

lumiHumanAll.db

July 17, 2024

lumiHumanAllACCNUM *Map Manufacturer identifiers to Accession Numbers*

Description

lumiHumanAllACCNUM is an R object that contains mappings between a manufacturer's identifiers and manufacturers accessions.

Details

For chip packages such as this, the ACCNUM mapping comes directly from the manufacturer. This is different from other mappings which are mapped onto the probes via an Entrez Gene identifier.

Each manufacturer identifier maps to a vector containing a GenBank accession number.

Mappings were based on data provided by: Entrez Gene <ftp://ftp.ncbi.nlm.nih.gov/gene/DATA> With a date stamp from the source of: 2013-Sep12

Examples

```
x <- lumiHumanAllACCNUM
# Get the probe identifiers that are mapped to an ACCNUM
mapped_probes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_probes])
if(length(xx) > 0) {
  # Get the ACCNUM for the first five probes
  xx[1:5]
  # Get the first one
  xx[[1]]
}
```

lumiHumanAllALIAS2PROBE

Map between Common Gene Symbol Identifiers and Manufacturer Identifiers

Description

lumiHumanAllALIAS is an R object that provides mappings between common gene symbol identifiers and manufacturer identifiers.

Details

Each gene symbol is mapped to a named vector of manufacturer identifiers. The name represents the gene symbol and the vector contains all manufacturer identifiers that are found for that symbol. An NA is reported for any gene symbol that cannot be mapped to any manufacturer identifiers.

This mapping includes ALL gene symbols including those which are already listed in the SYMBOL map. The SYMBOL map is meant to only list official gene symbols, while the ALIAS maps are meant to store all used symbols.

Mappings were based on data provided by: Entrez Gene <ftp://ftp.ncbi.nlm.nih.gov/gene/DATA> With a date stamp from the source of: 2013-Sep12

Examples

```
# Convert the object to a list
xx <- as.list(lumiHumanAllALIAS2PROBE)
if(length(xx) > 0){
  # Get the probe identifiers for the first two aliases
  xx[1:2]
  # Get the first one
  xx[[1]]
}
```

lumiHumanAll.db

Bioconductor annotation data package

Description

Welcome to the lumiHumanAll.db annotation Package. The purpose of this package is to provide detailed information about the lumiHumanAll platform. This package is updated biannually.

You can learn what objects this package supports with the following command:

```
ls("package:lumiHumanAll.db")
```

Each of these objects has their own manual page detailing where relevant data was obtained along with some examples of how to use it.

Examples

```
ls("package:lumiHumanAll.db")
```

```
lumiHumanAllCHR      Map Manufacturer IDs to Chromosomes
```

Description

lumiHumanAllCHR is an R object that provides mappings between a manufacturer identifier and the chromosome that contains the gene of interest.

Details

Each manufacturer identifier maps to a vector of chromosomes. Due to inconsistencies that may exist at the time the object was built, the vector may contain more than one chromosome (e.g., the identifier may map to more than one chromosome). If the chromosomal location is unknown, the vector will contain an NA.

Mappings were based on data provided by: Entrez Gene <ftp://ftp.ncbi.nlm.nih.gov/gene/DATA> With a date stamp from the source of: 2013-Sep12

Examples

```
x <- lumiHumanAllCHR
# Get the probe identifiers that are mapped to a chromosome
mapped_probes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_probes])
if(length(xx) > 0) {
  # Get the CHR for the first five probes
  xx[1:5]
  # Get the first one
  xx[[1]]
}
```

```
lumiHumanAllCHRENGTHS      A named vector for the length of each of the chromosomes
```

Description

lumiHumanAllCHRENGTHS provides the length measured in base pairs for each of the chromosomes.

lumiHumanAllENSEMBL *Map Ensembl gene accession numbers with Entrez Gene identifiers*

Description

lumiHumanAllENSEMBL is an R object that contains mappings between manufacturer identifiers and Ensembl gene accession numbers.

Details

This object is a simple mapping of manufacturer identifiers to Ensembl gene Accession Numbers.

Mappings were based on data provided by BOTH of these sources: <http://www.ensembl.org/biomart/martview/> <ftp://ftp.ncbi.nlm.nih.gov/gene/DATA>

For most species, this mapping is a combination of manufacturer to ensembl IDs from BOTH NCBI and ensembl. Users who wish to only use mappings from NCBI are encouraged to see the `ncbi2ensembl` table in the appropriate organism package. Users who wish to only use mappings from ensembl are encouraged to see the `ensembl2ncbi` table which is also found in the appropriate organism packages. These mappings are based upon the `ensembl` table which contains data from BOTH of these sources in an effort to maximize the chances that you will find a match.

