Package 'MetaboAnnotatoR'

November 12, 2026

Title Automated Annotation of All-Ion Fragmentation LC-MS Metabolomic Features

Version 1.0.0

Date 2026-01-30

Description Performs feature annotations on LC-MS All-ion fragmentation datasets using fragment ion libraries.

Depends R (>= 4.5.0), xcms (>= 3.0.0), MSnbase (>= 2.16.1), ProtGenerics

Imports ggplot2 (>= 3.3.3), gridExtra (>= 2.3)

biocViews MassSpectrometry, Metabolomics

License GPL-3

BugReports https://github.com/gggraca/MetaboAnnotatoR/issues

URL https://github.com/gggraca/MetaboAnnotatoR

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.3

Suggests rmarkdown, knitr, testthat (>= 3.0.0), BiocStyle, MsDataHub

LazyData false

VignetteBuilder knitr

Config/testthat/edition 3

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

git_branch RELEASE_3_23

git_last_commit 48ec722

git_last_commit_date 2026-04-28

Repository Bioconductor 3.23

Date/Publication 2026-11-11

Author Goncalo Graca [aut, cre] (ORCID:
<https://orcid.org/0000-0002-0876-3876>),
Yuheng (Rene) Cai [aut],
Timothy Ebbels [aut]

Maintainer Goncalo Graca <ggraca22@gmail.com>

Contents

acetaminophen . . . 2
annotateAIF . . . 3
annotateRC . . . 5
checkIsotope . . . 6
compFrag . . . 7
genFragEntry . . . 9
getPseudoMSMS . . . 10
LipidNeg . . . 12
LipidPos . . . 12
MetabolitesNeg . . . 13
MetabolitesPos . . . 14
mspToLib . . . 15
plotResultSpec . . . 16
pseudoMSMS . . . 17
rankScore . . . 18
RC . . . 19
rCandidates . . . 20
RCspec . . . 20
saveAnnotations . . . 21
searchLib . . . 23
xcmsSpec . . . 24
xset . . . 25

Index **26**

`acetaminophen` *MetaboAnnotatoR fragment database entry for Acetaminophen Positive mode*

Description

This dataset is an example of a MetaboAnnotatoR database entry. It is a data table containing the Acetaminophen fragments and respective score weights. The database entry was generated from one of the MassBank entries for Acetaminophen (MSBNK-Athens_Univ-AU276702) using function `MetaboAnnotatoR::genFragEntry()`.

Usage

```
data("acetaminophen")
```

Format

A list containing one data frame with the database entry.

Author(s)

Goncalo Graca

See Also

genFragEntry

---

annotateAIF *Annotate features from LC-MS AIF raw chromatograms*

Description

This function annotates features from raw LC-MS AIF chromatograms, by performing pseudo-MS/MS spectra deconvolution and then matching ions to metabolite/lipid fragment libraries.

Usage

```
annotateAIF(
  targets,
  xcmsOptions,
  libs = LipidPos,
  RTs = "none",
  nCE = 1,
  corThresh = 0.8,
  checkIsotope = TRUE,
  tolerance = 25,
  maxMZdiff = 0.01,
  matchWeight = 0.5
)
```

Arguments

targets	A data frame containing the features to annotate and the file paths to the raw data.
xcmsOptions	A data frame containing the XCMS centWave peak-picking parameters. An example of such table can be found in the data provided with MetaboAnnotatoR as XCMS_options.csv (see example for details).
libs	Fragment libraries to use. Either the built-in libraries can be specified (LipidPOS, LipidNEG, MetabolitesPOS, MetabolitesNEG) or the full path to user-defined libraries.
RTs	Optional data.frame with Lipid/metabolites classes Retention Times in seconds.

nCE	Number of Collision Energy levels depending on the MS system used Waters, Bruker (QToF) and Thermo Orbitrap = 1, Agilent (QToF) > 1, however, only the highest energy level will be considered.
corThresh	Pearson correlation coefficient for EIC correlation.
checkIsotope	Whether or not to check the isotope type; default is set to TRUE
tolerance	Tolerance in ppm for the candidate search.
maxMZdiff	Maximum m/z difference between candidate fragments and pseudo-MS/MS or AIF ions in Da.
matchWeight	weight of the fragment matches to the final score; value between 0 and 1; the remaining fraction of the weight comes from the candidate m/z error.

Value

For each feature in the targets data table the function will return a data frame with each feature rank 1 annotation and a table with the options used for the function, the data and time of annotation. In addition, lists with the ranked candidates matched to each feature (`rankedResult`), ranked matched spectra (`rankedSpectra`) and a list with pseudo-MS/MS spectra, in-source spectra, and AIF spectra and respective EIC objects (`pseudoMSMS`, see also `getPseudoMSMS`, documentation for details).

