

Package ‘hpar’

March 26, 2013

Type Package

Title Human Protein Atlas in R

Version 1.0.1

Author@R person(given = “Laurent”, family = “Gatto”, email =
“lg390@cam.ac.uk”, role = c(“aut”, “cre”))

Author Laurent Gatto

Maintainer Laurent Gatto <lg390@cam.ac.uk>

Description A simple interface to and data from the Human Protein Atlas project.

Depends R (>= 2.15)

Suggests org.Hs.eg.db, GO.db, knitr

biocViews Bioinformatics, Proteomics, Homo_sapiens, CellBiology

License Artistic-2.0

Collate environment.R hpa.R zzz.R

R topics documented:

getHpa	2
getHpaDate	3
getHparOptions	3
hpar	4
setHparOptions	5

Index	7
--------------	----------

getHpa *HPA gene query*

Description

Queries one if the HPA data sets with the id Ensembl gene identifier. The data set to be used is defined by the hpadata argument.

Usage

```
getHpa(id, hpadata = NULL, type = c("data", "details"))
```

Arguments

id	A Ensembl gene identifier.
hpadata	The data set to query. One of "NormalTissue", "Rna", "SubcellularLoc" (or any unambiguous prefix) or NULL (default). In the latter case, the default option is used. See setHparOptions and getHparOptions for more details.
type	A character defining what type data to return. One of data (default, for the data as dataframe) or details (to open the id's HPA webpage).

Value

A dataframe with the information corresponding to the respective id genes. If type is details, then the dataframe is returned invisibly and a web page is opened with [browseURL](#).

Author(s)

Laurent Gatto

Examples

```
id <- "ENSG00000000003"
## Define 'hpadata' data manually
getHpa(id, hpadata = "Subcell")
head(getHpa(id, hpadata = "NormalTissue"), )
head(getHpa(id, hpadata = "Rna"))
## Default 'hpadata' data
head(getHpa(id)) ## default 'hpadata' is "NormalTissue"
## Sets default to "SubcellularLoc"
setHparOptions(hpadata = "SubcellularLoc")
getHpa(id) ## now uses "SubcellularLoc"
setHparOptions() ## reset to "NormalTissue"
head(getHpa(id))
## multiple ids
getHpa(id = c("ENSG00000000003", "ENSG00000000005"),
       hpadata = "SubcellularLoc")
## Not run:
## opens a browser with http://www.proteinatlas.org/ENSG00000163435
getHpa("ENSG00000163435",
       type = "details")

## End(Not run)
```

getHpaDate	<i>Obtain HPA release information</i>
------------	---------------------------------------

Description

Obtain release information about the HPA data, including version, release data and Ensembl version the data is based on. Detailed release information is accessed on-line using [browseURL](#).

Usage

```
getHpaDate()  
getHpaVersion()  
getHpaEnsembl()
```

Value

A character.

Author(s)

Laurent Gatto <lg390@cam.ac.uk>

References

HPA release history: <http://www.proteinatlas.org/about/releases>

Examples

```
getHpaVersion()  
getHpaDate()  
getHpaEnsembl()
```

getHparOptions	<i>Returns the hpar option.</i>
----------------	---------------------------------

Description

Returns the hpar options. Currently, there is only one, hpadata. See [setHparOptions](#) for details and examples.

Usage

```
getHparOptions()
```

Value

A list with the hpar option(s).

Author(s)

Laurent Gatto

Description

This package provides a simple interface to the Human Protein Atlas. From the Human Protein Atlas Project page: The Swedish Human Protein Atlas project, funded by the Knut and Alice Wallenberg Foundation, has been set up to allow for a systematic exploration of the human proteome using Antibody-Based Proteomics. This is accomplished by combining high-throughput generation of affinity-purified antibodies with protein profiling in a multitude of tissues and cells assembled in tissue microarrays. Confocal microscopy analysis using human cell lines is performed for more detailed protein localization. The program hosts the Human Protein Atlas portal with expression profiles of human proteins in tissues and cells.

