Package 'ExpressionAtlas'

April 8, 2025

Title Search, download and visualise datasets from EMBL-EBI Expression Atlas and Single-Cell Expression Atlas

Version 1.99.4

Date 2025-04-07

Description This package is for searching for datasets in EMBL-EBI Expression Atlas, and downloading them into R for further analysis. Each Expression Atlas dataset is represented as a SimpleList object with one element per platform. Sequencing data is contained in a SummarizedExperiment object, while microarray data is contained in an ExpressionSet or MAList object. SingleCellExperiment object is used for Single-Cell Expression Atlas datasets.

biocViews ExpressionData, ExperimentData, SequencingData, MicroarrayData, ArrayExpress, SingleCellData

Depends R (>= 4.4.0), methods, Biobase

Imports utils, XML, httr, reshape2, tidyr, yaml, SummarizedExperiment, limma, SingleCellExperiment, viridis, HDF5Array, S4Vectors, xml2, jsonlite, BiocStyle, zellkonverter, gplots, genefilter, edgeR, ggplot2, dplyr, ComplexHeatmap, circlize

Suggests knitr, testthat, rmarkdown

VignetteBuilder knitr

VignetteEngine knitr::rmarkdown

Collate functions.R

Encoding UTF-8

License GPL (>= 3)

LazyLoad yes

NeedsCompilation no

URL https://github.com/ebi-gene-expression-group/bioconductor-ExpressionAtlas

BugReports https://github.com/ebi-gene-expression-group/bioconductor-ExpressionAtlas/issues

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

git_branch devel

git_last_commit 74f3a2a

git_last_commit_date 2025-04-08

Repository Bioconductor 3.21

Date/Publication 2025-04-08

Author Pedro Madrigal [aut, cre] (ORCID:
<https://orcid.org/0000-0003-1959-8199>),
Anil Thanki [aut] (ORCID: <https://orcid.org/0000-0002-8941-444X>),
Maria Keays [aut] (ORCID: <https://orcid.org/0000-0003-2034-601X>)

Maintainer Pedro Madrigal <`pmadrigal@ebi.ac.uk`>

Contents

`allExps` *A SimpleList containing some dummy Expression Atlas datasets*

Description

This is a SimpleList object containing dummy data from some Expression Atlas experiments, to demonstrate a typical example of the reesults of using the `getAtlasData()` function. The dummy data objects only contain the first 10 rows of the real data, to save disk space. To get the real data, download it using `getAtlasData()`.

Usage

```
data( "allExps" )
```

Value

A SimpleList with one element per Expression Atlas dataset.

`atlasRes` *A DataFrame listing some Expression Atlas experiments*

Description

This is a DataFrame object listing some Expression Atlas experiments, to demonstrate a typical example of the results of using the `searchAtlasExperiments()` function.

Usage

```
data( "atlasRes" )
```

Value

A DataFrame with one row per Expression Atlas dataset.

`dotPlotSCAtlasExperiment` *Dot-plot for a Single-Cell Expression Atlas experiment*

Description

This function generates a Dot-Plot from a SingleCellExperiment object coming from a Single-Cell Expression Atlas experiment.

Usage

```
dotPlotSCAtlasExperiment( singleCellExperiment, genes, sel.K=NULL, scaleNormExp=FALSE )
```

Arguments

`singleCellExperiment`	A SingleCellExperiment object from Single Cell Expression Atlas
`genes`	Vector of gene IDs to include in the dot-plot.
`sel.K`	Number of clusters for Single Cell Expression Atlas Experiment. If both `sel.K` is `NULL`, the function will use default clustering for this experiment.
`scaleNormExp`	Logical indicating whether to scale normalized expression values, which adjusts the data to have a mean of zero and a standard deviation of one. Default is `FALSE`.

Value

A Heatmap-class object from the ComplexHeatmap package. This object can be further customized or combined with other heatmaps using the ComplexHeatmap package’s functions.

