of proteins. If users input nothing, it shows 'NO data here. Please paste the genes/IDs, or load the example data to check first.'
- B.2. Expression data. Here shows the matched results of the two lists of proteins.
- B.3. Correlation plot. Here shows the Correlation plot based on the matched results.

# 2.5.3. Example Output

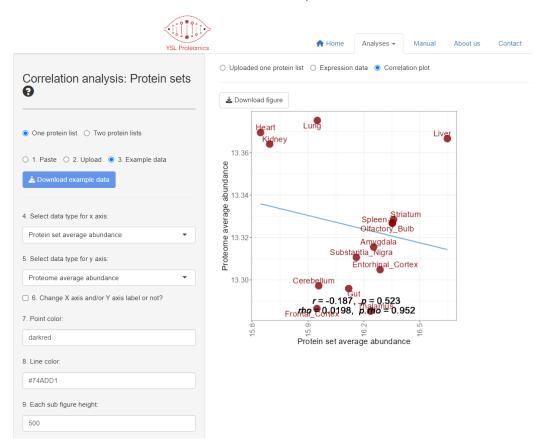
If users choose 'One protein list', and click '3. Example data', the 'Uploaded one protein list' will be shown as below, which means the example dataset contains 17 proteins:



Then users can click 'Expression data', below shows the matched results for the input 17 proteins.

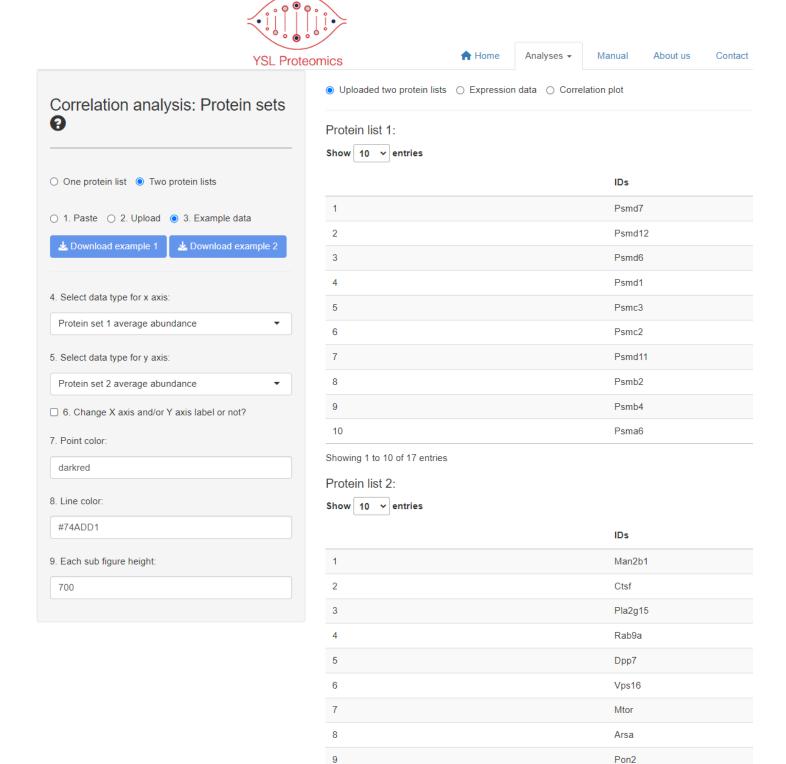


When users click 'Correlation plot', here shows the correlation plot as below. Users can choose the parameter "6. Change X axis and/or Y axis label or not?" to rename the protein lists.



*r* means Pearson correlation coefficient, *p* means the p value of the correlation test with a pearson method. *rho* means Spearman correlation coefficient, *p.rho* means the p value of the correlation test with a spearman method.

If users choose 'Two protein lists', and click '3. Example data', the 'Uploaded two protein lists' will be shown as below, there are two example datasets here:

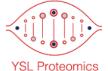


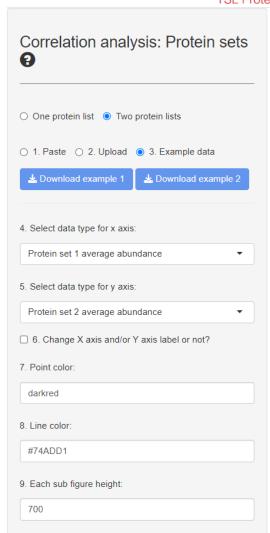
Showing 1 to 10 of 22 entries

Lamp1

10

Then users can click 'Expression data', below shows the matched results for the two lists of proteins.







