Package ‘APackOfTheClones’

May 9, 2023

Type Package

Title Visualization of T-Cell Clonal Expansion

Version 0.1.2

Maintainer Qile Yang <qile0317@gmail.com>

Description Visualize T-cell clonal expansion via circle-packing. 'APackOfTheClones' can integrate single-cell RNA sequencing and T-cell receptor libraries from the 10X genomics Single Cell Immune Profiling and Cell Ranger pipeline, to produce a simple, publication-ready visualization of the clonal expansion. The method was originally implemented by Murray Christian and Ben Murrell in the following immunology study: Ma et al. (2021) <doi:10.1126/sciimmunol.abg6356>.

BugReports https://github.com/Qile0317/APackOfTheClones/issues/

Depends R (>= 3.5.0)

License MIT + file LICENSE

Encoding UTF-8

Language en-US

LazyData false

Suggests covr, knitr, rmarkdown, spelling, testthat (>= 3.0.0), vdiffr

Config/testthat/edition 3

Imports data.table, dplyr, ggforce, ggplot2, grDevices, hash, R6, rlang, Seurat, stats, utils

RoxygenNote 7.2.3

URL https://qile0317.github.io/APackOfTheClones/

NeedsCompilation no

Author Qile Yang [cre, aut, cph]

Repository CRAN

Date/Publication 2023-05-09 08:50:08 UTC
clonal_expansion_plot

**R topics documented:**

- clonal_expansion_plot ........................................... 2
- count_clone_sizes .............................................. 4
- integrate_tcr ................................................... 5
- mini_clonotype_data ........................................... 6
- mini_seurat_obj ................................................ 7

**Index**

8

---

**clonal_expansion_plot**  
*Visualize T cell clonal expansion with a ball-packing plot.*

**Description**

Integrates a cell ranger T cell library into a Seurat object with a UMAP reduction. Then gets sizes of unique clones and utilizes a circle-packing algorithm to pack circles representing individual clones in approximately the same UMAP coordinates and clusters into a ggplot object.

**Usage**

```r
clonal_expansion_plot(
  seurat_obj,
  tcr_df = "seurat_obj_already_integrated",
  res = 360,
  clone_scale_factor = 0.1,
  rad_scale_factor = 0.95,
  ORDER = TRUE,
  try_place = FALSE,
  verbose = TRUE,
  repulse = FALSE,
  repulsion_threshold = 1,
  repulsion_strength = 1,
  max_repulsion_iter = 10,
  use_default_theme = TRUE,
  show_origin = FALSE,
  retain_axis_scales = FALSE,
  add_size_legend = TRUE,
  legend_sizes = c(1, 5, 50),
  legend_position = "top_left",
  legend_buffer = 1.5,
  legend_color = "#808080",
  legend_spacing = 0.4
)
```
Arguments

seurat_obj  Seurat object with at least a UMAP reduction. Can either already have been integrated with a T cell library via `integrate_tcr(seurat_obj, tcr_df)' in which case the subsequent `tcr_df' argument can be left empty. Else, the object must be a regular seurat object and a T cell library must be inputted in the following `tcr_df' argument

tcr_df  If left empty, `seurat_obj' is assumed to be already integrated. Otherwise, should be a `data.frame' of the T cell library generated by 10X genomics' Cell Ranger. The dataframe has to at least have the `barcode' and `raw_clonotype_id' columns.

res  The number of points on the generated path per full circle. From plot viewers, if circles seem slightly too pixelated, it is highly recommended to first try to export the plot as an `.svg' before increasing `res'

clone_scale_factor  numeric. Decides how much to scale each circle. Usually should be kept at around 0.01 to somewhat maintain UMAP structure for large datasets. However, if the plot that is displayed is ever blank, first try increasing this value.

rad_scale_factor  numeric. indicates how much the radii of the clones should decrease to add a slight gap between all of them. Defaults to 1 but 0.85-0.95 values are recommended. Both `rad_scale_factor' and `clone_scale_factor' may need to be repeatedly readjusted

ORDER logical. Decides if the largest clones should be at the cluster centroids

try_place  If `TRUE', always minimizes distance from a newly placed circle to the origin

verbose  logical. Decides if visual cues print to the R console of the packing progress

repulse  If `TRUE', will attempt to push overlapping clusters away from each other.

repulsion_threshold  numeric. The radius that cluster overlap is acceptable

repulsion_strength  numeric. The smaller the value the less the clusters repulse each other

max_repulsion_iter  numeric. The number of repulsion iterations, note that increasing this value may occasionally even lead to worse looking plots as clusters may repulse eachother too much

use_default_theme  If `TRUE', the resulting plot will have the same theme as the seurat UMAP. Else, the plot will simply have a blank background

show_origin  logical. If `TRUE', only the centers of each circle will be plotted

retain_axis_scales  If `TRUE', approximately maintains the axis scales of the original UMAP. However, it will only attempt to extend the axes and never shorten.

add_size_legend  If `TRUE', adds a legend to the plot titled "'Clone sizes" indicating the relative sizes of clones.

legend_sizes  numeric vector. Indicates the circle sizes to be displayed on the legend and defaults to `c(1, 5, 10)'.
count_clone_sizes

legend_position
character. Can be set to either "top_left", "top_right", "bottom_left", "bottom_right" and places the legend roughly in the corresponding position.

legend_buffer
numeric. Indicates how much to "push" the legend towards the center of the plot from the selected corner. If negative, will push away.

legend_color
character. Indicates the hex color of the circles displayed on the legend. Defaults to the hex code for gray.

legend_spacing
numeric. Indicates the horizontal distance between each stacked circle on the size legend. Usually should be kept below 0.75-ish depending on plot size.

