Package ‘metaboData’

July 22, 2025

Type Package

Title Example Metabolomics Data Sets

Version 0.6.3

Description Data sets from a variety of biological sample matrices,
analysed using a number of mass spectrometry based metabolomic analytical techniques.
The example data sets are stored remotely using GitHub releases
<https://github.com/aberHRML/metaboData/releases> which can be accessed from R using the package.
The package also includes the 'abr1' FIE-MS data set from the 'FIEmspro' package <https://users.aber.ac.uk/jhd/> <doi:10.1038/nprot.2007.511>.

Depends R (>= 3.4.0)

Imports dplyr, fs, magrittr, piggyback, purrr, readr, rlang, stringr,
tibble, yaml

License GPL (>= 3)

Encoding UTF-8

LazyData true

RoxygenNote 7.1.2

Suggests testthat, covr, knitr, rmarkdown, prettydoc

URL https://aberhrml.github.io/metaboData/

BugReports https://github.com/aberHRML/metaboData/issues

VignetteBuilder knitr

NeedsCompilation no

Author Jasen Finch [aut, cre] (ORCID: <https://orcid.org/0000-0002-6070-7476>),
Manfred Beckmann [ctb],
David Enot [ctb],
Wanchang Lin [ctb]

Maintainer Jasen Finch <jsf9@aber.ac.uk>

Repository CRAN

Date/Publication 2022-02-15 23:00:02 UTC

Contents

`abr1` *The abr1 dataset*

Description

Real world FIE-MS dataset.

Usage

```
data(abr1)
```

Details

FIE-MS data matrices developed from analysis of samples representing a time course of pathogen attack in a model plant species (Brachypodium distachyon). The data was developed in a single batch with all samples randomised using a Thermo LTQ linear ion trap processed using `fiems_ltq_main`. Both positive and negative ion mode are given (`abr1$pos` and `abr1$neg`). To avoid confusions, variable names are given with a letter corresponding to the ionisation mode followed by the actual nominal mass value (e.g. P130 corresponds to the nominal mass 130 in the positive mode).

Experimental factors are given in the `abr1$fact` data frame:

- `injorder`: sample injection order
- `name`: sample name
- `rep`: biological replicate for a given class
- `day`: number of days following infection after which the sample has been harvested - Level H corresponds to an healthy plant.
- `class`: identical to day except that `class=6` when `day=H`
- `pathcdf,filecdf,name.org,remark`: are generated from profile processing and are kept for traceability purposes.

Factor of interest for classification are contained in `abr1$fact$day`. There are 20 biological replicates in each class has

Value

A list with the following elements:

`fact`	A data frame containing experimental meta-data.
`pos`	A data frame for positive data with 120 observations and 2000 variables.
`neg`	A data frame for negative data with 120 observations and 2000 variables.

Author(s)

Manfred Beckmann, David Enot and Wanchang Lin

Source

The FIEmspro package https://github.com/aberHRML/FIEmspro

Examples

```
# Load data set
data(abr1)

# Select data set
dat <- abr1$neg

# number of observations and variables in the negative mode matrix
dim(dat)

# names of the variables
dimnames(dat)[[2]] %>%
 head()

# print out the experimental factors
abr1$fact %>%
 head()

# check out the repartition of class
table(abr1$fact$class)
```

`availableDataSets` *Available data sets*

Description

Displays available data sets.

Usage

```
availableDataSets(dataSetDir = "DataSets", internalDir = TRUE)
```

Arguments

`dataSetDir`	directory containing local data set store. If `internalDir = TRUE` the full directory path would be relative to the package installation location.
`internalDir`	stored the data set internally to the package installation location

Value

A tibble containing available data set information.

Examples

```
## Not run:
availableDataSets()

## End(Not run)
```

`availableFiles` *Available files for a data set*

Description

Display the available files for a given data set.

Usage

```
availableFiles(technique, dataSet, dataSetDir = "DataSets", internalDir = TRUE)
```

Arguments

`technique`	metabolomic technique name
`dataSet`	data set name
`dataSetDir`	directory containing local data set store. If `internalDir = TRUE` the full directory path would be relative to the package installation location.
`internalDir`	stored the data set internally to the package installation location

Value

A tibble containing available file information.

Examples

```
## Not run:
availableFiles('FIE-HRMS','BdistachyonTechnical')

## End(Not run)
```

`dataSets` *Available data sets for a given technique*

Description

Return a vector of available data sets for a given metabolomic technique.

