Package ‘scviR’

March 12, 2024

Date 2023-03-07
Title experimental interface from R to scvi-tools
Version 1.2.0
Description This package defines interfaces from R to scvi-tools. A vignette works through the totalVI tutorial for analyzing CITE-seq data. Another vignette compares outputs of Chapter 12 of the OSCA book with analogous outputs based on totalVI quantifications. Future work will address other components of scvi-tools, with a focus on building understanding of probabilistic methods based on variational autoencoders.
License Artistic-2.0
Encoding UTF-8
Depends R (>= 4.3), basilisk, shiny, SingleCellExperiment
Imports reticulate, BiocFileCache, utils, pheatmap,
   SummarizedExperiment, S4Vectors, limma, scater, stats,
   MatrixGenerics
Suggests knitr, testthat, reshape2, ggplot2, rhdf5, BiocStyle
VignetteBuilder knitr
biocViews Infrastructure, SingleCell, DataImport
RoxygenNote 7.2.3
URL https://github.com/vjcitn/scviR
BugReports https://github.com/vjcitn/scviR/issues
git_url https://git.bioconductor.org/packages/scviR
git_branch RELEASE_3_18
git_last_commit 2a594ba
git_last_commit_date 2023-10-24
Repository Bioconductor 3.18
Date/Publication 2024-03-11
Author Vincent Carey [aut, cre] (<https://orcid.org/0000-0003-4046-0063>)
Maintainer Vincent Carey <stvjc@channing.harvard.edu>
R topics documented:

- adtProfiles
- anndataR
- bsklenv
- cacheCiteseq5k10kPbmcs
- cacheCiteseq5k10kTutvae
- clusters.adt
- clusters.rna
- exploreSubcl
- getCh12AllSce
- getCh12Sce
- getCiteseq5k10kPbmcs
- getCiteseqTutvae
- getPro5k10kAdata
- getSubcILM
- getSubclusteringFeatures
- getTotalVI5k10kAdata
- getTotalVINormalized5k10k
- pyHelp2
- scanpyHelper
- scanpyR
- scviHelper
- scviR

Description

 produce a heatmap from a specialized CITE-seq SingleCellExperiment

Usage

  adtProfiles(x, lb = -3, ub = 3, do_z = FALSE)

Arguments

- x          SingleCellExperiment instance that has an ‘se.averaged’ component in its metadata
- lb         numeric(1) lower bound on ‘breaks’ sequence for ComplexHeatmap::pheatmap, defaults to -3
- ub         numeric(1) upper bound on ‘breaks’ sequence for ComplexHeatmap::pheatmap, defaults to 3
- do_z       logical(1) if TRUE, divide the residuals by their standard deviation across clusters, defaults to false
anndataR

Value

ComplexHeatmap::pheatmap instance
side effect of pheatmap::pheatmap call

Note

See the OSCA book ch12.5.2 for the application.

Examples

\begin{verbatim}
ch12sce <- getCh12Sce()
adtProfiles(ch12sce)
adtProfiles(ch12sce, do_z = TRUE)
\end{verbatim}

anndataR basic interface to anndata

Description

basic interface to anndata

Usage

anndataR()

Value

basiliskRun result with import from reticulate, typically a Module

Examples

\begin{verbatim}
ad <- anndataR()
ad
ad$read
\end{verbatim}

bsklenv python declarations

Description

python declarations

Usage

bsklenv

Format

An object of class BasiliskEnvironment of length 1.
cacheCiteseq5k10kPbmcs

grab scvi-tools-processed PBMC CITE-seq data in anndata format (gzipped) from Open Storage Network

Description

grab scvi-tools-processed PBMC CITE-seq data in anndata format (gzipped) from Open Storage Network

Usage

```r
cacheCiteseq5k10kPbmcs()
```

Value

invisibly, the path to the .h5ad file

Note

Original h5ad files obtained using scvi-tools 0.18.0 scvi.data.pbmcs_10x_cite_seq, then processed according to steps in the scviR vignette, which follow the [scvi-tools tutorial](https://colab.research.google.com/github/scverse/scvi-tutorials/blob/0.18.0/totalVI.ipynb) by Gayoso et al.

It may be advantageous to set `options(timeout=3600)` or to allow an even greater time for internet downloads, if working at a relatively slow network connection.

