

org.Hs.ref.db

October 16, 2009

R topics documented:

org.Hs.ref_dbconn	1
org.Hs.ref.db	2
org.Hs.refDE	2
org.Hs.refGENEID	3
org.Hs.refGO	3
org.Hs.refKEGG	4
org.Hs.refMAPCOUNTS	5
org.Hs.refORGANISM	6
org.Hs.refPATH	6
org.Hs.refREFSEQ	7
org.Hs.refSEQ	7
org.Hs.refSYMBOL	8
Index	9

org.Hs.ref_dbconn *Collect information about the package annotation DB*

Description

Some convenience functions for getting a connection object to (or collecting information about) the package annotation DB.

Usage

```
org.Hs.ref_dbconn()  
org.Hs.ref_dbfile()  
org.Hs.ref_dbschema(file="", show.indices=FALSE)  
org.Hs.ref_dbInfo()
```

Arguments

`file` A connection, or a character string naming the file to print to (see the file argument of the cat function for the details).

`show.indices` The CREATE INDEX statements are not shown by default. Use show.indices=TRUE to get them.

Details

org.Hs.ref_dbconn returns a connection object to the package annotation DB. **IMPORTANT:** Don't call `dbDisconnect` on the connection object returned by `org.Hs.ref_dbconn` or you will break all the `AnnDbObj` objects defined in this package!

org.Hs.ref_dbfile returns the path (character string) to the package annotation DB (this is an SQLite file).

org.Hs.ref_dbschema prints the schema definition of the package annotation DB.

org.Hs.ref_dbInfo prints other information about the package annotation DB.

Examples

```
## Show the first three rows.
dbGetQuery(org.Hs.ref_dbconn(), "select * from basic limit 3")

## The connection object returned by org.Hs.ref_dbconn() was created with:
dbConnect(SQLite(), dbname=org.Hs.ref_dbfile(), cache_size=64000, synchronous=0)

org.Hs.ref_dbschema()

org.Hs.ref_dbInfo()
```

org.Hs.ref.db	<i>annotation data package</i>
---------------	--------------------------------

Description

Welcome to the org.Hs.ref.db annotation Package. The annotation package was built using a downloadable R package - PAnnBuilder (download and build your own). The purpose is to provide detailed information about the reference protein in NCBI RefSeq database: ftp://ftp.ncbi.nih.gov/refseq/H_sapiens/mRNA_Prot/human.protein.faa.gz March10, 2008

Each of these objects has their own manual page detailing where relevant data was obtained along with examples of how to use it. Many of these objects also have a reverse map available. When this is true, expect to usually find relevant information on the same manual page as the forward map.

Examples

```
# You can learn what objects this package supports with the following command:
ls("package:org.Hs.ref.db")
```

org.Hs.refDE	<i>Maps protein identifier to textural descriptions</i>
--------------	---

Description

org.Hs.refDE maps protein identifiers to their descriptive information.

Details

Mappings were based on data provided by: NCBI (<ftp://ftp.ncbi.nih.gov/gene/DATA/gene2refseq.gz>; ftp://ftp.ncbi.nih.gov/gene/DATA/gene_info.gz) on

Examples

```
x <- org.Hs.refDE
# Get the protein identifiers that are mapped to textural descriptions.
mapped_proteins <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_proteins])
  if(length(xx) > 0){
    # Get the value of the first key
    xx[[1]]
    # Get the values for a few keys
    if(length(xx) >= 3){
      xx[1:3]
    }
  }
}
```

org.Hs.refGENEID *Map protein identifier to Entrez gene identifier*

Description

org.Hs.refGENEID maps protein identifiers to Entrez Gene identifiers.

Details

Mappings were based on data provided by: NCBI (<ftp://ftp.ncbi.nih.gov/gene/DATA/gene2refseq.gz>; ftp://ftp.ncbi.nih.gov/gene/DATA/gene_info.gz) on

Examples

```
x <- org.Hs.refGENEID
# Get the protein identifiers that are mapped to Entrez gene identifier.
mapped_proteins <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_proteins])
  if(length(xx) > 0){
    # Get the value of the first key
    xx[[1]]
    # Get the values for a few keys
    if(length(xx) >= 3){
      xx[1:3]
    }
  }
}
```

org.Hs.refGO *Map protein identifier to GO*

Description

org.Hs.refGO maps protein identifiers to Gene Ontology identifiers .

