Package 'MetabolomiQCsR'

August 12, 2016

Type Package

Title QC for LC-MS data

Version 1.0

Date 2016-07-18

Author Jan Stanstrup

Maintainer Jan Stanstrup <stanstrup@gmail.com>

Description Functions for performing QC on LC-MS.

Depends R (>= 3.0)

Imports tibble, magrittr, dplyr, purrr, xcms, massageR, tidyr, scales, ggplot2, ini, stringr, RCurl, readr, utils, WGCNA

Suggests knitr, rmarkdown, faahKO, plotly, svglite, chemhelper, RColorBrewer, heatmaply, listviewer

License GPL (>= 2)

URL None

biocViews MassSpectrometry, Metabolomics

VignetteBuilder knitr

RoxygenNote 5.0.1

Remotes stanstrup/massageR, stanstrup/chemhelper

R topics documented:

closest_match *Match list of standards to peak table*

Description

This function will match a table of standard compounds and a peak table by m/z and retention time. If there is more than one possible hit the highest intensity peak will be chosen.

Usage

```
closest_match(stds, peakTable, rt_tol = 0.25, mz_ppm = 30, rt_col = "rt",
  mz_col = "mz", int_col = "into")
```

Arguments

stds	tibble of standards to match to a peak table
peakTable	tibble containing peak table supplied by findPeaks (but converted to tibble/data.frame).
rt_tol	Retention time tolerance for matching peaks. Pay attention to the unit of your tables. rt_tol should match and stds and peakTable should use same units (i.e. minutes of seconds).
mz_ppm	ppm for matching peaks.
rt_col	Character string giving the column containing the retention times. Must be same in standards and peak table.
mz_col	Character string giving the column containing the m/z values. Must be same in standards and peak table.
int_col	Character string giving the column containing the intensities in the peak table.

Value

A vector having the length equivalent to the number of rows in stds giving the indices of the hits in peakTable.

EIC_contaminants *Find EICs that behave like contaminants*

Description

This function looks in raw LC-MS data for "features"/EICs that behave like contaminants. Behaving like contaminants in this case means that a certain m/z values is present in more than min_time above intensity min_int.

Usage

```
EIC_contaminants(raw, bin_ppm = 30, interval_ppm = 30, min_time = 5,
  merge_corr = 0.9, merge_ppm = 30, min_int = 5000)
```

Arguments

raw	xcmsRaw object to profile
bin_ppm	Tolerance (ppm) for initial binning of m/z values
interval_ppm	Tolerance for creating final EICs after merging similar bins
min_time	Minimum time (minutes) an EIC should be above min_int to be considered a contaminant.
merge_corr	Minimum correlation between EICs to be merged.
merge_ppm	Maximum difference (ppm) between EICs to be merged.
min_int	Minimum intensity that the EIC needs to be above for a a minimum of min_time.

Value

A tibble containing the columns:

- **mz:** m/z of the proposed contaminant
- **EIC:** EIC of the m/z.

---

extract_polarity *Extract polarity from xcmsRaw object.*

Description

Extracts polarity from an xcmsRaw object. The polarity found in the majority of scans is returned.

Usage

```
extract_polarity(xraw)
```

Arguments

xraw	The xcmsRaw object to extract polarity from.

Value

A character string giving the polarity. Can be "positive", "negative", or "unknown".

get_cont_list *Get list of known contaminants*

Description

Get list of known contaminants

Usage

```
get_cont_list(polarity = c("positive", "negative", "unknown"), type = "URL")
```

Arguments

polarity	The polarity to get contaminants for. Can be "positive", "negative" or "unknown". If "unknown" the list specified in the MetabolomiQCsR.conf is used. MetabolomiQCsR.conf can be in the working folder or the home folder. If those are not found the package default is used (unknown will used the positive mode list).
type	If using local or remote. Only "URL" implemented which downloads a list from https://github.com/stanstrup/common_mz

Value

tbl A tibble containing the columns:

- **Monoisotopic ion mass (singly charged):** m/z of the contaminant
- **Ion type:** Notation for adduct/fragment type
- **Formula for M or subunit or sequence:** Molecular formula
- **Compound ID or species:** Name of the compound
- **Possible origin and other comments:** Suggestion for the origin of the contaminant
- **References:** Reference for the contaminant

References

https://github.com/stanstrup/common_mz

get_EICs *Get EICs from xcmsRaw object*

Description

Takes an xcmsRaw object and extracts EICs. Can do multiple ranges and exclude certain masses unlike getEIC. Can be used to extract the TIC too.