Author(s)

Goncalo Graca (Imperial College London)

Examples

```
# Set a directory to save the example .mzML file
userDir <- tempdir()
# Download the example .mzML from zenodo website into the specified
#directory as "Lipid_Positive_QC.mzML"...define file.path first:
fpath <- file.path(userDir, "Lipid_Positive_QC.mzML")
download.file(
"https://zenodo.org/records/17408169/files/Lipid_Positive_QC.mzML?download=1",
fpath)
# create a new targetTable with one feature to annotate
# the Sample.name is the path to the mzML file
targets <- data.frame(feature.mz=520.3408533, feature.rt=100.6238759,
Sample.name=fpath)
# read the default xcms parameters on the XCMS_options.csv file and modify
# the noise threshold parameter
xcmsOptionsPath <- system.file("extdata", "XCMS_options.csv",
package="MetaboAnnotatoR")
xcmsOptions <- read.csv(xcmsOptionsPath)
xcmsOptions[2,2] <- 1000
# Read the default lipid positive libraries
data("LipidPos")
# Run the annotation using the built-in lipid POS library:
annotations <- annotateAIF(targets, xcmsOptions,
libs="LipidPos", RTs="none", nCE=1, corThresh=0.8,
checkIsotope=TRUE)
```

annotateRC *Annotate features from LC-MS AIF datasets processed with RAMClustR*

Description

This function annotates features from LC-MS AIF using pseudo-MS/MS spectra obtained using RAMClustR, by matching ions to metabolite/lipid fragment libraries.

Usage

```
annotateRC(
  targets,
  xcmsObject,
  ramclustObj,
  libs = LipidPos,
  RTs = "none",
  checkIsotope = TRUE,
  tolerance = 25,
  maxMZdiff = 0.01,
  matchWeight = 0.5
)
```

Arguments

targets	A data.frame containing the features to annotate and the file paths to the raw data.
xcmsObject	XCMS object containing the processed AIF datasets.
ramclustObj	RAMClustR object with parent-fragment reconstructions (clusters). See RAMClustR paper for more details (https://pubs.acs.org/doi/10.1021/ac501530d).
libs	Fragment libraries to use. Specify one of default libraries provided as data object with the package (LipidPos, LipidNeg, MetabolitesPos and MetabolitesNeg) or the full path to user-defined libraries.
RTs	Optional data.frame with Lipid/metabolites classes Retention Times in seconds.
checkIsotope	Whether or not to check the isotope type; default is set to TRUE.
tolerance	Tolerance in ppm for the candidate search.
maxMZdiff	Maximum m/z difference between candidate fragments and pseudo-MS/MS or AIF ions in Da.
matchWeight	weight of the fragment matches to the final score; value between 0 and 1; the remaining fraction of the weight comes from the candidate m/z error.

Value

For each feature in the targets data frame the function will return a list containing: a data frame with with rank 1 annotations (`global`), the the date and time of annotation, a data frame with the annotation options. For each feature the following lists are returned: ranked annotations for each feature (`rankedResults`), the corresponding ranked matched spectra (`rankedSpectra`), the pseudo-MS/MS spectra (`pseudoMSMS`), in-source spectra (`inSourceSpectra`) and AIF spectrum (`AIFspectra`).

Author(s)

Goncalo Graca & Yuheng (Rene) Cai (Imperial College London)

Examples

```
# Read RAMClustR (RC) and XCMS processed example data lipid positive
# LC-MS data
data("RC")
data("xset")
# read the table containing features to annotate
tfile <- system.file("extdata", "targetTable.csv",
package="MetaboAnnotatoR")
targets <- read.csv(tfile)
# Read the default lipid positive libraries
data("LipidPos")
# Run the annotation procedure
annotations <- annotateRC(targets, xcmsObject=xset, ramclustObj=RC,
libs="LipidPos", RTs="none", checkIsotope=TRUE)
```

`checkIsotope` *Isotopologue type determination*

Description

Checks the type of isotope of an LC-MS feature (e.g. M+0, M+1, M+2, ...).

Usage

```
checkIsotope(fmz, frt, spec, mztol = 0.01)
```

Arguments

`fmz`	Feature m/z.
`frt`	Feature RT in seconds.
`spec`	A data frame containing XCMS peaks ("raw") or a RAMClustR pseudo-MS/MS spectrum ("cluster").
`mztol`	Absolute tolerance for feature m/z search in Da (default is 0.01).

Value

A "tag" of the isotope from an isotopic series as 0, 1, 2 or 3 for M+0, M+1, M+2 and M+3, respectively.

Author(s)

Goncalo Graca and Yuheng (Rene) Cai (Imperial College London)

Examples

```
# create a data frame with test spectra
spObject <- data.frame(mz=c(703.5769, 704.5799, 705.5812, 706.5721),
into=c(205458624, 85536216, 22717336, 5887723))
# check the isotope of feature 703.5769 m/z and 70 s
iso <- checkIsotope(fmz=703.5769, frt=70, spec=spObject)
```

---

compFrag *Match (pseudo-)MS/MS spectra ions to library fragments*

Description

This function compares pseudo-MS/MS and high-collision-energy spectra peaks with fragments from candidates of the metabolite/lipid fragment libraries.

