Package 'AnanseSeurat'

November 11, 2023

Title Construct ANANSE GRN-Analysis Seurat

Version 1.2.0

Description Enables gene regulatory network (GRN) analysis on single cell clusters, using the GRN analysis software 'ANANSE', Xu et al.(2021) <doi:10.1093/nar/gkab598>. Export data from 'Seurat' objects, for GRN analysis by 'ANANSE' implemented in 'snakemake'. Finally, incorporate results for visualization and interpretation.

License Apache License (>= 2)

BugReports https://github.com/JGASmits/AnanseSeurat/issues

URL https://github.com/JGASmits/AnanseSeurat/

Encoding UTF-8

RoxygenNote 7.2.3

Imports dplyr, ggplot2, ggpubr, magrittr, patchwork, png, purrr, rlang, Seurat, stringr, utils,

Suggests covr, knitr, rmarkdown, Signac, testthat (>= 3.0.0)

Config/testthat/edition 3

VignetteBuilder knitr

Depends R (>= 3.50)

NeedsCompilation no

Author Jos Smits [aut, cre, cph] (<https://orcid.org/0000-0001-5858-3905>), Huiqing Jo Zhou [aut, fnd, cph] (<https://orcid.org/0000-0002-2434-3986>)

Maintainer Jos Smits <Jsmits@science.ru.nl>

Repository CRAN

Date/Publication 2023-11-11 21:43:17 UTC

R topics documented:

`config_scANANSE` *config_scANANSE*

Description

This functions generates a sample file and config file for running Anansnake based on the seurat object

Usage

```
config_scANANSE(
  seurat_object,
  output_dir,
  min_cells = 50,
  cluster_id = "seurat_clusters",
  genome = "./scANANSE/data/hg38",
  additional_contrasts = c()
)
```

Arguments

`seurat_object`	seurat object
`output_dir`	directory where the files are outputted
`min_cells`	minimum of cells a cluster needs to be exported
`cluster_id`	ID used for finding clusters of cells
`genome`	genomepy name or location of the genome fastq file
`additional_contrasts`	additional contrasts to add between clusters within cluster_ID

Value

None, outputs snakemake config file in the output directory

Examples

```
sce_small <- readRDS(system.file("extdata","sce_small.Rds",package = 'AnanseSeurat'))
config_scANANSE(sce_small, min_cells = 2, output_dir = tempdir())
```

DEGS_scANANSE *DEGS_scANANSE*

Description

Calculate the differential genes needed for ananse influence

Usage

```
DEGS_scANANSE(
  seurat_object,
  output_dir,
  min_cells = 50,
  cluster_id = "seurat_clusters",
  genome = "./scANANSE/data/hg38",
  RNA_count_assay = "RNA",
  additional_contrasts = "None"
)
```

Arguments

seurat_object	seurat object
output_dir	directory where the files are outputted
min_cells	minimum of cells a cluster needs to be exported
cluster_id	ID used for finding clusters of cells
genome	path to the genome folder used for the anansnake config file
RNA_count_assay	assay containing the RNA data
additional_contrasts	additional contrasts to add between clusters within cluster_ID

Value

None, outputs DEG files in the output directory

Examples

```
sce_small <- readRDS(system.file("extdata","sce_obj_tiny.Rds",package = 'AnanseSeurat'))
DEGS_scANANSE(sce_small, min_cells = 2, output_dir = tempdir())
```

export_ATAC_maelstrom *export_seurat_Maelstrom*

Description

normalize and export the peak table of a seurat object based on clusters

Usage

```
export_ATAC_maelstrom(
  seurat_object,
  output_dir,
  min_cells = 50,
  ATAC_peak_assay = "peaks",
  cluster_id = "seurat_clusters",
  select_top_rows = TRUE,
  n_top_rows = 1e+05
)
```

Arguments

seurat_object	object
output_dir	directory where the files are outputted
min_cells	minimum of cells a cluster needs to be exported
ATAC_peak_assay	assay of the seurat object containing the peaks and peakcounts
cluster_id	ID used for finding clusters of cells
select_top_rows	only output the top variable rows, or all rows if false
n_top_rows	amount of variable rows to export

