Package 'schex'

November 13, 2024

Type Package

Title Hexbin plots for single cell omics data

Version 1.20.0

Description Builds hexbin plots for variables and dimension reduction stored
in single cell omics data such as SingleCellExperiment. The
ideas used in this package are based on the excellent work of Dan Carr,
Nicholas Lewin-Koh, Martin Maechler and Thomas Lumley.

Depends SingleCellExperiment (>= 1.7.4), ggplot2 (>= 3.2.1)

Imports hexbin, stats, methods, cluster, dplyr, entropy, ggforce,
grid, rlang, concaveman

Suggests ggrepel, knitr, rmarkdown, testthat (>= 2.1.0), covr,
TENxPBMCData, scater, Seurat, shinydashboard, iSEE, igraph,
scran, tibble, scuttle

URL <https://github.com/SaskiaFreytag/schex>

BugReports <https://github.com/SaskiaFreytag/schex/issues>

License GPL-3

Encoding UTF-8

RoxygenNote 7.2.3

biocViews Software, Sequencing, SingleCell, DimensionReduction,
Visualization, ImmunoOncology, DataImport

VignetteBuilder knitr

git_url https://git.bioconductor.org/packages/schex

git_branch RELEASE_3_20

git_last_commit 5da0507

git_last_commit_date 2024-10-29

Repository Bioconductor 3.20

Date/Publication 2024-11-12

Author Saskia Freytag [aut, cre],
Wancheng Tang [ctb],
Zimo Peng [ctb],
Jingxiu Huang [ctb]

Maintainer Saskia Freytag <freytag.s@wehi.edu.au>

Contents

---

make_hexbin   *Bivariate binning of single cell data into hexagon cells.*

Description

make_hexbin returns a [SingleCellExperiment](#) object of binned hexagon cells.

Usage

```
make_hexbin(sce, nbins = 80, dimension_reduction = "UMAP", use_dims = c(1, 2))

## S4 method for signature 'SingleCellExperiment'
make_hexbin(sce, nbins = 80, dimension_reduction = "UMAP", use_dims = c(1, 2))
```

Arguments

sce	A [SingleCellExperiment](#) object.
nbins	The number of bins partitioning the range of the first component of the chosen dimension reduction.
dimension_reduction	
	A string indicating the reduced dimension result to calculate hexagon cell representation of.
use_dims	A vector of two integers specifying the dimensions used.

Details

This function bins observations with computed reduced dimension results into hexagon cells. For a [SingleCellExperiment](#) as a list in the @metadata. The list contains two items. The first item stores a vector specifying the hexagon ID for each observation. The second item stores a matrix with the x and y positions of the hexagon cells and the number of observations in each of them.

Value

A [SingleCellExperiment](#) object.

Functions

- make_hexbin(SingleCellExperiment): Bivariate binning of SingleCellExperiment into hexagon cells.

Examples

```
# For SingleCellExperiment object
library(TENxPBMCData)
library(scater)
tenx_pbmc3k <- TENxPBMCData(dataset = "pbmc3k")
rm_ind <- calculateAverage(tenx_pbmc3k) < 0.1
tenx_pbmc3k <- tenx_pbmc3k[!rm_ind, ]
tenx_pbmc3k <- logNormCounts(tenx_pbmc3k)
tenx_pbmc3k <- runPCA(tenx_pbmc3k)
tenx_pbmc3k <- make_hexbin(tenx_pbmc3k, 80, dimension_reduction = "PCA")
```

---

make_hexbin_label   *Group label plot position.*

Description

Group label plot position.

Usage

```
make_hexbin_label(sce, col)
```

Arguments

| sce | A [SingleCellExperiment](#) object. |
| col | The name referring to one column in meta data for which the label position on the plot is calculated for every level. The chosen column needs to be a factor. |

Value

A dataframe.