Usage

```
compFrag(
  candidate,
  lib,
  fmz,
  frt,
  iso,
  highCESpec,
  pseudoSpec,
  maxMZdiff = 0.01,
  matchWeight = 0.5,
  useMZerrorWeight = TRUE,
  NoMatchWeight = 0.5,
  additional = TRUE
)
```

Arguments

candidate	Library entry containing the candidate fragments.
lib	A list containing the metabolite library to use.
fmz	The m/z for the feature of interest.
frt	Retention time in seconds for the feature of interest.

iso	Isotope "tag" to add to the results.
highCESpec	MS2 peaks at the RT window of the feature of interest.
pseudoSpec	MS2 peaks related to the feature of interest.
maxMZdiff	Maximum m/z difference between candidate fragments and pseudo-MS/MS or AIF ions in Da (0.01 by default) .
matchWeight	weight of the fragment matches to the final score; value between 0 and 1; the remaining fraction of the weight comes from the candidate m/z error (0.5 by default).
useMZerrorWeight	Logical value to indicate if the m/z error between feature and candidate m/z is to be used for final scoring. Default is TRUE.
NoMatchWeight	Weight to give to the additional matches between the candidate fragments and the MS2 peaks at the RT window of the feature of interest (0.5 by default).
additional	Logical value to indicate if the fragments remaining unmatched to the pseudo-MS/MS are to be tested against the MS2 peaks at the RT window of the feature of interest (default is TRUE).

Value

A list containing one data frame with the summary result of the matching of a pseudo-MS/MS and fragments of a candidate and a data frame with the pseudo-MS/MS spectrum of matched ion fragments.

Author(s)

Goncalo Graca & Yuheng (Rene) Cai (Imperial College London)

Examples

```
# read a hypothetical pseudo-MS/MS spectrum of a feature 152.0720 m/z, 125s
# and assume it is a isotope M+0
fmz <- 152.0720
frt <- 125
iso <- 0
pseudoSpec <- data.frame(mz=c(59.0489, 65.0389, 66.0427, 67.0550, 70.0659,
73.0762, 82.0658, 92.0498, 93.0355, 93.0569, 109.0523,110.0622, 111.0452,
111.0647, 112.0476, 121.0408, 134.0611, 136.0762, 152.0716, 154.0781),
into=c(3228, 8696, 564, 1004, 432, 592, 2092, 4836, 832, 560, 448, 30696,
8516, 3400, 464, 804, 4480, 368, 65236, 464))
# assume high-colision energy spectrum is the same as pseudo-MS/MS spectrum
highCESpec <- pseudoSpec
# Load the small molecule ESI+ library of fragments
data("MetabolitesPos")
lib <- MetabolitesPos$lib
# Compare the pseudo-MS/MS of the test features with the fragments from
# Acetaminophen (entry 8 of the library)
candidate <- lib[[8]]
result <- compFrag(candidate, lib, fmz, frt, iso, highCESpec, pseudoSpec,
maxMZdiff=0.01, matchWeight=0.5, useMZerrorWeight=TRUE, NoMatchWeight=0.5,
additional = TRUE)
```

genFragEntry *Generate a fragment library entry*

Description

Function to generate metabolite database entries from MS/MS spectra obtained from from public databases, stored as a .txt file containing m/z and intensity values, and read imported into R as matrix.

Usage

```
genFragEntry(
  specObject,
  name,
  adduct,
  tmz,
  DirPath = "",
  filename,
  noise = 0.005,
  mpeaksScore = 0.9,
  mpeaksThres = 0.1,
  mzTol = 0.01
)
```

Arguments

specObject	a matrix or data frame object containing the MS/MS spectrum arranged in two columns: 'mz' and 'intensity'. Intensity can be provided in absolute or relative scale.
name	Metabolite name.
adduct	Type of adduct of the parent ion.
tmz	m/z value of the parent ion.
DirPath	Path to the user-defined folder, where the library entry will be saved.
filename	The name of file that will hold the the library entry.
noise	Noise intensity threshold expressed as a ratio to the peak with the highest intensity.
mpeaksScore	The occurrence score to be attributed to the most intense peaks of the MS/MS spectrum which should correspond to the most characteristic fragmentation ions from the metabolite (or 'marker' peaks). These will be the peaks above mpeaksThres value. This score is divided by the number of peaks above mpeaksThres threshold. By default this value is defined at 0.9, which means that peaks below mpeaksThres threshold will be given an occurrence score of 0.1, so that the sum of all fragment occurrence scores is 1.
mpeaksThres	Intensity threshold to select peaks of the MS/MS spectrum considered to be highest intensity, expressed as a ratio to the peak with the highest intensity.
mzTol	Absolute tolerance for feature m/z search in Da.

Value

A .csv file containing fragment and parent m/z values and corresponding occurrence scores.

Author(s)

Goncalo Graca (Imperial College London)

Examples

```
# create a specObject containing the MS/MS spectra to generate a fragment
# database entry from (Panthotenic acid)
specObject <- data.frame(V1=c(70.0298, 85.0652, 90.0556, 98.024, 116.0353,
124.0766, 184.0981, 202.1085, 220.1185), V2=c(13.965907, 13.534607, 100.0,
26.165537, 15.383036, 25.231054, 28.578764, 43.017047, 64.962005))
# Choose a folder to store the result
userDir <- tempdir()
# Generate fragment entry
genFragEntry(specObject, "Pantothenic acid", "[M+H]+", 220.1179,
DirPath=userDir, "Pantothenic_acid_pos", noise=0.005, mpeaksScore=0.9,
mpeaksThres=0.1, mzTol=0.01)
```

---

getPseudoMSMS *Obtain pseudo-MS/MS spectra for an LC-MS feature of interest*

Description

Function to obtain in-source MS and pseudo MS/MS spectra from a feature of interest from All-ion fragmentation experiments (e.g. MSe, bbCID, AIF).

