Package ‘rgoslin’

May 4, 2024

**Type**  Package
**Title**  Lipid Shorthand Name Parsing and Normalization
**Version**  1.8.0
**Date**  2024-04-12

**Description**  The R implementation for the Grammar of Succint Lipid Nomenclature parses different short hand notation dialects for lipid names. It normalizes them to a standard name. It further provides calculated monoisotopic masses and sum formulas for each successfully parsed lipid name and supplements it with LIPID MAPS Category and Class information. Also, the structural level and further structural details about the head group, fatty acyls and functional groups are returned, where applicable.

**License**  MIT + file LICENSE
**Imports**  Rcpp (&gt;= 1.0.3), dplyr
**LinkingTo**  Rcpp
**Suggests**  testthat (&gt;= 2.1.0), BiocStyle, knitr, rmarkdown, kableExtra, BiocManager, stringr, stringi, ggplot2, tibble, lipidr

**RoxygenNote**  7.1.2
**Encoding**  UTF-8
**VignetteBuilder**  knitr

**biocViews**  Software, Lipidomics, Metabolomics, Preprocessing, Normalization, MassSpectrometry

**BugReports**  https://github.com/lifs-tools/rgoslin/issues
**URL**  https://github.com/lifs-tools/rgoslin
**git_url**  https://git.bioconductor.org/packages/rgoslin
**git_branch**  RELEASE_3_19
**git_last_commit**  4152af5
**git_last_commit_date**  2024-04-30
**Repository**  Bioconductor 3.19
**Date/Publication**  2024-05-03
rgoslin-package

Author Nils Hoffmann [aut, cre] (<https://orcid.org/0000-0002-6540-6875>), Dominik Kopczynski [aut] (<https://orcid.org/0000-0001-5885-4568>)

Maintainer Nils Hoffmann <nils.hoffmann@cebitec.uni-bielefeld.de>

Contents

  rgoslin-package .............................................. 2
  isValidLipidName ........................................... 3
  listAvailableGrammars ..................................... 3
  parseLipidNames ............................................ 4

Index 5

rgoslin-package rgoslin: Lipid Shorthand Name Parsing and Normalization

Description

The R implementation for the Grammar of Succint Lipid Nomenclature parses different shorthand notation dialects for lipid names. It normalizes them to a standard name. It further provides calculated monoisotopic masses and sum formulas for each successfully parsed lipid name and supplements it with LIPID MAPS Category and Class information. Also, the structural level and further structural details about the head group, fatty acyls and functional groups are returned, where applicable.

Details

rgoslin is the R implementation of the "grammar of succint lipid nomenclature". It provides implementations for parsing of shorthand lipid names from LipidMAPS, SwissLipids and other common sources of lipid names.

Author(s)

Maintainer: Nils Hoffmann <nils.hoffmann@cebitec.uni-bielefeld.de> (ORCID)

Authors:

• Dominik Kopczynski <dominik.kopczynski@univie.ac.at> (ORCID)

References

NA

See Also

Useful links:

• https://github.com/lifs-tools/rgoslin

• Report bugs at https://github.com/lifs-tools/rgoslin/issues
isValiLipidName  

Description

isValiLipidName checks the provided lipid name against the built-in grammars. Will return FALSE if none of the parsers was able to parse the provided name successfully. Will stop execution via stop if non character input is detected.

Usage

isValidLipidName(lipidName)

Arguments

lipidName The lipid name to check.

Value

TRUE if the lipidName could be parsed, FALSE otherwise.

Examples

isValidLipidName("PC 32:1")
isValidLipidName("PC(32:1)")
isValidLipidName("PCX(32:1)")

listAvailableGrammars  Return the list of grammars supported by goslin.

Description

listAvailableGrammars returns the list of grammars that the underlying cppgoslin library supports.

Usage

listAvailableGrammars()

Value

the list of grammars

Examples

listAvailableGrammars()
parseLipidNames

Parse multiple lipid names and return a data frame with the results.

Description

parseLipidNames reads the provided lipid names vector and returns structural information as a data frame. Will return a cell with the "Grammar" column set to "NOT_PARSEABLE" if none of the parsers was able to parse the provided name successfully. Will stop execution via `stop` if invalid non character input is detected or fatal errors are encountered during parsing.

Usage

```r
parseLipidNames(lipidNames, grammar = NULL)
```

Arguments

- `lipidNames`: The vector of lipid names to parse.
- `grammar`: The grammar to use. One of "Goslin", "GoslinFragments", "SwissLipids", "LipidMaps", "HMDB", "FattyAcids". Call `listAvailableGrammars()` for a complete list of available grammars. If `grammar` is omitted or `NULL` is passed as a parameter, all available grammars / parsers will be tested. The first successful one will win. If all parsers fail, the "Messages" column in the returned data frame will contain the last parsers message.

Value

Data frame where each row reports the parsing result of each element in lipidNames.

Examples

```r
parseLipidNames(c("PC 32:1", "LPC 34:1", "TG(18:1_18:0_16:1)"))
parseLipidNames(c("Cer(d18:1(8Z)/24:0)", grammar = "LipidMaps"))
```
Index

* internal
  rgoslin-package, 2

isValidLipidName, 3
listAvailableGrammars, 3
parseLipidNames, 4
rgoslin (rgoslin-package), 2
rgoslin-package, 2