Details

Each Protein identifier is mapped to a list. Each component contain : GO ID, Evidence and Ontology (C,F,P). NAs are assigned to probe identifiers that can not be mapped to any Gene Ontology information.

The Evidence element contains a code indicating what kind of evidence supports the association of the GO id to the protein id. The evidence codes in use include: IMP: inferred from mutant phenotype IGI: inferred from genetic interaction IPI: inferred from physical interaction ISS: inferred from sequence similarity IDA: inferred from direct assay IEP: inferred from expression pattern IEA: inferred from electronic annotation TAS: traceable author statement NAS: non-traceable author statement ND: no biological data available IC: inferred by curator

Mappings were based on data provided by: NCBI (<ftp://ftp.ncbi.nih.gov/gene/DATA/gene2go.gz>) on

Examples

```
x <- org.Hs.refGO
# Get the protein identifiers that are mapped to a GO ID
mapped_genes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_genes])
if(length(xx) > 0) {
  # Try the first one
  got <- xx[[1]]
  got[[1]][["GOID"]]
  got[[1]][["Ontology"]]
  got[[1]][["Evidence"]]
}
# For the reverse map:
xx <- as.list(org.Hs.refGO2GI)
if(length(xx) > 0){
  goids <- xx[2:3]
}
```

org.Hs.refKEGG

Map protein identifier to KEGG gene identifiers

Description

org.Hs.refKEGG maps protein identifiers to KEGG gene identifiers.

Details

Each protein identifier maps to KEGG gene identifiers.

Mappings were based on data provided by: KEGG (ftp://ftp.genome.jp/pub/kegg/hsa/hsa_pathway.list; ftp://ftp.genome.jp/pub/kegg/hsa/hsa_ncbi-geneid.list) on

Examples

```
x <- org.Hs.refKEGG
# Get the protein identifiers that are mapped to KEGG gene identifiers.
mapped_proteins <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_proteins])
  if(length(xx) > 0){
    # Get the value of the first key
    xx[[1]]
  }
```

org.Hs.refMAPCOUNTS

Number of mapped keys for the maps in package org.Hs.ref.db

Description

org.Hs.refMAPCOUNTS provides the "map count" (i.e. the count of mapped keys) for each map in package org.Hs.ref.db.

Details

This "map count" information is precalculated and stored in the package annotation DB. This allows some quality control and is used by the [checkMAPCOUNTS](#) function defined in AnnotationDbi to compare and validate different methods (like `count.mappedkeys(x)` or `sum(!is.na(as.list(x)))`) for getting the "map count" of a given map.

See Also

[mappedkeys](#), [count.mappedkeys](#), [checkMAPCOUNTS](#)

Examples

```
org.Hs.refMAPCOUNTS
mapnames <- names(org.Hs.refMAPCOUNTS)
org.Hs.refMAPCOUNTS[mapnames[1]]
x <- get(mapnames[1])
sum(!is.na(as.list(x)))
count.mappedkeys(x) # much faster!

## Check the "map count" of all the maps in package org.Hs.ref.db
checkMAPCOUNTS("org.Hs.ref.db")
```

org.Hs.refORGANISM *The Organism for org.Hs.ref.db*

Description

org.Hs.refORGANISM is an R object that contains a single item: a character string that names the organism for which org.Hs.ref.db was built.

Details

Although the package name is suggestive of the organism for which it was built, org.Hs.refORGANISM provides a simple way to programmatically extract the organism name.

Examples

```
org.Hs.refORGANISM
```

org.Hs.refPATH *Map protein identifier to KEGG pathway*

Description

org.Hs.refPATH maps protein identifiers to KEGG pathway identifiers.

Details

Each protein identifier maps to KEGG pathway identifiers.