Usage

```
getPseudoMSMS(
  fmz,
  frt,
  xcmsF1,
  xcmsF2,
  peaksF1,
  peaksF2,
  cthres1 = 0.9,
  cthres2 = 0.8
)
```

Arguments

fmz	The m/z for the feature of interest.
frt	Retention time in seconds for the feature of interest.
xcmsF1	MSn object containing the LC-MS no-collision energy scans.

xcmsF2	MSn object containing the LC-MS all-ion fragmentation scans. Should be set to NULL to obtain only the in-source fragmentation (ISF) pseudo-MS/MS.
peaksF1	LC-MS picked peaks from xcmsF1 dataset using XCMS.
peaksF2	LC-MS picked peaks from xcmsF2 dataset using XCMS. Should be set to NULL to obtain only the in-source fragmentation (ISF) pseudo-MS/MS.
cthres1	Correlation threshold for the selection of in-source ions related to the feature of interest.
cthres2	Correlation threshold for the selection of all-ion fragment ions related to the feature of interest.

Value

A list containing several objects: insource, all MS1 peaks related to the feature of interest; aif, all MS2 peaks related to the feature; ms1_peaks, all MS1 peaks at the feature RT; ms2_peaks, all MS2 peaks at the feature RT; ms2_eic, all EICs for the AIF features in the RT window of the feature of interest; mz_ms2, vector of m/z values for the MS2 ions in the RT window of the feature of interest; feic, EIC of the feature of interest; feic_aif, the EICs of all MS2 ions correlated with the feature of interest. If xcmsF2 is set to NULL the in-source pseudo-MS/MS spectrum will be saved instead of the AIF pseudo-MS/MS and similarly the EICs from the MS1 ions correlated with the feature of interest.

Author(s)

Goncalo Graca (Imperial College London)

Examples

```
# obtain the pseudo-MS/MS of one feature from the
# MS1 scans (in-source fragments)
# read the example LC-MS data from the MsDataHub package
library(MsDataHub)
filePath <- filePath <- PestMix1_SWATH.mzML()
xcmsF1 <- MSnbase::readMSData(filePath, msLevel.=1, mode="onDisk")
# perform peak-picking
cwp <- CentWaveParam(snthresh=5, noise=100, ppm=10, peakwidth=c(3, 30))
peaksF1 <- xcms::findChromPeaks(xcmsF1, msLevel = 1L, param = cwp)
# feature m/z and RT
fmz <- 304.1124
frt <- 423.945
# obtain the pseudo-MS/MS from MS1 scans (in-source fragments spectrum):
pseudoMSMS <- getPseudoMSMS(fmz, frt, xcmsF1, xcmsF2=NULL, peaksF1,
peaksF2=NULL, cthres1=0.9, cthres2=0.8)
```

LipidNeg *MetaboAnnotatoR Lipid Negative fragment libraries*

Description

MetaboAnnotatoR default Lipid Negative fragment libraries, comprising 44 lipid class/adducts combinations, corresponding to a total of 15729 lipid entries. The libraries consist of records that include parent ion m/z and expected fragments from negative electrospray ionisation (ESI) collision-induced decay (CID) MS/MS experiments. The lipid libraries were adapted from those of the LipidMatch R package which is a library of theoretical m/z values for experimentally observed lipid fragments. The libraries were adapted to retain only fragments that were commonly observed experimentally in ESI MS/MS spectra and well-documented in the literature.

Usage

```
data("LipidNeg")
```

Format

A list containing two large lists: 1) one library (data frame) per lipid class; 2) A vector with the original libraries filenames.

Details

Each fragment entry from contains the m/z values of each lipid fragment and their respective score weights, which were attributed based on experimental intensity observed. The full details on the libraries are described in the MetaboAnnotatoR paper. This object contains both the libraries for the different lipid classes/adducts and the filepaths of the original libraries .csv files.

References

Graca G., Cai Y., Lau C-H. E., Vorkas P.A., Lewis M.R., Want E.J., Herrington D., Ebbels, T.M.D. Automated Annotation of Untargeted All-Ion Fragmentation LC-MS Metabolomics Data with MetaboAnnotatoR. *Analytical Chemistry*, 2022, 94(8), 3446-3455. DOI: 10.1021/acs.analchem.1c03032

LipidPos *MetaboAnnotatoR Lipid Positive fragment libraries*

Description

MetaboAnnotatoR default Lipid Positive fragment libraries, comprising 47 lipid class/adducts combinations, corresponding to a total of 74786 lipid entries. The libraries consist of records that include parent ion m/z and expected fragments from positive electrospray ionisation (ESI) collision-induced decay (CID) MS/MS experiments. The lipid libraries were adapted from those of the LipidMatch R package which is a library of theoretical m/z values for experimentally observed lipid fragments. The libraries were adapted to retain only fragments that were commonly observed experimentally in ESI MS/MS spectra and well-documented in the literature.

