

# Package ‘RforProteomics’

July 4, 2024

**Type** Package

**Title** Companion package to the 'Using R and Bioconductor for proteomics data analysis' publication

**Version** 1.43.0

**Depends** R (>= 3.5), MSnbase (>= 2.5.3)

**Imports** R.utils, biocViews, BiocManager

**Suggests** AnnotationDbi, rpx (>= 2.0.3), DT, knitr, rmarkdown, BiocStyle, mzR, xcms, msdata, MALDIquant (>= 1.12), MALDIquantForeign, readBrukerFlexData, Rdisop, OrgMassSpecR, SummarizedExperiment, BRAIN, rols, hpar, GO.db, org.Hs.eg.db, e1071, biomaRt, RColorBrewer, ggplot2, reshape2, xtable, lattice, mzID, pRoloc, pRolocdata, MSnID, msmsTests, msmsEDA, DEP, corrplot, beanplot, Heatplus, gplots, VennDiagram, protViz, genefilter, plotly, gridExtra, dplyr, lubridate, magick, cleaver

**Description** This package contains code to illustrate the 'Using R and Bioconductor for proteomics data analysis' and 'Visualisation of proteomics data using R and Bioconductor' manuscripts. The vignettes describe the code and data needed to reproduce the examples and figures described in the paper and functionality for proteomics visualisation. It also contain various function to discover R software for mass spectrometry and proteomics.

**URL** <http://lgatto.github.com/RforProteomics/>

**biocViews** ExperimentData, MassSpectrometryData, ReproducibleResearch

**License** Artistic-2.0

**VignetteBuilder** knitr

**RoxygenNote** 7.1.1

**git\_url** <https://git.bioconductor.org/packages/RforProteomics>

**git\_branch** devel

**git\_last\_commit** e219147

**git\_last\_commit\_date** 2024-04-30  
**Repository** Bioconductor 3.20  
**Date/Publication** 2024-07-04  
**Author** Laurent Gatto [aut, cre],  
          Sebastian Gibb [ctb],  
          Vlad Petyuk [ctb],  
          Thomas Pedersen Lin [ctb]  
**Maintainer** Laurent Gatto <laurent.gatto@uclouvain.be>

Contents

downloadData . . . . .	2
getPackagesInBiocView . . . . .	3
getPXD000001mzData . . . . .	4
getPXD000001mzTab . . . . .	4
getPXD000001mzXML . . . . .	5
getThermoHelaPRTC . . . . .	5
id . . . . .	6
new_ions . . . . .	7
packageDF . . . . .	8
proteomicsPackages . . . . .	8
qnt . . . . .	9
RforProteomics . . . . .	10
<b>Index</b>	<b>12</b>

---

downloadData	<i>Download a file</i>
--------------	------------------------

---

Description

Unless already present, downloads src in the destdir directory.

Usage

downloadData(src, destdir = ".", unpack = TRUE, ...)

Arguments

- |         |  |
|---------|--|
| src     | The url of the file to download.                               |
| destdir | The destination directory. Default is ".".                     |
| unpack  | Should src be uncompressed? Default is TRUE.                   |
| ...     | Additional paramters passed to <a href="#">download.file</a> . |

**Value**

Invisible returns the full path of the downloaded file.

**Author(s)**

Laurent Gatto

---

*getPackagesInBiocView Packages in a biocView*

---

**Description**

Finds the package names that have a specific biocView.

**Usage**

```
getPackagesInBiocView(  
  view,  
  rep = c("BioCsoft", "BioCann", "BioCexp", "BioCextra"),  
  biocVersion  
)
```

**Arguments**

<code>view</code>	The biocView of interest. For example "Proteomics".
<code>rep</code>	Repository of interest. One of "BioCsoft", "BioCann", "BioCexp" or "BioCextra".
<code>biocVersion</code>	A character with the Bioconductor version of interest. For example "2.14".

**Value**

An instance of class BiocView. NULL if the biocView was not found.

**Author(s)**

Laurent Gatto

getPXD000001mzData      *Download the PXD000001 mzTab file*

---

**Description**

Unless already present, downloads the PXD000001 mzData file in the destdir directory. The resulting file is named PRIDE\_Exp\_Complete\_Ac\_22134.xml

**Usage**

```
getPXD000001mzData(destdir = ".")
```

**Arguments**

destdir                  A character with the destination folder.

**Value**

Invisibly returns the name of the downloaded file.

**Author(s)**

Laurent Gatto

---

getPXD000001mzTab      *Download the PXD000001 mzTab file*

---

**Description**

Unless already present, downloads the PXD000001 mzTab file in the destdir directory. The resulting file is named F063721.dat-mztab.txt.

**Usage**

```
getPXD000001mzTab(destdir = ".")
```

**Arguments**

destdir                  A character with the destination folder.