Usage

```
hpaRna  
hpaNormalTissue  
hpaSubcellularLoc
```

Details

Three data tables are distributed by the HPA project and available within the package as dataframes. The description below is adapted from the HPA site:

1. **Normal tissue data** (hpaNormalTissue) Expression profiles for proteins in human tissues based on immunohistochemistry using tissue micro arrays. The dataframe includes Ensembl gene identifier ("Gene"), tissue name ("Tissue"), annotated cell type ("Cell.type"), expression value ("Level"), the type of annotation (annotated protein expression (APE), based on more than one antibody, or staining, based on one antibody only) ("Expression.type"), and the reliability or validation of the expression value ("Reliability").
2. **Subcellular location data** (hpaSubcellularLoc) Subcellular localisation of proteins based on immunofluorescently stained cells. The dataframe includes Ensembl gene identifier ("Gene"), main subcellular location of the protein ("Main.location"), other locations ("Other.location"), the type of annotation (annotated protein expression (APE), based on more than one antibody, or staining, based on one antibody only) ("Expression.type"), and the reliability or validation of the expression value ("Reliability"). Note that the Gene Ontology and UniProt are both importing part of this data into their respective databases. UniProt cite the source with *source:HPA*.
3. **RNA data** (hpaRna) RNA levels in three different cell lines, based on RNAseq. The dataframe includes Ensembl gene identifier ("Gene"), analysed cell line ("Cell.line"), number of reads per kilobase gene model and million reads ("RPKM"), and abundance class ("Abundance").

Detailed description for gene entries and images and not included in the package but can be accessed from within the R environment through a web browser while on-line.

The full data sets can be individually loaded using the `data` function (see example below). Data about individual genes of interest can be retrieved with the `getHpa` function.

HPA data usage policy: The use of data and images from this site in publications and presentations is permitted provided that the following conditions are met:

1. The publication and/or presentation are solely for informational and non-commercial purposes.
2. The source of the data and/or image is referred to this site (www.proteinatlas.org) and/or one or more of our publications are cited.

Author(s)

Laurent Gatto <lg390@cam.ac.uk>

References

See the Human Protein Atlas Project page <http://www.proteinatlas.org/> for more details and documentation.

Uhlen et al (2010). Towards a knowledge-based Human Protein Atlas. Nat Biotechnol. 28(12):1248-50.

Berglund et al (2008). A gene-centric Human Protein Atlas for expression profiles based on antibodies. Mol Cell Proteomics. 7(10):2019-27.

Uhlen et al (2005). A human protein atlas for normal and cancer tissues based on antibody proteomics. Mol Cell Proteomics. 4(12):1920-1932.

Ponten et al (2008). The Human Protein Atlas - a tool for pathology. J Pathology. 216(4):387-93.

See Also

[getHpaDate](#) for release information. Gene-specific information should be accessed using the [getHpa](#) function.

The package vignette can be accessed with `vignette("hpar")`.

Examples

```
data(hpaRna)
head(hpaRna)
dim(hpaRna)
data(hpaNormalTissue)
head(hpaNormalTissue)
dim(hpaNormalTissue)
data(hpaSubcellularLoc)
head(hpaSubcellularLoc)
dim(hpaSubcellularLoc)
```

setHparOptions

Sets the 'hpar' options

Description

Function used to set the hpdata hpar option.

Usage

```
setHparOptions(hpdata = c("NormalTissue", "Rna", "SubcellularLoc"))
```

Arguments

`hpdata` A character defining the default data to be queries. Choices are "NormalTissue", "Rna" or "SubcellularLoc", or any unambiguous prefix.

Value

Returns the set value of the hpar option. Used for its side effect.

Author(s)

Laurent Gatto

Examples

```
oldOpt <- getHparOptions()
oldOpt$hpdata
setHparOptions("Rna")
getHparOptions()
setHparOptions(oldOpt$hpdata)
```

Index

*Topic **datasets**

hpar, 4

browseURL, 2, 3

data, 4

getHpa, 2, 4, 5

getHpaDate, 3, 5

getHpaEnsembl (getHpaDate), 3

getHparOptions, 2, 3

getHpaVersion (getHpaDate), 3

hpaNormalTissue (hpar), 4

hpar, 4

hpaRna (hpar), 4

hpaSubcellularLoc (hpar), 4

setHparOptions, 2, 3, 5