

isobar for developers

Florian P. Breitwieser, Jacques Colinge

May 19, 2021

Contents

1	Introduction	1
2	Classes	1
2.1	IBSpectra	1
2.2	ProteinGroup	2
2.3	NoiseModel	3
3	Session Information	4

1 Introduction

This documents highlights the structure of the S4 classes and methods in the `isobar` package.

```
> library(isobar)
```

2 Classes

2.1 IBSpectra

```
> getClass("IBSpectra")
```

Virtual Class "IBSpectra" [package "isobar"]

Slots:

Name:	proteinGroup	reporterTagNames	reporterTagMasses
Class:	ProteinGroup	character	numeric

Name:	isotopeImpurities	log	assayData
Class:	matrix	matrix	AssayData

Name:	phenoData	featureData	experimentData
Class:	AnnotatedDataFrame	AnnotatedDataFrame	MIAxE

```

Name:          annotation      protocolData  .__classVersion__
Class:         character AnnotatedDataFrame      Versions

```

Extends:

```

Class "eSet", directly
Class "VersionedBiobase", by class "eSet", distance 2
Class "Versioned", by class "eSet", distance 3

```

Known Subclasses:

```

Class "iTRAQspectra", directly
Class "TMTspectra", directly
Class "iTRAQ4plexspectra", by class "iTRAQspectra", distance 2
Class "iTRAQ8plexspectra", by class "iTRAQspectra", distance 2
Class "TMT2plexspectra", by class "TMTspectra", distance 2
Class "TMT6plexspectra", by class "TMTspectra", distance 2
Class "TMT6plexspectra2", by class "TMTspectra", distance 2
Class "TMT10plexspectra", by class "TMTspectra", distance 2

```

identifications and quantitative values. Spectrums are identified as stemming from distinct peptides, and quantitative information of each spectrum are extracted from a certain m/z region.

IBSpectra class holds this qualitative and quantitative information. It is a virtual class. It extends `eSet` from Biobase to store meta-information of spectrum identifications and quantitative information (m/z and `intensity`) of reporter tags. `eSet` is extended by slots for protein grouping, tag names, tag masses and isotope impurity correction matrix.

`ProteinGroup` objects store the mapping and grouping of peptide level identifications to protein identifications.

IBSpectra is a virtual class. Currently used isobaric tagging kits iTRAQ 4plex and 8plex, and TMT 2plex and 6plex are implemented in the `iTRAQ4plexspectra`, `iTRAQ8plexspectra`, `TMT2plexspectra`, `TMT6plexspectr` and `TMT10plexspectr`, respectively. These are subclasses of `iTRAQspectra` and `TMTspectra`, resp. which in turn are virtual subclasses of `IBSpectra`.

2.2 ProteinGroup

```
> getClass("ProteinGroup")
```

```
Class "ProteinGroup" [package "isobar"]
```

Slots:

```

Name:          spectrumToPeptide      spectrumId
Class:         character              data.frame

```

```

Name:          peptideSpecificity      peptideNProtein
Class:         data.frame              matrix

```

Name:	indistinguishableProteins	proteinGroupTable
Class:	character	data.frame

Name:	overlappingProteins	isoformToGeneProduct
Class:	matrix	data.frame

Name:	proteinInfo	peptideInfo
Class:	data.frame	data.frame

Name:	.__classVersion__
Class:	Versions

Extends:

Class "VersionedBiobase", directly

Class "Versioned", by class "VersionedBiobase", distance 2

mapped back to proteins. This mapping leads to protein groups, which explain the observed peptides according to the parsimony law.

A `ProteinGroup` object is generated when a `IBSpectra` object is created by `readIBSpectra`. Protein to peptide to spectrum mapping is extracted from a suitable identification format¹

2.3 NoiseModel

```
> getClass("NoiseModel")
```

Virtual Class "NoiseModel" [package "isobar"]

Slots:

Name:	na.region	low.intensity	f	parameter
Class:	numeric	numeric	function	numeric

Name:	.__classVersion__
Class:	Versions

Extends:

Class "VersionedBiobase", directly

Class "Versioned", by class "VersionedBiobase", distance 2

Known Subclasses: "ExponentialNoANoiseModel", "ExponentialNoiseModel", "InverseNoiseModel", "InverseNoANoiseModel", "GeneralNoiseModel"

in the spectrum-level ratios of a certain experimental setup.

¹IBSpectra CSV, and MzIdentML format. Mascot DAT and Phenyx pidres.xml format converters to IBSpectra format are provided.

3 Session Information

The version number of R and packages loaded for generating the vignette were:

- R version 4.1.0 RC (2021-05-10 r80283), x86_64-apple-darwin17.0
- Locale: C/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
- Running under: macOS Mojave 10.14.6
- Matrix products: default
- BLAS:
/Library/Frameworks/R.framework/Versions/4.1/Resources/lib/libRblas.dylib
- LAPACK:
/Library/Frameworks/R.framework/Versions/4.1/Resources/lib/libRlapack.dylib
- Base packages: base, datasets, grDevices, graphics, methods, parallel, stats, utils
- Other packages: Biobase 2.52.0, BiocGenerics 0.38.0, isobar 1.38.0
- Loaded via a namespace (and not attached): AnnotationDbi 1.54.0, BiocFileCache 2.0.0, Biostrings 2.60.0, DBI 1.1.1, GenomeInfoDb 1.28.0, GenomeInfoDbData 1.2.6, IRanges 2.26.0, KEGGREST 1.32.0, MASS 7.3-54, R6 2.5.0, RCurl 1.98-1.3, RSQLite 2.2.7, Rcpp 1.0.6, S4Vectors 0.30.0, XML 3.99-0.6, XVector 0.32.0, assertthat 0.2.1, biomaRt 2.48.0, bit 4.0.4, bit64 4.0.5, bitops 1.0-7, blob 1.2.1, cachem 1.0.5, colorspace 2.0-1, compiler 4.1.0, crayon 1.4.1, curl 4.3.1, dbplyr 2.1.1, digest 0.6.27, distr 2.8.0, dplyr 1.0.6, ellipsis 0.3.2, fansi 0.4.2, fastmap 1.1.0, filelock 1.0.2, generics 0.1.0, ggplot2 3.3.3, glue 1.4.2, grid 4.1.0, gtable 0.3.0, hms 1.1.0, httr 1.4.2, lifecycle 1.0.0, magrittr 2.0.1, memoise 2.0.0, munsell 0.5.0, pillar 1.6.1, pkgconfig 2.0.3, plyr 1.8.6, png 0.1-7, prettyunits 1.1.1, progress 1.2.2, purrr 0.3.4, rappdirs 0.3.3, rlang 0.4.11, rstudioapi 0.13, scales 1.1.1, sfsmisc 1.1-11, startupmsg 0.9.6, stats4 4.1.0, stringi 1.6.2, stringr 1.4.0, tibble 3.1.2, tidysselect 1.1.1, tools 4.1.0, utf8 1.2.1, vctrs 0.3.8, zlibbioc 1.38.0