Package ‘pmd’

August 22, 2019

Type Package

Title Paired Mass Distance Analysis for GC/LC-MS Based Non-Targeted Analysis

Version 0.1.5

Date 2019-08-21

Maintainer Miao YU <yufreecas@gmail.com>

Description Paired mass distance (PMD) analysis proposed in Yu, Olkowski and Pawliszyn (2018) \(\text{<doi:10.1016/j.aca.2018.10.062>}\) for gas/liquid chromatography-mass spectrometry (GC/LC-MS) based non-targeted analysis. PMD analysis including GlobalStd algorithm and structure/reaction directed analysis. GlobalStd algorithm could find independent peaks in m/z-retention time profiles based on retention time hierarchical cluster analysis and frequency analysis of paired mass distances within retention time groups. Structure directed analysis could be used to find potential relationship among those independent peaks in different retention time groups based on frequency of paired mass distances. A GUI for PMD analysis is also included as a 'shiny' application.

URL https://yufree.github.io/pmd

BugReports https://github.com/yufree/pmd/issues

License GPL-2

Encoding UTF-8

LazyData true

Suggests knitr, enviGCMS, igraph

VignetteBuilder knitr

biocViews

Depends R (>= 3.5.0)

Imports RColorBrewer, shiny, rmarkdown, rcdk, stats, utils

RoxygenNote 6.1.1

NeedsCompilation no

Author Miao YU [aut, cre] (<https://orcid.org/0000-0002-2804-6014>)

Repository CRAN

Date/Publication 2019-08-22 15:50:02 UTC
`getchain` Get reaction chain for specific mass to charge ratio

**Description**

Get reaction chain for specific mass to charge ratio

**Usage**

`getchain(list, diff, mass, accuracy = 4, ...)`

**Arguments**

- `list`: a list with mzrt profile
- `diff`: paired mass distance(s) of interests
- `mass`: a specific mass for known compound or a vector of masses
- `accuracy`: measured mass or mass to charge ratio in digits, default 4
- `...`: other parameters for getpmd
Value

a list with mzrt profile and reaction chain dataframe

Examples

data(spmeinvivo)
  # check metabolites of C18H39N0
  pmd <- getchain(spmeinvivo, diff = c(2.02, 14.02, 15.99), mass = 286.3101)

getcluster(list, corcutoff = NULL, accuracy = 4)

Arguments

list a list from getstd function

corcutoff cutoff of the correlation coefficient, default NULL

accuracy measured mass or mass to charge ratio in digits, default 4

Value

list with Pseudo-Spectrum index

See Also

getpaired, getstd, plotstd

Examples

data(spmeinvivo)
  re <- getpaired(spmeinvivo)
  re <- getstd(re)
  cluster <- getcluster(re)
getcorcluster

Get Pseudo-Spectrum as peaks cluster based on correlation analysis.

Description

Get Pseudo-Spectrum as peaks cluster based on correlation analysis.

Usage

getcode_cluster(list, corcutoff = 0.9, rtcutoff = 10, accuracy = 4)

Arguments

- list: a list with peaks intensity
- corcutoff: cutoff of the correlation coefficient, default 0.9
- rtcutoff: cutoff of the distances in cluster, default 10
- accuracy: measured mass or mass to charge ratio in digits, default 4

Value

list with Pseudo-Spectrum index

Examples

```r
data(spmeinvivo)
cluster <- getcorcluster(spmeinvivo)
```

getpaired

Filter ions/peaks based on retention time hierarchical clustering, paired mass distances(PMD) and PMD frequency analysis.

Description

Filter ions/peaks based on retention time hierarchical clustering, paired mass distances(PMD) and PMD frequency analysis.

Usage

getpaired(list, rtcutoff = 10, ng = NULL, digits = 2, accuracy = 4)
getpmd

Arguments

- **list**: a list with mzrt profile
- **rtcutoff**: cutoff of the distances in retention time hierarchical clustering analysis, default 10
- **ng**: cutoff of global PMD’s retention time group numbers, default NULL
- **digits**: mass or mass to charge ratio accuracy for pmd, default 2
- **accuracy**: measured mass or mass to charge ratio in digits, default 4

Value

list with tentative isotope, multi-chargers, adducts, and neutral loss peaks’ index, retention time clusters.