Details

Check out the web-only user vignette at ‘https://qile0317.github.io/APackOfTheClones/articles/web_only/Clonal_expansion_plotting.html’ for a walkthrough on using this function, and additional details.

Value

Returns a ggplot2 object of the ball packing plot. Can be operated on like normal ggplot objects.

See Also

integrate_tcr

Examples

library(Seurat)
library(APackOfTheClones)
data("mini_clonotype_data", "mini_seurat_obj")

# produce and show the ball-packing plot by integrating the data
ball_pack_plot <- clonal_expansion_plot(mini_seurat_obj, mini_clonotype_data)
ball_pack_plot

# it's also possible to input an integrated Seurat object
integrated_seurat_object <- integrate_tcr(mini_seurat_obj, mini_clonotype_data)
ball_pack_plot <- clonal_expansion_plot(integrated_seurat_object)
ball_pack_plot

---

count_clone_sizes
count the number of clonotype sizes per cell cluster in a seurat object integrated with a TCR library

Description

count the number of clonotype sizes per cell cluster in a seurat object integrated with a TCR library
Usage

count_clone_sizes(integrated_seurat_obj)

Arguments

integrated_seurat_obj

Seurat object that has been integrated with a T-cell receptor library with integrate_tcr. More specifically, in the metadata, there must at least be the elements 'seurat_clusters' and 'raw_clonotype_id'.

Value

Returns a list of 'table' objects, where each element is tabled clonotype frequencies for the seurat cluster corresponding to the same index - 1. For example, the 5th element is a tabled frequency of counts that corresponds to the 4th seurat cluster (as seurat clusters are 0-indexed). If an element is 'NULL', it indicates that there were no corresponding T-cell receptor barcode for the cells in the cluster.

See Also

integrate_tcr

Examples

library(Seurat)
library(APackOfTheClones)
data("mini_clonotype_data","mini_seurat_obj")

# produce an integrated seurat_object
integrated_seurat_object <- integrate_tcr(mini_seurat_obj, mini_clonotype_data)
clonotype_counts <- count_clone_sizes(integrated_seurat_object)
clonotype_counts
### Arguments

- **seurat_obj**  
  Seurat object
- **tcr_file**  
  `data.frame` of the T cell library generated by Cell Ranger. It is very important that the row with cell barcodes is strictly named “barcode”, which is the default name of barcodes in 10X’s `all_contig_annotations.csv` file.
- **verbose**  
  if ‘TRUE’, will display a progress bar to the R console.

### Details

Columns from cells (barcodes) that had duplicates in another row are concatenated into strings, separated by ‘__’ in the metadata element. Barcodes from the TCR library that had no matches to barcodes in the ‘seurat_obj’ will add ‘NA’s for all elements of the same index.

### Value

Returns a new Seurat object with new elements in the metadata

### References


### Examples

```r
library(Seurat)
library(APackOfTheClones)
data("mini_clonotype_data","mini_seurat_obj")

# integrate the TCR data into new seurat object
integrated_seurat_object <- integrate_tcr(mini_seurat_obj, mini_clonotype_data)
integrated_seurat_object
```

---

### mini_clonotype_data

*Artificially generated T cell receptor library*

### Description

A generated dataframe of a T-cell receptor (TCR) library generated from single cell immune profiling. It is a subset the full dataframe which would usually have up to 18 columns containing different data, because the intended purpose of this object is to test various functions in 'APackOfTheClones'. The dataframe compliments ‘mini_seurat_obj’ and can be integrated into it with ‘integrate_tcr’.

### Usage

```r
data("mini_clonotype_data")
```
Format

'\text{data.frame}' A data frame with 80 rows and 2 columns:

- **barcode** barcodes corresponding to each sequenced cell
- **raw\_clonotype\_id** clonotype information for each cell

Details

Note that the clonotypes in the 'raw\_clonotype\_id' column actually do not contain all of clonotype'1'...clonotype'n'.

See Also

mini\_seurat\_obj

---

**mini\_seurat\_obj** *Artificially generated Seurat object*

Description

A generated 'SeuratObject' of a small sc-RNAseq experiment. Has a corresponding T-cell receptor library generated from single cell immune profiling, named "mini\_clonotype\_data".

Usage

data("mini\_seurat\_obj")

Format

A Seurat object with the following slots filled

- **assays** Currently only contains one assay ("RNA" - scRNA-seq expression data)
  - counts - Raw expression data
  - • data - Normalized expression data
  - • scale.data - Scaled expression data
  - • var.features - names of the current features selected as variable
  - • meta.features - Assay level metadata such as mean and variance

- **meta\_data** Cell level metadata
- **active\_assay** Current default assay
- **active\_ident** Current default idents
- **graphs** Neighbor graphs computed, currently stores the SNN
- **reductions** Dimensional reductions: PCA, UMAP, and tSNE
- **version** Seurat version used to create the object
- **commands** Command history

See Also

mini\_clonotype\_data
Index

* datasets
  mini_clonotype_data, 6
  mini_seurat_obj, 7
* sc_RNAseq
  integrate_tcr, 5

clonal_expansion_plot, 2
count_clone_sizes, 4
integrate_tcr, 4, 5, 5
mini_clonotype_data, 6, 7
mini_seurat_obj, 7, 7