**Value**

Invisibly returns the name of the downloaded file.

**Author(s)**

Laurent Gatto

---

getPXD000001mzXML	<i>Download the PXD000001 mzXML file</i>
-------------------	--

---

**Description**

Unless already present, downloads the PXD000001 mzXML file in the destdir directory. The resulting file is named TMT\_Erwinia\_1uLSike\_Top10HCD\_isol2\_45stepped\_60min\_01.mzXML.

**Usage**

```
getPXD000001mzXML(destdir = ".")
```

**Arguments**

destdir	A character with the destination folder.
---------	--

**Value**

Invisibly returns the name of the downloaded file.

**Author(s)**

Laurent Gatto

---

getThermoHelaPRTC	<i>Download Thermo Hela PRTC data</i>
-------------------	---------------------------------------

---

**Description**

Downloads on of multiple Thermo Hela/PRTC data files.

**Usage**

```
getThermoHelaPRTC(src, destdir = ".")
```

**Arguments**

src	The name of the file to be downloaded. If missing, a vector of possible filenames is returned. If "all", all files are downloaded. Alternatively, a pattern can be used to grep the files from the output getThermoHelaPRTC() the files to be downloaded.
destdir	Destination directory. Default is ".".

**Value**

Invisibly return the path of the downloaded files.

**Author(s)**

Laurent Gatto

**See Also**

downloadData

**Examples**

```
## Not run:
getThermoHelaPRTC()
getThermoHelaPRTC("design")
getThermoHelaPRTC("all")

## End(Not run)
```

---

id	<i>An mzIdentML file</i>
----	--------------------------

---

**Description**

This file has been generated by searching the raw mzXML file of the ProteomeXchange PXD000001 data set against the erwinia\_carotovora.fasta using the MSGF+ search engine: `java -jar ~/bin/MSGFPlus.20140630/MSGFPlus.jar -s TMT_Erwinia_1uLSike_Top10HCD_isol2_45stepped_60min_01.mzXML -d erwinia_carotovora.fasta -inst 1 -protocol 2`

**Examples**

```
## source files to repeat the search
library("rpx")
px <- PXDataset("PXD000001")
pxfiles(px)

f <- dir(system.file("extdata", package = "RforProteomics"),
        pattern = "mzid", full.names=TRUE)
library("mzID")
id <- mzID(f)
id
```

---

new_ions	Create, analyse and detect ions
----------	---------------------------------

---

## Description

This is the constructor function to generate a set of ions that can later be analysed with ‘analyse()’ and detected with ‘detect()’.

## Usage

```
new_ions(npeaks = 10, mzrange = c(100, 1000), nimg = 100)

analyse(x, sleep = 0.1)

analyze(x, sleep = 0.1)

detect(x, new = FALSE)

spectrum(x, ...)
```

## Arguments

npeaks	A ‘numeric’ scalar defining the number of unique peaks (M/Z values). Default is 10.
mzrange	A ‘numeric’ of length 2 defining the range of possible M/Z values. Default is ‘c(100, 1000)’.
nimg	A ‘numeric’ scalar. When analysing the ions, their separation along their M/Z values will be split along a sequence of length ‘nimg’. Default is 100.
x	An object of class ‘ions’.
sleep	How much time to wait before producing the next plot.
new	A ‘logical’ scalar, indicating if the separated ions (last frame of calling ‘analyse’) should be plotting, or whether the detection should be overlaid. Default is ‘FALSE’, to add the plot on top of the opened device.
...	Additional arguments passed to [graphics::plot()].

## Value

An object of class ‘ions’.

‘analyse’, ‘detect’ and ‘spectrum’ are used for their side effect or producing plots. They all invisibly return ‘NULL’.

## Author(s)

Laurent Gatto

**Examples**

```
set.seed(1L)
x <- new_ions(nimg = 5)
x
analyse(x)
detect(x)
spectrum(x)
```

---

packageDF	<i>Package descriptions</i>
-----------	-----------------------------

---

**Description**

Format a BiocView as a data.frame.

**Usage**

```
packageDF(x, nsub = TRUE, version = TRUE)
```

**Arguments**

x	An instance of class BiocView, as produced by getPackagesInBiocView.
nsub	A logical indicating "\n" are to be replaced by a space.
version	A logical specifying if the package version should be added.

**Value**

A data.frame with package information.

**Author(s)**

Laurent Gatto

---

proteomicsPackages	<i>Proteomics and MS biocView packages</i>
--------------------	--

---

**Description**

Searches for all the packages with the "Proteomics" (software), "MassSpectrometry" (software) and "MassSpectrometryData" (data) packages and return their names, titles and versions as a data.frame. The (unexported but documented) underlying functions are RforProteomics::getPackagesInBiocView (to find relevant package) and RforProteomics::packageDF (data.frame formatting).



