Package ‘HiResTEC’

February 27, 2025

Type Package

Title Non-Targeted Fluxomics on High-Resolution Mass-Spectrometry Data

Version 0.63.1

Date 2025-02-21

Maintainer Jan Lisec <jan.lisec@bam.de>

Description Identifying labeled compounds in a 13C-tracer experiment in non-targeted fashion is a cumbersome process. This package facilitates such type of analyses by providing high level quality control plots, deconvoluting and evaluating spectra and performing a multitude of tests in an automatic fashion. The main idea is to use changing intensity ratios of ion pairs from peak list generated with 'xcms' as candidates and evaluate those against base peak chromatograms and spectra information within the raw measurement data automatically. The functionality is described in Hoffmann et al. (2018) <doi:10.1021/acs.analchem.8b00356>.

License GPL-3

URL https://github.com/janlisec/HiResTEC, https://janlisec.github.io/HiResTEC/

Depends R (>= 3.5)

Imports CorrectOverloadedPeaks (>= 1.3.5), plyr

Encoding UTF-8

Language en-US

LazyData TRUE

RoxygenNote 7.3.2

Suggests beeswarm, CorMID, InterpretMSSpectrum, knitr, openxlsx, rmarkdown

VignetteBuilder knitr

NeedsCompilation yes

Author Jan Lisec [aut, cre] (<https://orcid.org/0000-0003-1220-2286>), Friederike Hoffmann [aut]

Repository CRAN
Date/Publication 2025-02-27 09:50:05 UTC

Contents

DeconvoluteSpectrum 2
EvaluateCandidateListAgainstRawData 3
EvaluatePairsFromXCMSSet 5
GenerateCandXLSX 6
GenerateQCPlots 7
getMultipleBPC 8
plotBPC 9
plotMID 10
raw 12

Index **13**

`DeconvoluteSpectrum` *Deconvolute a Mass Spectrum from a list of raw data files.*

Description

`DeconvoluteSpectrum` will evaluate a list of ‘xcmsRaw‘ or ‘xcmsRawLike‘ objects at a given time (rt) and potential mass (mz1). The main purpose is to deconvolute the mass spectrum at rt including mz1.

Usage

```
DeconvoluteSpectrum(
  dat = NULL,
  rt = NULL,
  rt_dev = 3,
  mz1 = NULL,
  mz_dev = 0.003,
  use.mz.adjust = FALSE,
  ionization = c("APCI", "ESI")[1],
  smooth = 0
)
```

Arguments

`dat`	A list of ‘xcmsRaw‘ or ‘xcmsRawLike‘ objects.
`rt`	Retention time of the expected peak.
`rt_dev`	Allowed retention time deviation.
`mz1`	If specified, ensure that this mass is included in the spectrum (assumed base peak). Can be NULL otherwise in which case the most intense peak at rt will be selected as mz1.

mz_dev	Allowed mz deviation [Da].
use.mz.adjust	Will adjust mz on an experiment wide basis.
ionization	Either APCI or ESI. Choice will modify some internal parameters and checks performed.
smooth	Smoothing parameter passed on to getMultipleBPC.

Details

The specific advantage of DeconvoluteSpectrum is, that it does not deconvolute signals within a single measurement file but uses correlation tests over a set of measurements to improve statistical power. It will test all mz around a specified rt to co-apex with some mz1, have a low rt difference and consistent intensity ratio over all samples.

Value

A pseudo spectrum at rt (containing mz1 if specified). Effectively a 2-column matrix (mz, int) with rt as attribute.

Examples

```
# The example measurement data provided with HiResTEC contain a peak at 1026s
raw <- HiResTEC::raw
HiResTEC::DeconvoluteSpectrum(raw, rt = 1026)
```

EvaluateCandidateListAgainstRawData

Evaluate m/z pairs against raw data.

Description

EvaluateCandidateListAgainstRawData will compare the result of function EvaluatePairsFromXCMSSet against raw data files.

Usage

```
EvaluateCandidateListAgainstRawData(
  x = NULL,
  tp = NULL,
  gr = NULL,
  dat = NULL,
  dmz = 0.025,
  drt = 1,
  dEcut = 1,
  Pcut = 0.01,
  Icut = 1000,
  method = c("APCI", "ESI")[1],
```

```
  rolp = c("non", "pos", "neg", "all")[2],
  smooth = 0
)
```

Arguments

x	Dataframe of results (output of EvaluatePairsFromXCMSet).
tp	Timepoint.
gr	group, e.g. different genotypes or concentrations.
dat	list of xcmsRaw's for deconvolution and plotting.
dmz	Allowed mass deviation in Da (for BPC extraction).
drt	Allowed rt deviation in seconds (for get extraction).
dEcut	Minimum required change in enrichment before a candidate ID is assigned.
Pcut	Maximum allowed P value before a candidate ID is assigned.
Icut	Minimum required median peak intensity before a candidate ID is assigned.
method	Either APCI or ESI. Choice will modify some internal parameters and checks performed.
rolp	RemoveOverLappingPeaks parameter, overlapping means from a deconvoluted spectrum where another peak was already evaluated.
smooth	Smoothing parameter passed to getMultipleBPC.

