Package 'accucor'

October 12, 2022

Type Package

Title Natural Abundance Correction of Mass Spectrometer Data

Version 0.3.0

Description An isotope natural abundance correction algorithm that is needed especially for high resolution mass spectrometers. Supports correction for 13C, 2H and 15N. Su X, Lu W and Rabinowitz J (2017) <doi:10.1021/acs.analchem.7b00396>.

URL https://github.com/XiaoyangSu/AccuCor

BugReports https://github.com/XiaoyangSu/AccuCor/issues

License MIT + file LICENSE

Encoding UTF-8

Imports nnls, dplyr, stringr, readxl, readr, rlang, tibble, writexl, CHNOSZ

RoxygenNote 7.1.2

Suggests testthat

NeedsCompilation no

Author Xiaoyang Su [aut] (<https://orcid.org/0000-0001-8081-1396>), Lance Parsons [aut, cre] (<https://orcid.org/0000-0002-8521-714X>), Yujue Wang [ctb] (<https://orcid.org/0000-0001-7088-1729>), Princeton University [cph]

Maintainer Lance Parsons <lparsons@princeton.edu>

Repository CRAN

Date/Publication 2021-11-17 11:40:12 UTC

R topics documented:

`accucor` *accucor: A package for natural abundance correction of mass spectrometer data*

Description

AccuCor is an isotope natural abundance correction algorithm that is needed especially for high resolution mass spectrometers. AccuCor supports correction for 13C, 2H and 15N.

AccuCor functions

`natural_abundance_correction`

`carbon_isotope_correction` *Natural Abundance carbon isotope correction for one metabolite*

Description

Natural Abundance carbon isotope correction for one metabolite

Usage

```
carbon_isotope_correction(
  formula,
  datamatrix,
  label,
  Resolution,
  ResDefAt = 200,
  purity = 0.99,
  ReportPoolSize = TRUE
)
```

Arguments

`formula`	String representing molecular formula
`datamatrix`	Matrix of abundances for each sample for each isotope
`label`	vector of integer labels
`Resolution`	For Exactive, the Resolution is 100000, defined at Mw 200
`ResDefAt`	Resolution defined at (in Mw), e.g. 200 Mw
`purity`	Carbon 13 purity, default: 0.99
`ReportPoolSize`	default: TRUE

Value

Named list of matrices: 'Corrected', 'Normalized', 'PoolBeforeDF', and 'PoolAfterDF'.

Examples

```
## Not run:
carbon_isotope_correction(formula = "C6H13O9P",
                          datamatrix = DataMatrix,
                          label = c(0, 1, 2, 3, 4, 5),
                          Resolution = 100000)

## End(Not run)
```

`clean_data_frame` *Standardize data frame columns and data types*

Description

Standardize data frame columns and data types

Usage

```
clean_data_frame(df, columns_to_skip = NULL)
```

Arguments

`df`	Data frame to clean
`columns_to_skip`	Specify column heading to skip. All other columns not named 'compound', 'formula', and 'isotopelabel' will be assumed to be sample names.

Value

"cleaned" data.frame which with columns 'compound', 'formula', 'isotope_label', label_index', followed by columns for each sample

deuterium_isotope_correction *Natural Abundance deuterium isotope correction for one metabolite*

Description

Natural Abundance deuterium isotope correction for one metabolite

Usage

```
deuterium_isotope_correction(
  formula,
  datamatrix,
  label,
  Resolution,
  ResDefAt = 200,
  purity = 0.99,
  ReportPoolSize = TRUE
)
```

Arguments

formula	String representing molecular formula
datamatrix	Matrix of abundances for each sample for each isotope
label	vector of integer labels
Resolution	For Exactive, the Resolution is 100000, defined at Mw 200
ResDefAt	Resolution defined at (in Mw), e.g. 200 Mw
purity	Deuterium purity, default: 0.99
ReportPoolSize	default: TRUE

Value

Named list of matrices: 'Corrected', 'Normalized', 'PoolBeforeDF', and 'PoolAfterDF'.

Examples

```
## Not run:
deuterium_isotope_correction(formula = "C6H13O9P",
                             datamatrix = DataMatrix,
                             label = c(0, 1),
                             Resolution = 100000)

## End(Not run)
```

natural_abundance_correction *Natural Abundance correction for mass spectrometry data*

Description

natural_abundance_correction returns the corrected and normalized intensities of isotopically labeled mass spectrometry data. It was designed to work with input data from El-MAVEN and MAVEN software.

