Package 'SiPSiC'

October 19, 2024

Type Package

Title Calculate Pathway Scores for Each Cell in scRNA-Seq Data

Version 1.5.4

Date 2022-12-29

Description Infer biological pathway activity of cells from single-cell RNA-sequencing data by calculating a pathway score for each cell (pathway genes are specified by the user). It is recommended to have the data in Transcripts-Per-Million (TPM) or Counts-Per-Million (CPM) units for best results. Scores may change when adding cells to or removing cells off the data. SiPSiC stands for Single Pathway analysis in Single Cells.

License file LICENSE

URL https://www.genome.org/cgi/doi/10.1101/gr.278431.123

biocViews Software, DifferentialExpression, GeneSetEnrichment, BiomedicalInformatics, CellBiology, Transcriptomics, RNASeq, SingleCell, Transcription, Sequencing, ImmunoOncology, DataImport

Depends Matrix, SingleCellExperiment

Suggests knitr, rmarkdown, BiocStyle

VignetteBuilder knitr

BugReports https://github.com/DanielDavis12/SiPSiC/issues

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

git_branch devel

git_last_commit 7d3747e

git_last_commit_date 2024-08-06

Repository Bioconductor 3.20

Date/Publication 2024-10-18

Author Daniel Davis [aut, cre] (<https://orcid.org/0000-0001-7863-1521>), Yotam Drier [aut]

Maintainer Daniel Davis <DanielDavis000@gmail.com>

Contents

`SiPSiC-package` *Calculate Pathway Scores for Each Cell in scRNA-Seq Data*

Description

Infer biological pathway activity of cells from single-cell RNA-sequencing data by calculating a pathway score for each cell (pathway genes are specified by the user). It is recommended to have the data in Transcripts-Per-Million (TPM) or Counts-Per-Million (CPM) units for best results. Scores may change when adding cells to or removing cells off the data. SiPSiC stands for Single Pathway analysis in Single Cells.

Details

Use this package to calculate per-cell scores for a biological pathway of your choice, from single-cell RNA-seq data. Transcripts-Per-Million (TPM) or Counts-Per-Million (CPM) units of the data are recommended for best results.

Author(s)

Daniel Davis, Yotam Drier. Maintainer: Daniel Davis

References

https://medicine.ekmd.huji.ac.il/en/research/yotamd/Pages/default.aspx

See Also

`getPathwayScores`

`getPathwayScores` *Calculate Pathway Scores for all Cells in a scRNA-Seq Data*

Description

Calculate a pathway score for each cell included in the input scRNA-seq data, using the cell's transcription level of the pathway's genes.

Usage

```
getPathwayScores(dataMatrix, pathwayGenes, percentForNormalization)
```

Arguments

a list of three items:

`dataMatrix`	a matrix whose rows and columns represent genes and cells, respectively, containing the scRNA-seq data; Counts-Per-Million (CPM) or Transcripts-Per-Million (TPM) units should be used for best results. The matrix should be of type sparse matrix (dgCMatrix), otherwise errors might be raised.
`pathwayGenes`	a character vector of the gene names of which the relevant biological pathway consists.
`percentForNormalization`	The percent of top cells for each gene whose median is used to normalize the gene's expression values (5 by default).

Value

`pathwayScores`	an array (type double) of the pathway score of each cell in the input dataMatrix, less than two of the pathway genes are found in the data, in which case NA is returned.
`index`	a numeric array of the row indices in the dataMatrix where genes of the pathway were found.

Author(s)

Daniel Davis, Yotam Drier

References

https://medicine.ekmd.huji.ac.il/en/research/yotamd/Pages/default.aspx

Examples

```
library(SiPSiC)
geneCountsMatrix <- matrix(rpois(16, lambda = 10), ncol = 4, nrow = 4)
geneCountsMatrix <- as(geneCountsMatrix, "dgCMatrix")
rownames(geneCountsMatrix) <- c("Gene1", "Gene2", "Gene3", "Gene4")
colnames(geneCountsMatrix) <- c("Cell1", "Cell2", "Cell3", "Cell4")
assayData <- SingleCellExperiment(assays = list(counts = geneCountsMatrix))
pathwayGenesList <- c("Gene1", "Gene2", "Gene4")
percentForNormalization <- 7
scoresAndIndices <- getPathwayScores(counts(assayData), pathwayGenesList, percentForNormalization)
pathwayScoresOfCells <- scoresAndIndices$pathwayScores
pathwayGeneIndices <- scoresAndIndices$index
```

---

normalizeCountsForGene

Gene counts normalization

Description

Get the counts of a single gene normalized by the median of the top 5 percent cells, unless it's zero; In this case, the counts are all divided by the maximum value across all cells. If all counts are zeros, they are returned untouched.

Usage

```
normalizeCountsForGene(expressionValues, percentForNormalization)
```

Arguments

`expressionValues`
: An array of type double, containing the counts (in any units, e.g. CPM or TPM) of a single gene across different cells.

`percentForNormalization`
: The percent of top cells for each gene whose median is used to normalized the gene's expression values.

Value

An array (type double) of the normalized input counts.

Author(s)

Daniel Davis, Yotam Drier

Index