Package 'netSmooth'

October 16, 2019

Type Package

Title Network smoothing for scRNAseq

Version 1.4.0

Description netSmooth is an R package for network smoothing of single cell RNA sequencing data. Using bio networks such as protein-protein interactions as priors for gene co-expression, netsmooth improves cell type identification from noisy, sparse scRNAseq data.

biocViews Network, GraphAndNetwork, SingleCell, RNASeq, GeneExpression, Sequencing, Transcriptomics, Normalization, Preprocessing, Clustering, DimensionReduction

URL https://github.com/BIMSBbioinfo/netSmooth

BugReports https://github.com/BIMSBbioinfo/netSmooth/issues

License GPL-3

Encoding UTF-8

LazyData true

Depends R (>= 3.5), scater (>= 1.9.20), clusterExperiment (>= 2.1.6)

Imports entropy, SummarizedExperiment, SingleCellExperiment, Matrix, cluster, data.table, stats, methods, DelayedArray, HDF5Array

Suggests knitr, testthat, Rtsne, biomaRt, igraph, STRINGdb, NMI, pheatmap, ggplot2, BiocStyle, rmarkdown, BiocParallel

VignetteBuilder knitr

RoxygenNote 6.1.1

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

git_branch RELEASE_3_9

git_last_commit 679ff5e

git_last_commit_date 2019-05-02

Date/Publication 2019-10-15

Author Jonathan Ronen [aut, cre], Altuna Akalin [aut]

Maintainer Jonathan Ronen <yablee@gmail.com>

R topics documented:

human.ppi 2
mouse.ppi 2
netSmooth,matrix-method 3
pickDimReduction,matrix-method 4
robustClusters,SummarizedExperiment-method 5
smallPPI 6
smallscRNAseq 6

Index **7**

`human.ppi` *Human Protein-Protein interaction graph*

Description

An adjacency matrix of the 10 percent highest confidence interactions between human proteins on STRINGdb.

Usage

```
human.ppi
```

Format

A square matrix where A_ij=1 if gene i interacts with gene j

Details

See the script in ‘system.file(package="netSmooth", "data-raw", "make_ppi_from_string.R")‘ for full details of how this object was made.

Source

http://www.string-db.org/

`mouse.ppi` *Mouse Protein-Protein interaction graph*

Description

An adjacency matrix of the 10 percent highest confidence interactions between mouse proteins on STRINGdb.

Usage

```
mouse.ppi
```

Format

A square matrix where A_ij=1 if gene i interacts with gene j

Details

See the script in ‘system.file(package="netSmooth", "data-raw", "make_ppi_from_string.R")‘ for full details of how this object was made.

Source

http://www.string-db.org/

`netSmooth,matrix-method`

Perform network smoothing of gene expression or other omics data

---

Description

Perform network smoothing of gene expression or other omics data

Usage

```
## S4 method for signature 'matrix'
netSmooth(x, adjMatrix, alpha = "auto",
  normalizeAdjMatrix = c("rows", "columns"),
  autoAlphaMethod = c("robustness", "entropy"), autoAlphaRange = 0.1 *
  (seq_len(9)), autoAlphaDimReduceFlavor = "auto", is.counts = TRUE,
  bpparam = BiocParallel::SerialParam(), ...)

## S4 method for signature 'SummarizedExperiment'
netSmooth(x, ...)

## S4 method for signature 'SingleCellExperiment'
netSmooth(x, ...)

## S4 method for signature 'Matrix'
netSmooth(x, adjMatrix, alpha = "auto",
  normalizeAdjMatrix = c("rows", "columns"),
  autoAlphaMethod = c("robustness", "entropy"), autoAlphaRange = 0.1 *
  (seq_len(9)), autoAlphaDimReduceFlavor = "auto", is.counts = TRUE,
  bpparam = BiocParallel::SerialParam(), ...)

## S4 method for signature 'DelayedMatrix'
netSmooth(x, adjMatrix, alpha = "auto",
  normalizeAdjMatrix = c("rows", "columns"),
  autoAlphaMethod = c("robustness", "entropy"), autoAlphaRange = 0.1 *
  (seq_len(9)), autoAlphaDimReduceFlavor = "auto", is.counts = TRUE,
  bpparam = BiocParallel::SerialParam(), chunk.size = 1,
  filepath = NULL, ...)
```

Arguments

x	matrix or SummarizedExperiment
adjMatrix	adjacency matrix of gene network to use

alpha	numeric in [0,1] or 'audo'. if 'auto', the optimal value for alpha will be automatically chosen among the values specified in 'autoAlphaRange', using the strategy specified in 'autoAlphaMethod'
normalizeAdjMatrix	how to normalize the adjacency matrix possible values are 'rows' (in-degree) and 'columns' (out-degree)
autoAlphaMethod	if 'robustness', pick alpha that gives the highest proportion of samples in robust clusters if 'entropy', pick alpha that gives highest Shannon entropy in 2D PCA embedding
autoAlphaRange	if 'alpha='optimal'', search these values for the best alpha
autoAlphaDimReduceFlavor	algorithm for dimensionality reduction that will be used to pick the optimal value for alpha. Either the 2D embedding to calculate the Shannon entropy for (if 'autoAlphaMethod='entropy''), or the dimensionality reduction algorithm to be used in robust clustering (if 'autoAlphamethod='robustness'')
is.counts	logical: is the assay count data
bpparam	instance of bpparam, for parallel computation with the 'alpha='auto'' option. See the BiocParallel manual.
...	arguments passed on to 'robustClusters' if using the robustness criterion for optimizing alpha
chunk.size	integer in [1,length(colnames[x])]. Number of columns that processed at the same time when using disk based DelayedMatrix. Will be ignored when regular matrices or SummarizedExperiment are used as input.
filepath	String: Path to location where hdf5 output file is supposed to be saved. Will be ignored when regular matrices or SummarizedExperiment are used as input.