Usage

```
natural_abundance_correction(
  data,
  sheet = NULL,
  compound_database = NULL,
  output_base = NULL,
  output_filetype = "xlsx",
  columns_to_skip = NULL,
  resolution,
  resolution_defined_at = 200,
  purity = NULL,
  report_pool_size_before_df = FALSE,
  path = NULL
)
```

Arguments

data	Path to input data file (xlsx, xls, csv, txt, or tsv) OR dataframe. If dataframe is specified, specify output_base to output files automatically written.
sheet	Name of sheet in xlsx file with columns 'compound', 'formula', 'isotopelabel', and one column per sample. Defaults to the first sheet.
compound_database	Path to compound database in csv format. Only used for classic MAVEN style input when formula is not specified.
output_base	Path to basename of output file, default is the basename of the input path. '_corrected' will be appended. If 'FALSE' then no output file is written.
output_filetype	Filetype of the output file, one of: 'xls', xlsx', 'csv', or 'tsv'. The default is 'xlsx'.
columns_to_skip	Specify column heading to skip. All other columns not named 'compound', 'formula', and 'isotopelabel' will be assumed to be sample names.
resolution	For Exactive, the resolution is 100000, defined at Mw 200
resolution_defined_at	Mw at which the resolution is defined, default 200 Mw

purity — Isotope purity, default: Carbon 0.99; Deuterium 0.98; Nitrogen 0.99

report_pool_size_before_df — Report PoolSizeBeforeDF, default = FALSE

path — Deprecated. Specify path to input data file (alias for 'data').

Details

C13, H2, and N15 isotopes are supported. The isotopes are detected from the `isotopeLabel` column of the input file. The expected label text is `C13-label-#`. `D-label-#`. or `N15-label-#`. Parent (unlabeled) compounds are specified by `C12 PARENT`.

Value

Named list of matrices: 'Corrected', 'Normalized', 'PoolBeforeDF', and 'PoolAfterDF'.

Examples

```
## Not run:
natural_abundance_correction("inst/extdata/C_Sample_Input_Simple.xlsx",
  Resolution = 100000, ResDefAt = 200
)

## End(Not run)
```

---

nitrogen_isotope_correction

Natural Abundance deuterium isotope correction for one metabolite

---

Description

Natural Abundance deuterium isotope correction for one metabolite

Usage

```
nitrogen_isotope_correction(
  formula,
  datamatrix,
  label,
  Resolution,
  ResDefAt = 200,
  purity = 0.99,
  ReportPoolSize = TRUE
)
```

Arguments

formula	String representing molecular formula
datamatrix	Matrix of abundances for each sample for each isotope
label	vector of integer labels
Resolution	For Exactive, the Resolution is 100000, defined at Mw 200
ResDefAt	Resolution defined at (in Mw), e.g. 200 Mw
purity	Nitrogen purity, default: 0.99
ReportPoolSize	default: TRUE

Value

Named list of matrices: 'Corrected', 'Normalized', 'PoolBeforeDF', and 'PoolAfterDF'.

Examples

```
## Not run:
nitrogen_isotope_correction(formula = "C23H38N7O17P3S",
                            datamatrix = DataMatrix,
                            label = c(0, 1, 2, 3, 4, 5, 6, 7),
                            Resolution = 140000)

## End(Not run)
```

---

read_elmaven *Natural Abundance correction for Carbon labeled samples*

Description

Natural Abundance correction for Carbon labeled samples

Usage

```
read_elmaven(
  path,
  sheet = NULL,
  compound_database = NULL,
  columns_to_skip = NULL,
  filetype = NULL,
  ...
)
```

Arguments

path	Path to input file.
sheet	Name of sheet in xlsx file with columns 'compound', 'formula', 'isotopelabel', and one column per sample. Defaults to the first sheet.
compound_database	Path to compound database in csv format. Only used for classic MAVEN style input when formula is not specified.
columns_to_skip	Specify column heading to skip. All other columns not named 'compound', 'formula', and 'isotopelabel' will be assumed to be sample names.
filetype	Specify file type, default is to determine by file extension.
...	Pass additional parameters to readxl::read_excel

Value

List containing three items: "original" data.frame which is result of read_excel, "cleaned" data.frame which with columns 'compound', 'formula', 'isotope_label', label_index', followed by columns for each sample, and "isotope" which is a character indicating the isotope

Examples

```
## Not run:
read_elmaven_xlsx("ExcelFile", "Sheet1")

## End(Not run)
```

Index