Package ‘SlimR’

July 29, 2025

Title Marker-Based Package for Single-Cell and Spatial-Transcriptomic Annotation

Version 1.0.3

Description

Annotating single-cell and spatial-transcriptomic (ST) data based on the Marker dataset. It supports the creation of a unified marker list, Markers_list, using sources including: the package's built-in curated species-specific cell type and marker reference databases (e.g., 'Cellmarker2', 'PanglaoDB'), Seurat objects containing cell label information, or user-provided Excel tables mapping cell types to markers.

Based on the Markers_list, 'SlimR' can iterate through different cell types to generate corresponding annotation reference plots (e.g., 'Markers_Dotplot', 'Metric_Heatmap', 'Mean_expression_Box_plot'). Furthermore, it enables one-click generation of an annotation heatmap ('Annotation_Heatmap') visualizing the relationship between input cell types and the reference marker list. For more details see Kabacoff (2015, ISBN:9781617291388) and Hu et al. (2023) <doi:10.1093/nar/gkac947> and Franzén et al. (2019) <doi:10.1093/

License GPL-3

Encoding UTF-8

RoxygenNote 7.3.2

Depends R (>= 3.5)

LazyData true

Date 2025-07-27

Imports cowplot, dplyr, ggplot2, patchwork, pheatmap, readxl, scales, Seurat, tidyr, tools, magrittr, tibble

Suggests crayon

NeedsCompilation no

Author Zhao qing Wang [aut, cre] (ORCID: <https://orcid.org/0000-0001-8348-7245>)

Maintainer Zhao qing Wang <851091628@qq.com>

Repository CRAN

Date/Publication 2025-07-29 13:20:02 UTC

Contents

`calculate_probability` *Calculate gene set expression and infer probabilities with control datasets (Use in package)*

Description

Calculate gene set expression and infer probabilities with control datasets (Use in package)

Usage

```
calculate_probability(
  object,
  features,
  assay = NULL,
  cluster_col = NULL,
  min_expression = 0.1,
  specificity_weight = 3
)
```

Arguments

`object`	Enter a Seurat object.
`features`	Enter one or a set of markers.
`assay`	Enter the assay used by the Seurat object, such as "RNA". Default parameters use "assay = NULL".

`cluster_col`	Enter the meta.data column in the Seurat object to be annotated, such as "seurat_cluster". Default parameters use "cluster_col = NULL".
`min_expression`	The min_expression parameter defines a threshold value to determine whether a cell's expression of a feature is considered "expressed" or not. It is used to filter out low-expression cells that may contribute noise to the analysis. Default parameters use "min_expression = 0.1".
`specificity_weight`	The specificity_weight parameter controls how much the expression variability (standard deviation) of a feature within a cluster contributes to its "specificity score." It amplifies or suppresses the impact of variability in the final score calculation.Default parameters use "specificity_weight = 3".

Value

Average expression of genes in the input "Seurat" object given "cluster_col" and given "features".

`Cellmarker2` *Cellmarker2 dataset*

Description

A dataset containing marker genes for different cell types from Cellmarker2

Usage

```
Cellmarker2
```

Format

A data frame with 8 columns:

Details

This dataset is used to filter and create a standardized marker list. The dataset can be filtered based on species, tissue class, tissue type, cancer type, and cell type to generate a list of marker genes for specific cell types.

Source

`http://117.50.127.228/CellMarker/`

`Cellmarker2_raw` *Cellmarker2 raw dataset*

Description

A dataset containing marker genes for different cell types from Cellmarker2

Usage

```
Cellmarker2_raw
```

Format

A data frame with 20 columns contined in the Cellmarker2 database:

Details

This dataset is used to filter and create a standardized marker list. The dataset can be filtered based on species, tissue class, tissue type, cancer type, and cell type to generate a list of marker genes for specific cell types.

Source

`http://117.50.127.228/CellMarker/`

`Cellmarker2_table` *Cellmarker2 table*

Description

A dataset containing marker genes for different cell types from Cellmarker2

Usage

```
Cellmarker2_table
```

Format

A list contain different types like species, tissue_class, tissue_type, cancer_type, cell_type

Details

This list is used to choose filters for creation of standardized marker list.