Details

This function will evaluate candidate mz pairs found within an 'xcmsSet' object or any peak list by EvaluatePairsFromXCMSSet against the raw measurement data. This step is required to minimize redundancy and false positive results. It will allow to generate a number of informative quality control plots. As quite some input data is required for this function, please have a look in the vignette for an example. A special parameter in this function is 'rolp' which can be set to 'non', 'pos', 'neg' or 'all'. It will influence the time performance of the function by determining how many peaks are effectively tested. If 'rolp' is set to 'non', no overlapping peaks will be skipped, every individual mz-pair will be sequentially evaluated (slow but most informative). If it is set to 'pos' or 'neg', overlapping peaks (determined by experiment wide deconvolution) will not be tested additionally for positive or negative hits ('neg' is standard). If set to 'all' overlapping peaks will always be removed from the list of mz-pairs to be tested (fast).

Value

A list of evaluation results.

EvaluatePairsFromXCMSSet

Identify and evaluate mz pairs from a peak list.

Description

EvaluatePairsFromXCMSSet will analyze an ‘xcmsSet‘ result or a generic peak list from a mass spectrometry experiment for mass pairs (mz1, mz2) with changes due to any tracer incorporation.

Usage

```
EvaluatePairsFromXCMSSet(
  xg = NULL,
  tp = NULL,
  gr = NULL,
  drt = 1,
  dmz = 0.025,
  mz_iso = 1.00335,
  n = 6,
  method = c("APCI", "ESI")[1],
  specific_row = NULL,
  testing = FALSE,
  silent = FALSE
)
```

Arguments

xg	xcmsSet object with group information. Alternatively, can be a numeric matrix containing ‘mz‘ and ‘rt‘ information in the first two columns followed by peak intensities of all samples in the same order as in parameters ‘tp‘ and ‘gr‘.
tp	Time point information for all samples (obviously required, internally converted to factor).
gr	Group information for all samples, e.g. different genotypes or concentrations (optional, factor).
drt	Allowed rt deviation in time units of xcmsSet (usually seconds) to test for candidates.
dmz	Allowed mass deviation in Da.
mz_iso	Mass defect of the isotope under investigation (use 1.00335 for ^13^C experiments.
n	Number of maximal incorporated carbons to test.
method	Currently APCI or ESI. If APCI, dmz will be modified depending on n (see details).
specific_row	A single row of the peak list to process.
testing	Stop in function using browser() if specific_row is specified; can be a isotope number, i.e. 3 will stop at third isotope.
silent	Suppress warnings and console output if TRUE.

Details

Using 'APCI' as method assumes that (i) you analyze TMS-derivatized compounds and (ii) your MS resolution does not allow to separate Si and C isotopes but reports an intermediate mass as m/z. In this case you will find carbon isotopes below there expected masses, i.e. M+1 would be 1.001 mDa apart from M+0 instead of 1.003. The effect is increased with isotope number, i.e. M+6 will be ~20 mDa below the expected value. Hence, selecting method 'APCI' will combine your selected dmz with a allowed deviation due to mass shifts caused by Si isotopes. Use 'ESI' if you are not sure if this effect takes place in your settings.

Value

A dataframe with all observable pairs within the provided xg object (usually an 'xcmsSet' peak) list including mean group intensities and P values.

Examples

```
# The example measurement data provided with HiResTEC contain a peak at 1026s
raw <- HiResTEC::raw
sam <- HiResTEC::sam
mz <- c(556.26, 561.26, 564.26)

# extract the peak intensities for 3 m/z of this peak
int <- sapply(raw, function(x) {
  tmp <- getMultipleBPC(x = x, mz = mz, mz_dev = 0.04, rt = 1026)
  tmp[attr(tmp, "maxBPC"),]
})
colnames(int) <- sam$ID; rownames(int) <- NULL
xg <- data.frame(
 "mz" = mz,
 "rt" = rep(1026.5, 3),
 int
)

# evaluate this peak list for interesting pairs
EvaluatePairsFromXCMSSet(xg=xg, tp=sam$TP, gr=sam$Group, silent=TRUE, n=8)
```

GenerateCandXLSX — *Generate a table for the candidates obtained in EvaluateCandidateListAgainstRawData.*

Description

GenerateCandXLSX will produce a XLSX of a list containing test results objects.