Examples

```
# For SingleCellExperiment object
library(TENxPBMCData)
library(scater)
tenx_pbmc3k <- TENxPBMCData(dataset = "pbmc3k")
rm_ind <- calculateAverage(tenx_pbmc3k) < 0.1
tenx_pbmc3k <- tenx_pbmc3k[!rm_ind, ]
tenx_pbmc3k <- logNormCounts(tenx_pbmc3k)
tenx_pbmc3k <- runPCA(tenx_pbmc3k)
tenx_pbmc3k <- make_hexbin(tenx_pbmc3k, 80, dimension_reduction = "PCA")
tenx_pbmc3k$random <- factor(sample(1:3, ncol(tenx_pbmc3k), replace = TRUE))
make_hexbin_label(tenx_pbmc3k, col = "random")
```

---

plot_hexbin_bivariate   *Plot of feature expression and uncertainty of single cells in bivariate hexagon cells.*

Description

Plot of feature expression and uncertainty of single cells in bivariate hexagon cells.

Usage

```
plot_hexbin_bivariate(
  sce,
  mod = "RNA",
  type,
  feature,
  fan = FALSE,
  title = NULL,
  xlab = NULL,
  ylab = NULL
)
```

Arguments

sce	A `SingleCellExperiment` object.
mod	A string referring to the name of the modality used for plotting. For RNA modality use "RNA". For other modalities use name of alternative object for the `SingleCellExperiment` object.
type	A string referring to the type of assay in the `SingleCellExperiment` object.
feature	A string referring to the name of one feature.
fan	Logical indicating whether to plot uncertainty surprising palette.
title	A string containing the title of the plot.
xlab	A string containing the title of the x axis.
ylab	A string containing the title of the y axis.

Details

This function plots the mean expression and a measure of uncertainty of any feature in the hexagon cell representation calculated with `make_hexbin` using a bivariate scale. When fan=FALSE, the standard deviation and the mean expression are plotted. When fan=TRUE, the mean expression and coefficient of variation are calculated. The coefficient of variation is converted to a percentage of uncertainty. When using fan=TRUE the raw count data should be used. In order to enable the calculation of the coefficient of variation a pseduo-count of 1 is added to every count.

Value

A `ggplot2`{ggplot} object.

Examples

```
# For SingleCellExperiment object
library(TENxPBMCData)
library(scater)
tenx_pbmc3k <- TENxPBMCData(dataset = "pbmc3k")
rm_ind <- calculateAverage(tenx_pbmc3k) < 0.1
tenx_pbmc3k <- tenx_pbmc3k[!rm_ind, ]
tenx_pbmc3k <- logNormCounts(tenx_pbmc3k)
tenx_pbmc3k <- runPCA(tenx_pbmc3k)
tenx_pbmc3k <- make_hexbin(tenx_pbmc3k, 80, dimension_reduction = "PCA")
plot_hexbin_bivariate(tenx_pbmc3k, type = "counts", feature = "ENSG00000135250")
plot_hexbin_bivariate(tenx_pbmc3k, type = "counts", feature = "ENSG00000135250", fan = TRUE)
```

---

plot_hexbin_density	*Plot of density of observations from single cell data in bivariate hexagon cells.*

Description

Plot of density of observations from single cell data in bivariate hexagon cells.

Usage

```
plot_hexbin_density(sce, title = NULL, xlab = NULL, ylab = NULL)
```

Arguments

sce	A [SingleCellExperiment](#) object.
title	A string containing the title of the plot.
xlab	A string containing the title of the x axis.
ylab	A string containing the title of the y axis.

Value

A [ggplot2](#){ggplot} object.

Examples

```
# For SingleCellExperiment object
library(TENxPBMCData)
library(scater)
tenx_pbmc3k <- TENxPBMCData(dataset = "pbmc3k")
rm_ind <- calculateAverage(tenx_pbmc3k) < 0.1
tenx_pbmc3k <- tenx_pbmc3k[!rm_ind, ]
tenx_pbmc3k <- logNormCounts(tenx_pbmc3k)
tenx_pbmc3k <- runPCA(tenx_pbmc3k)
tenx_pbmc3k <- make_hexbin(tenx_pbmc3k, 10, dimension_reduction = "PCA")
plot_hexbin_density(tenx_pbmc3k)
```

| plot_hexbin_fc | *Plot of fold change of selected gene in single cell data using bivariate hexagon cells.* |

Description

Plot of fold change of selected gene in single cell data using bivariate hexagon cells.