Value

None, outputs maelstrom peak counts table in the output directory

Examples

```
sce_small <- readRDS(system.file("extdata","sce_small.Rds",package = 'AnanseSeurat'))
config_scANANSE(sce_small, min_cells = 2, output_dir = tempdir())
```

export_ATAC_scANANSE *export_ATAC_scANANSE*

Description

This functions exports ATAC values from a seurat object

Usage

```
export_ATAC_scANANSE(
  seurat_object,
  output_dir,
  min_cells = 50,
  ATAC_peak_assay = "peaks",
  cluster_id = "seurat_clusters"
)
```

Arguments

seurat_object	object
output_dir	directory where the files are outputted
min_cells	minimum of cells a cluster needs to be exported
ATAC_peak_assay	assay of the seurat object containing the peaks and peakcounts
cluster_id	ID used for finding clusters of cells

Value

None, outputs ATAC peak count file in the output directory

Examples

```
sce_small <- readRDS(system.file("extdata","sce_small.Rds",package = 'AnanseSeurat'))
export_ATAC_scANANSE(sce_small, min_cells = 2, output_dir = tempdir())
```

export_CPM_scANANSE *export_CPM_scANANSE*

Description

This functions exports CPM values from a seurat object

Usage

```
export_CPM_scANANSE(
  seurat_object,
  output_dir,
  min_cells = 50,
  RNA_count_assay = "RNA",
  cluster_id = "seurat_clusters"
)
```

Arguments

seurat_object	the seurat object used to export the CPM values from
output_dir	directory where the files are outputted
min_cells	minimum of cells a cluster needs to be exported
RNA_count_assay	assay of the seurat object containing the RNA count data
cluster_id	ID used for finding clusters of cells

Value

None, outputs CPM and counts files in the output directory

Examples

```
sce_small <- readRDS(system.file("extdata","sce_small.Rds",package = 'AnanseSeurat'))
export_CPM_scANANSE(sce_small, min_cells = 2, output_dir = tempdir())
```

Factor_Motif_Plot *Factor_Motif_Plot*

Description

plot both expression of a TF, and the motif accessibility of the associated motif. Finally, fetch the motif logo from the Maelstrom directory.

Usage

```
Factor_Motif_Plot(
  seurat_object,
  TF_list,
  assay_RNA = "RNA",
  assay_maelstrom = "MotifTFanticor",
  logo_dir = "~/maelstrom/logos",
  col = c("darkred", "white", "darkgrey"),
  dim_reduction = "umap"
)
```

Arguments

seurat_object	seurat object
TF_list	list of TFs to plot the expression and linked motif Z-score for
assay_RNA	RNA_count_assay assay containing the RNA data
assay_maelstrom	maelstrom assay used for zscore vizualization, often either TFcor or TFanticor
logo_dir	directory containing motif logos generated by gimme maelstrom
col	colours used for zscore vizualization
dim_reduction	dimensionality reduction method to use

Value

patchwork plot containing a expression dimreduction plot, a maelstrom motif score dimreduction plot, and a png image of the motif

Examples

```
sce_small <- readRDS(system.file("extdata","sce_small.Rds",package = 'AnanseSeurat'))
logos_dir_path <- system.file("extdata","maelstrom","logos",package = 'AnanseSeurat')
sce_small <- Factor_Motif_Plot(sce_small,
  c('gene1', 'gene2'),
  dim_reduction = 'pca',
  logo_dir = logos_dir_path)
```

---

import_seurat_maelstrom *import_seurat_Maelstrom*

Description

load Maelstrom enriched motifs

Usage

```
import_seurat_maelstrom(
  seurat_object,
  cluster_id = "seurat_clusters",
  maelstrom_file = "~/final.out.txt",
  return_df = FALSE
)
```

Arguments

seurat_object	object
cluster_id	ID used for finding clusters of cells
maelstrom_file	maelstrom final.out.txt file
return_df	return both the seurat object and a dataframe with maelstrom scores as a list

