

Package ‘hpar’

April 12, 2018

Type Package

Title Human Protein Atlas in R

Version 1.20.0

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)

Imports utils

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

VignetteBuilder knitr

License Artistic-2.0

Collate environment.R hpa.R zzz.R

biocViews Proteomics, Homo_sapiens, CellBiology

RoxygenNote 5.0.1

NeedsCompilation no

R topics documented:

getHpa	2
getHpaDate	3
getHparOptions	3
hpar	4
setHparOptions	6

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 `hpdata` argument.

Usage

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

Arguments

<code>id</code>	A Ensembl gene identifier.
<code>hpdata</code>	A character with the data set to query. One of available hpar datasets, available by calling allHparData (or any unambiguous prefix), or NULL (default). In the latter case, the default option is used. See setHparOptions and getHparOptions for more details.
<code>type</code>	A character defining what type data to return. One of <code>data</code> (default, for the data as <code>data.frame</code>) or <code>details</code> (to open the <code>id</code> 's HPA webpage).

Value

A `data.frame` 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 'hpdata' data manually
getHpa(id, hpadata = "hpaSubcellularLoc")
head(getHpa(id, hpadata = "hpaNormalTissue"), )
head(getHpa(id, hpadata = "rnaGeneTissue"))
head(getHpa(id, hpadata = "rnaGeneCellLine"))
head(getHpa(id, hpadata = "hpaCancer"))
## Sets default to "SubcellularLoc"
setHparOptions(hpdata = "hpaSubcellularLoc")
getHpa(id)      ## now uses "hpaSubcellularLoc"
setHparOptions() ## reset to "hpaNormalTissue"
head(getHpa(id))
## multiple ids
getHpa(id = c("ENSG00000000003", "ENSG00000000005"),
      hpadata = "hpaSubcellularLoc")
## 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.

Details

Several flat files are distributed by the HPA project and available within the package as `data.frames`. The description below is taken from the HPA site:

hpaNormalTissue Normal tissue data: Expression profiles for proteins in human tissues based on immunohistochemistry using tissue micro arrays. The comma-separated file 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").

hpaCancer Cancer tumor data: Staining profiles for proteins in human tumor tissue based on immunohistochemistry using tissue micro arrays. The comma-separated file includes Ensembl gene identifier ("Gene"), tumor name ("Tumor"), staining value ("Level"), the number of patients that stain for this staining value ("Count patients"), the total amount of patients for this tumor type ("Total patients") and the type of annotation staining ("Expression type").

rnaGeneTissue RNA gene data: RNA levels in 45 cell lines and 32 tissues based on RNA-seq. The comma-separated file includes Ensembl gene identifier ("Gene"), analysed sample ("Sample"), fragments per kilobase of transcript per million fragments mapped ("Value" and "Unit"), and abundance class ("Abundance").

rnaGeneCellLine RNA gene data: RNA levels in 45 cell lines and 32 tissues based on RNA-seq. The comma-separated file includes Ensembl gene identifier ("Gene"), analysed sample ("Sample"), fragments per kilobase of transcript per million fragments mapped ("Value" and "Unit"), and abundance class ("Abundance").

hpaSubcellularLoc Subcellular location data: Subcellular localization of proteins based on immunofluorescently stained cells. The comma-separated file 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").

hpaSubcellularLoc14 Same as above, for version 14.

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/> and <http://www.proteinatlas.org/about/download> 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

```
allHparData()

data(hpaNormalTissue)
head(hpaNormalTissue)

data(rnaGeneCellLine)
head(rnaGeneCellLine)

data(rnaGeneTissue)
head(rnaGeneTissue)

data(hpaSubcellularLoc)
head(hpaSubcellularLoc)

data(hpaSubcellularLoc14)
head(hpaSubcellularLoc14)

data(hpaCancer)
head(hpaCancer)
```

setHparOptions	<i>Sets the 'hpar' options</i>
----------------	--------------------------------

Description

Function used to set the hpadata hpar option.

Usage

```
setHparOptions(hpdata = allHparData())
```

Arguments

hpdata	A character defining the default data to be queries. Choices are any of available hpar datasets, available by calling allHparData , 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$hpadata
setHparOptions("hpaNormalTissue")
getHparOptions()
setHparOptions(oldOpt$hpadata)
```

Index

*Topic **datasets**

hpar, 4

allHparData, 2, 6

allHparData (hpar), 4

browseURL, 2, 3

data, 4

getHpa, 2, 4, 5

getHpaDate, 3, 5

getHpaEnsembl (getHpaDate), 3

getHparOptions, 2, 3

getHpaVersion (getHpaDate), 3

hpaCancer (hpar), 4

hpaNormalTissue (hpar), 4

hpar, 4

hpaSubcellularLoc (hpar), 4

hpaSubcellularLoc14 (hpar), 4

rnaGeneCellLine (hpar), 4

rnaGeneTissue (hpar), 4

setHparOptions, 2, 3, 6