Usage

```
data("LipidPos")
```

Format

A list containing two large lists: 1) one library (data frame) per lipid class; 2) A vector with the original libraries filenames.

Details

Each fragment entry from contains the m/z values of each lipid fragment and their respective score weights, which were attributed based on experimental intensity observed. The full details on the libraries are described in the MetaboAnnotatoR paper. This object contains both the libraries for the different lipid classes/adducts and the filepaths of the original libraries .csv files.

References

Graca G., Cai Y., Lau C-H. E., Vorkas P.A., Lewis M.R., Want E.J., Herrington D., Ebbels, T.M.D. Automated Annotation of Untargeted All-Ion Fragmentation LC-MS Metabolomics Data with MetaboAnnotatoR. *Analytical Chemistry*, 2022, 94(8), 3446-3455. DOI: 10.1021/acs.analchem.1c03032

MetabolitesNeg *MetaboAnnotatoR Metabolite Negative fragment libraries*

Description

MetaboAnnotatoR default nonlipid small-molecule negative electrospray ionisation mode (ESI-) fragment libraries, comprising 79 metabolite class/adducts combinations corresponding to a total of 158 entries. The libraries consist of records that include parent ion m/z and expected fragments from collision-induced decay (CID) MS/MS experiments. The nonlipid small-molecule ESI- library was generated from experimental CID MS/MS spectra from deprotonated ions corresponding to metabolites commonly found in human biofluids, such as urine and blood serum or plasma deposited in MassBank and GNPS GNPS databases.

Usage

```
data("MetabolitesNeg")
```

Format

A list containing two large lists: 1) one library (data frame) per metabolite/adduct combination; 2) A vector with the original libraries filenames.

Details

Each fragment entry from contains the fragment m/z values of each metabolite and respective score weights, which were attributed based on the experimental intensity of the reference MS/MS spectra from MassBank and GNPS. The full details on the libraries are described in the MetaboAnnotatoR paper. This object contains both the libraries for the different metabolite classes/adducts and the filepaths of the original libraries .csv files.

References

Graca G., Cai Y., Lau C-H. E., Vorkas P.A., Lewis M.R., Want E.J., Herrington D., Ebbels, T.M.D. Automated Annotation of Untargeted All-Ion Fragmentation LC-MS Metabolomics Data with MetaboAnnotatoR. *Analytical Chemistry*, 2022, 94(8), 3446-3455. DOI: 10.1021/acs.analchem.1c03032

`MetabolitesPos` *MetaboAnnotatoR Metabolite Positive fragment libraries*

Description

MetaboAnnotatoR default nonlipid small-molecule positive electrospray ionisation mode (ESI+) fragment libraries, comprising 102 metabolite class/adducts combinations corresponding to a total of 255 entries. The libraries consist of records that include parent ion m/z and expected fragments from collision-induced decay (CID) MS/MS experiments. The nonlipid small-molecule ESI+ library was generated from experimental CID MS/MS spectra from proton or sodium adduct ions corresponding to metabolites commonly found in human biofluids, such as urine and blood serum or plasma deposited in MassBank and GNPS databases.

Usage

```
data("MetabolitesPos")
```

Format

A list containing two large lists: 1) one library (data frame) per metabolite/adduct combination; 2) A vector with the original libraries filenames.

Details

Each fragment entry from contains the fragment m/z values of each metabolite and respective score weights, which were attributed based on the experimental intensity of the reference MS/MS spectra from MassBank and GNPS. The full details on the libraries are described in the MetaboAnnotatoR paper. This object contains both the libraries for the different metabolite classes/adducts and the filepaths of the original libraries .csv files.

References

Graca G., Cai Y., Lau C-H. E., Vorkas P.A., Lewis M.R., Want E.J., Herrington D., Ebbels, T.M.D. Automated Annotation of Untargeted All-Ion Fragmentation LC-MS Metabolomics Data with MetaboAnnotatoR. *Analytical Chemistry*, 2022, 94(8), 3446-3455. DOI: 10.1021/acs.analchem.1c03032

mspToLib *Generate metabolite entries from MS/MS spectra (.msp files)*

Description

Function to generate metabolite database entries from MS/MS spectra obtained from from an .msp file.

Usage

```
mspToLib(
  msp_file,
  LibDir = "",
  noise = 0.005,
  mpeaksScore = 0.9,
  mpeaksThres = 0.1
)
```

Arguments

msp_file	an MS/MS spectral library for spectra from one or both polarities.
LibDir	Custom library folder path under which library files will be saved. will be created and where the respective library entries will be stored.
noise	Noise intensity threshold expressed as a ratio to the peak with the highest intensity.
mpeaksScore	The occurrence score to be attributed to the most intense peaks of the MS/MS spectrum which should correspond to the most characteristic fragmentation ions from the metabolite (or 'marker' peaks). These will be the peaks above mpeaksThres value. This score is divided by the number of peaks above mpeaksThres threshold. By default this value is defined at 0.9, which means that peaks below mpeaksThres threshold will be given an occurrence score of 0.1, so that the sum of all fragment occurrence scores is 1.
mpeaksThres	Intensity threshold to select peaks of the MS/MS spectrum considered to be highest intensity, expressed as a ratio to the peak with the highest intensity.

