Package ‘SCOPRO’

July 21, 2025

Type Package

Title Score Projection Between in 'Vivo' and in 'Vitro' Datasets

Version 0.1.0

Author Gabriele Lubatti

Maintainer Gabriele Lubatti <`gabriele.lubatti@helmholtz-muenchen.de`>

Description Assigns a score projection from 0 to 1 between a given in 'vivo' stage and each single cluster from an in 'vitro' dataset. The score is assigned based on the the fraction of specific markers of the in 'vivo' stage that are conserved in the in 'vitro' clusters <`https://github.com/ScialdoneLab`>.

License GPL-3

Depends R (>= 4.0)

Imports ggplot2

Suggests CIARA, rmarkdown, testthat (>= 3.0.0), knitr

VignetteBuilder knitr

biocViews software

BugReports `https://github.com/ScialdoneLab/SCOPRO/issues`

Encoding UTF-8

RoxygenNote 7.1.1

Config/testthat/edition 3

NeedsCompilation no

Repository CRAN

Date/Publication 2022-08-23 08:10:05 UTC

Contents

`filter_in_vitro` *filter_in_vitro*

Description

For a given gene in in *marker_all*, if the fraction of cells in one or more clusters with an expression above *threshold* is greater than *fraction*, then the gene is kept

Usage

```
filter_in_vitro(
  norm_vitro,
  cluster_vitro,
  marker_all,
  fraction = 0.1,
  threshold = 0
)
```

Arguments

`norm_vitro`	Norm count matrix (n_genes X n_cells) for in vitro dataset
`cluster_vitro`	cluster for in vitro dataset
`marker_all`	First element of the list given as output by the function *select_top_markers*
`fraction`	Numeric value.
`threshold`	Numeric value

Value

Character vector with the names of kept genes

Author(s)

Gabriele Lubatti <`gabriele.lubatti@helmholtz-muenchen.de`>

`plot_score` *plot_score*

Description

plot_score

Usage

```
plot_score(
  SCOPRO_output,
  marker_stages,
  marker_stages_filter,
  selected_stages,
  name_vivo,
  y_name,
  fill_name,
  title_name
)
```

Arguments

`SCOPRO_output`	output given by function *SCOPRO*
`marker_stages`	Second element of the list given as output by the function *select_top_markers*
`marker_stages_filter`	output from the function *filter_in_vitro*
`selected_stages`	In vivo stages for which the markers where computed with the function *select_top_markers*
`name_vivo`	name of the in vivo stage on which SCOPRO is run
`y_name`	Character value
`fill_name`	Character value.
`title_name`	Character value.

Value

ggplot2::ggplot2 object.

Author(s)

Gabriele Lubatti <`gabriele.lubatti@helmholtz-muenchen.de`>

`plot_score_genes` *plot_score_genes*

Description

plot_score_genes

Usage

```
plot_score_genes(
  markers_to_plot,
  label_1,
  label_2,
  norm_vitro,
  norm_vivo,
  cluster_vitro,
  cluster_vivo,
  final_name,
  max_size = 9,
  text_size = 9.5,
  title_name
)
```

Arguments

`markers_to_plot`	Character vector with the names of the genes to plot.
`label_1`	Character value. Label for the in vitro dataset
`label_2`	Character value. Label for the in vivo dataset
`norm_vitro`	Norm count matrix (n_genes X n_cells) for in vitro dataset
`norm_vivo`	Norm count matrix (n_genes X n_cells) for in vivo dataset
`cluster_vitro`	cluster for in vitro dataset
`cluster_vivo`	cluster for in vivo dataset
`final_name`	Character vector with the names of the genes to show in the plot.
`max_size`	Numeric value, specifying the size of the dot.
`text_size`	Numeric value, specifying the size of the text in the plot.
`title_name`	Character value.

Value

ggplot2::ggplot2 object showing balloon plot.

Author(s)

Gabriele Lubatti `<gabriele.lubatti@helmholtz-muenchen.de>`

`SCOPRO` *SCOPRO*

Description

The mean expression profile of *marker_stages_filter* genes is computed for each cluster in the in vivo and in vitro dataset. For a given cluster, a connectivity matrix is computed with number of rows and number of columns equal to the length of *marker_stages_filter*. Each entry (i,j) in the matrix can be 1 if the fold_change between gene i and gene j is above *fold_change*. Otherwise is 0. Finally the connectivity matrix of a given *name_vivo* stage and all the clusters in the in vitro dataset are compared. A gene i is considered to be conserved between *name_vivo* and an in vitro cluster if the jaccard index of the links of gene i is above *threshold*.

Usage

```
SCOPRO(
  norm_vitro,
  norm_vivo,
  cluster_vitro,
  cluster_vivo,
  name_vivo,
  marker_stages_filter,
  threshold = 0.1,
  number_link = 1,
  fold_change = 3,
  threshold_fold_change = 0.1,
  marker_stages,
  selected_stages
)
```

Arguments

`norm_vitro`	Norm count matrix (n_genes X n_cells) for in vitro dataset
`norm_vivo`	Norm count matrix (n_genes X n_cells) for in vivo dataset
`cluster_vitro`	cluster for in vitro dataset
`cluster_vivo`	cluster for in vivo dataset
`name_vivo`	name of the in vivo stage on which SCOPRO is run
`marker_stages_filter`	output from the function *filter_in_vitro*
`threshold`	Numeric value. For a given gene, the jaccard index between the links from the in vivo and in vitro datasets is computed. If the jaccard index is above *threshold*, then the gene is considered to be conserved between the two datasets.
`number_link`	Numeric value. For a given gene in the in vivo dataset with links above *number_link*, the jaccard index between the links from in vitro and in vivo dataset is computed.