**Usage**

```
proteomicsPackages(biocv, cache = FALSE)
```

```
massSpectrometryPackages(biocv, cache = FALSE)
```

```
massSpectrometryDataPackages(biocv, cache = FALSE)
```

**Arguments**

**biocv** A character with the Bioconductor version to search for relevant packages. If missing, the running version is used.

**cache** A logical indicating whether cached package information should be used. Default is FALSE. All except development versions are up-to-date.

**Value**

A data.frame with the respective package names, titles and versions.

**Author(s)**

Laurent Gatto

**Examples**

```
head(pp <- proteomicsPackages("3.0"))
ppc <- proteomicsPackages("3.0", cache = TRUE)
all.equal(pp, ppc)
```

---

qnt

---

*PXD000001 example MSnSet*


---

**Description**

In this TMT 6-plex experiment, four exogenous proteins were spiked into an equimolar *Erwinia carotovora* lysate with varying proportions in each channel of quantitation; yeast enolase (ENO) at 10:5:2.5:1:2.5:10, bovine serum albumin (BSA) at 1:2.5:5:10:5:1, rabbit glycogen phosphorylase (PHO) at 2:2:2:2:1:1 and bovin cytochrome C (CYT) at 1:1:1:1:1:2. Proteins were then digested, differentially labelled with TMT reagents, fractionated by reverse phase nanoflow UPLC (nanoACQUITY, Waters), and analysed on an LTQ Orbitrap Velos mass spectrometer (Thermo Scientific). Files in multiple format will be used to illustrate the input/output capabilities that are available to the proteomics audience. The companion package provides dedicated functions to directly download the data.

The data has been downloaded from the ProteomeXchange repository and imported into R as illustrated in the example. It is of class `MSnSet`. See also the `MSnbase`-demo vignette for more details.

**Usage**

```
data("qnt")
```

## Format

An instance of class MSnSet

## References

Laurent Gatto (2014). RforProteomics: Companion package to the 'Using R and Bioconductor for proteomics data analysis' publication R package version 1.3.1.

Gatto L, Christoforou A. Using R and Bioconductor for proteomics data analysis. Biochim Biophys Acta. 2013 May 18. doi:pii: S1570-9639(13)00186-6. 10.1016/j.bbapap.2013.04.032. [Epub ahead of print] PubMed PMID: 23692960.

## Examples

```
## Not run:
library("rpx")
px1 <- PXDataset("PXD000001")
mztab <- pxget(px1, "PXD000001_mztab.txt")
library("MSnbase")
qnt <- readMzTabData(mztab, what = "PEP")
sampleNames(qnt) <- reporterNames(TMT6)
qnt$conditions <- factor(c("A", "A", "B", "B", "B", "A"))
qnt <- filterNA(qnt)

selA <- qnt$conditions == "A"

fData(qnt)$log2FC <-
  log(rowMeans(exprs(qnt)[, selA]), 2) -
  log(rowMeans(exprs(qnt)[, !selA]), 2)
fData(qnt)$baseMean <- log(rowMeans(exprs(qnt)), 10)

## End(Not run)

library("RforProteomics")
library("MSnbase")
data(qnt)
class(qnt)
head(exprs(qnt))
head(fData(qnt))
```

---

RforProteomics

*Opens the package vignettes*


---

## Description

Opens the package and visualisation vignettes.

**Usage**

```
RforProteomics()
```

```
RProtVis()
```

```
RProtViz()
```

**Value**

An instance of class `vignette`. Used for its side effect, opening a vignette.

**Author(s)**

Laurent Gatto

# Index

## \* datasets

id, [6](#)

qnt, [9](#)

analyse (new\_ions), [7](#)

analyze (new\_ions), [7](#)

detect (new\_ions), [7](#)

download.file, [2](#)

downloadData, [2](#)

getPackagesInBiocView, [3](#)

getPXD000001mzData, [4](#)

getPXD000001mzTab, [4](#)

getPXD000001mzXML, [5](#)

getThermoHelaPRTC, [5](#)

id, [6](#)

massSpectrometryDataPackages  
(proteomicsPackages), [8](#)

massSpectrometryPackages  
(proteomicsPackages), [8](#)

MSnSet, [9](#)

mstut (new\_ions), [7](#)

new\_ions, [7](#)

packageDF, [8](#)

proteomicsPackages, [8](#)

qnt, [9](#)

RforProteomics, [10](#)

RProtVis (RforProteomics), [10](#)

RProtViz (RforProteomics), [10](#)

spectrum (new\_ions), [7](#)

TMT\_Erwinia\_1uLSike\_Top10HCD\_isol2\_45stepped\_60min\_01.mzid  
(id), [6](#)