See Also

getstd, getsda, plotpaired

Examples

data(spmeinvivo)
pmd <- getpaired(spmeinvivo)

getpmd(list, pmd, rtcutoff = 10, corcutoff = NULL, digits = 2, accuracy = 4)

Description

Get pmd for specific reaction

Usage

getpmd(list, pmd, rtcutoff = 10, corcutoff = NULL, digits = 2, accuracy = 4)

Arguments

- **list**: a list with mzrt profile
- **pmd**: a specific paired mass distances
- **rtcutoff**: cutoff of the distances in retention time hierarchical clustering analysis, default 10
- **corcutoff**: cutoff of the correlation coefficient, default NULL
- **digits**: mass or mass to charge ratio accuracy for pmd, default 2
- **accuracy**: measured mass or mass to charge ratio in digits, default 4
getrda

Perform structure/reaction directed analysis for mass only.

Usage

getrda(mz, freqcutoff = 10, digits = 3, top = 20, formula = NULL)

Arguments

mz numeric vector for independant mass or mass to charge ratio. Mass to charge ratio from GlobalStd algorithm is suggested. Isomers would be excluded automatically
freqcutoff pmd frequency cutoff for structures or reactions, default 10
digits mass or mass to charge ratio accuracy for pmd, default 3
top top n pmd frequency cutoff when the freqcutoff is too small for large data set
formula vector for formula when you don’t have mass or mass to charge ratio data

Value

logical matrix with row as the same order of mz or formula and column as high frequency pmd group

See Also

getsda

Examples

data(spmeinvivo)
pmd <- getpmd(spmeinvivo,pmd=15.99)
sda <- getrda(spmeinvivo$mz[std$stdmassindex])
getreact

Get quantitative paired peaks list for specific reaction/pmd

Description

Get quantitative paired peaks list for specific reaction/pmd

Usage

getreact(list, pmd, rtcutoff = 10, digits = 2, accuracy = 4, ratiocv = 30, ...)

Arguments

- **list**: a list with mzrt profile and data
- **pmd**: a specific paired mass distances
- **rtcutoff**: cutoff of the distances in retention time hierarchical clustering analysis, default 10
- **digits**: mass or mass to charge ratio accuracy for pmd, default 2
- **accuracy**: measured mass or mass to charge ratio in digits, default 4
- **ratiocv**: ratio cv cutoff for quantitative paired peaks, default 30
- **...**: other parameters for getpmd

Value

list with quantitative paired peaks.

See Also

getpaired, getstd, getsda, getrda, getpmd,

Examples

data(spmeinvivo)
pmd <- getreact(spmeinvivo, pmd=15.99)
getsva

Perform structure/reaction directed analysis for peaks list.

Description

Perform structure/reaction directed analysis for peaks list.

Usage

getsva(list, rtcutoff = 10, freqcutoff = 10, top = 50,
corcutoff = NULL, digits = 2, accuracy = 4)

Arguments

- list: a list with mzrt profile
- rtcutoff: cutoff of the distances in retention time hierarchical clustering analysis, default 10
- freqcutoff: cutoff of frequency of PMDs between RT cluster for peaks, default 10
- top: top n pmd frequency cutoff when the freqcutoff is too small for large data set, default 50
- corcutoff: cutoff of the correlation coefficient, default NULL
- digits: mass or mass to charge ratio accuracy for pmd, default 2
- accuracy: measured mass or mass to charge ratio in digits, default 4

Value

list with tentative isotope, adducts, and neutral loss peaks’ index, retention time clusters.

See Also

getpaired, getstd, plotpaired

Examples

data(spmeinvivo)
pmd <- getpaired(spmeinvivo)
std <- getstd(pmd)
sda <- getsva(std)
getstd

**Description**

Find the independent ions for each retention time hierarchical clustering based on PMD relationship within each retention time cluster and isotope and return the index of the std data for each retention time cluster.