Examples

```
egeod6552 <- getAtlasSCExperiment( "E-GEOD-36552" )

dotPlotSCAtlasExperiment(egeod6552, genes=c('ENSG00000166681','ENSG00000178928', 'ENSG00000142182' , 'ENSG000
```

`getAnalyticsDifferentialAtlasExpression` *Download differential expression analytics data for an Expression Atlas experiment*

Description

This function downloads and returns differential expression analytics data for an Expression Atlas experiment, based on the ArrayExpress accession of the experiment.

Usage

```
getAnalyticsDifferentialAtlasExpression( experimentAccession )
```

Arguments

`experimentAccession`	ArrayExpress experiment accession e.g. "E-MTAB-10104"

Value

A dataframe with `p.value` and `log2foldchange` values of the genes for each assay group contrast.

Examples

```
# Download the differential expression analytics data for experiment E-MTAB-10104
mtab10104_dea <- getAnalyticsDifferentialAtlasExpression( "E-MTAB-10104" )

# See the entries available
dim(mtab10104_dea)
# Prints out the following:
# [1] 56748    10
```

`getAtlasData` *Download data from multiple Expression Atlas experiments*

Description

This function downloads Expression Atlas experiment summary SimpleList objects based on a vector of ArrayExpress/BioStudies experiment accessions, and returns a list containing these objects.

Usage

```
getAtlasData( experimentAccessions )
```

Arguments

`experimentAccessions`	Vector of ArrayExpress/BioStudies accessions for experiments to be downloaded.

Value

A list with one entry per experiment summary SimpleList object. Entries are named using the ArrayExpres/BioStudies accession of the respective experiment.

Examples

```
# Download some Expression Atlas data into a list.
myExperimentSummaries <- getAtlasData(
    c(
        "E-GEOD-11175",
        "E-MTAB-3007",
        "E-GEOD-21070"
    )
)
```

`getAtlasExperiment` *Download data for an Expression Atlas experiment*

Description

This function downloads and returns a SimpleList object representing a single Expression Atlas experiment, based on the ArrayExpress accession of the experiment.

Usage

```
getAtlasExperiment( experimentAccession )
```

Arguments

`experimentAccession`	ArrayExpress experiment accession e.g. "E-GEOD-11175"

Value

A SimpleList object representing a single Expression Atlas experiment. The SimpleList contains one entry per platform used in the experiment. For sequencing experiments, there is a single entry in the list. For microarray experiments, there is one entry per array design used. Currently Expression Atlas does not support multi-technology (e.g. microarray and RNA-seq) experiments.

For a single-channel microarray experiment, each entry of the list is an ExpressionSet object. For a sequencing experiment, the single entry is a SummarizedExperiment object. Please refer to the relevant documentation on these classes for more information about them.

RNA-seq data

Each SummarizedExperiment object contains the following:

- Matrix of raw counts (not normalized), in the assays slot, in a counts element.
- Sample annotations, in the colData slot.
- Brief outline of methods, from QC of FASTQ files to production of raw counts, in the exptData slot.

Single-channel microarray data

Each ExpressionSet object contains the following:

- Matrix of normalized intensity values, in the assayData, accessed via: exprs(expressionSet)
- Sample annotations, in the phenoData, accessed via: pData(expressionSet)
- Brief outline of normalization method applied, in the experimentData slot, accessed via: preproc(experimentData(expressionSet))

Examples

```
# Download the experiment summary for E-GEOD-11175
geod11175 <- getAtlasExperiment( "E-GEOD-11175" )

# See the entries available (in this case array design accessions)
names( geod11175 )
# Prints out the following:
# [1] "A-AFFY-126"

# Get the only ExpressionSet object from this experiment.
eset <- geod11175[[ "A-AFFY-126" ]]
```

`getAtlasSCExperiment` *Download data for a Single-Cell Expression Atlas experiment*

Description

This function downloads and returns a SingleCellExperiment object representing a Single-Cell Expression Atlas experiment.

Usage

```
getAtlasSCExperiment( experimentAccession )
```

Arguments

`experimentAccession`	Single-cell experiment accession e.g. "E-ENAD-19"

Value

A SingleCellExperiment object representing a single-Cell Expression Atlas experiment.

