

Package: phenomenalist (via r-universe)

October 24, 2024

Type Package

Title Analysis Toolkit for PhenoCycler (CODEX) Data in R

Version 0.0.0.9003

Description A collection of tools for cleaning, clustering, and plotting PhenoCycler (CODEX) data.

License MIT + file LICENSE

URL <https://igordot.github.io/phenomenalist/>

BugReports <https://github.com/igordot/phenomenalist/issues>

Depends ggplot2, R (>= 4.1), SpatialExperiment

Imports ComplexHeatmap, cowplot, data.table, dplyr, FNN, ggsci, glue, igraph (>= 1.2.7), janitor (>= 1.2.0), methods, RColorBrewer, readr, rlang (>= 0.4.11), scattermore, scan, scuttle, SingleCellExperiment, stringr, SummarizedExperiment, tibble, tidyr, tidyselect, uwot

Suggests covr, knitr, remotes, rmarkdown, roxygen2, testthat (>= 3.0.0)

VignetteBuilder knitr

Config/testthat/edition 3

Encoding UTF-8

RoxygenNote 7.3.1

Roxygen list(markdown = TRUE)

Repository <https://igordot.r-universe.dev>

RemoteUrl <https://github.com/igordot/phenomenalist>

RemoteRef HEAD

RemoteSha c35f531d0035ac6ea879bfe1b0340c6b5cc6b09e

Contents

clean_col_names	2
cluster	2
create_object	3
detect_exprs_cols	4
plot_distribution	5
plot_dr	5
plot_heatmap	6
plot_scatter	7
plot_spatial	8
run_umap	8
transform	9

Index	11
--------------	-----------

clean_col_names	<i>Clean up data frame column names</i>
-----------------	---

Description

Clean up data frame column names

Usage

```
clean_col_names(x)
```

Arguments

x A data frame.

Value

A data frame.

cluster	<i>Perform clustering</i>
---------	---------------------------

Description

Cluster similar cells based on their expression profiles.

Usage

```
cluster(  
  x,  
  method = c("leiden"),  
  resolution = 1,  
  n_neighbors = 50,  
  out_dir = NULL  
)
```

Arguments

x	A SpatialExperiment object.
method	Clustering method (only leiden is implemented currently).
resolution	Value of the parameter controlling the coarseness of the clusters (higher resolution yields more clusters).
n_neighbors	The maximum number of nearest neighbors to compute.
out_dir	Name of the output analysis directory. If specified, the object and the corresponding plots will be saved there.

Value

A [SpatialExperiment](#) object.

create_object	<i>Create a SpatialExperiment object</i>
---------------	--

Description

Create a [SpatialExperiment](#) object (used in the same manner as [SingleCellExperiment](#)) that will store all of the project data.

Usage

```
create_object(  
  x,  
  expression_cols = NULL,  
  metadata_cols = NULL,  
  skip_cols = NULL,  
  clean_names = TRUE,  
  transformation = NULL,  
  out_dir = NULL  
)
```

Arguments

x	A data frame or a path to a file (can be gzipped).
expression_cols	Column names of markers/antibody expression as a vector or a grep pattern. Auto-detected if not specified.
metadata_cols	Column names of cell metadata (not markers/antibodies) as a vector or a grep pattern. Auto-detected if not specified.
skip_cols	Column names to ignore (such as bad antibodies) as a vector or a grep pattern.
clean_names	A logical scalar. Clean the data frame column names to remove problematic characters and make them unique.
transformation	A character string indicating which transformation method should be used. See transform() .
out_dir	Name of the output analysis directory. If specified, the object and the corresponding plots will be saved there.

Value

A [SpatialExperiment](#) object.

Examples

```
tonsil_csv <- system.file("extdata", "tonsil-akoya-2018-500.csv", package = "phenomenalist")
tonsil_spe <- create_object(tonsil_csv, skip_cols = "DAPI|Blank", transformation = "z")
```

detect_exprs_cols *Detect marker/antibody column names*

Description

The input is generally a table of both expression values and various metadata columns. This function attempts to determine which of the columns contain expression values.

Usage

```
detect_exprs_cols(x)
```

Arguments

x	A data frame.
---	---------------

Value

A vector of column names.

plot_distribution	<i>Plot expression distribution</i>
-------------------	-------------------------------------

Description

Plot distribution of expression intensities, split by marker.

Usage

```
plot_distribution(x, assay = "exprs", out_dir = NULL)
```

Arguments

x	A SpatialExperiment object.
assay	A character string indicating which values should be used.
out_dir	Name of the output analysis directory. If specified, the plots will be saved there.

Value

.

plot_dr	<i>Plot dimensionality reduction</i>
---------	--------------------------------------

Description

Plot cell-level reduced dimension results stored in a [SpatialExperiment](#) object.

