Package 'IsotopicLabelling'

March 18, 2016

Type Package

Title A package for the analysis of MS isotopic patterns in labelling experiments

Version 1.0

Date 2016-01-26

Author Ruggero Ferrazza, Pietro Franceschi

Maintainer Ruggero Ferrazza <r.ferrazza@unitn.it>

Description The IsotopicLabelling package allows to analyse the isotopic patterns in MS data obtained following isotopic labelling experiments; by looking at the isotopic patterns, the package estimates the isotopic abundance of the stable isotope employed in the labelling experiment (either 2H or 13C) within the specified compound.

Depends xcms, ecipex, stringr, gsubfn

License GPL-2

Suggests knitr, rmarkdown

VignetteBuilder knitr

NeedsCompilation no

R topics documented:

`IsotopicLabelling-package` *A package for the analysis of MS isotopic patterns in labelling experiments*

Description

The `IsotopicLabelling` package allows to analyse the isotopic patterns in MS data obtained following isotopic labelling experiments; by looking at the isotopic patterns, the package estimates the isotopic abundance of the stable isotope employed in the labelling experiment (either ^{2}H or ^{13}C) within the specified compound.

Details

Given a data frame of LC-MS or GC-MS peak intensities or areas (one column for each sample to analyse), the `IsotopicLabelling` package first extracts the isotopic patterns of the specified compound, and then performs an isotopic pattern analysis to estimate the isotopic abundance of the labelling isotope. This is obtained through a fitting procedure by which the abundance is iteratively changed and the resulting expected patterns are compared to the experimental one; the final estimated isotopic abundance is the value for which the theoretical pattern best reproduces the experimental one. The theoretical patterns are computed using the `ecipex` R package.

The isotopic pattern analysis can be divided into the following steps:

1. Starting from a class `xcmsSet` object (from the `xcms` R package), generate a data frame of peak signal intensities or areas, with each column corresponding to a sample. This step can be avoided if the data frame is already available (obtained by other means);
2. Extract from the data frame the experimental isotopic patterns of the specified compound (one pattern for each sample). In the chemical formula of the compound, the element whose abundance is unknown is called "X";
3. Normalise the patterns and find the theoretical patterns that best fit them, by iteratively changing the isotopic abundance of the labelling isotope;
4. Summarize the results.

Author(s)

Ruggero Ferrazza, Pietro Franceschi

Maintainer: Ruggero Ferrazza <r.ferrazza@unitn.it>

Examples

```
data(xcms_obj)
peak_table <- table_xcms(xcms_obj)
fitted_abundances <- main_labelling(peak_table, compound="X40H77NO8P", labelling="C",
mass_shift=0.05, RT=285, RT_shift=20, chrom_width=7, initial_abundance=NA)
summary(fitted_abundances)
plot(fitted_abundances, type="patterns", saveplots=FALSE)
plot(fitted_abundances, type="residuals", saveplots=FALSE)
plot(fitted_abundances, type="summary", saveplots=FALSE)
save_labelling(fitted_abundances)

# Other possible lipid compounds include:
```

```
# [PC34:1 + H]+. compound="X42H83NO8P", RT=475, chrom_width=10
# [TAG47:3 + NH4]+ (a minor species). compound="X50H94NO6", RT=891, chrom_width=7
```

`find_abundance` *Fit experimental isotopic patterns*

Description

Function that takes each of the provided experimental MS isotopic patterns, and fits the best theoretical pattern that reproduces it by iteratively changing the isotopic abundance of the labelling isotope X.

Usage

```
find_abundance(patterns, info, initial_abundance)
```

Arguments

`patterns`	A matrix of experimental isotopic patterns (one column for each sample), with the first two columns representing *m/z* and retention time of the corresponding peaks.
`info`	Named list containing isotopic information, output of the `isotopic_information` function.
`initial_abundance`	Either NA, or a numeric vector of length equal to the number of samples, with the initial guesses on the percentage isotopic abundance of the labelling isotope (denoted as X, it can be either ^{2}H or ^{13}C). If provided, numbers between 0 and 100.