Usage

```
dataSets(technique)
```

Arguments

`technique`	metabolomic technique name

Value

A character vector of available data sets.

Examples

```
## Not run:
dataSets(techniques()[1])

## End(Not run)
```

`description` *Data set description*

Description

Return list containing experiment description elements

Usage

```
description(
  technique,
  dataSet,
  dataSetDir = "DataSets",
  internalDir = TRUE,
  ...
)
```

Arguments

`technique`	metabolomic technique name
`dataSet`	data set name
`dataSetDir`	directory to store local data sets. If `internalDir` = `TRUE` this is a directory relative to the library location.
`internalDir`	Logical, should the directory for storing local data sets be internal to the package location.
`...`	arguments to pass to `downloadDataSet()`

Value

A list containing data set descriptors.

Examples

```
## Not run:
description(
   techniques()[1],
   dataSets(techniques()[1])[1])

## End(Not run)
```

`downloadDataSet` *Download a data set*

Description

Download an example data set.

Usage

```
downloadDataSet(
  technique,
  dataSet,
  dataSetDir = "DataSets",
  internalDir = TRUE,
  ask = TRUE,
  ...
)
```

Arguments

`technique`	metabolomic technique name
`dataSet`	data set name
`dataSetDir`	directory to store local data sets. If `internalDir` = `TRUE` this is a directory relative to the library location.

`internalDir`	stored the data set internally to the package installation location
`ask`	ask before downloading
`...`	arguments to pass to `piggyback::pb_download()`

Value

A list object containing the API request response results for the downloaded files.

Examples

```
## Not run:
downloadDataSet('FIE-HRMS','BdistachyonTechnical')

## End(Not run)
```

`downloadFiles` *Download specific files from a data set*

Description

Download specific files for a given example data set.

Usage

```
downloadFiles(
  files,
  technique,
  dataSet,
  dataSetDir = "DataSets",
  internalDir = TRUE,
  ask = TRUE,
  ...
)
```

Arguments

`files`	character vector of file paths to download
`technique`	metabolomic technique name
`dataSet`	data set name
`dataSetDir`	directory to store local data sets. If `internalDir = TRUE` this is a directory relative to the library location.
`internalDir`	stored the data set internally to the package installation location
`ask`	ask before downloading
`...`	arguments to pass to `piggyback::pb_download()`

Value

A list object containing the API request response results for the downloaded files.

Examples

```
## Not run:
downloadFiles(c('31.mzML.gz','32.mzML.gz'),'FIE-HRMS','BdistachyonTechnical')

## End(Not run)
```

`filePaths` *Data set file paths*

Description

Return a vector of file paths for a given data set of a given metabolomic technique.

Usage

```
filePaths(technique, dataSet, dataSetDir = "DataSets", internalDir = TRUE, ...)
```

Arguments

`technique`	metabolomic technique name
`dataSet`	data set name
`dataSetDir`	directory to store local data sets. If `internalDir = TRUE` this is a directory relative to the library location.
`internalDir`	Logical, should the directory for storing local data sets be internal to the package location.
`...`	arguments to pass to `downloadDataSet()`

Value

A character vector of file paths.

Examples

```
## Not run:
files <- filePaths(
   techniques()[1],
   dataSets(techniques()[1])[1])

head(files)

## End(Not run)
```

`runinfo` *Data set sample information*

Description

Return tibble of run information for a given data set of a given metabolomics technique.

Usage

```
runinfo(technique, dataSet, dataSetDir = "DataSets", internalDir = TRUE, ...)
```

Arguments

`technique`	metabolomic technique name
`dataSet`	data set name
`dataSetDir`	directory to store local data sets. If `internalDir` = `TRUE` this is a directory relative to the library location.
`internalDir`	Logical, should the directory for storing local data sets be internal to the package location.
`...`	arguments to pass to `downloadDataSet()`

Value

A tibble containing sample information.

Examples

```
## Not run:
info <- runinfo(
   techniques()[1],
   dataSets(techniques()[1])[1])

head(info)

## End(Not run)
```

`techniques` *Available data set techniques*

Description

Return a vector of available metabolomic techniques.

Usage

```
techniques()
```

Value

A character vector of available techniques.

Examples

```
## Not run:
techniques()

## End(Not run)
```

Index