`fold_change`	Numeric value. For a given gene, the fold change between all the other genes is computed. If fold change is above *fold_change*, then there is a link with weight 1 between the two genes.
`threshold_fold_change`	Numeric value. Above *threshold* the fold change between genes is computed. Below *threshold* the difference between genes is computed.
`marker_stages`	Second element of the list given as output by the function *select_top_markers*
`selected_stages`	In vivo stages for which the markers where computed with the function *select_top_markers*

Value

A list with five elements:

`common_link`	Vector with the names of the genes conserved between *name_vivo* and all the clusters in the vitro dataset
`no_common_link`	Vector with the names of the genes not conserved between *name_vivo* and the clusters in the vitro dataset
`link_kept`	List with the names of the genes conserved between *name_vivo* and each single cluster in the vitro dataset
`link_no_kept`	List with the names of the genes not conserved between *name_vivo* and each single cluster in the vitro dataset
`final_score`	Numeric value, given by the fraction of conserved markers of *name_vivo* and each single cluster in the in vitro dataset

Author(s)

Gabriele Lubatti <`gabriele.lubatti@helmholtz-muenchen.de`>

Examples

```
load(system.file("extdata", "norm_es_vitro_small.Rda", package = "SCOPRO"))
n_es= norm_es_vitro_small
load(system.file("extdata", "norm_vivo_small.Rda", package = "SCOPRO"))
n_v = norm_vivo_small
load(system.file("extdata", "cluster_es_vitro_small.Rda", package = "SCOPRO"))
c_es=cluster_es_vitro_small
load(system.file("extdata", "cluster_vivo_small.Rda", package = "SCOPRO"))
c_v=cluster_vivo_small
load(system.file("extdata", "marker_stages_filter.Rda", package = "SCOPRO"))
m_s_f = marker_stages_filter
load(system.file("extdata", "marker_stages.Rda", package = "SCOPRO"))
m_s = marker_stages
stages = c("Late_2_cell","epiblast_4.5","epiblast_5.5","epiblast_6.5")
output_SCOPRO = SCOPRO(n_es,n_v,c_es,c_v,"Late_2_cell",m_s_f,0.1,1,3,0.1,m_s,stages)
plot_score(output_SCOPRO,m_s,m_s_f,stages,"Late_2_cell","Score","Cluster","2-cells")
```

`select_common_genes` *select_common_genes*

Description

select_common_genes

Usage

```
select_common_genes(
  SCOPRO_output,
  marker_stages,
  selected_stages,
  name_vivo,
  cluster_vitro,
  name_vitro
)
```

Arguments

`SCOPRO_output`	output given by function *SCOPRO*
`marker_stages`	Second element of the list given as output by the function *select_top_markers*
`selected_stages`	In vivo stages for which the markers where computed with the function *select_top_markers*
`name_vivo`	name of the in vivo stage on which SCOPRO is run
`cluster_vitro`	cluster for in vitro dataset
`name_vitro`	name of the in vitro stage for which we want to know the conserved markers with the *name_vivo* stage

Value

Character vector with the names of the conserved markers of *name_vivo* stage in the *name_vitro* stage

Author(s)

Gabriele Lubatti `<gabriele.lubatti@helmholtz-muenchen.de>`

`select_no_common_genes` *select_no_common_genes*

Description

select_no_common_genes

Usage

```
select_no_common_genes(
  SCOPRO_output,
  marker_stages,
  selected_stages,
  name_vivo,
  cluster_vitro,
  name_vitro
)
```

Arguments

`SCOPRO_output`	output given by function *SCOPRO*
`marker_stages`	Second element of the list given as output by the function *select_top_markers*
`selected_stages`	In vivo stages for which the markers where computed with the function *select_top_markers*
`name_vivo`	name of the in vivo stage on which SCOPRO is run
`cluster_vitro`	cluster for in vitro dataset
`name_vitro`	name of the in vitro stage for which we want to know the non-conserved markers with the *name_vivo* stage

Value

Character vector with the names of the non-conserved markers of *name_vivo* stage in the *name_vitro* stage

Author(s)

Gabriele Lubatti <`gabriele.lubatti@helmholtz-muenchen.de`>

`select_top_markers` *select_top_markers*

Description

For each stage in *selected_stages*, starting from the markers given by *markers_cluster_seurat* function of the package CIARA, only the markers with a median above *threshold* in the stage and below *threshold* in all the other stages are kept.

Usage

```
select_top_markers(
  selected_stages,
  cluster_vivo,
  norm_vivo,
  markers_small,
  max_number = 100,
  threshold = 0.1
)
```

Arguments

`selected_stages`	Character vector with the name of the selected in vivo stages
`cluster_vivo`	cluster for in vivo dataset
`norm_vivo`	Norm count matrix (n_genes X n_cells) for in vivo dataset
`markers_small`	Output given by the function *markers_cluster_seurat* of the package CIARA
`max_number`	Numeric value. Maximum number of top markers to consider for each stage in *selected_stages*
`threshold`	Numeric value.

Value

A list with two elements:

`marker_all`	Vector with the union of all the *top_number* markers for each stage in *selected_stages*
`marker_stages`	List with length equal to number of stages in *selected_stages* . Each element contains the *top_number* markers for a given stage in *selected_stages*

Author(s)

Gabriele Lubatti <`gabriele.lubatti@helmholtz-muenchen.de`>

See Also

`https://CRAN.R-project.org/package=CIARA`

Index