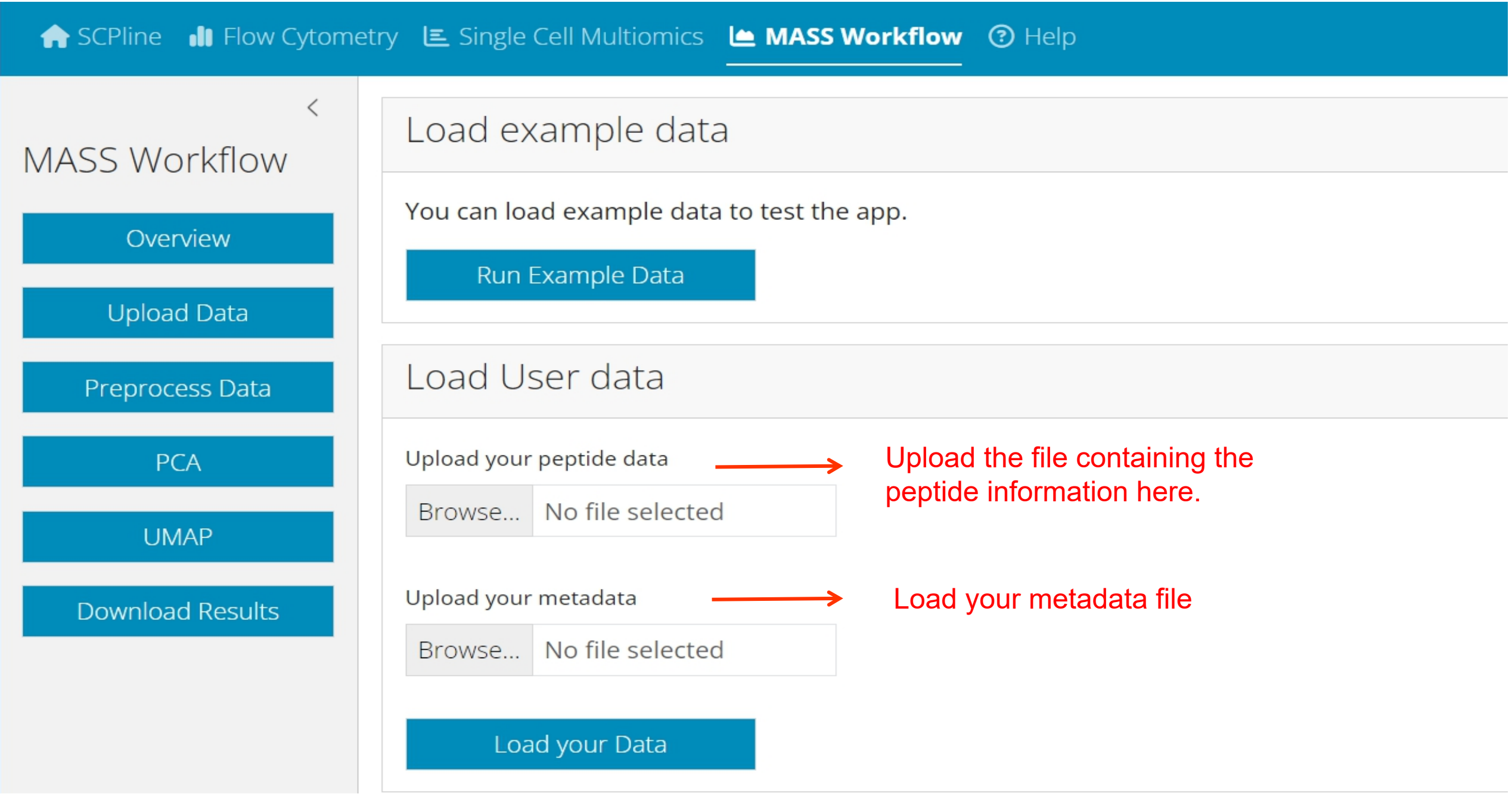
SCPline

Make single-cell analysis available to every investigator

SCPline:An interactive shiny framework for the study of single-cell proteomics data preprocessing

SCPline is an integrated one-stop single-cell proteomics data preprocessing analysis tool developed by Guohua Wang Laboratory, aiming to provide efficient data preprocessing analysis for laboratory researchers.

Moudle 3 : MASS Workflow



Moudle 3 : MASS Workflow

SCPline | Flow Cytometry | Single Cell Multiomics | **MASS Workflow** | Help

MASS Workflow

- Overview
- Upload Data
- Preprocess Data
- PCA
- UMAP
- Download Results

Preprocess Data

Preprocess Method

median

The data preprocessing includes the following steps:

- Replace zeros with NA values for compatibility with downstream analysis.
- Filter out data with more than 99% missing values to ensure data quality.
- Count unique features at peptide and protein levels.
- Aggregate peptide data to the protein level using various aggregation functions (sums, means, medians, and median polish).
- Normalize data by centering columns and rows with median and mean values.
- Impute missing values using k-nearest neighbors (kNN) method.