Usage

```
plot_dr(  
  x,  
  dr,  
  color_by,  
  assay = "logcounts",  
  smooth = FALSE,  
  range = c(0.01, 0.99),  
  out_dir = NULL  
)
```

Arguments

x	A SpatialExperiment object.
dr	.
color_by	Column metadata field(s) or feature(s) to color by.
assay	A character string indicating which expression values should be used.
smooth	A logical scalar. Smooth values. Helps to visualize expression patterns in a plot with many overlapping points.
range	A vector of 2 values indicating the minimum and maximum percentiles for the color range. Helps to visualize expression patterns when extreme outliers are present. For example, <code>c(0, 0.99)</code> will not expand the color scale above 99th percentile.
out_dir	Name of the output analysis directory. If specified, the plots will be saved there.

Value

.

plot_heatmap

Generate a heatmap

Description

Generate a heatmap

Usage

```
plot_heatmap(x, group_by, assay = "logcounts", out_dir = NULL)
```

Arguments

x	A SpatialExperiment object.
group_by	Column metadata field(s) to group by.
assay	A character string indicating which expression values should be used.
out_dir	Name of the output analysis directory. If specified, the plots will be saved there.

Value

A [Heatmap](#) object.

`plot_scatter`*Generate a generic scatter plot*

Description

Create a scatter plot to display the relationship between two continuous variables. This is essentially a wrapper for `ggplot2::geom_point()`.

Usage

```
plot_scatter(  
  data,  
  x,  
  y,  
  color_by,  
  smooth = FALSE,  
  range = c(0.01, 0.99),  
  title = "",  
  aspect_ratio = 1  
)
```

Arguments

<code>data</code>	A data frame.
<code>x</code>	.
<code>y</code>	.
<code>color_by</code>	Column metadata field(s) or feature(s) to color by.
<code>smooth</code>	A logical scalar. Smooth values. Helps to visualize expression patterns in a plot with many overlapping points.
<code>range</code>	A vector of 2 values indicating the minimum and maximum percentiles for the color range. Helps to visualize expression patterns when extreme outliers are present. For example, <code>c(0, 0.99)</code> will not expand the color scale above 99th percentile.
<code>title</code>	Plot title.
<code>aspect_ratio</code>	Aspect ratio of the panel.

Value

A ggplot object.

plot_spatial	<i>Plot cells on their spatial coordinates</i>
--------------	--

Description

Plot cell-level physical coordinates stored in a SpatialExperiment object.

Usage

```
plot_spatial(
  x,
  color_by,
  assay = "logcounts",
  smooth = FALSE,
  range = c(0.01, 0.99),
  out_dir = NULL
)
```

Arguments

x	A SpatialExperiment object.
color_by	Column metadata field(s) or feature(s) to color by.
assay	A character string indicating which expression values should be used.
smooth	A logical scalar. Smooth values. Helps to visualize expression patterns in a plot with many overlapping points.
range	A vector of 2 values indicating the minimum and maximum percentiles for the color range. Helps to visualize expression patterns when extreme outliers are present. For example, c(0, 0.99) will not expand the color scale above 99th percentile.
out_dir	Name of the output analysis directory. If specified, the plots will be saved there.

Value

.

run_umap	<i>Run UMAP dimensionality reduction</i>
----------	--

Description

Run UMAP dimensionality reduction

Usage

```
run_umap(
  x,
  assay = "exprs",
  n_neighbors = 50,
  min_dist = 0.01,
  n_threads = NULL,
  out_dir = NULL
)
```

Arguments

x	A SpatialExperiment object.
assay	.
n_neighbors	The size of local neighborhood (in terms of number of neighboring sample points) used for manifold approximation. Larger values result in more global views of the manifold, while smaller values result in more local data being preserved. In general values should be in the range 2 to 100.
min_dist	The effective minimum distance between embedded points. Smaller values will result in a more clustered/clumped embedding where nearby points on the manifold are drawn closer together, while larger values will result on a more even dispersal of points. The value should be set relative to the spread value, which determines the scale at which embedded points will be spread out.
n_threads	Number of threads to use (except during stochastic gradient descent). Default is half the number of concurrent threads supported by the system. For nearest neighbor search, only applies if <code>nn_method = "annoy"</code> . If <code>n_threads > 1</code> , then the Annoy index will be temporarily written to disk in the location determined by tempfile .
out_dir	Name of the output analysis directory. If specified, the object will be saved there.

Value

A [SpatialExperiment](#) object.

transform	<i>Transform/normalize expression data</i>
-----------	--

Description

Transform/normalize expression data. The options are currently:

- log: $\log_{10} + 1$
- z: Z normalization of each marker (recommended by Hickey et al.)

Usage

```
transform(x, method = c("log", "z"), out_dir = NULL)
```

Arguments

x	A SpatialExperiment object.
method	A character string indicating which transformation method should be used.
out_dir	Name of the output directory. If specified, the density plots will be saved there.

Value

A [SpatialExperiment](#) object with transformed values in the exprs assay.

Index

`clean_col_names`, 2
`cluster`, 2
`create_object`, 3

`detect_exprs_cols`, 4

`ggplot2::geom_point()`, 7

Heatmap, 6

`plot_distribution`, 5
`plot_dr`, 5
`plot_heatmap`, 6
`plot_scatter`, 7
`plot_spatial`, 8

`run_umap`, 8

SingleCellExperiment, 3
SpatialExperiment, 3–6, 8–10

`tempfile`, 9
`transform`, 9
`transform()`, 4