Source

`http://117.50.127.228/CellMarker/`

`Celltype_annotation_Box`
Uses "marker_list" to generate Box plot for cell annotation

Description

Uses "marker_list" to generate Box plot for cell annotation

Usage

```
Celltype_annotation_Box(
  seurat_obj,
  gene_list,
  species,
  cluster_col = "seurat_clusters",
  assay = "RNA",
  save_path = NULL,
  metric_names = NULL
)
```

Arguments

`seurat_obj`	Enter the Seurat object with annotation columns such as "seurat_cluster" in meta.data to be annotated.
`gene_list`	A list of cells and corresponding gene controls, the name of the list is cell type, and the first column of the list corresponds to markers. Lists can be generated using functions such as "Markers_filter_Cellmarker2 ()", "Markers_filter_PanglaoDB ()", "read_excel_markers ()", "read_seurat_markers ()", etc.
`species`	This parameter selects the species "Human" or "Mouse" for standard gene format correction of markers entered by "Marker_list".
`cluster_col`	Enter annotation columns such as "seurat_cluster" in meta.data of the Seurat object to be annotated. Default parameters use "cluster_col = "seurat_clusters"".
`assay`	Enter the assay used by the Seurat object, such as "RNA". Default parameters use "assay = "RNA"".
`save_path`	The output path of the cell annotation picture. Default parameters use "save_path = "./SlimR/Celltype_annotation_Bar/"".
`metric_names`	Warning: Do not enter information. This parameter is used to check if "Marker_list" conforms to the SlimR output.

Value

The cell annotation picture is saved in "save_path".

Examples

```
## Not run: Celltype_annotation_Box(seurat_obj = sce,
        gene_list = Markers_list,
        species = "Human",
        cluster_col = "seurat_clusters",
        assay = "RNA",
        save_path = file.path(tempdir(),"SlimR_Celltype_annotation_Box")
        )

## End(Not run)
```

`Celltype_annotation_Cellmarker2`

Uses "marker_list" from Cellmarker2 for cell annotation

Description

Uses "marker_list" from Cellmarker2 for cell annotation

Usage

```
Celltype_annotation_Cellmarker2(
  seurat_obj,
  gene_list,
  species,
  cluster_col = "seurat_clusters",
  assay = "RNA",
  save_path = NULL,
  min_counts = 1
)
```

Arguments

`seurat_obj`	Enter the Seurat object with annotation columns such as "seurat_cluster" in meta.data to be annotated.
`gene_list`	Enter the standard "Marker_list" generated by the Cellmarker2 database for the SlimR package, generated by the "Markers_filter_Cellmarker2 ()" function.
`species`	This parameter selects the species "Human" or "Mouse" for standard gene format correction of markers entered by "Marker_list".
`cluster_col`	Enter annotation columns such as "seurat_cluster" in meta.data of the Seurat object to be annotated. Default parameters use "cluster_col = "seurat_clusters"".
`assay`	Enter the assay used by the Seurat object, such as "RNA". Default parameters use "assay = "RNA"".
`save_path`	The output path of the cell annotation picture. Default parameters use "save_path = "./SlimR/Celltype_annotation_Cellmarker2/"".

`min_counts`	The minimum number of counts of genes in "Marker_list" entered. This number represents the number of the same gene in the same species and the same location in the Cellmarker2 database used for annotation of this cell type. Default parameters use "min_counts = 1".

Value

The cell annotation picture is saved in "save_path".

Examples

```
## Not run: Celltype_annotation_Cellmarker2(seurat_obj = sce,
        gene_list = Markers_list_Cellmarker2,
        species = "Human",
        cluster_col = "seurat_clusters",
        assay = "RNA",
        save_path = file.path(tempdir(),"SlimR_Celltype_annotation_Cellmarker2")
        )

## End(Not run)
```

`Celltype_annotation_Excel` *Uses "marker_list" from Excel input for cell annotation*

Description

Uses "marker_list" from Excel input for cell annotation

Usage

```
Celltype_annotation_Excel(
  seurat_obj,
  gene_list,
  species,
  cluster_col = "seurat_clusters",
  assay = "RNA",
  save_path = NULL,
  metric_names = NULL
)
```

Arguments

`seurat_obj`	Enter the Seurat object with annotation columns such as "seurat_cluster" in meta.data to be annotated.
`gene_list`	Enter the standard "Marker_list" generated by the Excel files database for the SlimR package, generated by the "read_excel_markers()" function.