**Usage**

```r
getstd(list, corcutoff = NULL, digits = 2, accuracy = 4)
```

**Arguments**

- `list`: a list from `getpaired` function
- `corcutoff`: cutoff of the correlation coefficient, default NULL
- `digits`: mass or mass to charge ratio accuracy for pmd, default 2
- `accuracy`: measured mass or mass to charge ratio in digits, default 4

**Value**

list with std mass index

**See Also**

`getpaired`, `getsda`, `plotstd`

**Examples**

```r
data(spmeinvivo)
pmd <- getpaired(spmeinvivo)
std <- getstd(pmd)
```
gettarget

Get multiple injections index for selected retention time

Description

Get multiple injections index for selected retention time

Usage

gettarget(rt, drt = 10, n = 6)

Arguments

rt retention time vector for peaks in seconds
drt retention time drift for targeted analysis in seconds, default 10.
n max ions numbers within retention time drift windows

Value

index for each injection

Examples

data(spmeinvivo)
pmd <- getpaired(spmeinvivo)
std <- getstd(pmd)
index <- gettarget(std$rt[std$stdmassindex])
table(index)

globalstd

GlobalStd algorithm with structure/reaction directed analysis

Description

GlobalStd algorithm with structure/reaction directed analysis

Usage

globalstd(list, rtcutoff = 10, ng = 10, corcutoff = NULL,
 freqcutoff = 10, top = 50, digits = 2, accuracy = 4)
Arguments

- list: a peaks list with mass to charge, retention time and intensity data
- rtcutoff: cutoff of the distances in cluster, default 10
- ng: cutoff of global PMD’s retention time group numbers
- corcutoff: cutoff of the correlation coefficient, default NULL
- freqcutoff: cutoff of frequency of PMDs between RT cluster for independent peaks, default 10
- top: top n pmd frequency cutoff when the freqcutoff is too small for large data set, default 50
- digits: mass or mass to charge ratio accuracy for pmd, default 2
- accuracy: measured mass or mass to charge ratio in digits, default 4

Value

list with GlobalStd algorithm processed data.

See Also

getpaired, getstd, getsda, plotstd, plotstdsda, plotstdrt

Examples

data(spmeinvivo)
re <- globalstd(spmeinvivo)

hmdb

A dataframe containing HMDB top 10000 unique accurate mass pmd and related reactions

Description

A dataframe containing HMDB top 10000 unique accurate mass pmd and related reactions

Usage

data(hmdb)

Format

A dataframe with atoms numbers of C, H, O, N, P, S

- percentage: accuracy of atom numbers prediction
- pmd: pmd with two digits
**hmdbp**

A list dataset containing HMDB unique accurate mass pmd analysis results

**Description**

A list dataset containing HMDB unique accurate mass pmd analysis results

**Usage**

```r
data(hmdbp)
```

**Format**

A list with two vectors

- **massp**: all unique hmdb mass probability across all pmds
- **pmdp**: pmds probability across all unique hmdb mass

---

**keggrall**

A dataframe containing reaction related accurate mass pmd and related reaction formula with KEGG ID

**Description**

A dataframe containing reaction related accurate mass pmd and related reaction formula with KEGG ID

**Usage**

```r
data(keggrall)
```

**Format**

A dataframe with KEGG reaction, their related pmd and atoms numbers of C, H, O, N, P, S

- **ID**: KEGG reaction ID
- **pmd**: pmd with three digits
omics

A dataframe containing multiple reaction database ID and their related accurate mass pmd and related reactions

Description

A dataframe containing multiple reaction database ID and their related accurate mass pmd and related reactions

Usage

data(omics)

Format

A dataframe with reaction and their related pmd

KEGG  KEGG reaction ID
RHEA_ID  RHEA_ID
DIRECTION  reaction direction
MASTER_ID  master reaction RHEA ID
ec  ec reaction ID
ecocyc  ecocyc reaction ID
macie  macie reaction ID
metacyc  metacyc reaction ID
reactome  reactome reaction ID
compounds  reaction related compounds
pmd  pmd with two digits
pmd2  pmd with three digits

pcasf

Compare matrices using PCA similarity factor

Description

Compare matrices using PCA similarity factor

Usage

pcasf(x, y, dim = NULL)
Arguments

- **x**: Matrix with sample in column and features in row
- **y**: Matrix is compared to x.
- **dim**: number of retained dimensions in the comparison. Defaults to all.