14.3848529362463

14.2328356359768

12.7470735702032

16.7779213232078

14.2087521428589

15.2380045628869

13.5021439335074

17.3028013956936

14.450898183

14.663150131

13.621022236

17.188529636

Showing 1 to 10 of 22 entries

Q9JLN9

P50428

Q62086

P11438;Q9DC13

Mtor

Arsa

Pon2

Lamp1

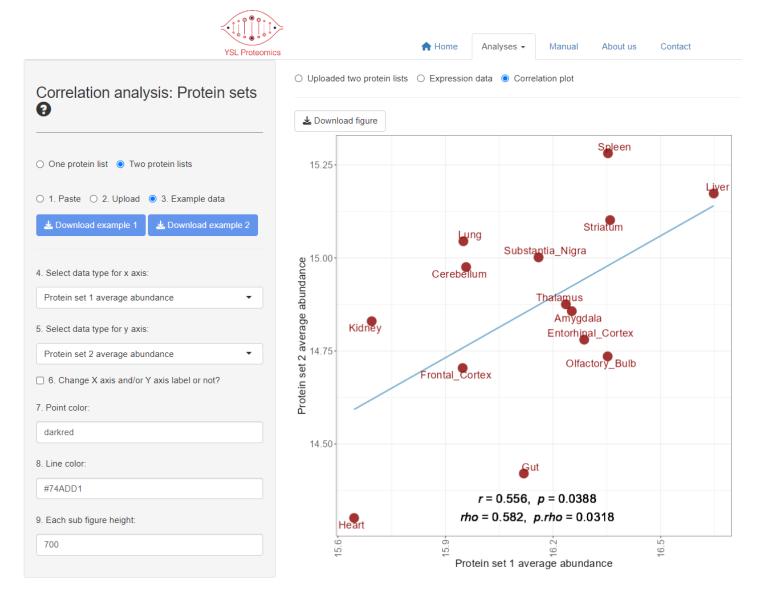
6264

894

8142

5337

When users click 'Correlation plot', here shows the correlation plot as below. Users can choose the parameter "6. Change X axis and/or Y axis label or not?" to rename the protein lists.



*r* means Pearson correlation coefficient, *p* means the p value of the correlation test with a pearson method. *rho* means Spearman correlation coefficient, *p.rho* means the p value of the correlation test with a spearman method.

# 3. How to run this tool locally?

Tissue-PPT is an open source application and all codes can be also obtained on our GitHub: https://github.com/yslproteomics/tissuePPT. If users want to run Tissue-PPT on their own computer, they should operate as below:

As this tool was developed with R, you may:

- a) Install R. You can download R from here: https://www.r-project.org/.
- b) Install RStudio. (Recommendatory but not necessary). You can download RStudio from here: https://www.rstudio.com/.
- c) Check packages. After installing R and RStudio, you should check whether you have installed these packages (shiny, shinyjs, shinyWidgets, shinyBS, DT, data.table, ggsci, ggplot2, ggrepel, patchwork, dplyr, openxlsx, cowplot, grid, ggpubr, impute, pheatmap, ggplotify). You may run the codes below to check them:

if(!require(pacman)) install.packages("pacman") pacman::p\_load(shiny, shinyjs, shinyWidgets, shinyBS, DT, data.table, ggsci, ggplot2, ggrepel, patchwork, dplyr, openxlsx, cowplot, grid, ggpubr, impute, pheatmap, ggplotify)

Please note, if you find some packages cannot be installed directly using the above command, you can find them in the GitHub source and install them by, for example:

library(devtools) install\_github("yslproteomics/tissuePPT")

d) Run this tool locally

if(!require(tissuePPT)) devtools::install\_github("yslproteomics/tissuePPT")
library(tissuePPT)
tissuePPT app()