Usage

```
GenerateCandXLSX(res_list = NULL, xlsx_file = NULL, rejected = FALSE)
```

Arguments

`res_list`	A list of result objects (each testing an individual mz pair).
`xlsx_file`	File name.
`rejected`	Logical. Prepare table of rejected candidates if TRUE.

Details

Just a wrapper, to get the important information in a tabular layout.

Value

Candidate table as data.frame.

Examples

```
# load evaluation result of example data and
# generate table within R (use parameter xlsx_file to write to file)
x <- GenerateCandXLSX(HiResTEC::res_list)
str(x)
x[,1:5]
```

`GenerateQCPlots` *Generate quality control plots for mz pair candidates.*

Description

`GenerateQCPlots` will produce QC plots for a list containing test results objects (generated by EvaluateCandidateListAgainstRawData.

Usage

```
GenerateQCPlots(
  res_list = NULL,
  pdf_file = NULL,
  mfrow = NULL,
  skip_plots = NULL
)
```

Arguments

`res_list`	A list of result objects (each testing an individual mz pair).
`pdf_file`	The path to a writable file.
`mfrow`	If NULL automatically determined, otherwise useful to specify a layout.
`skip_plots`	NULL or numeric vector in which case plots with numbers in skip_plots will be empty.

Details

If you provide a single candidate (list of length = 1) an output to a screen plotting device is reasonable, otherwise a target PDF file should be specified in parameter ‘pdf_file‘.

Value

Figure output.

Examples

```
## Not run:
  # load evaluation result of example data
  data(res_list)
  # generate Figures on screen (use PDF output for multiple candidates)
  GenerateQCPlots(res_list[1])

## End(Not run)
```

---

getMultipleBPC *Extract multiple ion chromatograms from mass spectrometry data.*

Description

`getMultipleBPC` will extract multiple BPCs from an ‘xcmsRaw‘ or ‘xcmsRawLike‘ object for a vector of mz within the limits given by rt, rt_dev and mz_dev.

Usage

```
getMultipleBPC(
  x,
  mz = NULL,
  mz_dev = 0.005,
  rt = NULL,
  rt_dev = 2,
  zeroVal = NA,
  smooth = 0,
  returnEIC = FALSE
)
```

Arguments

`x`	‘xcmsRaw‘ or ‘xcmsRawLike‘ object.
`mz`	Numeric vector of masses or NULL (default) to return the overall BPC.
`mz_dev`	Allowed mass deviations (can be a single numeric, a vector, a matrix with one row (lower bound, upper bound) or a matrix with `length(mz)` rows giving lower and upper bound for each mz).

rt	Target retention time or NULL (default) to use full time range.
rt_dev	Allowed time deviation (if rt is specified).
zeroVal	Set values <=0 to NA or keep as is with NULL.
smooth	Window size for moving average smoother, 0 = no smoothing.
returnEIC	Return EIC instead of BPC?

Details

While there are other functions to extract BPC information from raw data, this one is particularly useful to get all traces belonging to a isotopologue group. It will attach several derived values to the results object, i.e. describing the observed mass shift (deviation from expected value) which is helpful in QC for non-targeted tracer analyses. While the ‘mz‘ and ‘mz_dev‘ parameters can be vectorized, the ‘rt‘ and ‘rt_dev‘ values will be consistently used for all ion traces.

Value

A matrix with scan wise (rows) intensities for all requested masses (columns) as either EIC or BPC.

References

Uses C code modified from XCMS (see `citation("xcms")`).

Examples

```
raw <- HiResTEC::raw
# search for mz = 556.263 and its isotopic traces
mz <- 556.263 + c(0:3) * 1.0034
getMultipleBPC(x = raw[[1]], mz = mz, mz_dev = 0.04, rt = 1026)
```

---

plotBPC — *Plot base peak chromatograms for multiple high resolution masses in multiple samples.*

Description

`plotBPC` will plot for each item of a list of result-objects from getMultipleBPC the BPC traces and the spectrum at the scan where the summed intensity of all ions is max.

Usage

```
plotBPC(
  bpc = NULL,
  mfrow = NULL,
  skip_plots = NULL,
  ylim = NULL,
  col = NULL,
```

```
  ids = NULL,
  type = "both",
  ann = c("mdev", "mz", "none")
)
```

Arguments

bpc	A bpc object (list of intensity matrices, rt x mz, including several attributes as attached by getMultipleBPC).
mfrow	Specify mfrow explicitly (is optimized internally if NULL to cover n=length(bpc)).
skip_plots	Allows to block certain subplots in the mfrow matrix to better align replicates.
ylim	Can be specified specifically, will be adjusted to overall min/max otherwise.
col	Specific color vector for masses may be provided.
ids	Specific plot ids may be explicitly provided.
type	Switch between co-plot of BPC and Spectrum ("both") or BPC alone ("bpc").
ann	Select value to annotate peaks in spectrum. Usually the mass deviation from the expected value in mDa.