Value

Ratio of projected variance to total variance

Author(s)

Edgar Zanella Alvarenga

References


Examples

c1 <- matrix(rnorm(16), nrow=4)
c2 <- matrix(rnorm(16), nrow=4)
plotpaired(c1, c2)

Description

Plot the mass pairs and high frequency mass distances

Usage

```
plotpaired(list, index = NULL, …)
```

Arguments

- **list**: a list from getpaired function
- **index**: index for PMD value
- **…**: other parameters for plot function

See Also

globalstd, getpaired
plotrtg

Examples

```r
data(spmeinvivo)
pmd <- getpaired(spmeinvivo)
plotpaired(pmd)
```

plotrtg  
*Plot the retention time group*

Description

Plot the retention time group

Usage

```r
plotrtg(list, ...)
```

Arguments

- `list`: a list from `getpaired` function
- `...`: other parameters for plot function

See Also

`getpaired`, `globalstd`

Examples

```r
data(spmeinvivo)
pmd <- getpaired(spmeinvivo)
plotrtg(pmd)
```

plotnda

*Plot the specific structure directed analysis (SDA) groups*

Description

Plot the specific structure directed analysis (SDA) groups

Usage

```r
plotnda(list, ...)
```

Arguments

- `list`: a list from `getpmd` function
- `...`: other parameters for plot function
See Also

getstd, globalstd, plotstd, plotpaired, plotstdrt

Examples

data(spmeinvivo)
re <- getpmd(spmeinvivo,pmd=78.9)
plotsda(re)

plotstd

plot the std mass from GlobalStd algorithm

Description

Plot the std mass from GlobalStd algorithm

Usage

plotstd(list)

Arguments

list a list from getstd function

See Also

getstd, globalstd

Examples

data(spmeinvivo)
pmd <- getpaired(spmeinvivo)
std <- getstd(pmd)
plotstd(std)
### plotstdrt

**Plot the std mass from GlobalStd algorithm in certain retention time groups**

**Description**

Plot the std mass from GlobalStd algorithm in certain retention time groups

**Usage**

```r
plotstdrt(list, rtcluster, ...)
```

**Arguments**

- `list` a list from getstd function
- `rtcluster` retention time group index
- `...` other parameters for plot function

**See Also**

`getstd`, `globalstd`, `plotstd`, `plotpaired`, `plotstdsda`

**Examples**

```r
data(spmeinvivo)
pmd <- getpaired(spmeinvivo)
std <- getstd(pmd)
plotstdrt(std, rtcluster = 6)
```

### plotstdsda

**Plot the std mass from GlobalStd algorithm in structure directed analysis(SDA) groups**

**Description**

Plot the std mass from GlobalStd algorithm in structure directed analysis(SDA) groups

**Usage**

```r
plotstdsda(list, index = NULL, ...)
```

**Arguments**

- `list` a list from getsda function
- `index` index for PMD value
- `...` other parameters for plot function
See Also
getstd, globalstd, plotstd, plotpaired, plotstdrt

Examples

data(spmeinvivo)
re <- globalstd(spmeinvivo)
plotstdsda(re)

runPMD

Shiny application for PMD analysis

Description
Shiny application for PMD analysis

Usage
runPMD()

sda

A dataset containing common Paired mass distances of substructure, ions replacements, and reaction

Description
A dataset containing common Paired mass distances of substructure, ions replacements, and reaction

Usage
data(sda)

Format
A data frame with 94 rows and 4 variables:

- **PMD**  Paired mass distances
- **origin**  potential sources
- **Ref.**  references
- **mode**  positive, negative or both mode to find corresponding PMDs
Description

A peaks list dataset containing 9 samples from 3 fish with triplicates samples for each fish from LC-MS.

Usage

data(spmeinvivo)

Format

A list with 4 variables from 1459 LC-MS peaks:

- **mz**  mass to charge ratios
- **rt**  retention time
- **data** intensity matrix
- **group** group information
Index

*Topic **datasets**
  - hmdb, 11
  - hmdbp, 12
  - keggrall, 12
  - omics, 13
  - sda, 18
  - spmeinvivo, 19

getchain, 2
getcluster, 3
getcorcluster, 4
getpaired, 3, 4, 6–9, 11, 14, 15
getpmd, 5, 7
getrda, 6, 6, 7
getreact, 7
getsda, 5–7, 8, 9, 11
getstd, 3, 5–8, 9, 11, 16–18
gettarget, 10
globalstd, 10, 14–18

hmdb, 11
hmdbp, 12
keggrall, 12
omics, 13

pcasf, 13
plotpaired, 5, 8, 14, 16–18
plotrtg, 15
plotsda, 15
plotstd, 3, 9, 11, 16, 17, 18
plotstdrt, 11, 16, 17, 18
plotstdsda, 11, 17, 17
runPMD, 18

sda, 18
spmeinvivo, 19