Examples

```r
h5path <- cacheCiteseq5k10kPbmcs()
cmeta <- rhdf5::h5ls(h5path)
dim(cmeta)
head(cmeta, 17)
```

---

cacheCiteseq5k10kTutvae

grab scvi-tools VAE instance built on the PBMC datasets following the tutorial

Description

grab scvi-tools VAE instance built on the PBMC datasets following the tutorial

Usage

```r
cacheCiteseq5k10kTutvae()
```
Value

invisibly, the path to the .zip file holding the fitted VAE and associated data

Note


It may be advantageous to set `options(timeout=3600)` or to allow an even greater time for internet downloads, if working at a relatively slow network connection.

Examples

```r
zpath <- cacheCiteseq5k10kTutvae()
td <- tempdir()
utils::unzip(zpath, exdir = td)
vaedir <- paste0(td, "/vae2_ov")
scvi <- scviR()
adm <- anndataR()
hpath <- cacheCiteseq5k10kPbmcs()
adata <- adm$read(hpath)
mod <- scvi$model$"_totalvi"$TOTALVI$load(vaedir, adata, use_gpu = FALSE)
mod
```

---

### clusters.adt

**ADT-based cluster labels for 7472 cells in OSCA chapter 12 analysis**

**Description**

ADT-based cluster labels for 7472 cells in OSCA chapter 12 analysis

**Usage**

```r
clusters.adt
```

**Format**

- factor
clusters.rna  mRNA-based cluster labels for 7472 cells in OSCA chapter 12 analysis

Description
mRNA-based cluster labels for 7472 cells in OSCA chapter 12 analysis

Usage
clusters.rna

Format
factor

exploreSubcl  app to explore diversity in RNA-subclusters within ADT clusters

Description
app to explore diversity in RNA-subclusters within ADT clusters

Usage
exploreSubcl(sce, inlist, adtcls)

Arguments

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>sce</td>
<td>a SingleCellExperiment with altExp with ADT quantification</td>
</tr>
<tr>
<td>inlist</td>
<td>list of SingleCellExperiments (SCEs) formed by scran::quickSubCluster</td>
</tr>
<tr>
<td>adtcls</td>
<td>vector of ADT cluster assignments</td>
</tr>
</tbody>
</table>

Value
shinyApp instance

Note
TSNE should already be available in `altExp(sce)`; follow OSCA book 12.5.2. If using example, set `ask=FALSE`.

Examples

```r
cce <- getCh12Sce()
all.sce <- getCh12AllSce()
data(clusters.adt)
runApp(exploreSubcl(sce, all.sce, clusters.adt)) # trips up interactive pkgdown?)```
Description
get list of cluster-specific SCE for 10k PBMC annotated as in OSCA book chapter 12

Usage
getCh12AllSce()

Value
SimpleList of SingleCellExperiment instances

Note
This is a list of SingleCellExperiment instances with data on a total of 7472 cells from a 10x CITE-seq experiment. An altExp component in each list element includes antibody-derived tag (ADT) counts on 17 proteins. The data are acquired and processed as described in ch 12 of the OSCA book, circa February 2023. List elements correspond to mRNA-based sub-clusters of ADT-based clusters.

Examples
ch12_allsc <- getCh12AllSce()
vapply(ch12_allsc, ncol, numeric(1))

getCh12Sce
get SCE for 10k PBMC annotated as in OSCA book chapter 12

Description
get SCE for 10k PBMC annotated as in OSCA book chapter 12

Usage
getCh12Sce(clear_cache = FALSE)

Arguments
clear_cache logical(1) will delete relevant entries in available cache before continuing, defaults to FALSE
Value

SingleCellExperiment instance

Note

This is a SingleCellExperiment instance with data on 7472 cells from a 10x CITE-seq experiment. An altExp component includes antibody-derived tag (ADT) counts on 17 proteins. The data are acquired and processed as described in ch 12 of the OSCA book, circa February 2023. A metadata element (se.averaged) includes the result of averaging protein abundance estimates within ADT-based clusters, as is done to give rise to Figure 12.8 of the OSCA book.

Examples

```r
ch12sce <- getCh12Sce()
ch12sce
```

---

cGetCiteseq5k10kPbmcs  helper to get the processed anndata for CITE-seq PBMCs from scvi-tools tutorial

Description

helper to get the processed anndata for CITE-seq PBMCs from scvi-tools tutorial

Usage

```r
cGetCiteseq5k10kPbmcs()
```

Value

python reference to anndata

Note

It may be advantageous to set ‘options(timeout=3600)’ or to allow an even greater time for internet downloads, if working at a relatively slow network connection.

Examples