Each SCE object contains the following:

- Assays, containing the raw counts (X), filtered counts (filtered), and normalised and filtered counts (normalised).
- Column data (cell-level metadata)
- Different dimensionality reductions (e.g. PCA, t-SNE, UMAP)

Examples

```
# Download the experiment summary for E-GEOD-11175
enad19 <- getAtlasSCExperiment( ”E-ENAD-19” )

enad19
# Prints out the following:
#
# class: SingleCellExperiment
# dim: 3644 26
# metadata(0):
# assays(3): X filtered normalised
# rownames: NULL
# rowData names(1): PCs
# colnames(26): ERR2192535 ERR2192536 ... ERR2192559 ERR2192560
# colData names(27): age cell_type ... louvain_resolution_1.0
#   louvain_resolution_2.0
# reducedDimNames(21): X_pca X_tsne_perplexity_1 ...
#   X_umap_neighbors_n_neighbors_5 X_umap_neighbors_n_neighbors_50
# mainExpName: NULL
# altExpNames(0):
```

`getNormalisedAtlasExpression` *Download normalised expression data for an Expression Atlas experiment*

Description

This function downloads and returns normalised expression data for an Expression Atlas experiment, based on the ArrayExpress accession of the experiment.

Usage

```
getNormalisedAtlasExpression( experimentAccession, normalisation )
```

Arguments

`experimentAccession`	ArrayExpress experiment accession e.g. "E-MTAB-4045"
`normalisation`	Normalisation method type e.g. "tpm", "fpkm", "cpm"

Value

A dataframe with expression values of the genes for each assay group. Entries are mean values of replicates.

Examples

```
# Download the TPM Expression values for experiment E-MTAB-4045
mtab4045_tpm <- getNormalisedAtlasExpression( "E-MTAB-4045", "tpm" )

# See the entries available
dim(mtab4045_tpm)
# Prints out the following:
# [1] 57147    45

# Download the CPM Expression values for experiment E-MTAB-4045
mtab4045_cpm <- getNormalisedAtlasExpression( "E-MTAB-4045", "cpm" )

# See the entries available
dim(mtab4045_cpm)
# Prints out the following:
# [1] 57147    275
```

`heatmapAtlasExperiment` *Generate HeatMap from normalised expression data for an Expression Atlas experiment*

Description

This function generates a heatmap for an Expression Atlas experiment, from normalised expression data.

Usage

```
heatmapAtlasExperiment( df, filename = "heatmap", save_pdf = FALSE, show_plot = TRUE, palette = "viri
```

Arguments

`df`	Dataframe in the output format of getNormalisedAtlasExpression.
`filename`	Filename for the output heatmap file. Default is "heatmap.pdf".
`save_pdf`	Save PDF file with the heatmap. Default is `FALSE`.
`show_plot`	Show plot in the R console. Default is `TRUE`.
`palette`	Viridis colour palette for the heatmap. Default is "viridis". Other options are: magma, plasma, inferno, cividis, rocket, mako, turbo
`top_n`	Number of top most variable genes to be shown in the heatmap. Default is 100.
`scaled`	Scale the normaliased gene expression data. Default is `FALSE`.
`show_heatmap_title`	Whether to show a title in the heatmap. Default is `TRUE`.

Value

A PDF file with the generated heatmap.

Examples

```
 # Download the TPM Expression values for experiment E-MTAB-4045
 mtab4045_tpm <- getNormalisedAtlasExpression( "E-MTAB-4045", "tpm" )

 # Generate heatmap with default parameters
heatmapAtlasExperiment( mtab4045_tpm, "heatmap", TRUE, FALSE, "viridis", 100, FALSE, TRUE )

 # Check if the PDF file was created
 file.exists("heatmap.pdf")

# Download the CPM Expression values for experiment E-MTAB-4045 with custom parameters
 mtab4045_cpm <- getNormalisedAtlasExpression( "E-MTAB-4045", "cpm" )

 # Generate heatmap with default parameters
```

```
heatmapAtlasExperiment( mtab4045_cpm, "E-MTAB-4045-cpm-heatmap", TRUE, FALSE, "viridis", 100, FALSE, TRUE )

 # Check if the PDF file was created
 file.exists("E-MTAB-4045-cpm-heatmap.pdf")
```

---

heatmapSCAtlasExperiment

Visualise heatmap for Single-Cell Expression Atlas experiment

---

Description

This function plots a heatmap from a SingleCellExperiment object coming from a Single-Cell Expression Atlas experiment.

