

# Package ‘Single.mTEC.Transcriptomes’

June 28, 2022

**Type** Package

**Title** Single Cell Transcriptome Data and Analysis of Mouse mTEC cells

**Version** 1.24.0

**Date** 2021-11-21

**Author** Alejandro Reyes

**Maintainer** Alejandro Reyes <alejandro.reyes.ds@gmail.com>

**Description** This data package contains the code used to analyse the single-cell RNA-seq and the bulk ATAC-seq data from the manuscript titled: Single-cell transcriptome analysis reveals coordinated ectopic-gene expression patterns in medullary thymic epithelial cells. This paper was published in Nature Immunology 16,933-941(2015). The data objects provided in this package has been pre-processed: the raw data files can be downloaded from ArrayExpress under the accession identifiers E-MTAB-3346 and E-MTAB-3624. The vignette of this data package provides a documented and reproducible workflow that includes the code that was used to generate each statistic and figure from the manuscript.

**License** LGPL

**biocViews** ExperimentData

**VignetteBuilder** knitr

**Suggests** DESeq2, GenomicRanges, GenomicFeatures, genefilter, statmod, gdata, RColorBrewer, ggplot2, gplots, cluster, clue, grid, gridExtra, ggbio, Gviz, geneplotter, matrixStats, pheatmap, BiocStyle, knitr, BiocParallel

**RoxygenNote** 5.0.1

**NeedsCompilation** no

**git\_url** <https://git.bioconductor.org/packages/Single.mTEC.Transcriptomes>

**git\_branch** RELEASE\_3\_15

**git\_last\_commit** f6796f1

**git\_last\_commit\_date** 2022-04-26

**Date/Publication** 2022-06-28

**R topics documented:**

Single.mTec.Transcriptomes-package	2
aireDependentSansom	3
biotype	3
biotypesHuman	3
cealCoexpression	4
corMatsNoMarker	4
deGenesNone	5
deGenesSansom	5
dxdATAC	5
fantom	6
geneNames	6
geneNamesHuman	6
geneRanges	7
mTECdxd	7
muc1Coexpression	7
nomarkerCellsClustering	8
percentsGG	8
permutationResults	8
scLVM_output	9
tras	9
<b>Index</b>	<b>10</b>

---

Single.mTec.Transcriptomes-package  
*Single-cell transcriptome data of medullary thymic epithelial cells*

---

**Description**

This document contains all the code used to analyse the single-cell RNA-seq and the bulk ATAC-seq data from the manuscript by Brennecke et al, 2015. The purpose of this package is to provide full reproducibility of the results presented in the manuscript. This package provides a documented and reproducible workflow of the code that was used to generate each number and figure from the manuscript.

**References**

Brennecke et al. Single-cell transcriptome analysis reveals coordinated ectopic gene-expression patterns in medullary thymic epithelial cells. *Nature Immunology* 16,933-941 (2015)

---

aireDependentSansom	<i>List of Aire-dependent genes</i>
---------------------	-------------------------------------

---

**Description**

A character vector of ensembl gene identifiers defined by Sansom et al, 2014. This list was downloaded from the supplementary material of the manuscript.

**Usage**

```
data(aireDependentSansom)
```

**References**

Sansom et al. Population and single-cell genomics reveal the Aire dependency, relief from Polycomb silencing, and distribution of self-antigen expression in thymic epithelia. *Genome Res.* 24, 1918–1931 (2014).

---

biotype	<i>Mouse biotypes</i>
---------	-----------------------

---

**Description**

A character vector defining biotype for each mouse ensembl gene identifier. Biotypes were queried using **biomaRt**.

**Usage**

```
data(biotypes)
```

---

biotypesHuman	<i>Human biotypes</i>
---------------	-----------------------

---

**Description**

A character vector defining biotype for each human ensembl gene identifier. Biotypes were queried using **biomaRt**.

**Usage**

```
data(biotypesHuman)
```

---

cea1Coexpression      *Cea1 co-expression group data*

---

**Description**

A data frame containing information about the Cea1 co-expression group defined in Pinto et al, 2013.

**Usage**

```
data(cea1Coexpression)
```

**References**

Pinto et al. Overlapping gene coexpression patterns in human medullary thymic epithelial cells generate self-antigen diversity. Proc. Natl. Acad. Sci. U.S.A. 110, E3497–3505, (2013).

---

corMatsNoMarker      *Gene-gene correlation network*

---

**Description**

A matrix containing the gene-gene Spearman correlation across single unselected cells.

A matrix containing the gene-gene Spearman correlation across single unselected cells, using the data from Sansom et al, 2014.

**Usage**

```
data(corMatsNoMarker)
```

```
data(corMatsSansom)
```

**References**

Sansom et al. Population and single-cell genomics reveal the Aire dependency, relief from Polycomb silencing, and distribution of self-antigen expression in thymic epithelia. Genome Res. 24, 1918–1931 (2014).

---

deGenesNone	<i>Highly variable genes.</i>
-------------	-------------------------------

---

**Description**

List of highly variable genes according to the method by Brennecke et al, 2013.

**Usage**

```
data(deGenesNone)
```

**References**

Brennecke et al. Accounting for technical noise in single-cell RNA-seq experiments. Nat. Methods. 10, 1093-1095 (2013).

---

deGenesSansom	<i>Highly variable genes from Sansom et al.</i>
---------------	---

---

**Description**

List of highly variable genes according to the method by Brennecke et al, 2013, using the data by Santom et al.