Usage

```
plot_hexbin_fc(
  sce,
  col,
  mod = "RNA",
  type,
  feature,
  title = NULL,
  xlab = NULL,
  ylab = NULL,
  colors
)
```

Arguments

sce	A [SingleCellExperiment](#) object.
col	A string referring to the name of one column in the meta data of sce by which to compare. Note this factor can only contain two levels.
mod	A string referring to the name of one column in the meta data of sce by which to compare. Note this factor can only contain two levels.
type	A string referring to the name of one column in the meta data of sce by which to compare. Note this factor can only contain two levels.
feature	A string referring to the name of one feature.
title	A string containing the title of the plot.
xlab	A string containing the title of the x axis.
ylab	A string containing the title of the y axis.
colors	A vector of strings specifying which colors to use for plotting the different levels in the selected column of the meta data.

Details

This function plots fold change within each hexagon, which are calculated with [make_hexbin](#). Note that the fold change is only accurate if the condition investigated is within the same individual. For conditions across different individuals different methods that account for individual-specific effects are required.

Value

A [ggplot2](#){ggplot} object.

Examples

```
# For SingleCellExperiment
library(TENxPBMCData)
library(scater)
tenx_pbmc3k <- TENxPBMCData(dataset = "pbmc3k")
rm_ind <- calculateAverage(tenx_pbmc3k) < 0.1
tenx_pbmc3k <- tenx_pbmc3k[!rm_ind, ]
colData(tenx_pbmc3k) <- cbind(colData(tenx_pbmc3k), perCellQCMetrics(tenx_pbmc3k))
tenx_pbmc3k <- logNormCounts(tenx_pbmc3k)
tenx_pbmc3k <- runPCA(tenx_pbmc3k)
tenx_pbmc3k <- make_hexbin(tenx_pbmc3k, 20, dimension_reduction = "PCA")
tenx_pbmc3k$random <- factor(sample(1:2, ncol(tenx_pbmc3k), replace = TRUE))
plot_hexbin_fc(tenx_pbmc3k, col = "random", feature = "ENSG00000187608", type = "counts")
```

---

plot_hexbin_feature   *Plot of feature expression of single cells in bivariate hexagon cells.*

Description

Plot of feature expression of single cells in bivariate hexagon cells.

Usage

```
plot_hexbin_feature(
  sce,
  mod = "RNA",
  type,
  feature,
  action,
  title = NULL,
  xlab = NULL,
  ylab = NULL,
  lower_cutoff = 0,
  upper_cutoff = 1
)
```

Arguments

sce	A [SingleCellExperiment](#) object.
mod	A string referring to the name of the modality used for plotting. For RNA modality use "RNA". For other modalities use name of alternative object for the [SingleCellExperiment](#) object.
type	A string referring to the type of assay in the [SingleCellExperiment](#) object.
feature	A string referring to the name of one feature.
action	A strings pecifying how meta data of observations in binned hexagon cells are to be summarized. Possible actions are prop_0, mode, mean and median (see details).
title	A string containing the title of the plot.
xlab	A string containing the title of the x axis.

ylab	A string containing the title of the y axis.
lower_cutoff	For mode, mean and median actions, remove expression values below this quantile. Expressed as decimal. Default: 0
upper_cutoff	For mode, mean and median actions, remove expression values above this quantile. Expressed as decimal. Default: 1

Details

This function plots the expression of any feature in the hexagon cell representation calculated with [make_hexbin](). The chosen gene expression is summarized by one of four actions prop_0, mode, mean and median:

prop_0 Returns the proportion of observations in the bin greater than 0. The associated meta data column needs to be numeric.

mode Returns the mode of the observations in the bin. The associated meta data column needs to be numeric.

mean Returns the mean of the observations in the bin. The associated meta data column needs to be numeric.

median Returns the median of the observations in the bin. The associated meta data column needs to be numeric.

Value

A [ggplot2]{ggplot} object.

