Package 'emtdata'

November 10, 2026

Title An ExperimentHub Package for data sets with an Epithelial to Mesenchymal Transition (EMT)

Version 1.20.0

Description This package provides pre-processed RNA-seq data where the epithelial to mesenchymal transition was induced on cell lines. These data come from three publications Cursons et al. (2015), Cursons etl al. (2018) and Foroutan et al. (2017). In each of these publications, EMT was induces across multiple cell lines following treatment by TGFb among other stimulants. This data will be useful in determining the regulatory programs modified in order to achieve an EMT. Data were processed by the Davis laboratory in the Bioinformatics division at WEHI.

biocViews ExperimentHub, Homo_sapiens_Data, RNASeqData, ExpressionData

Encoding UTF-8

LazyData true

Roxygen list(markdown = TRUE)

RoxygenNote 7.1.1

Depends R (>= 4.1)

Imports edgeR, ExperimentHub, SummarizedExperiment

Suggests testthat (>= 3.0.0), stringr, plyr, prettydoc, BiocStyle, Homo.sapiens, RColorBrewer, rmarkdown, BiocFileCache, knitr

URL https://github.com/DavisLaboratory/emtdata

BugReports https://github.com/DavisLaboratory/emtdata/issues

NeedsCompilation no

VignetteBuilder knitr

Config/testthat/edition 3

License GPL-3

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

git_branch RELEASE_3_23

git_last_commit e57905c

git_last_commit_date 2026-04-28

Repository Bioconductor 3.23

Date/Publication 2026-11-10

Author Malvika D. Kharbanda [aut, cre] (ORCID:
<https://orcid.org/0000-0001-9726-3023>),
Chin Wee Tan [aut] (ORCID: <https://orcid.org/0000-0001-9695-7218>),
Dharmesh D. Bhuva [aut] (ORCID:
<https://orcid.org/0000-0002-6398-9157>)

Maintainer Malvika D. Kharbanda <kharbanda.m@wehi.edu.au>

Contents

asDGEList 2
cursons2015_se 3
cursons2018_se 4
foroutan2017_se 4

Index **6**

`asDGEList` *Convert a SummarizedExperiment object to a DGEList object*

Description

This function converts a SummarizedExperiment object to a DGEList object to enhance differential expression analysis using the edgeR package.

Usage

```
asDGEList(se, assay_name = "counts")
```

Arguments

`se`	a SummarizedExperiment object.
`assay_name`	a character, specifying the assay to retrieve.

Value

a DGEList object

Examples

```
library(ExperimentHub)

eh = ExperimentHub()
query(eh, 'emtdata')

cursons2018_se = eh[['EH5440']]
```

```
cursons2018_dge = asDGEList(cursons2018_se)
```

`cursons2015_se` *EGF or hypoxia treatment of Breast Cancer Cell lines from Cursons et al. (2015)*

Description

Dataset used in Cursons et al. 2015. where parallel 'deep sequencing' of RNA (RNA-Seq) were conducted to examine the changes in expression profiles between breast cancer cell lines PMC42-ET, PMC42-LA and MDA-MB-468 cells.

Format

A SummarizedExperiment object, containing gene expression data of different sub-lines of human breast cancer cell lines. The `SummarizedExperiment::colData()` function can be used to access the sample annotations.

Details

For PMC42 cell lines, cells were treated for 3 or 7 days in the presence or absence of 10 ng/ml EGF. For MDA-MB-468 cells, cells were treated for 7 days in the presence or absence of either 10 ng/ml EGF or kept under Hypoxia.

There are 3 biological replicates per condition which were summed. Samples were sequenced on the Illumina HiSeq 2000, 100bp paired end.

Data from this publication were downloaded from EMBL-EBI ENA and processed into a SummarizedExperiment object. Sample annotations were modified from the original publication and SRA portal.

References

Cursons, J., Leuchowius, KJ., Waltham, M., Tomaskovic-Crook, E., Foroutan, M., Bracken, CP., Redfern,A., Crampin, EJ., Street, I., Davis, MJ. & Thompson,EW. (2015) Stimulus-dependent differences in signalling regulate epithelial-mesenchymal plasticity and change the effects of drugs in breast cancer cell lines. Cell Commun Signal 13, 26 (2015).

Examples

```
library(ExperimentHub)
eh <- ExperimentHub()
emt_datasets <- query(eh, "emtdata")
```

cursons2018_se *Combinatorial miRNAs in breast cancer EMT from Cursons et al. (2018)*

Description

Cursons et al. used the Human Mammary Epithelial Cells (HMLE) cell line dataset. A mesenchymal HMLE (mesHMLE) phenotype was induced following treatment with transforming growth factors (TGFb). The mesHMLE subline was then treated with mir200c to reinduce an epithelial phenotype. All mRNA RNA-seq was collected using the Illumina HiSeq 2500 with a paired end read length of 100bp.

Format

A SummarizedExperiment object, containing gene expression data of the Human Mammary Epithelial Cells (HMLE) cell line. The `SummarizedExperiment::colData()` function can be used to access the sample annotations.

Details

Data from this publication were downloaded from the European Nucleotide Archive (ENA) and processed into a SummarizedExperiment object. Sample annotations were modified from the original publication.

References

Cursons, J., Pillman, K. A., Scheer, K. G., Gregory, P. A., Foroutan, M., Hediyeh-Zadeh, S., ... & Davis, M. J. (2018). Combinatorial targeting by microRNAs co-ordinates post-transcriptional control of EMT. Cell systems, 7(1), 77-91.

Examples

```
library(ExperimentHub)
eh <- ExperimentHub()
emt_datasets <- query(eh, "emtdata")
```

foroutan2017_se *TGFb stimulated cell lines from Foroutan et al. (2017)*

Description

Foroutan et al. collected data from multiple different studies where various cell lines (from various tissues of origin) were stimulated with transforming growth factor beta (TGFb) to induce an epithelial to mesenchymal transition (EMT). Since the data were from different studies (and platforms), the surrogate variable analysis (SVA) and ComBat approaches were applied to correct for batch effects.

Format

A SummarizedExperiment object, containing gene expression data of human cell lines treated with TGFb. The `SummarizedExperiment::colData()` function can be used to access the sample annotations.

Details

Data from this publication were downloaded from figshare and processed into a SummarizedExperiment object. Sample annotations were retrieved from the original publication.

References

Foroutan, M., Cursons, J., Hediyeh-Zadeh, S., Thompson, E. W., & Davis, M. J. (2017). A transcriptional program for detecting TGFb-induced EMT in Cancer. Molecular Cancer Research, 15(5), 619-631.

Examples

```
library(ExperimentHub)
eh <- ExperimentHub()
emt_datasets <- query(eh, "emtdata")
```

Index

asDGEList, 2

cursons2015_se, 3
cursons2018_se, 4

foroutan2017_se, 4

SummarizedExperiment::colData(), *3–5*