Value

An object of class `labelling`, which is a list containing the results of the fitting procedure:

`compound`	Character vector specifying the chemical formula of the compound of interest, with X being the element with unknown isotopic distribution (to be fitted).
`best_estimate`	Numeric vector of length equal to the number of samples, containing the estimated percentage abundances of the labelling isotope X (either ^{2}H or ^{13}C). Numbers between 0 and 100.
`std_error`	Numeric vector with the standard errors of the estimates, output of the `nls` fitting procedure.
`dev_percent`	Numeric vector with the percentage deviations between best fitted and related experimental patterns.
`x_scale`	Numeric vector containing the m/z values relative to the signals of the experimental patterns.
`y_exp`	Matrix of normalised experimental isotopic patterns (one column for each sample). The most intense signal of each pattern is set to 100.
`y_theor`	Matrix of normalised fitted theoretical isotopic patterns (one column for each sample). The most intense signal of each pattern is set to 100.
`residuals`	Matrix of residuals: each column is the difference between experimental and best fitted theoretical patterns.
`warnings`	Character vector with possible warnings from the `nls` fitting procedure.

See Also

isotopic_information

`isotopic_information` *Get useful isotopic information*

Description

This function gathers essential isotopic information required by the other functions of the `IsotopicLabelling` package.

Usage

```
isotopic_information(compound, labelling)
```

Arguments

`compound`	Character vector specifying the chemical formula of the compound of interest, with X being the element with unknown isotopic distribution (to be fitted).
`labelling`	Character, either "H" or "C", specifying the labelling element.

Details

The specified compound is not the neutral molecular species of interest, but the adduct observed by ESI-MS (such as protonated or sodiated species).

In the chemical formula, the element with unknown abundance should be denoted by X. For example, the proton adduct of TAG 52:2, C55H103O6, should be written X55H103O6 for ^{13}C labelling experiments, and C55X102HO6 for ^{2}H labelling experiments. Note that in this last case only 102 hydrogen atoms have unknown isotopic distribution, since the one giving rise to the adduct comes from the solvent, and is considered to have fixed natural abundance.

Value

A named list containing the following information:

`compound`	The same as input.
`isotopes`	Table containing the natural isotopic abundances of the elements present in compound (numbers between 0 and 1). The two isotopes of element X are given NA values.
`target`	Named vector with the exact masses of all the possible isotopologues arising from the labelling isotope. M+0 is the monoisotopic mass (sum of the masses of the atoms using the lightest isotope for each element, X included); in M+1 one light isotope is replaced by its heaviest counterpart, and so forth.
`nX`	The number of X atoms. In other words, the number of atoms with unknown isotopic distribution.
`nTOT`	The total number of atoms of the labelling element (either H+X or C+X).

isotopic_pattern *Extract experimental isotopic patterns from a table of MS peaks*

Description

Function that extracts the experimental isotopic patterns of a specified compound from a data frame containing MS peak intensities or areas.

Usage

```
isotopic_pattern(peak_table, info, mass_shift, RT, RT_shift, chrom_width)
```

Arguments

peak_table	Data frame of experimental MS peak intensities or areas (one column for each sample), with the first two columns representing *m/z* and retention times of the peaks. If not already available, the table can be obtained from an xcmsSet object (output of the xcms R package) through the table_xcms function.
info	Named list containing isotopic information, output of the isotopic_information function.
mass_shift	Maximum difference between theoretical and experimental mass. In other words, the mass accuracy.
RT	Expected retention time of the compound of interest.
RT_shift	Maximum difference between expected and experimental retention time of the peaks.
chrom_width	An estimate of the chromatographic peak width.

Value

A matrix of extracted experimental isotopic patterns (one column for each sample), with the first two columns representing the exact *m/z* and the retention times of the peaks.

See Also

table_xcms , isotopic_information

labelling-class *Class labelling*

Description

A class labelling object is a list containing the results of the isotopic pattern analysis. Such object is the output of the functions find_abundance and main_labelling.