```r
cGetCiteseq5k10kPbmcs()
```
**getDescription**

*helper to get the tutorial VAE for PBMCs from scvi-tools tutorial*

**Usage**

ggetCiteseqTutvae(use_gpu = FALSE)

**Arguments**

use_gpu logical(1), defaulting to FALSE, passed to TOTALVI.load

**Value**

python reference to anndata

**Examples**

ggetCiteseqTutvae()

**getDescription**

*get an anndata reference to 5k10k protein after totalVI from tutorial*

**Usage**

ggetPro5k10kAdata()

**Value**

python reference to anndata

**Note**

It may be advantageous to set `options(timeout=3600)` or to allow an even greater time for internet downloads, if working at a relatively slow network connection.

**Examples**

ggetPro5k10kAdata()
### Description

get lmFit for heterogeneity across subclusters

### Usage

```r
getSubclLM(inlist, clname)
```

### Arguments

- `inlist`: list of SingleCellExperiments (SCEs) formed by scran::quickSubCluster
- `clname`: character(1) name of cluster SCE to assess

### Value

limma::lmFit output

### Note

It is assumed that 'logcounts' is an assay element, and that 'subcluster' is a colData element of each SCE in inlist

### Examples

```r
all.sce <- getCh12AllSce()
lm3 <- getSubclLM(all.sce, "3")
names(lm3)
```

---

### Description

get lmFit F-stat based collection of n genes most varying in mean across subclusters

### Usage

```r
getSubclusteringFeatures(inlist, clname, n = 20)
```
Arguments

inlist  list of SingleCellExperiments (SCEs) formed by scran::quickSubCluster
cname  character(1) name of cluster SCE to assess
n  numeric(1) number to preserve

Value

list with two elements, feat = rowData corresponding to variable genes, stats = topTable result

Note

Symbol will be taken from feat and placed in stats component if available

Examples

all.sce <- getCh12AllSce()
scl <- getSubclusteringFeatures(all.sce, "3", 10)
names(scl)

full <- getTotalVI5k10kAdata()
getTotalVINormalized5k10k

get matrices of normalized quantifications from full totalVI 5k10k from tutorial

Description
get matrices of normalized quantifications from full totalVI 5k10k from tutorial

Usage
getTotalVINormalized5k10k()

Value
list of matrices

Examples
nmlist <- getTotalVINormalized5k10k()
vapply(nmlist, dim, numeric(2))

pyHelp2

helper to get text from python help utility – may need handling through basilisk

Description
helper to get text from python help utility – may need handling through basilisk

Usage
pyHelp2(object)

Arguments
object a reference to a python module typically with class 'python.builtin.module'

Value
character vector of lines from python help result
scanpyHelper

---

**scanpyHelper**

*shiny app that helps access documentation on python-accessible components*

---

**Description**

shiny app that helps access documentation on python-accessible components

**Usage**

`scanpyHelper()`

**Value**

shinyApp instance

---

**scanpyR**

*basic interface*

---

**Description**

basic interface

**Usage**

`scanpyR()`

**Value**

basiliskRun result with import from reticulate, typically a Module

**Examples**

```r
sc <- scanpyR()
sca
sc
sc$pp
```
scviHelper

**shiny app that helps access documentation on python-accessible components**

**Description**

shiny app that helps access documentation on python-accessible components

**Usage**

scviHelper()

**Value**

shinyApp instance

---

scviR

**basic interface**

**Description**

basic interface

**Usage**

scviR()

**Value**

basiliskRun result with import from reticulate, typically a Module

**Examples**

scvi <- scviR()
scvi
scvi$model
Index

* datasets
  bsklenv, 3
  clusters.adt, 5
  clusters.rna, 6
  adtProfiles, 2
  anndataR, 3
  bsklenv, 3
  cacheCiteseq5k10kPbmcs, 4
  cacheCiteseq5k10kTutvae, 4
  clusters.adt, 5
  clusters.rna, 6
  exploreSubcl, 6
  getCh12AllSce, 7
  getCh12Sce, 7
  getCiteseq5k10kPbmcs, 8
  getCiteseqTutvae, 9
  getPro5k10kAdata, 9
  getSubcLM, 10
  getSubclusteringFeatures, 10
  getTotalVIS5k10kAdata, 11
  getTotalVINormalized5k10k, 12
  pyHelp2, 12
  scanpyHelper, 13
  scanpyR, 13
  scviHelper, 14
  scviR, 14