Package 'aggregateBioVar'

November 11, 2026

Type Package

Title Differential Gene Expression Analysis for Multi-subject scRNA-seq

Version 1.22.0

URL https://github.com/jasonratcliff/aggregateBioVar

BugReports https://github.com/jasonratcliff/aggregateBioVar/issues

Description For single cell RNA-seq data collected from more than one subject (e.g. biological sample or technical replicates), this package contains tools to summarize single cell gene expression profiles at the level of subject. A SingleCellExperiment object is taken as input and converted to a list of SummarizedExperiment objects, where each list element corresponds to an assigned cell type. The SummarizedExperiment objects contain aggregate gene-by-subject count matrices and inter-subject column metadata for individual subjects that can be processed using downstream bulk RNA-seq tools.

License GPL-3

Encoding UTF-8

LazyData true

VignetteBuilder knitr

Roxygen list(markdown = TRUE)

RoxygenNote 7.1.1

biocViews Software, SingleCell, RNASeq, Transcriptomics, Transcription, GeneExpression, DifferentialExpression

Depends R (>= 4.0)

Imports stats, methods, S4Vectors, SummarizedExperiment, SingleCellExperiment, Matrix, tibble, rlang

Suggests BiocStyle, magick, knitr, rmarkdown, testthat, BiocGenerics, DESeq2, magrittr, dplyr, ggplot2, cowplot, ggtext, RColorBrewer, pheatmap, viridis

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

git_branch RELEASE_3_23

git_last_commit 2886618

git_last_commit_date 2026-04-28

Repository Bioconductor 3.23

Date/Publication 2026-11-11

Author Jason Ratcliff [aut, cre] (ORCID:
<https://orcid.org/0000-0001-7079-334X>),
Andrew Thurman [aut],
Michael Chimenti [ctb],
Alejandro Pezzulo [ctb]

Maintainer Jason Ratcliff <jason-ratcliff@uiowa.edu>

Contents

aggregateBioVar 2
countsByCell 3
countsBySubject 4
scSubjects 5
small_airway 5
subjectMetaData 6
summarizedCounts 7

Index **8**

`aggregateBioVar` *Aggregate subject-level biological variation*

Description

Given an input gene-by-cell count matrix from a SingleCellExperiment object, sum within-subject gene counts into an aggregate gene-by-subject count matrix. Column metadata accessed by colData are collated by subjectMetaData to remove variables with inter-cell intrasubject variation, effectively retaining between-subject variation. The summary operations are performed across all cell types and within each cell type. A list of SummarizedExperiment objects is returned each with aggregate gene-by-subject count matrix and inter-subject metadata.

Usage

```
aggregateBioVar(scExp, subjectVar, cellVar)
```

Arguments

scExp	SingleCellExperiment object containing (at minimum) gene counts and column metadata describing sample identifiers and cell types.
subjectVar	Metadata column name assigning biological sample identity to aggregate within-subject feature counts.
cellVar	Metadata column name assigning cell type. Used for aggregating gene-by-subject count matrices by cell type.

Value

List of SummarizedExperiment objects with gene-by-subject count matrices and variable inter-subject column metadata across and within cell types.

Examples

```
## Aggregate gene-by-subject count matrix and inter-subject metadata
aggregateBioVar(
    scExp=small_airway,
    subjectVar="orig.ident", cellVar="celltype"
)
```

---

countsByCell *Within cell type gene-by-subject matrices*

Description

Given a vector of unique cell types, calculate a gene-by-subject matrix and inter-subject metadata for a differential expression design matrix.

Usage

```
countsByCell(scExp, subjectVar, cellVar)
```

Arguments

scExp	SingleCellExperiment object containing (at minimum) gene counts and column metadata describing sample identifiers and cell types.
subjectVar	Metadata column name assigning biological sample identity to aggregate within-subject feature counts.
cellVar	Metadata column name assigning cell type. Used for aggregating gene-by-subject count matrices by cell type.

Value

List of gene-by-subject and design matrices for each cell type.

See Also

summarizedCounts for aggregate counts and metadata SummarizedExperiment object.

Examples

```
## Return list of `SummarizedExperiments` with gene-by-subject count matrices
## and subject metadata for each unique `SingleCellExperiment` cell type.
countsByCell(
    scExp=small_airway,
    subjectVar="orig.ident", cellVar="celltype"
)
```

---

countsBySubject *Gene-by-subject count matrix*

Description

Convert gene-by-cell count matrix to gene-by-subject count matrix. Row sums are calculated for each feature (i.e. gene) across cells by subject.

Usage

```
countsBySubject(scExp, subjectVar)
```

Arguments

scExp	SingleCellExperiment object containing (at minimum) gene counts and column metadata describing sample identifiers and cell types.
subjectVar	Metadata column name assigning biological sample identity to aggregate within-subject feature counts.