Usage

```
heatmapSCAtlasExperiment( singleCellExperiment, genes=NULL, sel.K=NULL, scaleNormExp=FALSE, show_r
```

Arguments

`singleCellExperiment`	A `SingleCellExperiment` object from Single Cell Expression Atlas
`genes`	Vector of gene IDs to include in the heatmap. If `NULL`, all genes will be shown for the default clustering. Default is `NULL`.
`sel.K`	Number of clusters for Single Cell Expression Atlas Experiment. If `NULL`, a gene list must be provided. If both `sel.K` and `genes` are `NULL`, the function will use default clustering for all markers genes for this experiment.
`scaleNormExp`	Logical indicating whether to scale normalized expression values, which adjusts the data to have a mean of zero and a standard deviation of one. Default is `FALSE`.
`show_row_names`	Logical indicating whether to show gene names in rows. Default is `FALSE`.

Value

A Heatmap-class object from the ComplexHeatmap package. This object can be further customized or combined with other heatmaps using the ComplexHeatmap package's functions.

Examples

```
 egeod6552 <- getAtlasSCExperiment( "E-GEOD-36552" )

heatmap6552 <- heatmapSCAtlasExperiment( egeod6552, genes=NULL, sel.K=NULL, scaleNormExp=FALSE, show_row_names

 # Example with a specific gene list
heatmap6552_genes <- heatmapSCAtlasExperiment( egeod6552, genes=c('ENSG00000151611','ENSG00000020577', 'ENSG0
```

`plotDimRedSCAtlasExperiment` *Plot Dimensionality reductions of a Single-Cell Expression Atlas experiment*

Description

This function plots dimensionality reduction (PCA, t-sne, UMAP) of a SingleCellExperiment object representing a Single-Cell Expression Atlas experiment.

Usage

```
plotDimRedSCAtlasExperiment( sceObject, dimRed, colorby )
```

Arguments

`sceObject`	Single-cell experiment object, representing a Single-Cell Expression Atlas experiment, obtained with the function `getAtlasSCExperiment`.
`dimRed`	Dimensionality reduction to plot. Options are: "X_pca", "X_tsne_perplexity_10", "X_umap_neighbors_20", etc.
`colorby`	Column name in `colData` to color the cells by.

Value

A ggplot2 geom point visualisation, which can be further customised by the user.

Examples

```
# Example of plotting PCA dimensionality reduction colored by age

# Download the experiment summary for E-ENAD-19
enad19 <- getAtlasSCExperiment( "E-ENAD-19" )

plotDimRedSCAtlasExperiment(enad19, dimRed = "X_pca", colorby = "age" )
```

`rnaseqExps` *A SimpleList containing a dummy RNA-seq Expression Atlas dataset*

Description

This is a SimpleList object containing dummy data from an RNA-seq Expression Atlas experiment, to demonstrate a typical example of the results of using the `getAtlasData()` function for a subset of results obtained using `searchAtlasExperiments`. This object contains the first 10 rows of the original data, to save disk space. To get the full dataset, download it using `getAtlasData()`.

Usage

```
data( "rnaseqExps" )
```

Value

A SimpleList with one element per Expression Atlas dataset.