`species`	This parameter selects the species "Human" or "Mouse" for standard gene format correction of markers entered by "Marker_list".
`cluster_col`	Enter annotation columns such as "seurat_cluster" in meta.data of the Seurat object to be annotated. Default parameters use "cluster_col = "seurat_clusters"".
`assay`	Enter the assay used by the Seurat object, such as "RNA". Default parameters use "assay = "RNA"".
`save_path`	The output path of the cell annotation picture. Default parameters use "save_path = "./SlimR/Celltype_annotation_Excel/"".
`metric_names`	Warning: Do not enter information. This parameter is used to check if "Marker_list" conforms to the Excel files output.

Value

The cell annotation picture is saved in "save_path".

Examples

```
## Not run: Celltype_annotation_Excel(seurat_obj = sce,
         gene_list = Markers_list_Excel,
         species = "Human",
         cluster_col = "RNA_snn_res.0.4",
         assay = "RNA",
         save_path = file.path(tempdir(),"SlimR_Celltype_annotation_Excel")
         )

## End(Not run)
```

`Celltype_annotation_Heatmap`

Uses "marker_list" to generate heatmaps for cell annotation

Description

Uses "marker_list" to generate heatmaps for cell annotation

Usage

```
Celltype_annotation_Heatmap(
  seurat_obj,
  gene_list,
  species,
  cluster_col = "seurat_clusters",
  assay = "RNA",
  min_expression = 0.1,
  specificity_weight = 3
)
```

Arguments

`seurat_obj`	Enter the Seurat object with annotation columns such as "seurat_cluster" in meta.data to be annotated.
`gene_list`	A list of cells and corresponding gene controls, the name of the list is cell type, and the first column of the list corresponds to markers. Lists can be generated using functions such as "Markers_filter_Cellmarker2 ()", "Markers_filter_PanglaoDB ()", "read_excel_markers ()", "read_seurat_markers ()", etc.
`species`	This parameter selects the species "Human" or "Mouse" for standard gene format correction of markers entered by "Marker_list".
`cluster_col`	Enter annotation columns such as "seurat_cluster" in meta.data of the Seurat object to be annotated. Default parameters use "cluster_col = "seurat_clusters"".
`assay`	Enter the assay used by the Seurat object, such as "RNA". Default parameters use "assay = "RNA".
`min_expression`	The min_expression parameter defines a threshold value to determine whether a cell's expression of a feature is considered "expressed" or not. It is used to filter out low-expression cells that may contribute noise to the analysis. Default parameters use "min_expression = 0.1".
`specificity_weight`	The specificity_weight parameter controls how much the expression variability (standard deviation) of a feature within a cluster contributes to its "specificity score." It amplifies or suppresses the impact of variability in the final score calculation.Default parameters use "specificity_weight = 3".

Value

The heatmap of the comparison between "cluster_col" in the Seurat object and the given gene set "gene_list" needs to be annotated.

Examples

```
## Not run: Celltype_annotation_Heatmap(seurat_obj = sce,
        gene_list = Markers_list,
        species = "Human",
        cluster_col = "seurat_clusters",
        assay = "RNA",
        min_expression = 0.1,
        specificity_weight = 3
        )

## End(Not run)
```

`Celltype_annotation_PanglaoDB`

Uses "marker_list" from PanglaoDB for cell annotation

Description

Uses "marker_list" from PanglaoDB for cell annotation

Usage

```
Celltype_annotation_PanglaoDB(
  seurat_obj,
  gene_list,
  species,
  cluster_col = "seurat_clusters",
  assay = "RNA",
  save_path = NULL,
  metric_names = NULL
)
```

Arguments

`seurat_obj`	Enter the Seurat object with annotation columns such as "seurat_cluster" in meta.data to be annotated.
`gene_list`	Enter the standard "Marker_list" generated by the PanglaoDB database for the SlimR package, generated by the "Markers_filter_PanglaoDB ()" function.
`species`	This parameter selects the species "Human" or "Mouse" for standard gene format correction of markers entered by "Marker_list".
`cluster_col`	Enter annotation columns such as "seurat_cluster" in meta.data of the Seurat object to be annotated. Default parameters use "cluster_col = "seurat_clusters"".
`assay`	Enter the assay used by the Seurat object, such as "RNA". Default parameters use "assay = "RNA"".
`save_path`	The output path of the cell annotation picture. Default parameters use "save_path = "./SlimR/Celltype_annotation_PanglaoDB/"".
`metric_names`	Warning: Do not enter information. This parameter is used to check if "Marker_list" conforms to the PanglaoDB database output.

Value

The cell annotation picture is saved in "save_path".

