Package: nontargetData (via r-universe)

November 4, 2024

Type Package

Title Quantized simulation data of isotope pattern centroids

Version 1.1

Date 2014-07-22

Author Martin Loos, Francesco Corona

Maintainer Martin Loos <Martin.Loos@eawag.ch>

Description Data sets for isotope pattern grouping of LC-HRMS peaks
with package nontarget. Based on a vast set of unique PubChem
molecular formulas, quantized (a) m/z, (b) m/z differences, (c)
intensity ratios and (d) marker centroids of simulated centroid
pairs are listed for different instrument resolutions.

License GPL-3

Repository https://rickhelmus.r-universe.dev

RemoteUrl https://github.com/blosloos/nontargetData

RemoteRef HEAD

RemoteSha 75b2a92cd549f72cc71dfe7b86bd42947d62f2c4

Contents

---

nontargetData-package *Quantized simulation data of isotope pattern centroids*

Description

Data sets for isotope pattern grouping of LC-HRMS peaks with package nontarget. Based on a vast
set of PubChem molecular formulas, quantized (a) m/z, (b) m/z differences, (c) intensity ratios and
(d) marker centroids of simulated centroid pairs are listed for different instrument resolutions.

Details

Package:	nontargetData
Type:	Package
Version:	1.0
Date:	2014-07-22
License:	GPL-3

Author(s)

Martin Loos, Francesco Corona Maintainer: Martin Loos <Martin.Loos@eawag.ch>

OrbitrapXL_VelosPro_R60000at400_q

Quantized simulation data of isotope pattern centroids

Description

Data sets for isotope pattern grouping of LC-HRMS peaks with package nontarget. Based on a vast set of PubChem molecular formulas, quantized (a) m/z, (b) m/z differences, (c) intensity ratios and (d) marker centroids of simulated centroid pairs are listed for different instrument resolutions.

Usage

```
data(OrbitrapXL_VelosPro_R60000at400_q)
```

Format

List of 8

$ R_set : MS Instrument with a specific resolution=f(m/z) setting

$ R_param : Ignore, for debugging

$ isotope_key : Isotope categories

$ charge_key : Charge (z) categories

$ Quant_param : Quantization parameters

$ Quant : Quantized data in categories of dominant isotope, charge and presence/absence of a marker mass

$ mass_slots : m/z windows used for filtering

$ intrat_slots: m/z windows used for filtering

Details

The isotope pattern centroids of several million organic molecular formulas from the PubChem database were calculated for various classes of adducts. Molecular formulas were filtered to be unique and only to contain C, H, O, N, Cl, Br, K, Na, S, Si, F, P and/or I. The resulting >250 million centroid pairs from individual patterns were then categorized for their dominant isotopologues, charge and the possible presence of another centroid of higher intensity than that of the pair (marker peak). Within these categories, data on centroid pair (a) m/z, (b) m/z differences, (c) intensity ratios and (d) marker m/z was quantized by a recursive partitioning procedure. The resulting compressed data representation was extended by nearest neigbour estimates in the above dimensions (a) to (d) to account for queries with molecular formulas possibly not represented by the PubChem set. Query access to the data is provided through a tree-like space-partitioning structure for hyperrectangles.

Source

http://pubchem.ncbi.nlm.nih.gov/

Index