Value

A .csv file containing fragment and parent m/z values and corresponding occurrence scores.

Author(s)

Goncalo Graca (Imperial College London)

Examples

```
# read example.msp file and import as "custom" library
msp_path <- system.file("/Data/MassBank_example.msp",
package="MetaboAnnotatoR")
# Set the library directory to store the library files
userDir <- tempdir()
mspToLib(msp_path, LibDir=userDir, noise=0.005, mpeaksScore=0.9,
mpeaksThres=0.1)
```

---

plotResultSpec *Plot pseudo-MS/MS composed of candidate matched ions*

Description

Function to visualise the spectra containing the matched ions to each candidate annotation result.

Usage

```
plotResultSpec(annotations, feature, candidate)
```

Arguments

annotations	An output object from the annotateAIF or annotaterRC functions.
feature	Index of the annotated feature, as specified in the targets data frame.
candidate	Index of the candidate annotation as specified in the annotations$rankedResults data frame.

Value

A pseudo-MS/MS spectrum is plotted.

Author(s)

Goncalo Graca (Imperial College London)

Examples

```
# Read RAMClustR (RC) and XCMS processed example data:
data("RC")
data("xset")
# read table with features to annotate:
tfile <- system.file("extdata", "targetTable.csv",
package="MetaboAnnotatoR")
targets <- read.csv(tfile)
# read default lipid positive library
data("LipidPos")
# Run annotation of lipid features for positive LC-MS
# processed with RAMClustR:
```

```
annotations <- annotateRC(targets, xcmsObject=xset, ramclustObj=RC,
libs="LipidPos", RTs="none", checkIsotope=TRUE)
# plot the rank 1 candidate of the 3rd feature in the annotation$targets
plotResultSpec(annotations, 3, 1)
```

---

pseudoMSMS *MetaboAnnotatoR-generated example pseudoMSMS spectrum and related data*

Description

This dataset is an example of a MetaboAnnotatoR-generated pseudoMSMS spectrum for an LC-MS feature of interest (468.3095 m/z, 82.92 s) from a serum Lipidomics LC-MS dataset (Quality control sample). The raw data was acquired on a Waters Acquity UPLC system connected to a Waters Xevo-G2 Q-ToF system operated in the MSE mode (all-ion fragmentation mode).

Usage

```
data("pseudoMSMS")
```

Format

A list of 7 elements related to the feature of interest: 1) in-source fragment pseudo-MS/MS spectrum from MS1 scans (data frame); 2) all-ion fragmentation (AIF) pseudo-MS/MS spectrum (data frame); 3) MS1 spectrum (data frame); 4) AIF spectrum (data frame); 5) AIF extracted-ion chromatograms (EICs) (MChromatograms object); 6) m/z from the AIF spectrum (vector); 7) EIC from the feature of interest (MChromatograms object).

Details

The raw mzML file from which the data was extracted can be downloaded from zenodo: Lipid_Positive_QC.mzML.

Author(s)

Goncalo Graca

See Also

getPseudoMSMS MChromatograms

rankScore *Rank candidate metabolite annotations by score*

Description

Ranks the annotation results by final match score.

Usage

```
rankScore(result, specMatch)
```

Arguments

result	Results from fragment matching as data frame.
specMatch	Pseudo-MS/MS of ions matched to candidate fragments.

Value

Ranked results annotation table and ranked matched spectra as list.

Author(s)

Goncalo Graca & Yuheng (Rene) Cai (Imperial College London)

Examples

```
# Create a result object for 3 hypothetical metabolite matches
result <- data.frame(metabolite=c("Met A", "Met B", "Met C"),
                     feature.type=rep("parent",3),
                     ion.type=rep("[M+H]+"),
                     isotope=rep("M+0",3),
                     mz.metabolite=rep(152.0723, 3),
                     matched.mz=rep(152.0706, 3),
                     mz.error=rep(11, 3),
                     pseudoMSMS=rep("TRUE", 3),
                     fraction=c("2 of 5", "4 of 5","3 of 5"),
                     score=c(0.4, 0.9, 0.6))
# Create a list of matched spectra for the 3 hypothetical metabolite matches
specMatch <- list()
specMatch$`Met A` <- data.frame(
                      mz=c(152.0716, 134.0611, 59.0489, 65.0389, 66.0427),
                      into=c(432, 592, 2092, 4836, 832)
                      )
specMatch$`Met B` <- data.frame(
                      mz=c(152.0716, 134.0611, 110.0622, 109.0523, 59.0489),
                      into=c(65236, 4480, 30696, 448, 432)
                      )
specMatch$`Met C` <- data.frame(
                      mz=c(152.0716, 134.0611, 110.0622, 109.0523, 93.0569),
```

```
                    into=c(65236, 4480, 30696, 464, 804)
                    )
# Rank candidate metabolite result and spectraMatch by score
ranked <- rankScore(result, specMatch)
```

RC *Pseudo-MS/MS spectra from MESA Lipid Positive LC-MS serum samples*

Description

Pseudo-MS/MS spectra obtained by processing the MESA Lipid Positive LC-MS xcms data using (RAMClustR package). The xcms data consisted random selection of 100 LC-MS Lipid Positive chromatograms from human serum samples acquired for the Multi-Ethnic Study of Atherosclerosis (MESA).