For worms and flies however, this mapping is based only on sources from ensembl, as these organisms do not have ensembl to entrez gene mapping data at NCBI.

Examples

```
x <- lumiHumanAllENSEMBL
# Get the entrez gene IDs that are mapped to an Ensembl ID
mapped_genes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_genes])
if(length(xx) > 0) {
  # Get the Ensembl gene IDs for the first five genes
  xx[1:5]
  # Get the first one
  xx[[1]]
}
#For the reverse map ENSEMBL2PROBE:
# Convert to a list
xx <- as.list(lumiHumanAllENSEMBL2PROBE)
if(length(xx) > 0){
  # Gets the entrez gene IDs for the first five Ensembl IDs
  xx[1:5]
  # Get the first one
  xx[[1]]
}
```

lumiHumanAllENTREZID *Map between Manufacturer Identifiers and Entrez Gene*

Description

lumiHumanAllENTREZID is an R object that provides mappings between manufacturer identifiers and Entrez Gene identifiers.

Details

Each manufacturer identifier is mapped to a vector of Entrez Gene identifiers. An NA is assigned to those manufacturer identifiers that can not be mapped to an Entrez Gene identifier at this time.

If a given manufacturer identifier can be mapped to different Entrez Gene identifiers from various sources, we attempt to select the common identifiers. If a consensus cannot be determined, we select the smallest identifier.

Mappings were based on data provided by: Entrez Gene <ftp://ftp.ncbi.nlm.nih.gov/gene/DATA> With a date stamp from the source of: 2013-Sep12

References

<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene>

Examples

```
x <- lumiHumanAllENTREZID
# Get the probe identifiers that are mapped to an ENTREZ Gene ID
mapped_probes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_probes])
if(length(xx) > 0) {
  # Get the ENTREZID for the first five probes
  xx[1:5]
  # Get the first one
  xx[[1]]
}
```

lumiHumanAllENZYME *Maps between Manufacturer IDs and Enzyme Commission (EC) Numbers*

Description

lumiHumanAllENZYME is an R object that provides mappings between manufacturer identifiers and EC numbers. lumiHumanAllENZYME2PROBE is an R object that maps Enzyme Commission (EC) numbers to manufacturer identifiers.


```

# Now convert lumiHumanAllENZYME2PROBE to a list to see inside
xx <- as.list(lumiHumanAllENZYME2PROBE)
if(length(xx) > 0){
  # Get the probe identifiers for the first five enzyme
  #commission numbers
  xx[1:5]
  # Get the first one
  xx[[1]]
}

```

lumiHumanAllGENENAME *Map between Manufacturer IDs and Genes*

Description

lumiHumanAllGENENAME is an R object that maps manufacturer identifiers to the corresponding gene name.

Details

Each manufacturer identifier maps to a named vector containing the gene name. The vector name corresponds to the manufacturer identifier. If the gene name is unknown, the vector will contain an NA.

Gene names currently include both the official (validated by a nomenclature committee) and preferred names (interim selected for display) for genes. Efforts are being made to differentiate the two by adding a name to the vector.

Mappings were based on data provided by: Entrez Gene <ftp://ftp.ncbi.nlm.nih.gov/gene/DATA> With a date stamp from the source of: 2013-Sep12

Examples

```

x <- lumiHumanAllGENENAME
# Get the probe identifiers that are mapped to a gene name
mapped_probes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_probes])
if(length(xx) > 0) {
  # Get the GENENAME for the first five probes
  xx[1:5]
  # Get the first one
  xx[[1]]
}

```


mapped to the same manufacturer identifier more than once but the evidence code can be different. Mappings between Gene Ontology identifiers and Gene Ontology terms and other information are available in a separate data package named GO.

Whenever any of these mappings are cast as a data.frame, all the results will be output in an appropriate tabular form.

Mappings between manufacturer identifiers and GO information were obtained through their mappings to manufacturer identifiers. NAs are assigned to manufacturer identifiers that can not be mapped to any Gene Ontology information. Mappings between Gene Ontology identifiers and Gene Ontology terms and other information are available in a separate data package named GO.

All mappings were based on data provided by: Gene Ontology ftp://ftp.geneontology.org/pub/go/godatabase/archive/latest-lite/ With a date stamp from the source of: 20130907

References

<ftp://ftp.ncbi.nlm.nih.gov/gene/DATA/>

See Also

[lumiHumanAllGO2ALLPROBES](#).