Details

The methods available for class "labelling" objects are:

summary to produce a summary of the results contained in the `labelling` object;

plot to create graphical outputs.

See Also

`find_abundance`, `main_labelling`, `summary`, `plot`

`main_labelling` *Estimate isotopic abundances from experimental MS isotopic patterns*

Description

The main function of the `IsotopicLabelling` package, it estimates the isotopic abundance of the labelling isotope in the specified compound, given a data frame containing MS peak intensities or areas.

Usage

```
main_labelling(peak_table, compound, labelling, mass_shift,

 RT, RT_shift, chrom_width, initial_abundance)
```

Arguments

`peak_table`	Data frame of experimental MS peak intensities or areas (one column for each sample), with the first two columns representing *m/z* and retention times of the peaks. If not already available, the table can be obtained from an `xcmsSet` object (output of the `xcms` R package) through the `table_xcms` function.
`compound`	Character vector specifying the chemical formula of the compound of interest, with X being the element with unknown isotopic distribution (to be fitted).
`labelling`	Character, either "H" or "C", specifying the labelling element.
`mass_shift`	Maximum difference between theoretical and experimental mass. In other words, the mass accuracy.
`RT`	Expected retention time of the compound of interest.
`RT_shift`	Maximum difference between expected and experimental retention time of the peaks.
`chrom_width`	An estimate of the chromatographic peak width.
`initial_abundance`	Either NA, or a numeric vector of length equal to the number of samples, with the initial guesses on the percentage isotopic abundance of the labelling isotope (denoted as X, it can be either ^{2}H or ^{13}C). If provided, numbers between 0 and 100.

Value

An object of class labelling, which is a list containing the results of the fitting procedure:

compound	Character vector specifying the chemical formula of the compound of interest, with X being the element with unknown isotopic distribution (to be fitted).
best_estimate	Numeric vector of length equal to the number of samples, containing the estimated percentage abundances of the labelling isotope X (either ^{2}H or ^{13}C). Numbers between 0 and 100.
std_error	Numeric vector with the standard errors of the estimates, output of the nls fitting procedure.
dev_percent	Numeric vector with the percentage deviations between best fitted and related experimental patterns.
x_scale	Numeric vector containing the m/z values relative to the signals of the experimental patterns.
y_exp	Matrix of normalised experimental isotopic patterns (one column for each sample). The most intense signal of each pattern is set to 100.
y_theor	Matrix of normalised fitted theoretical isotopic patterns (one column for each sample). The most intense signal of each pattern is set to 100.
residuals	Matrix of residuals: each column is the difference between experimental and best fitted theoretical patterns.
warnings	Character vector with possible warnings from the nls fitting procedure.

Note

This function is made up of three important functions of the IsotopicLabelling package which are, in order: isotopic_information, isotopic_pattern and find_abundance. The user can therefore choose whether to carry out all the processing in a single step using this compact function, or to separately perform the three successive steps without using main_labelling.

See Also

table_xcms

Examples

```
data(xcms_obj)
peak_table <- table_xcms(xcms_obj)
fitted_abundances <- main_labelling(peak_table, compound="X40H77NO8P", labelling="C",
mass_shift=0.05, RT=285, RT_shift=20, chrom_width=7, initial_abundance=NA)
```

---

pattern_from_abundance

Compute theoretical isotopic patterns

Description

Function that computes the theoretical pattern of the specified compound, given the abundance of the labelling isotope X. The function returns a vector of normalised theoretical intensities, where the maximum signal is set to 100; the related masses are in the "target" vector contained in the input list.

Usage

```
pattern_from_abundance(abundance, info)
```

Arguments

abundance	Isotopic abundance of the labelling isotope X (either ^{2}H or ^{13}C); number between 0 and 1.
info	Named list containing isotopic information, output of the `isotopic_information` function.

Value

A vector representing the normalised isotopic pattern of the compound of interest, corresponding to the specified isotopic distribution.

References

The function makes use of the `ecipex` R package.