Value

seurat object with the maelstrom motif scores addes as an assay

Examples

```
sce_small <- readRDS(system.file("extdata","sce_small.Rds",package = 'AnanseSeurat'))
maelstromfile_path <- system.file("extdata","maelstrom","final.out.txt",package = 'AnanseSeurat')
sce_small <- import_seurat_maelstrom(sce_small, maelstrom_file = maelstromfile_path)
```

`import_seurat_scANANSE` *import_seurat_scANANSE*

Description

import the influences from a anansnake directory into a seurat object

Usage

```
import_seurat_scANANSE(
  seurat_object,
  cluster_id = "seurat_clusters",
  anansnake_inf_dir = "None",
  return_df = FALSE
)
```

Arguments

`seurat_object`	seurat object
`cluster_id`	ID used for finding clusters of cells
`anansnake_inf_dir`	influence directory generated by anansnake
`return_df`	return both the seurat object and a dataframe with influence scores as a list

Value

seurat object with the influence scores addes as an assay

Examples

```
sce_small <- readRDS(system.file("extdata","sce_small.Rds",package = 'AnanseSeurat'))
infdir <- system.file("extdata","influence",package = 'AnanseSeurat')
sce_small <- import_seurat_scANANSE(sce_small, anansnake_inf_dir = infdir)
```

Maelstrom_Motif2TF *Maelstrom_Motif2TF*

Description

create motif-factor links & export tables for printing motif score alongside its binding factor

Usage

```
Maelstrom_Motif2TF(
  seurat_object,
  mot_mat = NULL,
  m2f_df = NULL,
  cluster_id = "seurat_clusters",
  maelstrom_dir = "./maelstrom/",
  combine_motifs = "means",
  RNA_expression_assay = "RNA",
  RNA_expression_slot = "data",
  expr_tresh = 10,
  cor_tresh = 0.3,
  curated_motifs = FALSE,
  cor_method = "pearson",
  return_df = FALSE
)
```

Arguments

seurat_object	object
mot_mat	motif_matrix, if not provided extracts one from the single cell object from the maelstrom assay
m2f_df	motif to factor dataframe, if not provided extracts from the maelstrom directory
cluster_id	ID used for finding clusters of cells
maelstrom_dir	directory where the GimmeMotifs m2f table is stored
combine_motifs	means (take mean multiple motifscores), max_var (take motif with highest variance), or max_cor (take motif with best correlation to gene expression)
RNA_expression_assay	Seurat assay containing factor expression info
RNA_expression_slot	slot within assay used for calculating mean factor expression per cluster
expr_tresh	minimum sum of gene counts over all cells in RNA_expression_assay to filter genes by
cor_tresh	minimum value of to filter the cor() output by
curated_motifs	use only curated motifs (T), or all motifs in the database (F)
cor_method	specify one of the cor() methods
return_df	return both the seurat object and two dataframes with maelstrom scores and expression values as a list

Value

seurat object with two assays added, MotifTFcor for TFs with positive correlation to the linked motif, and MotifTFanticor for TFs with positive correlation to the linked motif

Examples

```
sce_small <- readRDS(system.file("extdata","sce_small.Rds",package = 'AnanseSeurat'))
maelstrom_dir_path <- system.file("extdata","maelstrom",package = 'AnanseSeurat')
sce_small <- Maelstrom_Motif2TF(sce_small, maelstrom_dir = maelstrom_dir_path)
```

---

per_cluster_df *per_cluster_df*

Description

generate a table of the assay score averages per cluster identifier cell

Usage

```
per_cluster_df(
  seurat_object,
  assay = "influence",
  cluster_id = "seurat_clusters"
)
```

Arguments

seurat_object	seurat object
assay	assay containing influence or motif scores generated from cluster pseudobulk
cluster_id	ID used for finding clusters of cells

Value

dataframe with assay scores, concatinating cells from each per cluster

Examples

```
sce_small <- readRDS(system.file("extdata","sce_small.Rds",package = 'AnanseSeurat'))
df <- per_cluster_df(sce_small)
```

Index