Mappings were based on data provided by: KEGG (ftp://ftp.genome.jp/pub/kegg/hsa/hsa_pathway.list; ftp://ftp.genome.jp/pub/kegg/hsa/hsa_ncbi-geneid.list) on

Examples

```
x <- org.Hs.refPATH
# Get the protein identifiers that are mapped to KEGG pathways.
mapped_proteins <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_proteins])
  if(length(xx) > 0){
    # Get the value of the first key
    xx[[1]]
  }

## For the reverse map org.Hs.refPATH2GI
xx <- as.list(org.Hs.refPATH2GI)
if(length(xx) > 0){
  ## Get the first one
  xx[[1]]
}
```

org.Hs.refREFSEQ *Map protein identifier to RefSeq identifiers*

Description

org.Hs.refREFSEQ maps protein identifiers to RefSeq identifiers.

Details

Each protein identifier maps to RefSeq identifiers.

Mappings were based on data provided by: NCBI (ftp://ftp.ncbi.nih.gov/refseq/H_sapiens/mRNA_Prot/human.protein.faa.gz) on March 10, 2008

Examples

```
x <- org.Hs.refREFSEQ
# Get the protein identifiers that are mapped to RefSeq identifiers.
mapped_proteins <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_proteins])
  if(length(xx) > 0){
    # Get the value of the first key
    xx[[1]]
  }

## For the reverse map org.Hs.refREFSEQ2GI
xx <- as.list(org.Hs.refREFSEQ2GI)
if(length(xx) > 0){
  ## Get the first one
  xx[[1]]
}
```

org.Hs.refSEQ *Map protein identifier to the protein sequence*

Description

org.Hs.refSEQ provides mappings between a protein identifier and the protein Sequence.

Details

Mappings were based on data provided by: RefSeq (ftp://ftp.ncbi.nih.gov/refseq/H_sapiens/mRNA_Prot/human.protein.faa.gz) on March 10, 2008

Examples

```
x <- org.Hs.refSEQ
# Get the protein identifiers that are mapped to protein sequences.
mapped_proteins <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_proteins])
# randomly display 10 proteins
sample(xx, 10)
```

org.Hs.refSYMBOL *Map protein identifier to gene symbols*

Description

org.Hs.refSYMBOL maps protein identifiers to gene symbols.

Details

Each protein identifier maps to an abbreviation for the corresponding gene. .

Mappings were based on data provided by: NCBI (<ftp://ftp.ncbi.nih.gov/gene/DATA/gene2refseq.gz>; ftp://ftp.ncbi.nih.gov/gene/DATA/gene_info.gz) on

Examples

```
x <- org.Hs.refSYMBOL
# Get the protein identifiers that are mapped to gene symbols.
mapped_proteins <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_proteins])
  if(length(xx) > 0){
    # Get the value of the first key
    xx[[1]]
  }
```


Index

*Topic **datasets**

- org.Hs.refMAPCOUNTS, 5
- AnnDbObj, 1
- checkMAPCOUNTS, 5
- count.mappedkeys, 5
- dbDisconnect, 1
- mappedkeys, 5
- org.Hs.ref.db, 2
- org.Hs.ref_dbconn, 1
- org.Hs.ref_dbfile
 - (org.Hs.ref_dbconn), 1
- org.Hs.ref_dbInfo
 - (org.Hs.ref_dbconn), 1
- org.Hs.ref_dbschema
 - (org.Hs.ref_dbconn), 1
- org.Hs.refDE, 2
- org.Hs.refGENEID, 3
- org.Hs.refGO, 3
- org.Hs.refGO2GI (org.Hs.refGO), 3
- org.Hs.refKEGG, 4
- org.Hs.refMAPCOUNTS, 5
- org.Hs.refORGANISM, 5
- org.Hs.refPATH, 6
- org.Hs.refPATH2GI
 - (org.Hs.refPATH), 6
- org.Hs.refREFSEQ, 6
- org.Hs.refREFSEQ2GI
 - (org.Hs.refREFSEQ), 6
- org.Hs.refSEQ, 7
- org.Hs.refSYMBOL, 7