Usage

```
get_EICs(xraw, range_tbl, exclude_mz = NULL, exclude_ppm = 30,
  range_tbl_cols = c("mz_lower", "mz_upper"), BPI = FALSE)
```

Arguments

xraw	xcmsRaw object to get EIC(s)/TIC from.
range_tbl	data.frame/tibble with columns for the lower and upper m/z boundaries of EIC slice(s).
exclude_mz	Masses to exclude from the EIC. Most useful to remove contaminants from TICs.
exclude_ppm	ppm tolerance of exclude_mz
range_tbl_cols	Which columns in range_tbl holds the lower and upper range. defaults to c("mz_lower","mz_upper").
BPI	Logical selecting to calculate TIC (FALSE) or BPI.

Value

tbl A tibble containing the columns:

- **scan:** scan number
- **scan_rt:** Retention time of scan
- **intensity:** The summed intensity for each scan in the given m/z interval

---

peak_factor *Calculate Tailing Factor and Asymmetry Factor*

Description

Calculate Tailing Factor and Asymmetry Factor

Usage

```
peak_factor(EIC, rt, factor = "TF")
```

Arguments

EIC	EIC containing the peak to calculate for. tibble as produced with get_EICs.
rt	Retention time of the center of the peak (Numeric)
factor	to calculate. Character string either "TF" (Tailing Factor) or "ASF" (Asymmetry Factor).

Value

Numeric

References

http://www.chromforum.org/viewtopic.php?t=20079

plotly_clean_tt *Make replacements in plotly tooltips*

Description

Make replacements in plotly tooltips

Usage

```
plotly_clean_tt(plotly, rep)
```

Arguments

plotly	A plotly object.
rep	A named character vector. Names are the text to replace and the string is the replacement string.

Value

A plotly object.

plot_chrom *Plot chromatogram*

Description

Plot chromatogram

Usage

```
plot_chrom(tbl, RT_col = "RT", Intensity_col = "Intensity")
```

Arguments

tbl	tbl with retention time and intensity to plot
RT_col	Name of the rentention time column
Intensity_col	Name of the intensity column

Value

a ggplot object.

plot_contaminants *Bar plot of contaminants*

Description

Bar plot of contaminants

Usage

```
plot_contaminants(data, title, x_var = "comp_name", y_var = "EIC_median")
```

Arguments

data	tbl with the contamination amounts
title	Plot title
x_var	Column name that holds the compound/contaminant names
y_var	Column name that holds the compound/contaminant values

Value

a ggplot object.

tbl2ROI *Convert a list of peaks (rt / m/z pairs) to a Region of Interest (ROI) list for use with findPeaks*

Description

Convert a list of peaks (rt / m/z pairs) to a Region of Interest (ROI) list for use with findPeaks

Usage

```
tbl2ROI(tbl, raw, ppm, rt_tol)
```

Arguments

tbl	tibble containing the columns "rt" and "mz". rt needs to be in seconds.
raw	xcmsRaw object to create ROI for. It needs to be a specific xcmsRaw to match retention times to scan nubmers.
ppm	ppm tolerance for the generated ROI.
rt_tol	Retention time tolerance (in sec!) for the generated ROI.

Value

List containing the ROIs. Each list contains mz, mzmin, mzmax, scmin, scmax, length (set to -1, not used by centWave) and intensity (set to -1, not used by centWave) columns.

`xcmsRaw_to_tbl` *Convert raw data into a tibble of xcmsRaw objects.*

Description

Convert raw data into a tibble of xcmsRaw objects.

Usage

```
xcmsRaw_to_tbl(files, ...)
```

Arguments

`files`	character vector of file names/paths.
`...`	further arguments to `xcmsRaw`.

Value

A `tibble` containing the columns:

- **file:** Filename without path.
- **polarity:** Character string of "positive", "negative", or "unknown".
- **raw:** The xcmsRaw objects.
- **path:** The input path (files).

Index