

# org.Rn.ref.db

October 16, 2009

## R topics documented:

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

<b>Index</b>	<b>9</b>
--------------	----------

---

org.Rn.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.Rn.ref_dbconn()  
org.Rn.ref_dbfile()  
org.Rn.ref_dbschema(file="", show.indices=FALSE)  
org.Rn.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.Rn.ref\_dbconn returns a connection object to the package annotation DB. **IMPORTANT:** Don't call `dbDisconnect` on the connection object returned by `org.Rn.ref_dbconn` or you will break all the `AnnDbObj` objects defined in this package!

org.Rn.ref\_dbfile returns the path (character string) to the package annotation DB (this is an SQLite file).

org.Rn.ref\_dbschema prints the schema definition of the package annotation DB.

org.Rn.ref\_dbInfo prints other information about the package annotation DB.

**Examples**

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

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

org.Rn.ref_dbschema()

org.Rn.ref_dbInfo()
```

---

org.Rn.ref.db                    *annotation data package*

---

**Description**

Welcome to the org.Rn.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/R\\_norvegicus/mRNA\\_Prot/rat.protein.faa.gz](ftp://ftp.ncbi.nih.gov/refseq/R_norvegicus/mRNA_Prot/rat.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.Rn.ref.db")
```

---

org.Rn.refDE                    *Maps protein identifier to textural descriptions*

---

**Description**

org.Rn.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](ftp://ftp.ncbi.nih.gov/gene/DATA/gene_info.gz)) on

## Examples

```
x <- org.Rn.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.Rn.refGENEID     *Map protein identifier to Entrez gene identifier*

---

## Description

org.Rn.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](ftp://ftp.ncbi.nih.gov/gene/DATA/gene_info.gz)) on

## Examples

```
x <- org.Rn.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.Rn.refGO     *Map protein identifier to GO*

---

## Description

org.Rn.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.Rn.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.Rn.refGO2GI)
if(length(xx) > 0){
  goids <- xx[2:3]
}
```

---

org.Rn.refKEGG

*Map protein identifier to KEGG gene identifiers*

---

## Description

org.Rn.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/rno/rno\\_pathway.list](ftp://ftp.genome.jp/pub/kegg/rno/rno_pathway.list); [ftp://ftp.genome.jp/pub/kegg/rno/rno\\_ncbi-geneid.list](ftp://ftp.genome.jp/pub/kegg/rno/rno_ncbi-geneid.list)) on

## Examples

```
x <- org.Rn.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.Rn.refMAPCOUNTS

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

---

## Description

org.Rn.refMAPCOUNTS provides the "map count" (i.e. the count of mapped keys) for each map in package org.Rn.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.Rn.refMAPCOUNTS
mapnames <- names(org.Rn.refMAPCOUNTS)
org.Rn.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.Rn.ref.db
checkMAPCOUNTS("org.Rn.ref.db")
```

---

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

---

### Description

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

### Details

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

### Examples

```
org.Rn.refORGANISM
```

---

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

---

### Description

org.Rn.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/rno/rno\\_pathway.list](ftp://ftp.genome.jp/pub/kegg/rno/rno_pathway.list); [ftp://ftp.genome.jp/pub/kegg/rno/rno\\_ncbi-geneid.list](ftp://ftp.genome.jp/pub/kegg/rno/rno_ncbi-geneid.list)) on

### Examples

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

---

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

---

### Description

org.Rn.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/R\\_norvegicus/mRNA\\_Prot/rat.protein.faa.gz](ftp://ftp.ncbi.nih.gov/refseq/R_norvegicus/mRNA_Prot/rat.protein.faa.gz)) on March 10, 2008

### Examples

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

---

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

---

### Description

org.Rn.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/R\\_norvegicus/mRNA\\_Prot/rat.protein.faa.gz](ftp://ftp.ncbi.nih.gov/refseq/R_norvegicus/mRNA_Prot/rat.protein.faa.gz)) on March 10, 2008

### Examples

```
x <- org.Rn.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.Rn.refSYMBOL    *Map protein identifier to gene symbols*

---

### Description

org.Rn.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](ftp://ftp.ncbi.nih.gov/gene/DATA/gene_info.gz)) on

### Examples

```
x <- org.Rn.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.Rn.refMAPCOUNTS, 5

AnnDbObj, 1

checkMAPCOUNTS, 5

count.mappedkeys, 5

dbDisconnect, 1

mappedkeys, 5

org.Rn.ref.db, 2

org.Rn.ref\_dbconn, 1

org.Rn.ref\_dbfile  
(org.Rn.ref\_dbconn), 1

org.Rn.ref\_dbInfo  
(org.Rn.ref\_dbconn), 1

org.Rn.ref\_dbschema  
(org.Rn.ref\_dbconn), 1

org.Rn.refDE, 2

org.Rn.refGENEID, 3

org.Rn.refGO, 3

org.Rn.refGO2GI (org.Rn.refGO), 3

org.Rn.refKEGG, 4

org.Rn.refMAPCOUNTS, 5

org.Rn.refORGANISM, 5

org.Rn.refPATH, 6

org.Rn.refPATH2GI  
(org.Rn.refPATH), 6

org.Rn.refREFSEQ, 6

org.Rn.refREFSEQ2GI  
(org.Rn.refREFSEQ), 6

org.Rn.refSEQ, 7

org.Rn.refSYMBOL, 7