Usage

```
data("RC")
```

Format

A RAMClustR object (list).

Details

One hundred serum samples from the MESA study, were randomly selected and the MS1 and MSE scans from Lipid Positive LC-MS chromatograms were separated in 2 different files for processing in xcms and RAMClustR. The raw data was acquired on a Waters Acquity UPLC system connected to a Waters Xevo-G2 Q-ToF system operated in the MSE mode (all-ion fragmentation mode), and processed in xcms. The full experimental details, xcms and RAMClustR parameter sets used for processing are described in the MetaboAnnotatoR paper.

References

Graca G., Cai Y., Lau C-H. E., Vorkas P.A., Lewis M.R., Want E.J., Herrington D., Ebbels, T.M.D. Automated Annotation of Untargeted All-Ion Fragmentation LC-MS Metabolomics Data with MetaboAnnotatoR. *Analytical Chemistry*, 2022, 94(8), 3446-3455. DOI: 10.1021/acs.analchem.1c03032

rCandidates *MetaboAnnotatoR-generated example of ranked annotation candidates*

Description

Ranked list of candidates from the Lipid Positive Library for feature 468.3095 m/z 82.92 s from the example Lipid Positive (QC sample). This output was generated by function `rankScore()`.

Usage

```
data("rCandidates")
```

Format

A list containing two elements: 1) a table with the ranked candidates and 2) matched spectra of the ranked candidates.

Details

The example raw mzML file from which the data was extracted can be downloaded from zenodo: Lipid_Positive_QC.mzML.

Author(s)

Goncalo Graca

See Also

`rankScore`

RCspec *Get RAMClustR pseudo-MS/MS spectra (cluster).*

Description

For a given feature, find out corresponding cluster (pseudo-MS/MS spectra) from RAMClustR object.

Usage

```
RCspec(fmz, frt, ramclustObj)
```

Arguments

fmz	The m/z for the feature of interest.
frt	Retention time in seconds for the feature of interest.
ramclustObj	RAMClustR object (list) with parent-fragment reconstructions. See RAMClustR paper for more details (https://pubs.acs.org/doi/10.1021/ac501530d).

Value

Pseudo-MS/MS spectrum for the feature of interest.

Author(s)

Goncalo Graca & Yuheng (Rene) Cai (Imperial College London)

Examples

```
# read  RamclustR example data
data("RC")
# obtain pseudo-MS/MS from the RC example data
spec <- RCspec(fmz=468.3094, frt=82.92, ramclustObj=RC)
```

---

saveAnnotations *Save annotation results*

Description

Saves all annotation related data, such as annotation table, options, and optionally plots of pseudo-MS/MS and matched fragments spectra.

Usage

```
saveAnnotations(
  annotations,
  DirPath = "",
  saveOptions = TRUE,
  saveXCMSoptions = FALSE,
  saveRanked = TRUE,
  saveRankedSpec = FALSE,
  savePseudoMSMS = FALSE
)
```

Arguments

annotations	Annotation object created from running annotation annotateAIF or annotateRC functions.
DirPath	Path to the directory where the plots (as .pdf), .csv tables and .mgf files will be saved. If no directory is specified, the files will be saved to the working directory.
saveOptions	If TRUE, will save the annotation options as .csv file.
saveXCMSoptions	Saves the XCMS options if the annotations originate from AIF raw files, if set to TRUE.
saveRanked	Option to save the ranked candidate table and respective pseudo-MS/MS. The default option is TRUE.
saveRankedSpec	Option to save the ranked candidate matched fragment pseudo-MS/MS spectra as .pdf files. If the annotations are from an single AIF raw file, both the EICs for each fragment an pseudo-MS/MS spectra will be save.
savePseudoMSMS	Option to save the pseudo-MS/MS for the features from the targets table. The default option is FALSE. If TRUE, the pseudo-MS/MS will be saved as .mgf file and the corresponding plots as .pdf.

Value

Global and candidate annotations as .csv files and pseudo-MS/MS spectra as .pdf and/or .mgf files.

Author(s)

Goncalo Graca & Yuheng (Rene) Cai (Imperial College London)

Examples

```
# Read RAMClustR (RC) and XCMS processed example data:
data("RC")
data("xset")
# Read the table of features to annotate:
tfile <- system.file("extdata", "targetTable.csv",
package="MetaboAnnotatoR")
targets <- read.csv(tfile)
# Read the default lipid positive libraries:
data("LipidPos")
# Run annotation of lipid features for positive LC-MS
# processed with RAMClustR:
annotations <- annotateRC(targets, xcmsObject=xset, ramclustObj=RC,
libs="LipidPos", RTs="none", checkIsotope=TRUE)
# Finally, save the results to a user-defined directory:
userDir <- tempdir()
saveAnnotations(annotations, DirPath=userDir, saveOptions=TRUE,
saveXCMSoptions=FALSE, saveRanked=TRUE,
saveRankedSpec=TRUE, savePseudoMSMS=TRUE)
```

searchLib *Search candidate metabolites from the fragments libraries*

Description

Search candidate metabolites from the fragments libraries for a given feature using m/z and RT (if metabolite RTs are known). If no match is found in the "parent" ions, for instance in the case of a feature corresponding to an in-source fragment, fragments are then searched. Fragment ions will only be considered if the parent ion is present in the same pseudo-MS spectrum (MS1).

