

# Package ‘ExpressionAtlas’

October 12, 2016

**Version** 1.0.3

**Date** 2016/07/07

**Title** Download datasets from EMBL-EBI Expression Atlas

**Author** Maria Keays

**Maintainer** Maria Keays <mkeays@ebi.ac.uk>

**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.

**biocViews** ExpressionData, ExperimentData, SequencingData,  
MicroarrayData, ArrayExpress

**Depends** R (>= 3.2.0), methods, Biobase, SummarizedExperiment, limma,  
S4Vectors, xml2

**Imports** utils, XML, httr

**Suggests** knitr, testthat, rmarkdown

**VignetteBuilder** knitr

**Collate** functions.R

**License** GPL (>= 3)

**LazyLoad** yes

**NeedsCompilation** no

## R topics documented:

allExps	2
atlasRes	2
getAtlasData	3
getAtlasExperiment	3
rnaseqExps	5
searchAtlasExperiments	5

<b>Index</b>	<b>7</b>
--------------	----------

---

allExps	<i>A SimpleList containing some dummy Expression Atlas datasets</i>
---------	---

---

**Description**

This is a SimpleList object containing dummy data from some Expression Atlas experiments, to demonstrate a typical example of the results 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	<i>A DataFrame listing some Expression Atlas experiments</i>
----------	--

---

**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.

---

getAtlasData	<i>Download data from multiple Expression Atlas experiments</i>
--------------	---

---

**Description**

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

**Usage**

```
getAtlasData( experimentAccessions )
```

**Arguments**

experimentAccessions  
Vector of ArrayExpress accessions for experiments to be downloaded.

**Value**

A list with one entry per experiment summary SimpleList object. Entries are named using the ArrayExpress 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	<i>Download data for an Expression Atlas experiment</i>
--------------------	---

---

**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" ]]
```

---

`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 vector of sample properties, and optionally a species name, and then searches for matching Expression Atlas experiments.

**Usage**

```
searchAtlasExperiments( properties, species = NULL )
```

**Arguments**

- |                         |   |
|-------------------------|---|
| <code>properties</code> | Character vector of sample properties to search Atlas for. These can be biological characteristics, experimental treatments, etc. |
| <code>species</code>    | Optional. The name of a species to limit results to. If not provided, search is performed across all species in Expression Atlas. |

**Value**

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

**Examples**

```
# Search for experiments on salt in rice.
atlasRes <- searchAtlasExperiments( "salt", "Oryza sativa" )

# Download data for all experiments found.
atlasData <- getAtlasData( atlasRes$Accession )
```

# Index

allExps (allExps), 2  
atlasRes (atlasRes), 2  
getAtlasData (getAtlasData), 3  
getAtlasExperiment  
    (getAtlasExperiment), 3  
rnaseqExps (rnaseqExps), 5  
searchAtlasExperiments  
    (searchAtlasExperiments), 5

allExps, 2  
atlasRes, 2

getAtlasData, 3  
getAtlasExperiment, 3

rnaseqExps, 5

searchAtlasExperiments, 5