Value

network-smoothed gene expression matrix or SummarizedExperiment object

Examples

```
x <- matrix(rnbinom(12000, size=1, prob = .1), ncol=60)
rownames(x) <- paste0('gene', seq_len(dim(x)[1]))

adj_matrix <- matrix(as.numeric(rnorm(200*200)>.8), ncol=200)
rownames(adj_matrix) <- colnames(adj_matrix) <- paste0('gene', seq_len(dim(x)[1]))
x.smoothed <- netSmooth(x, adj_matrix, alpha=0.5)
```

---

pickDimReduction,matrix-method

Pick the dimensionality reduction method for a dataset that gives the 2D embedding with the highest entropy

---

Description

Pick the dimensionality reduction method for a dataset that gives the 2D embedding with the highest entropy

Usage

```
## S4 method for signature 'matrix'
pickDimReduction(x, flavors = c("pca", "tsne",
  "umap"), is.counts = TRUE)

## S4 method for signature 'SummarizedExperiment'
pickDimReduction(x)

## S4 method for signature 'Matrix'
pickDimReduction(x, flavors = c("pca", "tsne",
  "umap"), is.counts = TRUE)

## S4 method for signature 'DelayedMatrix'
pickDimReduction(x, flavors = c("pca", "tsne",
  "umap"), is.counts = TRUE)
```

Arguments

x	matrix or SummarizedExperiment object [GENES x SAMPLES]
flavors	list of dimensionality reduction algorithms to try. Currently the options are "pca", "tsne" and "umap"
is.counts	logical: is exprs count data

Value

name of dimensionality reduction method that gives the highest 2d entropy

Examples

```
x <- matrix(rnbinom(60000, size=1, prob = .1), ncol=100)
pickDimReduction(x)
```

---

robustClusters,SummarizedExperiment-method

Perform robust clustering on dataset, and calculate the proportion of samples in robust clusters

---

Description

Perform robust clustering on dataset, and calculate the proportion of samples in robust clusters

Usage

```
## S4 method for signature 'SummarizedExperiment'
robustClusters(x,
  dimReduceFlavor = "auto", is.counts = TRUE, ...)

## S4 method for signature 'matrix'
robustClusters(x, ...)
```

Arguments

x	matrix or SummarizedExperiment object
dimReduceFlavor	algorithm for dimensionality reduction step of clustering procedure. May be 'pca', 'tsne', 'dm', 'umap' or 'auto', which uses shannon entropy to pick the algorithm.
is.counts	logical: is the data counts
...	arguments passed on to 'clusterExperimentWorkflow'

Value

list(clusters, proportion.robust)

Examples

```
data("smallscRNAseq")
robustClusters(smallscRNAseq, dimReduceFlavor='pca')
```

---

smallPPI — *A small human Protein-Protein interaction graph for use in examples.*

Description

Contains a synthetic PPI of human genes.

Usage

```
smallPPI
```

Format

An object of class `matrix` with 611 rows and 611 columns.

---

smallscRNAseq — *A small single cell RNA-seq dataset for use in examples.*

Description

Contains scRNAseq profiles of human blastomeres.

Usage

```
smallscRNAseq
```

Format

SingleCellExperiment

Source

https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE44183

Index

∗Topic **datasets**
human.ppi, 2
mouse.ppi, 2
smallPPI, 6
smallscRNAseq, 6

human.ppi, 2

mouse.ppi, 2

netSmooth *(*netSmooth,matrix-method*)*, 3
netSmooth,DelayedMatrix-method
*(*netSmooth,matrix-method*)*, 3
netSmooth,Matrix-method
*(*netSmooth,matrix-method*)*, 3
netSmooth,matrix-method, 3
netSmooth,SingleCellExperiment-method
*(*netSmooth,matrix-method*)*, 3
netSmooth,SummarizedExperiment-method
*(*netSmooth,matrix-method*)*, 3

pickDimReduction
*(*pickDimReduction,matrix-method*)*,
4
pickDimReduction,DelayedMatrix-method
*(*pickDimReduction,matrix-method*)*,
4
pickDimReduction,Matrix-method
*(*pickDimReduction,matrix-method*)*,
4
pickDimReduction,matrix-method, 4
pickDimReduction,SummarizedExperiment-method
*(*pickDimReduction,matrix-method*)*,
4

robustClusters
*(*robustClusters,SummarizedExperiment-method*)*,
5
robustClusters,matrix-method
*(*robustClusters,SummarizedExperiment-method*)*,
5
robustClusters,SummarizedExperiment-method,
5

smallPPI, 6
smallscRNAseq, 6