Value

S4 DataFrame of gene-by-subject count sums.

See Also

scSubjects for subjects values.

Examples

```
## Return cell count matrix aggregated by subject.
countsBySubject(scExp=small_airway, subjectVar="orig.ident")
```

scSubjects *SingleCellExperiment subject values*

Description

Extract unique values from SingleCellExperiment column (i.e. cell) metadata. Used to determine subject and cell type values.

Usage

```
scSubjects(scExp, ...)
```

Arguments

scExp	SingleCellExperiment object containing (at minimum) gene counts and column metadata describing sample identifiers and cell types.
...	Named metadata variables for subjects and cell types.

Value

List of character vectors with unique values from SingleCellExperiment column metadata variables.

Examples

```
## Examples of metadata column variable names.
names(SummarizedExperiment::colData(small_airway))

## Return list of subject and cell type values from experiment metadata.
scSubjects(scExp=small_airway, subjects="orig.ident", cellTypes="celltype")
```

small_airway *Small Airway* SingleCellExperiment *object*

Description

Small airway epithelium single-cell RNA sequencing data subset combined from 7 porcine individuals. Genotypes represent wildtype (Genotype WT; n=3) and CFTR-knockout subjects (Genotype CFTRKO; n=4) expressing a cystic fibrosis phenotype.

To access cell counts and column metadata:

- counts: `SummarizedExperiment::assay(small_airway, "counts")`
- metadata: `SummarizedExperiment::colData(small_airway)`

Cell types include:

- Secretory cell
- Endothelial cell
- Immune cell

Usage

```
small_airway
```

Format

A SingleCellExperiment object with feature counts of 1311 genes (rows) from 2687 individual cells (columns). Features were subset by gene ontology annotation "ion transport" with included child terms ("anion transport", "cation transport", etc.). Includes a count matrix accessed by assay and column metadata accessed by colData. Metadata variables are defined as:

orig.ident Biological sample identifier (i.e. subject)

nCount_RNA Number of UMIs (i.e. depth/size factor)

nFeature_RNA Features with non-zero counts (i.e. identified features)

Genotype Sample Genotype ('CF' == cystic fibrosis)

- WT = non-CF pig
- CFTRKO = CFTR knockout

Sample Airway region of sample

celltype Cell type labels

---

subjectMetaData *Collate metadata variation*

Description

Identify single cell experiment metadata variables that are identical within subject (e.g. genotype, treatment, cell line, sample preparation). Effectively excludes metadata variables containing between cell variation. Used as the design matrix for differential expression analysis.

Usage

```
subjectMetaData(scExp, subjectVar)
```

Arguments

scExp	SingleCellExperiment object containing (at minimum) gene counts and column metadata describing sample identifiers and cell types.
subjectVar	Metadata column name assigning biological sample identity to aggregate within-subject feature counts.

Value

Tibble data frame of metadata variables without intrasubject variation from single cell experiment metadata. Rows correspond to aggregated cells (i.e. subject / biological replicate) and columns to metadata attribute variables (e.g. genotype, treatment, cell line).

Examples

```
## Return experiment metadata sans intrasubject variation.
subjectMetaData(scExp=small_airway, subjectVar="orig.ident")
```

---

summarizedCounts *Aggregate feature counts and metadata by subject*

Description

Given an input sparse count matrix and corresponding column metadata, aggregate gene counts by subject level. Metadata variables with only inter-subject variation are retained; any variables with cell-level variation within a subject are dropped (e.g. feature / RNA count by cell).

Usage

```
summarizedCounts(scExp, subjectVar)
```

Arguments

scExp	SingleCellExperiment object containing (at minimum) gene counts and column metadata describing sample identifiers and cell types.
subjectVar	Metadata column name assigning biological sample identity to aggregate within-subject feature counts.

Value

SummarizedExperiment object with feature counts aggregated by subject and summarized inter-subject metadata.

Examples

```
## Construct SummarizedExperiment object with gene-by-subject count matrix
## and column metadata summarized to exclude intrasubject variation.
## See `SummarizedExperiment` accessor functions `assay()` and `colData()`
## to access the count matrix and column metadata for downstream analyses.
summarizedCounts(scExp=small_airway, subjectVar="orig.ident")
```

Index

∗ **datasets**
small_airway, 5

aggregateBioVar, 2
assay, *6*

colData, *2*, *6*
countsByCell, 3
countsBySubject, 4

scSubjects, *4*, 5
SingleCellExperiment, *2–7*
small_airway, 5
subjectMetaData, *2*, 6
summarizedCounts, *4*, 7
SummarizedExperiment, *2–4*, *7*