Examples

```
## Not run: Celltype_annotation_PanglaoDB(seurat_obj = sce,
        gene_list = Markers_list_panglaoDB,
        species = "Human",
        cluster_col = "seurat_clusters",
        assay = "RNA",
        save_path = file.path(tempdir(),"SlimR_Celltype_annotation_PanglaoDB")
        )

## End(Not run)
```

`Celltype_annotation_Seurat`

Uses "marker_list" from Seurat object for cell annotation

Description

Uses "marker_list" from Seurat object for cell annotation

Usage

```
Celltype_annotation_Seurat(
  seurat_obj,
  gene_list,
  species,
  cluster_col = "seurat_clusters",
  assay = "RNA",
  save_path = NULL,
  metric_names = NULL
)
```

Arguments

`seurat_obj`	Enter the Seurat object with annotation columns such as "seurat_cluster" in meta.data to be annotated.
`gene_list`	Enter the standard "Marker_list" generated by the Seurat object database for the SlimR package, generated by the "read_seurat_markers()" function.
`species`	This parameter selects the species "Human" or "Mouse" for standard gene format correction of markers entered by "Marker_list".
`cluster_col`	Enter annotation columns such as "seurat_cluster" in meta.data of the Seurat object to be annotated. Default parameters use "cluster_col = "seurat_clusters"".
`assay`	Enter the assay used by the Seurat object, such as "RNA". Default parameters use "assay = "RNA"".
`save_path`	The output path of the cell annotation picture. Default parameters use "save_path = "./SlimR/Celltype_annotation_Seurat/"".
`metric_names`	Warning: Do not enter information. This parameter is used to check if "Marker_list" conforms to the Seurat object output.

Value

The cell annotation picture is saved in "save_path".

Examples

```
## Not run: Celltype_annotation_Seurat(seurat_obj = sce,
        gene_list = Markers_list_Seurat,
        species = "Human",
        cluster_col = "seurat_clusters",
        assay = "RNA",
        save_path = file.path(tempdir(),"SlimR_Celltype_annotation_Seurat")
        )

## End(Not run)
```

`Markers_filter_Cellmarker2` *Create Marker_list from the Cellmarkers2 database*

Description

Create Marker_list from the Cellmarkers2 database

Usage

```
Markers_filter_Cellmarker2(
  df,
  species = NULL,
  tissue_class = NULL,
  tissue_type = NULL,
  cancer_type = NULL,
  cell_type = NULL,
  cell_name = NULL,
  marker = NULL,
  counts = NULL
)
```

Arguments

`df`	Standardized Cellmarkers2 database. It is read as data(Cellmarkers2) in the SlimR library.
`species`	Species information in Cellmarkers2 database. The default input is "Human" or "Mouse".The input can be retrieved by "Cellmarkers2_table". For more information,please refer to http://117.50.127.228/CellMarker/ on Cellmarkers2's official website.

`tissue_class`	Tissue_class information in Cellmarkers2 database. The input can be retrieved by "Cellmarkers2_table". For more information, please refer to http://117.50.127.228/CellMarker/ on Cellmarkers2's official website.
`tissue_type`	Tissue_type information in Cellmarkers2 database. The input can be retrieved by "Cellmarkers2_table". For more information, please refer to http://117.50.127.228/CellMarker/ on Cellmarkers2's official website.
`cancer_type`	Cancer_type information in Cellmarkers2 database. The input can be retrieved by "Cellmarkers2_table". For more information, please refer to http://117.50.127.228/CellMarker/ on Cellmarkers2's official website.
`cell_type`	Cell_type information in Cellmarkers2 database. The input can be retrieved by "Cellmarkers2_table". For more information, please refer to http://117.50.127.228/CellMarker/ on Cellmarkers2's official website.
`cell_name`	Warning: Do not enter information, this parameter is used to return results for saving
`marker`	Warning: Do not enter information, this parameter is used to return results for saving
`counts`	Warning: Do not enter information, this parameter is used to return results for saving

Value

The standardized "Marker_list" in the SlimR package

Examples

```
Cellmarker2 <- SlimR::Cellmarker2
          Markers_list_Cellmarker2 <- Markers_filter_Cellmarker2(
          Cellmarker2,
          species = "Human",
          tissue_class = "Intestine",
          tissue_type = NULL,
          cancer_type = NULL,
          cell_type = NULL,
          cell_name = NULL,
          marker = NULL,
          counts = NULL
          )
```