`searchAtlasExperiments` *Search for Expression Atlas experiments*

Description

This function accepts a query, and optionally a secondary filter, and then searches for matching Expression Atlas experiments in the EBI RESTful Web Services API

Usage

```
searchAtlasExperiments( query, secondaryFilter = NULL, detailed = FALSE  )
```

Arguments

`query`	Character of sample properties to search Atlas for. These can be biological characteristics, experimental treatments, species, etc.
`secondaryFilter`	Optional, a second filter.
`detailed`	If `TRUE`, it will perform a detailed search through Atlas experiment metadata information, and it might take 10-20 seconds. Default: `FALSE`.

Value

A DataFrame containing the Expression Atlas accessions, the species, experiment types, and titles of Expression Atlas experiments matching the query.

Examples

```
# Search for endoderm experiments (4 results)
atlasRes_1 <- searchAtlasExperiments( query="endoderm")

# Search for endoderm experiments in human (2 results)
atlasRes_2 <- searchAtlasExperiments( query="endoderm", secondaryFilter = "human" )
# it produces same results as:
# atlasRes_2 <- searchAtlasExperiments( query="human", secondaryFilter = "endoderm" )

# Download data for first experiment found.
if ( nrow( atlasRes_2 ) == 1 ) {
    atlasData <- getAtlasData( atlasRes_2$Accession )
} else {
```

```
    atlasData <- getAtlasData( atlasRes_2$Accession[1] )
}
```

---

searchSCAtlasExperiments

Search for Sinle-Cell Expression Atlas experiments

Description

This function accepts a query, and optionally a secondary filter, and then searches for matching Single-Cell Expression Atlas experiments in the EBI RESTful Web Services API. The function also performs ontology-based query search against the Single Cell Expression Atlas cell-type wheel.

Usage

```
searchSCAtlasExperiments( query, secondaryFilter = NULL  )
```

Arguments

`query`	Character of sample properties to search Atlas for. These can be biological characteristics, experimental treatments, species, etc.
`secondaryFilter`	Optional, a second filter. For instance, if the query is "lung", the secondary filter can be "human" to search for human lung experiments.

Value

A DataFrame containing the Expression Atlas accessions, the species, experiment types, and titles of Expression Atlas experiments matching the query.

Examples

```
# Search for pancreas experiments
atlasRes_pancreas <- searchSCAtlasExperiments( query="pancreas")

# Search for human lung experiments
atlasRes_human_lung <- searchSCAtlasExperiments( query="lung", secondaryFilter = "human")
```

`volcanoDifferentialAtlasExperiment` *Generate Volcano plots from analytics data for an Expression Atlas experiment*

Description

This function generates a volcano plots for an Expression Atlas experiment, from analytics data.

Usage

```
volcanoDifferentialAtlasExperiment( df, filename_prefix = "volcano-plot", save_pdf = FALSE, show_pl
```

Arguments

`df`	Dataframe in the output format of `getAnalysticsDifferentialAtlasExpression`.
`filename_prefix`	Filename for the output volcano plot images. Default is "volcano".
`save_pdf`	Save PDF file with the heatmap. Default is `FALSE`.
`show_plot`	Show plot in the R console. Default is `TRUE`.
`low_fc_colour`	Colour palette for the low foldchange. Default is "gray".
`high_fc_colour`	Colour palette for the high foldchange. Default is "blue".
`cutoff`	cutoff for the foldchange. Default is 1.
`show_volcanoplot_title`	Whether to show a title in the volcano plot. Default is `TRUE`.

Value

Image file(s) with the generated volcano plots.

Examples

```
 # Download the differential expression analytics data for experiment E-MTAB-10104
 mtab10104_dea <- getAnalyticsDifferentialAtlasExpression( "E-MTAB-10104" )

 # Generate volcano plots with default parameters
 volcanoDifferentialAtlasExperiment( mtab10104_dea )

 # Check if the image files were created
 # file.exists("heatmap.pdf")

 # Generate volcano plots with custom parameters
volcanoDifferentialAtlasExperiment( mtab10104_dea, "E-MTAB-10104", FALSE, TRUE, "Gray", "Blue", 1, TRUE )

 # Check if the image files were created
 # file.exists("E-MTAB-4045-cpm-heatmap.pdf")
```

Index