Examples

```
# For SingleCellExperiment object
library(TENxPBMCData)
library(scater)
tenx_pbmc3k <- TENxPBMCData(dataset = "pbmc3k")
rm_ind <- calculateAverage(tenx_pbmc3k) < 0.1
tenx_pbmc3k <- tenx_pbmc3k[!rm_ind, ]
colData(tenx_pbmc3k) <- cbind(
    colData(tenx_pbmc3k),
    perCellQCMetrics(tenx_pbmc3k)
)
tenx_pbmc3k <- logNormCounts(tenx_pbmc3k)
tenx_pbmc3k <- runPCA(tenx_pbmc3k)
tenx_pbmc3k <- make_hexbin(tenx_pbmc3k, 20, dimension_reduction = "PCA")
plot_hexbin_feature(tenx_pbmc3k,
    type = "logcounts",
    feature = "ENSG00000135250", action = "median"
)
plot_hexbin_feature(tenx_pbmc3k,
    type = "logcounts",
    feature = "ENSG00000135250", action = "mode"
)
```

plot_hexbin_feature_plus
                          Plot of gene expression and meta data of single cell data in bivariate
                          hexagon cells.

Description

Plot of gene expression and meta data of single cell data in bivariate hexagon cells.

Usage

```
plot_hexbin_feature_plus(
  sce,
  col,
  mod = "RNA",
  type,
  feature,
  action,
  colors = NULL,
  title = NULL,
  xlab = NULL,
  ylab = NULL,
  expand_hull = 3,
  ...
)
```

Arguments

sce	A [SingleCellExperiment]().
col	A string referring to the name of one column in the meta data of sce by which to colour the hexagons.
mod	A string referring to the name of the modality used for plotting. For RNA modality use "RNA". For other modalities use name of alternative object for the [SingleCellExperiment]() object.
type	A string referring to the type of assay in the [SingleCellExperiment]() object.
feature	A string referring to the name of one feature.
action	A string specifying how gene expression of observations in binned hexagon cells are to be summarized. Possible actions are prop_0, mode, mean and median (see details).
colors	A vector of strings specifying which colors to use for plotting the different levels in the selected column of the meta data.
title	A string containing the title of the plot.
xlab	A string containing the title of the x axis.
ylab	A string containing the title of the y axis.
expand_hull	A numeric value determining the expansion of the line marking different clusters.
...	Additional arguments passed on to [ggforce](){geom_mark_hull}.

Details

This function plots any gene expresssion in the hexagon cell representation calculated with [make_hexbin](#) as well as at the same time representing outlines of clusters. The chosen gene expression is summarized by one of four actions prop_0, mode, mean and median:

prop_0 Returns the proportion of observations in the bin greater than 0. The associated meta data column needs to be numeric.

mode Returns the mode of the observations in the bin. The associated meta data column needs to be numeric.

mean Returns the mean of the observations in the bin. The associated meta data column needs to be numeric.

median Returns the median of the observations in the bin. The associated meta data column needs to be numeric.

Value

A [ggplot2](#){ggplot} object.

Examples

```
# For SingleCellExperiment object
library(TENxPBMCData)
library(scater)
tenx_pbmc3k <- TENxPBMCData(dataset = "pbmc3k")
rm_ind <- calculateAverage(tenx_pbmc3k) < 0.1
tenx_pbmc3k <- tenx_pbmc3k[!rm_ind, ]
tenx_pbmc3k <- logNormCounts(tenx_pbmc3k)
tenx_pbmc3k <- runPCA(tenx_pbmc3k)
tenx_pbmc3k <- make_hexbin(tenx_pbmc3k, 10, dimension_reduction = "PCA")
tenx_pbmc3k$random <- factor(sample(1:3, ncol(tenx_pbmc3k), replace = TRUE))
plot_hexbin_feature_plus(tenx_pbmc3k,
    col = "random", type = "counts",
    feature = "ENSG00000135250", action = "mean"
)
```

---

plot_hexbin_interact   *Plot of interaction of expression of single cells in bivariate hexagon cells.*

Description

Plot of interaction of expression of single cells in bivariate hexagon cells.