Preprocess Data

Various aggregation functions are used to aggregate the peptide data to the protein level. In this system, you can choose the polymerization method of polypeptides, we integrate four common polymerization functions: sum, mean, median and medianPolish.

In this step, we go through a number of pre-processing steps, please read carefully to ensure that the quality of the data is suitable for downstream analysis.

MASS Workflow

- Overview
- Upload Data
- Preprocess Data
- PCA
- UMAP
- Download Results

The data preprocessing includes the following steps:

- Replace zeros with NA values for compatibility with downstream
- Filter out data with more than 99% missing values to ensure data
- Count unique features at peptide and protein levels.
- Aggregate peptide data to the protein level using various aggrega
- Normalize data by centering columns and rows with median and
- Impute missing values using k-nearest neighbors (kNN) method.

Preprocess Data

Batch Effect Correction

The batch effect correction includes the following steps:

- Select a batch effect correction field.
- Correct batch effects using the ComBat method.

Batch Effect Correction Field

batchlabel

Correct Batch Effects

Here, due to the large variability of the mass spectrum data, some of the data will contain batch effect metadata. SCPline In order to meet the broader needs, you can according to your metadata contains the "batch" related label column name, if so, the system will identify all the metadata column names, you can select the "batch" related column name to perform batch effect correction. If there is no batch effect label, you can ignore this step without affecting subsequent analysis.

Moudle 3 : MASS Workflow

SCPline | Flow Cytometry | Single Cell Multiomics | **MASS Workflow** | Help

MASS Workflow

- Overview
- Upload Data
- Preprocess Data
- PCA
- UMAP
- Download Results

PCA

The data analysis includes the following steps:

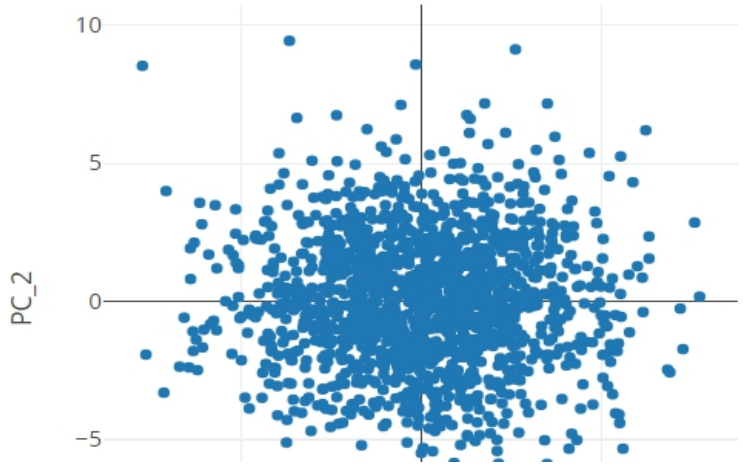
- Convert scp data to Seurat object.
- Perform principal component analysis (PCA) to visualize the data.

Run PCA

The system uses the Seurat package for data follow-up. Click this button to start PCA analysis.

PCA Plot

PCA is a dimensionality reduction technique that can be used for visualizing high-dimensional data in a low-dimensional space.



Moudle 3 : MASS Workflow

SCPline | Flow Cytometry | Single Cell Multiomics | **MASS Workflow** | Help

MASS Workflow

- Overview
- Upload Data
- Preprocess Data
- PCA
- UMAP
- Download Results

UMAP

UMAP is a dimensionality reduction technique that can be used for visualizing high-dimensional data in a low-dimensional space.

- Perform clustering analysis to cluster the data.
- Perform UMAP analysis to visualize the data.

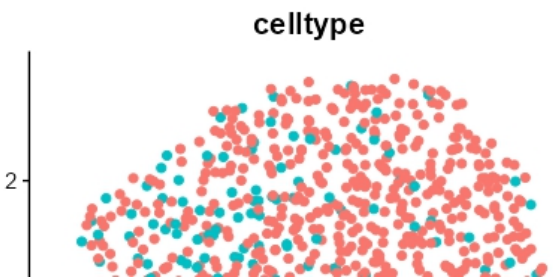
Run UMAP

UMAP Plot

UMAP is a dimensionality reduction technique that can be used for visualizing high-dimensional data in a low-dimensional space.

Group by

celltype



The UMAP shows the clustering of data in a low-dimensional space. You can select different properties, such as celltype, to see the clustering results of your data to better understand the distribution of different cell types.

Moudle 3 : MASS Workflow