See Also

`isotopic_information`, `ecipex`

---

plot.labelling *Plot method for class "labelling" objects*

Description

It is a method for the plot generic. Depending on the parameter "type", three operating conditions are possible: by default (type="patterns") this function plots, for each sample in the class `labelling` object, its normalised experimental pattern superimposed to its fitted theoretical pattern.

By setting type="residuals", the function plots the residuals (the differences between experimental and best fitted theoretical patterns).

Finally, with type="summary" a summary plot is produced, showing the estimated percentage abundances with related standard errors.

If the parameter "saveplots" is set to TRUE, the plots are saved as a *.pdf file in the working directory.

Usage

```
plot(fitted_abundances, type, saveplots)
```

Arguments

fitted_abundances	Object of class `labelling`.
type	The type of plots to produce: the default value is "patterns", but other two possible values are "residuals" and "summary" (see description above).
saveplots	A logical, with default value FALSE. If TRUE, the plots are saved to a single *.pdf file in the working directory.

Value

One or more plots.

`save_labelling` *Save class "labelling" objects to *.csv files*

Description

This function allows to save the results of the isotopic pattern analysis to a *.csv file.

Usage

```
save_labelling(fitted_abundances, path)
```

Arguments

`fitted_abundances`	Object of class `labelling`. It contains the results of the isotopic pattern analysis.
`path`	The directory where to save the *.csv file. If not specified, it is saved in the working directory.

Value

The "COMPOUND_Estimated_Abundances.csv" file, containing the results of the isotopic pattern analysis. For each sample (one for each row) there are four columns:

1. The estimated percentage abundance of the labelling isotope (either ^{2}H or ^{13}C);
2. The related standard error coming from the `nls` fitting procedure;
3. The percentage deviation between theoretical and experimental isotopic patterns;
4. The outcome message from the fitting procedure, to undersand whether there have been any convergence problems.

See Also

`main_labelling`

summary.labelling *Summarize the results contained in a class "labelling" object*

Description

This function summarizes the results contained in an object of class `labelling`, by showing the estimated percentage isotopic abundances of the labelling isotope, together with the related standard errors.

Usage

```
summary(fitted_abundances)
```

Arguments

`fitted_abundances`
: An object of class `labelling`, output of either `main_labelling` or `find_abundance` functions.

Value

A matrix summarizing the results contained in the class `labelling` object. It has two rows: in the first one are the estimated percentage isotopic abundances of the labelling isotope X (^{2}H or ^{13}C), in the second one are the related standard errors.

table_xcms *Create a data frame of peak intensities or areas from an xcmsSet object*

Description

The function generates a data frame of MS peak intensities or areas starting from an object of class `xcmsSet` (from the `xcms` R package).

Usage

```
table_xcms(xcms_obj)
```

Arguments

`xcms_obj`
: An object of class `xcmsSet`, output of the MS data processing with the `xcms` R package.

Value

A data frame of experimental MS peak intensities or areas (one column for each sample), with its first two columns containing the *m/z* and the retention times of the peaks.

Note

The output data frame, required by other functions of the IsotopicLabelling R package, can be obtained in a number of other independent ways, such as through proprietary software of the vendor of the MS instrument.

See Also

The xcms R package.

Examples

```
data(xcms_obj)
peak_table <- table_xcms(xcms_obj)
```

---

xcms_obj *Example data set, a class "xcmsSet" object*

Description

This data set is an example of LC-MS experimental data obtained following isotopic labelling experiments. The data was kindly provided by Dr. Jules Griffin and Dr. Nyasha Munjoma (Department of Biochemistry, University of Cambridge - UK).

It contains LC-MS data relative to ^{13}C labelling experiments on 8 samples: the first 4 represent lipid extracts from cell cultures grown under normal conditions (natural ^{13}C abundance), whereas in the last 4 samples the lipids were extracted from cells grown using uniformly-labelled ^{13}C glucose (99% ^{13}C labelling).

Usage

```
data("xcms_obj")
```

Index