Usage

```
plot_hexbin_interact(
  sce,
  mod,
  type,
  feature,
  interact,
```

```
    title = NULL,
    xlab = NULL,
    ylab = NULL
)
```

Arguments

sce	A [SingleCellExperiment](#) object.
mod	A vector of strings referring to the names of the modularities. For [SingleCellExperiment](#) use "RNA" to access the RNA expression data stored as the main experiment type.
type	A vector of strings referring to the types of assays in the [SingleCellExperiment](#) object.
feature	A vector of strings referring to the names of one features in the same order as the vector of modularities.
interact	A string specifying how interaction between features is calculated. Possible interaction measures are corr_spearman and mi (see details).
title	A string containing the title of the plot.
xlab	A string containing the title of the x axis.
ylab	A string containing the title of the y axis.

Details

This function plots the interaction between any features in the hexagon cell representation calculated with [make_hexbin](#). The interaction between the chosen features is calculated by one of two measurers corr_spearman, ratio and mi:

mi Returns the mutual information coefficient.

corr_spearman Returns the Spearman correlation.

fc Return the log fold change between the features.

Note that fc should be applied to log normalized values.

Value

A [ggplot2](#){ggplot} object.

Examples

```
# For SingleCellExperiment
library(TENxPBMCData)
library(scater)
tenx_pbmc3k <- TENxPBMCData(dataset = "pbmc3k")
rm_ind <- calculateAverage(tenx_pbmc3k) < 0.1
tenx_pbmc3k <- tenx_pbmc3k[!rm_ind, ]
colData(tenx_pbmc3k) <- cbind(colData(tenx_pbmc3k), perCellQCMetrics(tenx_pbmc3k))
tenx_pbmc3k <- logNormCounts(tenx_pbmc3k)
tenx_pbmc3k <- runPCA(tenx_pbmc3k)
tenx_pbmc3k <- make_hexbin(tenx_pbmc3k, 10, dimension_reduction = "PCA")
plot_hexbin_interact(tenx_pbmc3k,
    type = c("counts", "counts"), mod = c("RNA", "RNA"),
    feature = c("ENSG00000146109", "ENSG00000102265"), interact = "fc"
)
```

plot_hexbin_meta   *Plot of meta data of single cell data in bivariate hexagon cells.*

Description

Plot of meta data of single cell data in bivariate hexagon cells.

Usage

```
plot_hexbin_meta(
  sce,
  col,
  action,
  no = 1,
  colors = NULL,
  title = NULL,
  xlab = NULL,
  ylab = NULL,
  na.rm = FALSE
)
```

Arguments

sce	A [SingleCellExperiment](#) object.
col	A string referring to the name of one column in the meta data of sce by which to colour the hexagons.
action	A string specifying how meta data of observations in binned hexagon cells are to be summarized. Possible actions are majority, prop, prop_0, mode, mean and median (see details).
no	An integer specifying which level to plot of the column. Only in effect when action=prop.
colors	A vector of strings specifying which colors to use for plotting the different levels in the selected column of the meta data. Only in effect when the selected action="majority".
title	A string containing the title of the plot.
xlab	A string containing the title of the x axis.
ylab	A string containing the title of the y axis.
na.rm	Logical indicating whether NA values should be removed.

Details

This function plots any column of the meta data in the hexagon cell representation calculated with [make_hexbin](#). The chosen meta data column is summarized by one of six actions majority, prop, prop_0, mode, mean and median:

majority Returns the value of the majority of observations in the bin. The associated meta data column needs to be a factor or character.

prop Returns the proportion of each level or unique character in the bin. The associated meta data column needs to be a factor or character

prop_0 Returns the proportion of observations in the b factor or character in the bin greater than 0. The associated meta data column needs to be numeric.

mode Returns the mode of the observations in the bin. The associated meta data column needs to be numeric.

mean Returns the mean of the observations in the bin. The associated meta data column needs to be numeric.

median Returns the median of the observations in the bin. The associated meta data column needs to be numeric.

Value

A [ggplot2](ggplot) object.