**Usage**

```
data(deGenesSansom)
```

**References**

Brennecke et al. Accounting for technical noise in single-cell RNA-seq experiments. Nat. Methods. 10, 1093-1095 (2013). Sansom et al. Population and single-cell genomics reveal the Aire dependency, relief from Polycomb silencing, and distribution of self-antigen expression in thymic epithelia. Genome Res. 24, 1918–1931 (2014).

---

dxdATAC	<i>ATAC-seq summarized counts.</i>
---------	------------------------------------

---

**Description**

DESeqDataSet object summarizing the ATAC-seq data presented in the manuscript. This object contains the read counts of each sample over a window of 4Kb around transcription start sites.

**Usage**

```
data(dxdATAC)
```

---

fantom	<i>FANTOM dataset</i>
--------	-----------------------

---

**Description**

DESeqDataSet object of the read counts from selected tissues from the FANTOM dataset.

**Usage**

```
data(fantom)
```

**References**

Forrest et al. A promoter-level mammalian expression atlas. 24, 1918–1931. (2014).

---

geneNames	<i>Mouse gene names</i>
-----------	-------------------------

---

**Description**

Character vector of mouse gene names. The gene names were queried using **biomaRt**.

**Usage**

```
data(geneNames)
```

---

geneNamesHuman	<i>Human gene names</i>
----------------	-------------------------

---

**Description**

Character vector of human gene names. The gene names were queried using **biomaRt**.

**Usage**

```
data(geneNamesHuman)
```

---

geneRanges	<i>Genomic range coordinates.</i>
------------	-----------------------------------

---

**Description**

GenomicRanges object containing the coordinate ranges of mouse protein coding genes used in the manuscript.

**Usage**

```
data(geneRanges)
```

---

mTECdx	<i>Count data from the single-mTEC data.</i>
--------	--

---

**Description**

DESeqDataSet object containing the read counts from the single-cell RNA-seq dataset generated by Brennecke et al, 2015.

**Usage**

```
data(mTECdx)
```

**References**

Brennecke et al. Single-cell transcriptome analysis reveals coordinated ectopic gene-expression patterns in medullary thymic epithelial cells. *Nature Immunology* 16,933-941 (2015)

---

muc1Coexpression	<i>Muc1 co-expression group data</i>
------------------	--------------------------------------

---

**Description**

A data frame containing information about the Muc1 co-expression group defined in Pinto et al, 2013.

**Usage**

```
data(muc1Coexpression)
```

**References**

Pinto et al. Overlapping gene coexpression patterns in human medullary thymic epithelial cells generate self-antigen diversity. *Proc. Natl. Acad. Sci. U.S.A.* 110, E3497–3505, (2013).

nomarkerCellsClustering

*Clustering results*

---

**Description**

List containing the results from the clustering analysis from Figure 2 from the manuscript.

**Usage**

data(nomarkerCellsClustering)

---

percentsGG

*Mapping statistics*

---

**Description**

Data frame containing the mapping statistics from the single-cell RNA-seq data.

**Usage**

data(percentsGG)

---

permutationResults

*Permutation results*

---

**Description**

Result from the gene permutations from Figure 5 from the manuscript.

**Usage**

data(permutationResults)



---

scLVM_output	<i>Output from scLVM</i>
--------------	--------------------------

---

**Description**

Objects containing the output scLVM on the single-mTEC RNA-seq data.

**Usage**

```
data(scLVM_output)
```

**References**

Buettner et al. Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. Nat. Biotechnol. 2015.

---

tras	<i>Tissue restricted antigens lists.</i>
------	--

---

**Description**

Data frame containing the TRA classification from Pinto et al, 2013.

**Usage**

```
data(tras)
```

# Index

## \* datasets

- aireDependentSansom, 3
  - biotype, 3
  - biotypesHuman, 3
  - cea1Coexpression, 4
  - corMatsNoMarker, 4
  - deGenesNone, 5
  - deGenesSansom, 5
  - dx dATAC, 5
  - fantom, 6
  - geneNames, 6
  - geneNamesHuman, 6
  - geneRanges, 7
  - mTECdx d, 7
  - muc1Coexpression, 7
  - nomarkerCellsClustering, 8
  - percentsGG, 8
  - permutationResults, 8
  - scLVM\_output, 9
  - tras, 9
  - geneNames, 6
  - geneNamesHuman, 6
  - geneRanges, 7
  - h5GeneNames (scLVM\_output), 9
  - heter (scLVM\_output), 9
  - mTECdx d, 7
  - muc1Coexpression, 7
  - nomarkerCellsClustering, 8
  - percentsGG, 8
  - permsAllClusters (permutationResults), 8
  - permutationResults, 8
  - realAllClusters (permutationResults), 8
  - scLVM\_output, 9
  - Single.mTec.Transcriptomes-package, 2
  - tras, 9
  - vars (scLVM\_output), 9
  - Ycorr (scLVM\_output), 9
- aireDependentSansom, 3
- beta (scLVM\_output), 9
- beta0 (scLVM\_output), 9
- biotype, 3
- biotypesHuman, 3
- cea1Coexpression, 4
- corMatsNoMarker, 4
- corMatSp (corMatsNoMarker), 4
- corMatSpNoMarker (corMatsNoMarker), 4
- deGenesNone, 5
- deGenesSansom, 5
- dx d (mTECdx d), 7
- dx dATAC, 5
- dx dFANTOM (fantom), 6
- fantom, 6