`Markers_filter_PanglaoDB`

Create Marker_list from the PanglaoDB database

Description

Create Marker_list from the PanglaoDB database

Usage

```
Markers_filter_PanglaoDB(df, species_input, organ_input)
```

Arguments

`df`	Standardized PanglaoDB database. It is read as data(PanglaoDB) in the SlimR library.
`species_input`	Species information in PanglaoDB database. The default input is "Human" or "Mouse".The input can be retrieved by "PanglaoDB_table". For more information,please refer to https://panglaodb.se/ on PanglaoDB's official website.
`organ_input`	Organ type information in the PanglaoDB database. The input can be retrieved by "PanglaoDB_table".For more information, please refer to https://panglaodb.se/ on PanglaoDB's official website.

Value

The standardized "Marker_list" in the SlimR package

Examples

```
PanglaoDB <- SlimR::PanglaoDB
          Markers_list_panglaoDB <- Markers_filter_PanglaoDB(
          PanglaoDB,
          species_input = 'Human',
          organ_input = 'GI tract'
          )
```

`PanglaoDB` *PanglaoDB dataset*

Description

A dataset containing marker genes for different cell types from PanglaoDB

Usage

```
PanglaoDB
```

Format

A data frame with 9 columns:

Details

This dataset is used to filter and create a standardized marker list.'

Source

https://panglaodb.se/

`PanglaoDB_raw` *PanglaoDB raw dataset*

Description

A dataset containing marker genes for different cell types from PanglaoDB

Usage

```
PanglaoDB_raw
```

Format

A data frame with 14 columns contined in the PanglaoDB database:

Details

This dataset is used to filter and create a standardized marker list.'

Source

https://panglaodb.se/

`PanglaoDB_table` *PanglaoDB table*

Description

A dataset containing marker genes for different cell types from PanglaoDB

Usage

```
PanglaoDB_table
```

Format

A list contain different types like species, organ, cell type.

Details

This list is used to choose filters for creation of standardized marker list.

Source

https://panglaodb.se/

`plot_mean_expression` *Counts average expression of gene set and plots Boxplot (Use in package)*

Description

Counts average expression of gene set and plots Boxplot (Use in package)

Usage

```
plot_mean_expression(object, features, assay = NULL, cluster_col = NULL)
```

Arguments

`object`	Enter a Seurat object.
`features`	Enter one or a set of markers.
`assay`	Enter the assay used by the Seurat object, such as "RNA". Default parameters use "assay = NULL".
`cluster_col`	Enter the meta.data column in the Seurat object to be annotated, such as "seurat_cluster". Default parameters use "cluster_col = NULL".

Value

Average expression box plot of genes in the input "Seurat" object given "cluster_col" and given "features".

`read_excel_markers` *Create "Marker_list" from Excel files ".xlsx"*

Description

Create "Marker_list" from Excel files ".xlsx"

Usage

```
read_excel_markers(path)
```

Arguments

`path`	The path information of Marker files stored in ".xlsx" format. The Sheet name in the file is filled with cell type. The first line of each Sheet is the table head, the first column is filled with markers information, and the following column is filled with mertic information.

Value

The standardized "Marker_list" in the SlimR package.

Examples

```
## Not run: Markers_list_Excel <- read_excel_markers(
         "D:/Laboratory/Marker_load.xlsx"
         )

## End(Not run)
```

`read_seurat_markers` *Create "Marker_list" from Seurat object*

Description

Create "Marker_list" from Seurat object

Usage

```
read_seurat_markers(df, sort_by = "avg_log2FC", gene_filter = 20)
```

Arguments

`df`	Dataframe generated by "FindAllMarkers" function, recommend to use parameter "group.by = "Cell_type"" and "only.pos = TRUE".
`sort_by`	Marker sorting parameter, select "avg_log2FC" or "p_val_adj". Default parameters use "sort_by = "avg_log2FC"".
`gene_filter`	The number of markers left for each cell type based on the "sort_by" parameter's level of difference. Default parameters use "gene_fliter = 20"

Value

The standardized "Marker_list" in the SlimR package.

Examples

```
library(Seurat)
          data("pbmc_small")
          sce <- pbmc_small
          seurat_markers <- FindAllMarkers(sce,
                                           group.by = "RNA_snn_res.1",
                                           only.pos = TRUE)
          Markers_list_Seurat <- read_seurat_markers(seurat_markers,
          sort_by = "avg_log2FC",
          gene_filter = 10
          )
```

Index