Usage

```
searchLib(libraries, libfiles, fmz, frt, tolerance = 25, RTs, inSourceSpec)
```

Arguments

libraries	List object containing all loaded libraries.
libfiles	Path to the libraries files.
fmz	The m/z for the feature of interest.
frt	Retention time in seconds for the feature of interest.
tolerance	Tolerance for m/z candidate search in ppm.
RTs	Optional metabolites classes Retention Times in seconds. Default value is "none".
inSourceSpec	Data frame containing the pseudo-MS spectrum (MS1). This will be used to check for the "parent" ion when the feature of interest is matched to a fragment (in-source fragment).

Value

A list of data frames containing the candidates from the fragment libraries which will be used in the pseudo-MS/MS to fragment matching step.

Author(s)

Goncalo Graca & Yuheng (Rene) Cai (Imperial College London)

Examples

```
# load default libraries for Metabolites in Positive mode
data("MetabolitesPos")
libfiles <- MetabolitesPos$libfiles
lib <- MetabolitesPos$lib
# read feature mz and rt and in-source spectrum
fmz <- 152.0720
frt <- 125
inSourceSpec <- data.frame(mz = 152.0720, into = 1)
# Search the library for candidates
candidates <- searchLib(lib, libfiles, fmz, frt, tolerance=25,
RTs="none", inSourceSpec)
```

xcmsSpec *Extract the low- and high-collision-energy spectra from XCMS object (from AIF data)*

Description

Given a feature of interest (m/z RT pair), this function will extract the low-collision-energy (MS1) or high-collision-energy features (MS2) at the same RT window of the feature of interest from an XCMS output object obtained by processing both functions together from a set of LC-MS AIF chromatograms.

Usage

```
xcmsSpec(fmz, frt, xcmsObject, mztol = 0.01, rttol = 5, highCE = TRUE)
```

Arguments

fmz	The m/z for the feature of interest.
frt	Retention time in seconds for the feature of interest.
xcmsObject	An xcmsSet object resulting from the processing of one or multiple samples in XCMS.
mztol	Absolute tolerance for feature m/z search in Da.
rttol	Tolerance for feature RT search in seconds. The default (5 s) only applies to UPLC/UHPLC data.
highCE	Logic value. If TRUE the high collision-energy is extracted, otherwise if FALSE the "in-source" spectrum is returned.

Value

A data frame with ions (m/z and intensity) from the high collision-energy or low collision-energy features found at the same RT window as the feature of interest.

Author(s)

Goncalo Graca & Yuheng (Rene) Cai (Imperial College London)

Examples

```
# Extract the MS1 spectrum of feature 585.2692 m/z 72.8s using the xset
# test data
data("xset")
# obtain spectrum
spec <- xcmsSpec(fmz=585.2692, frt=72.8, mztol=0.01,
xset, rttol=1, highCE=FALSE)
```

xset *MESA cohort subset Lipid Positive LC-MS samples processed using XCMS*

Description

A random selection of 100 LC-MS Lipid Positive chromatograms from human serum samples acquired for the Multi-Ethnic Study of Atherosclerosis (MESA).

Usage

```
data("xset")
```

Format

An xcmsSet object containing the xcms processed data.

Details

One hundred serum samples from the MESA study, were randomly selected and the corresponding MS1 and MSE scans the from Lipid Positive LC-MS chromatograms were separated in 2 different files for processing in xcms. The raw data was acquired on a Waters Acquity UPLC system connected to a Waters Xevo-G2 Q-ToF system operated in the MSE mode (all-ion fragmentation mode). The full experimental details and xcms parameter sets used for processing are described in the MetaboAnnotatoR paper.

References

Graca G., Cai Y., Lau C-H. E., Vorkas P.A., Lewis M.R., Want E.J., Herrington D., Ebbels, T.M.D. Automated Annotation of Untargeted All-Ion Fragmentation LC-MS Metabolomics Data with MetaboAnnotatoR. *Analytical Chemistry*, 2022, 94(8), 3446-3455. DOI: 10.1021/acs.analchem.1c03032

See Also

xcmsSet

Index

acetaminophen, 2
annotateAIF, 3
annotateRC, 5

checkIsotope, 6
compFrag, 7

genFragEntry, *3*, 9
getPseudoMSMS, 10, *17*

LipidNeg, 12
LipidPos, 12

MChromatograms, *17*
MetabolitesNeg, 13
MetabolitesPos, 14
mspToLib, 15

plotResultSpec, 16
pseudoMSMS, 17

rankScore, 18, *20*
RC, 19
rCandidates, 20
RCspec, 20

saveAnnotations, 21
searchLib, 23

xcmsSet, *25*
xcmsSpec, 24
xset, 25