Examples

```
x <- lumiHumanAllGO
# Get the manufacturer 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:
# Convert to a list
xx <- as.list(lumiHumanAllGO2PROBE)
if(length(xx) > 0){
  # Gets the manufacturer ids for the top 2nd and 3rd GO identifiers
  goids <- xx[2:3]
  # Gets the manufacturer ids for the first element of goids
  goids[[1]]
  # Evidence code for the mappings
  names(goids[[1]])
}
# Convert lumiHumanAllGO2ALLPROBES to a list
xx <- as.list(lumiHumanAllGO2ALLPROBES)
if(length(xx) > 0){
  # Gets the manufacturer identifiers for the top 2nd and 3rd GO identifiers
  goids <- xx[2:3]
  # Gets all the manufacturer identifiers for the first element of goids
```

```
goids[[1]]
# Evidence code for the mappings
names(goids[[1]])
}
```

`lumiHumanAllMAP`*Map between Manufacturer Identifiers and cytogenetic maps/bands*

Description

`lumiHumanAllMAP` is an R object that provides mappings between manufacturer identifiers and cytoband locations.

Details

Each manufacturer identifier is mapped to a vector of cytoband locations. The vector length may be one or longer, if there are multiple reported chromosomal locations for a given gene. An NA is reported for any manufacturer identifiers that cannot be mapped to a cytoband at this time.

Cytogenetic bands for most higher organisms are labeled p1, p2, p3, q1, q2, q3 (p and q are the p and q arms), etc., counting from the centromere out toward the telomeres. At higher resolutions, sub-bands can be seen within the bands. The sub-bands are also numbered from the centromere out toward the telomere. Thus, a label of 7q31.2 indicates that the band is on chromosome 7, q arm, band 3, sub-band 1, and sub-sub-band 2.

Mappings were based on data provided by: Entrez Gene ftp://ftp.ncbi.nlm.nih.gov/gene/DATA With a date stamp from the source of: 2013-Sep12

References

<http://www.ncbi.nlm.nih.gov>

Examples

```
x <- lumiHumanAllMAP
# Get the probe identifiers that are mapped to any cytoband
mapped_probes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_probes])
if(length(xx) > 0) {
  # Get the MAP for the first five probes
  xx[1:5]
  # Get the first one
  xx[[1]]
}
```

`lumiHumanAllMAPCOUNTS` *Number of mapped keys for the maps in package lumiHumanAll.db*

Description

`lumiHumanAllMAPCOUNTS` provides the "map count" (i.e. the count of mapped keys) for each map in package `lumiHumanAll.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

```
lumiHumanAllMAPCOUNTS
mapnames <- names(lumiHumanAllMAPCOUNTS)
lumiHumanAllMAPCOUNTS[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 lumiHumanAll.db
checkMAPCOUNTS("lumiHumanAll.db")
```

`lumiHumanAllOMIM` *Map between Manufacturer Identifiers and Mendelian Inheritance in Man (MIM) identifiers*

Description

`lumiHumanAllOMIM` is an R object that provides mappings between manufacturer identifiers and OMIM identifiers.


```

    # Get the probe identifiers for the first two pathway identifiers
    xx[1:2]
    # Get the first one
    xx[[1]]
  }

```

lumiHumanAllPFAM	<i>Map Manufacturer IDs to Pfam IDs</i>
------------------	---

Description

lumiHumanAllPFAM is an R object that provides mappings between a manufacturer identifier and the associated Pfam identifiers.

Details

Each manufacturer identifier maps to a named vector of Pfam identifiers. The name for each Pfam identifier is the IPI accession number where this Pfam identifier is found.

If the Pfam is a named NA, it means that the associated Entrez Gene id of this manufacturer identifier is found in an IPI entry of the IPI database, but there is no Pfam identifier in the entry.

If the Pfam is a non-named NA, it means that the associated Entrez Gene id of this manufacturer identifier is not found in any IPI entry of the IPI database.

Mappings were based on data provided by: Uniprot <http://www.UniProt.org/> With a date stamp from the source of: Mon Sep 16 12:05:19 2013

Examples

```

x <- lumiHumanAllPFAM
# Get the probe identifiers that are mapped to any Pfam ID
mapped_probes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_probes])
# randomly display 10 probes
sample(xx, 10)

```

lumiHumanAllPMID	<i>Maps between Manufacturer Identifiers and PubMed Identifiers</i>
------------------	---

Description

lumiHumanAllPMID is an R object that provides mappings between manufacturer identifiers and PubMed identifiers. lumiHumanAllPMID2PROBE is an R object that provides mappings between PubMed identifiers and manufacturer identifiers.

