

# *isobar* for developers

Florian P. Breitwieser, Jacques Colinge

October 24, 2023

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Classes</b>	<b>1</b>
2.1	IBSpectra . . . . .	1
2.2	ProteinGroup . . . . .	2
2.3	NoiseModel . . . . .	3
<b>3</b>	<b>Session Information</b>	<b>4</b>

## 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 format<sup>1</sup>

## 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", "InverseNoANoiseModel", "InverseNoiseModel", "GeneralNoiseModel"

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

---

<sup>1</sup>IBSpectra 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 4.3.1 (2023-06-16 ucrt), x86\_64-w64-mingw32
- Locale: LC\_COLLATE=C, LC\_CTYPE=English\_United\_States.utf8, LC\_MONETARY=English\_United\_States.utf8, LC\_NUMERIC=C, LC\_TIME=English\_United\_States.utf8
- Time zone: America/New\_York
- TZcode source: internal
- Running under: Windows Server 2022 x64 (build 20348)
- Matrix products: default
- Base packages: base, datasets, grDevices, graphics, methods, stats, utils
- Other packages: Biobase 2.62.0, BiocGenerics 0.48.0, isobar 1.48.0
- Loaded via a namespace (and not attached): AnnotationDbi 1.64.0, BiocFileCache 2.10.0, Biostrings 2.70.0, DBI 1.1.3, GenomeInfoDb 1.38.0, GenomeInfoDbData 1.2.11, IRanges 2.36.0, KEGGREST 1.42.0, MASS 7.3-60, R6 2.5.1, RCurl 1.98-1.12, RSQLite 2.3.1, Rcpp 1.0.11, S4Vectors 0.40.0, XML 3.99-0.14, XVector 0.42.0, biomaRt 2.58.0, bit 4.0.5, bit64 4.0.5, bitops 1.0-7, blob 1.2.4, cachem 1.0.8, cli 3.6.1, colorspace 2.1-0, compiler 4.3.1, crayon 1.5.2, curl 5.1.0, dbplyr 2.3.4, digest 0.6.33, distr 2.9.2, dplyr 1.1.3, fansi 1.0.5, fastmap 1.1.1, filelock 1.0.2, generics 0.1.3, ggplot2 3.4.4, glue 1.6.2, grid 4.3.1, gtable 0.3.4, hms 1.1.3, httr 1.4.7, lifecycle 1.0.3, magrittr 2.0.3, memoise 2.0.1, munsell 0.5.0, pillar 1.9.0, pkgconfig 2.0.3, plyr 1.8.9, png 0.1-8, prettyunits 1.2.0, progress 1.2.2, rappdirs 0.3.3, rlang 1.1.1, scales 1.2.1, sfsmisc 1.1-16, startupmsg 0.9.6, stats4 4.3.1, stringi 1.7.12, stringr 1.5.0, tibble 3.2.1, tidyselect 1.2.0, tools 4.3.1, utf8 1.2.4, vctrs 0.6.4, xml2 1.3.5, zlibbioc 1.48.0