Examples

```
# For SingleCellExperiment object
library(TENxPBMCData)
library(scater)
tenx_pbmc3k <- TENxPBMCData(dataset = "pbmc3k")
rm_ind <- calculateAverage(tenx_pbmc3k) < 0.1
tenx_pbmc3k <- tenx_pbmc3k[-rm_ind, ]
colData(tenx_pbmc3k) <- cbind(
    colData(tenx_pbmc3k),
    perCellQCMetrics(tenx_pbmc3k)
)
tenx_pbmc3k <- logNormCounts(tenx_pbmc3k)
tenx_pbmc3k <- runPCA(tenx_pbmc3k)
tenx_pbmc3k <- make_hexbin(tenx_pbmc3k, 20, dimension_reduction = "PCA")
plot_hexbin_meta(tenx_pbmc3k, col = "total", action = "median")
```

---

plot_hexbin_meta_plus   *Plot of meta data with annotation of single cell data in bivariate hexagon cells.*

Description

Plot of meta data with annotation of single cell data in bivariate hexagon cells.

Usage

```
plot_hexbin_meta_plus(
  sce,
  col1,
  col2,
  action,
  no = 1,
  colors = NULL,
  title = NULL,
  xlab = NULL,
  ylab = NULL,
  expand_hull = 3,
  na.rm = FALSE,
```

```
    ...
 )
```

Arguments

sce	A [SingleCellExperiment](#) object.
col1	A string referring to the name of one column in the meta data of sce by which to make the outlines. Note that this should be a factor or a character.
col2	A string referring to the name of one column in the meta data of sce specifying what to color hexagons by.
action	A string specifying how meta data as specified in col2 of observations in binned hexagon cells are to be summarized. Possible actions are prop, mode, mean and median (see details).
no	An integer specifying which level to plot of the column. Only in effect when action=prop.
colors	A vector of strings specifying which colors to use for plotting the different levels in the selected column of the meta data.
title	A string containing the title of the plot.
xlab	A string containing the title of the x axis.
ylab	A string containing the title of the y axis.
expand_hull	A numeric value determining the expansion of the line marking different clusters.
na.rm	Logical indicating whether NA values should be removed.
...	Additional arguments passed on to [ggforce](#){geom_mark_hull}.

Details

This function plots any meta data in the hexagon cell representation calculated with [make_hexbin](#) as well as at the same time representing outlines of clusters. The chosen gene expression is summarized by one of four actions prop_0, mode, mean and median:

prop Returns the proportion of each level or unique character in the bin. The associated meta data column needs to be a factor or character.

mode Returns the mode of the observations in the bin. The associated meta data column needs to be numeric.

mean Returns the mean of the observations in the bin. The associated meta data column needs to be numeric.

median Returns the median of the observations in the bin. The associated meta data column needs to be numeric.

Value

A [ggplot2](#){ggplot} object.

Examples

```
# For SingleCellExperiment object
library(TENxPBMCData)
library(scater)
tenx_pbmc3k <- TENxPBMCData(dataset = "pbmc3k")
rm_ind <- calculateAverage(tenx_pbmc3k) < 0.1
tenx_pbmc3k <- tenx_pbmc3k[-rm_ind, ]
colData(tenx_pbmc3k) <- cbind(colData(tenx_pbmc3k), perCellQCMetrics(tenx_pbmc3k))
tenx_pbmc3k <- logNormCounts(tenx_pbmc3k)
tenx_pbmc3k <- runPCA(tenx_pbmc3k)
tenx_pbmc3k <- make_hexbin(tenx_pbmc3k, 20, dimension_reduction = "PCA")
tenx_pbmc3k$random <- factor(sample(1:3, ncol(tenx_pbmc3k), replace = TRUE))
tenx_pbmc3k$random <- as.factor(tenx_pbmc3k$random)
plot_hexbin_meta_plus(tenx_pbmc3k, col1 = "random", col2 = "total", action = "median")
```

---

schex   *schex: A package for plotting hexbin plots for single cell omics data.*

Description

Builds hexbin plots for variables and dimension reduction stored single cell omics data such as SingleCellExperiment. The ideas used in this package are based on the excellent work of Dan Carr, Nicholas Lewin-Koh, Martin Maechler and Thomas Lumley.

Details

Please see the help pages listed below:

- make_hexbin
- plot_hexbin_feature
- plot_hexbin_meta

Also see the vignettes for more usage examples.

Please report issues and suggest improvements at Github:

https://github.com/SaskiaFreytag/schex

Index