Details

When lumiHumanAllPMID is viewed as a list each manufacturer identifier is mapped to a named vector of PubMed identifiers. The name associated with each vector corresponds to the manufacturer identifier. The length of the vector may be one or greater, depending on how many PubMed identifiers a given manufacturer identifier is mapped to. An NA is reported for any manufacturer identifier that cannot be mapped to a PubMed identifier.

When lumiHumanAllPMID2PROBE is viewed as a list each PubMed identifier is mapped to a named vector of manufacturer identifiers. The name represents the PubMed identifier and the vector contains all manufacturer identifiers that are represented by that PubMed identifier. The length of the vector may be one or longer, depending on how many manufacturer identifiers are mapped to a given PubMed identifier.

Titles, abstracts, and possibly full texts of articles can be obtained from PubMed by providing a valid PubMed identifier. The pubmed function of annotate can also be used for the same purpose.

Mappings were based on data provided by: Entrez Gene <ftp://ftp.ncbi.nlm.nih.gov/gene/DATA> With a date stamp from the source of: 2013-Sep12

References

<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=PubMed>

Examples

```
x <- lumiHumanAllPMID
# Get the probe identifiers that are mapped to any PubMed ID
mapped_probes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_probes])
if(length(xx) > 0){
  # Get the PubMed identifiers for the first two probe identifiers
  xx[1:2]
  # Get the first one
  xx[[1]]
  if(interactive() && !is.null(xx[[1]]) && !is.na(xx[[1]]))
    && require(annotate){
      # Get article information as XML files
      xmls <- pubmed(xx[[1]], disp = "data")
      # View article information using a browser
      pubmed(xx[[1]], disp = "browser")
    }
}

# Now convert the reverse map object lumiHumanAllPMID2PROBE to a list
xx <- as.list(lumiHumanAllPMID2PROBE)
if(length(xx) > 0){
  # Get the probe identifiers for the first two PubMed identifiers
  xx[1:2]
  # Get the first one
  xx[[1]]
  if(interactive() && require(annotate)){
    # Get article information as XML files for a PubMed id
```

```
    xmls <- pubmed(names(xx)[1], disp = "data")
    # View article information using a browser
    pubmed(names(xx)[1], disp = "browser")
  }
}
```

lumiHumanAllPROSITE *Map Manufacturer IDs to PROSITE ID*

Description

lumiHumanAllPROSITE is an R object that provides mappings between a manufacturer identifier and the associated PROSITE identifiers.

Details

Each manufacturer identifier maps to a named vector of PROSITE identifiers. The name for each PROSITE identifier is the IPI accession number where this PROSITE identifier is found.

If the PROSITE is a named NA, it means that the associated Entrez Gene id of this manufacturer identifier is found in an IPI entry of the IPI database, but there is no PROSITE identifier in the entry.

If the PROSITE is a non-named NA, it means that the associated Entrez Gene id of this manufacturer identifier is not found in any IPI entry of the IPI database.

Mappings were based on data provided by: Uniprot <http://www.UniProt.org/> With a date stamp from the source of: Mon Sep 16 12:05:19 2013

Examples

```
x <- lumiHumanAllPROSITE
# Get the probe identifiers that are mapped to any PROSITE ID
mapped_probes <- mappedkeys(x)
# Convert to a list
xxx <- as.list(x[mapped_probes])
# randomly display 10 probes
xxx[sample(1:length(xxx), 10)]
```

lumiHumanAllREFSEQ *Map between Manufacturer Identifiers and RefSeq Identifiers*

Description

lumiHumanAllREFSEQ is an R object that provides mappings between manufacturer identifiers and RefSeq identifiers.

lumiHumanAllSYMBOL *Map between Manufacturer Identifiers and Gene Symbols*

Description

lumiHumanAllSYMBOL is an R object that provides mappings between manufacturer identifiers and gene abbreviations.

Details

Each manufacturer identifier is mapped to an abbreviation for the corresponding gene. An NA is reported if there is no known abbreviation for a given gene.

Symbols typically consist of 3 letters that define either a single gene (ABC) or multiple genes (ABC1, ABC2, ABC3). Gene symbols can be used as key words to query public databases such as Entrez Gene.