Details

plotBPC allows to get a quick overview of similar information from all samples of an experimental set. As it uses ‘mfrow‘ to arrange samples its output can not be used as subplot in other figures.

Value

A plot to the graphics device and NULL as invisible.

Examples

```
# load example raw data
bpc <- HiResTEC::res_list[[1]][["bpc"]][c(1:2, 13:14)]
plotBPC(bpc = bpc)
plotBPC(bpc = bpc, ann="mz", ids=LETTERS[1:4], mfrow=c(3,2), skip_plots=c(2,3))
```

---

plotMID *Plot a Mass Isotopomer Distribution (MID) for multiple samples.*

Description

plotMID will plot Mass Isotopomer Distributions (MIDs).

Usage

```
plotMID(
  mid = NULL,
  gr = NULL,
  name = "unknown",
  contr = NULL,
  stackedbars = FALSE,
  subplot_ylim = 100,
  ...
)
```

Arguments

mid	Matrix of corrected MIDs. Samples in columns, MID values in rows.
gr	Groups, a factor vector of length ncol(mid).
name	Name of analyte for annotation.
contr	Contrasts. Not yet clear if useful.
stackedbars	Alternative plotting layout using stacked bar plot.
subplot_ylim	Calculate ‘ylim‘ individually per subplot if 0, show full range in all subplots if 100 and limit to the minimal specified number otherwise.
...	Further arguments to ’boxplot’ or ’barplot’ (depending on parameter ’stackedbars’).

Details

Multiple styling options are available using the function parameters.

Value

An annotated barplot or boxplot.

Examples

```
mid <- matrix(c(seq(0, 0.3, 0.1), seq(1, 0.7, -0.1)), byrow = TRUE, nrow = 2)
gr <- gl(2, 2, labels = letters[1:2])
plotMID(mid = mid, gr = gr, name = "Metabolite X")
plotMID(mid = mid, gr = gr, stackedbars = TRUE, las = 1, xlab = "MID")
lt <- paste0("M", 0:1)
rownames(mid) <- lt
plotMID(mid = mid, gr = gr, stackedbars = TRUE, xlab = "MID", legend.text = lt)
plotMID(mid = mid[, 2, drop = FALSE], stackedbars = TRUE, col = c(3, 4))
colnames(mid) <- paste0("S", 1:4)
gr2 <- gl(n = 1, k = 1, labels = "bla")
plotMID(mid = mid[, 2, drop = FALSE], gr = gr2, stackedbars = TRUE, name = NULL)
plotMID(mid = mid, gr = factor(colnames(mid)), stackedbars = TRUE, name = NULL)
```

`raw` *Different example data files to be used in the help section of 'HiResTEC' functions.*

Description

'raw' contains a list of 'xcmsRawLike' objects, each containing a selected mass range for a small retention time window of 18 samples defined in sam.

'sam' provides a data frame containing the sample definition of 18 samples from a larger experiment.

'xcms_cand' contains a data frame with the analysis result of an 'xcmsSet' obtained using EvaluatePairsFromXCMSSet.

'res_list' is a list containing the evaluations results established based on processing example data with EvaluateCandidateListAgainstRawData.

'mz_shift_corrector' contains a list defining windows for high res APCI or ESI instrumentation.

Usage

```
data(raw)

data(sam)

data(xcms_cand)

data(res_list)

data(mz_shift_corrector)
```

Format

An object of class `list` of length 18.

An object of class `data.frame` with 18 rows and 8 columns.

An object of class `data.frame` with 72 rows and 19 columns.

An object of class `list` of length 14.

An object of class `list` of length 3.

Source

jan.lisec@bam.de

Index

∗ **data**
raw, 12

DeconvoluteSpectrum, 2

EvaluateCandidateListAgainstRawData, 3, *6*, *7*, *12*
EvaluatePairsFromXCMSSet, *3*, *4*, 5, *12*

GenerateCandXLSX, 6
GenerateQCPlots, 7
getMultipleBPC, *3*, *4*, 8, *9*, *10*

mz_shift_corrector *(raw)*, 12

plotBPC, 9
plotMID, 10

raw, 12
res_list *(raw)*, 12

sam *(raw)*, 12

xcms_cand *(raw)*, 12