Mappings were based on data provided by: Entrez Gene <ftp://ftp.ncbi.nlm.nih.gov/gene/DATA> With a date stamp from the source of: 2013-Sep12

References

<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene>

Examples

```
x <- lumiHumanAllSYMBOL
# Get the probe identifiers that are mapped to a gene symbol
mapped_probes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_probes])
if(length(xx) > 0) {
  # Get the SYMBOL for the first five probes
  xx[1:5]
  # Get the first one
  xx[[1]]
}
```

lumiHumanAllUNIGENE *Map between Manufacturer Identifiers and UniGene cluster identifiers*

Description

lumiHumanAllUNIGENE is an R object that provides mappings between manufacturer identifiers and UniGene identifiers.

Details

Each manufacturer identifier is mapped to a UniGene identifier. An NA is reported if the manufacturer identifier cannot be mapped to UniGene at this time.

A UniGene identifier represents a cluster of sequences of a gene. Using UniGene identifiers one can query the UniGene database for information about the sequences or the Entrez Gene database for information about the genes.

Mappings were based on data provided by: Entrez Gene <ftp://ftp.ncbi.nlm.nih.gov/gene/DATA> With a date stamp from the source of: 2013-Sep12

References

<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene>

Examples

```
x <- lumiHumanAllUNIGENE
# Get the probe identifiers that are mapped to an UNIGENE ID
mapped_probes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_probes])
if(length(xx) > 0) {
  # Get the UNIGENE for the first five probes
  xx[1:5]
  # Get the first one
  xx[[1]]
}
```

`lumiHumanAllUNIPROT` *Map Uniprot accession numbers with Entrez Gene identifiers*

Description

`lumiHumanAllUNIPROT` is an R object that contains mappings between the manufacturer identifiers and Uniprot accession numbers.

Details

This object is a simple mapping of manufacturer identifiers to Uniprot Accessions.

Mappings were based on data provided by NCBI (link above) with an exception for fly, which required retrieving the data from ensembl <http://www.ensembl.org/biomart/martview/>

Examples

```
x <- lumiHumanAllUNIPROT
# Get the entrez gene IDs that are mapped to an Uniprot ID
mapped_genes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_genes])
if(length(xx) > 0) {
  # Get the Uniprot IDs for the first five genes
  xx[1:5]
  # Get the first one
  xx[[1]]
}
```

lumiHumanAll_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

```
lumiHumanAll_dbconn()
lumiHumanAll_dbfile()
lumiHumanAll_dbschema(file="", show.indices=FALSE)
lumiHumanAll_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

lumiHumanAll_dbconn returns a connection object to the package annotation DB. IMPORTANT: Don't call `dbDisconnect` on the connection object returned by lumiHumanAll_dbconn or you will break all the [AnnDbObj](#) objects defined in this package!

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

lumiHumanAll_dbschema prints the schema definition of the package annotation DB.

lumiHumanAll_dbInfo prints other information about the package annotation DB.

Value

`lumiHumanAll_dbconn`: a `DBIConnection` object representing an open connection to the package annotation DB.

`lumiHumanAll_dbfile`: a character string with the path to the package annotation DB.

`lumiHumanAll_dbschema`: none (invisible `NULL`).

`lumiHumanAll_dbInfo`: none (invisible `NULL`).

See Also

[dbGetQuery](#), [dbConnect](#), [dbconn](#), [dbfile](#), [dbschema](#), [dbInfo](#)

Examples

```
## Count the number of rows in the "probes" table:
dbGetQuery(lumiHumanAll_dbconn(), "SELECT COUNT(*) FROM probes")

## The connection object returned by lumiHumanAll_dbconn() was
## created with:
dbConnect(SQLite(), dbname=lumiHumanAll_dbfile(), cache_size=64000,
synchronous=0)

lumiHumanAll_dbschema()

lumiHumanAll_dbInfo()
```


lumiHumanAllPATH2PROBE
 (lumiHumanAllPATH), 14
lumiHumanAllPFAM, 15
lumiHumanAllPMID, 15
lumiHumanAllPMID2PROBE
 (lumiHumanAllPMID), 15
lumiHumanAllPROSITE, 17
lumiHumanAllREFSEQ, 17
lumiHumanAllSYMBOL, 19
lumiHumanAllUNIGENE, 19
lumiHumanAllUNIPROT, 20
lumiHumanAllUNIPROT2PROBE
 (lumiHumanAllUNIPROT), 20

mappedkeys, 12