mogene10stprobeset.db

October 16, 2009

R topics documented:

mogene10stprobesetACCNUM  2
mogene10stprobesetALIAS2PROBE  2
mogene10stprobeset.db  3
mogene10stprobesetCHRLengths  3
mogene10stprobesetCHRLoc  4
mogene10stprobesetCHR  5
mogene10stprobeset_ENSEMBL  5
mogene10stprobeset_ENTREZID  7
mogene10stprobesetENZYMENamed2Probes  8
mogene10stprobesetENZYMENAME  9
mogene10stprobesetGENENAME  10
mogene10stprobesetGO2AllProbes  11
mogene10stprobesetGO2Probe  12
mogene10stprobesetGO  13
mogene10stprobesetMAPCounts  14
mogene10stprobesetMAP  15
mogene10stprobesetMGI  16
mogene10stprobesetORGANISM  16
mogene10stprobesetPATH2Probe  17
mogene10stprobesetPATH  18
mogene10stprobesetPFAM  18
mogene10stprobesetPMID2Probe  19
mogene10stprobesetPMID  20
mogene10stprobesetPROSITE  21
mogene10stprobesetREFSEQ  22
mogene10stprobeset_SYMBOL  23
mogene10stprobesetUNIGENE  23
mogene10stprobesetUNIPROTK
mogene10stprobesetACCNUM

Map Manufacturer identifiers to GenBank Accession Numbers

Description

mogene10stprobesetACCNUM is an R object that contains mappings between a manufacturer’s identifiers and GenBank accession numbers.

Details


Examples

```r
x <- mogene10stprobesetACCNUM
# 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]]
}
```

mogene10stprobesetALIAS2PROBE

Map between Common Gene Symbol Identifiers and Manufacturer Identifiers

Description

mogene10stprobesetALIAS 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.

Examples

```r
# Convert the object to a list
xx <- as.list(mogene10stprobesetALIAS2PROBE)
if(length(xx) > 0){
  # Get the probe identifiers for the first two aliases
  xx[1:2]
  # Get the first one
  xx[[1]]
}
```

mogene10stprobeset.db

Bioconductor annotation data package

Description

Welcome to the mogene10stprobeset.db annotation Package. The purpose of this package is to provide detailed information about the mogene10stprobeset platform. This package is updated biannually.

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

```r
ls("package:mogene10stprobeset.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

```r
ls("package:mogene10stprobeset.db")
```

mogene10stprobesetCHRLENGTHS

A named vector for the length of each of the chromosomes

Description

mogene10stprobesetCHRLENGTHS provides the length measured in base pairs for each of the chromosomes.

Details

This is a named vector with chromosome numbers as the names and the corresponding lengths for chromosomes as the values.

Total lengths of chromosomes were derived by calculating the number of base pairs on the sequence string for each chromosome.
mogene10stprobesetCHRLOC

Map Manufacturer IDs to Chromosomal Location

Description

mogene10stprobesetCHRLOC is an R object that maps manufacturer identifiers to the starting position of the gene. The position of a gene is measured as the number of base pairs.

The CHRLOCEND mapping is the same as the CHRLOC mapping except that it specifies the ending base of a gene instead of the start.

Details

Each manufacturer identifier maps to a named vector of chromosomal locations, where the name indicates the chromosome. Due to inconsistencies that may exist at the time the object was built, these vectors may contain more than one chromosome and/or location. If the chromosomal location is unknown, the vector will contain an NA.

Chromosomal locations on both the sense and antisense strands are measured as the number of base pairs from the p (5’ end of the sense strand) to q (3’ end of the sense strand) arms. Chromosomal locations on the antisense strand have a leading “-” sign (e. g. -1234567).

Since some genes have multiple start sites, this field can map to multiple locations.

Mappings were based on data provided by: UCSC Genome Bioinformatics (Mus musculus) (ftp://hgdownload.cse.ucsc.edu/goldenPath/currentGenomes/Mus_musculus ) on 2008-Sep3

Examples

```r
x <- mogene10stprobesetCHRLOC
# Get the probe identifiers that are mapped to chromosome locations
mapped_probes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_probes])
if(length(xx) > 0) {
    # Get the CHRLOC for the first five probes
    xx[1:5]
    # Get the first one
    xx[[1]]
}
```
mogene10stprobesetCHR

Map Manufacturer IDs to Chromosomes

Description

mogene10stprobesetCHR 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.


Examples

```r
x <- mogene10stprobesetCHR
# 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]]
}
```

mogene10stprobeset_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

```r
mogene10stprobeset_dbconn()
mogene10stprobeset_dbfile()
mogene10stprobeset_dbschema(file="", show.indices=FALSE)
mogene10stprobeset_dbInfo()
```
mogene10stprobesetENSEMBL

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

- `mogene10stprobeset_dbconn` returns a connection object to the package annotation DB. IMPORTANT: Don’t call `dbDisconnect` on the connection object returned by `mogene10stprobeset_dbconn` or you will break all the AnnDbObj objects defined in this package!
- `mogene10stprobeset_dbfile` returns the path (character string) to the package annotation DB (this is an SQLite file).
- `mogene10stprobeset_dbschema` prints the schema definition of the package annotation DB.
- `mogene10stprobeset_dbInfo` prints other information about the package annotation DB.

Value

- `mogene10stprobeset_dbconn`: a DBIConnection object representing an open connection to the package annotation DB.
- `mogene10stprobeset_dbfile`: a character string with the path to the package annotation DB.
- `mogene10stprobeset_dbschema`: none (invisible NULL).
- `mogene10stprobeset_dbInfo`: none (invisible NULL).

See Also

- `dbGetQuery`, `dbConnect`, `dbconn`, `dbfile`, `dbschema`, `dbInfo`

Examples

```r
## Count the number of rows in the "probes" table:
dbGetQuery(mogene10stprobeset_dbconn(), "SELECT COUNT(*) FROM probes")

## The connection object returned by `mogene10stprobeset_dbconn()` was created with:
dbConnect(SQLite(), dbname=mogene10stprobeset_dbfile(), cache_size=64000, synchronous=0)

mogene10stprobeset_dbschema()
mogene10stprobeset_dbInfo()
```
Details


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

This mapping is a combination of NCBI to ensembl IDs from \texttt{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 is contains data from \texttt{BOTH} of these sources in an effort to maximize the chances that you will find a match.

Examples

```r
x <- mogene10stprobesetENSEMBL
# 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(mogene10stprobesetENSEMBL2PROBE)
if(length(xx) > 0) {
    # Gets the entrez gene IDs for the first five Ensembl IDs
    xx[1:5]
    # Get the first one
    xx[[1]]
}
```

mogene10stprobesetENTREZID

\textit{Map between Manufacturer Identifiers and Entrez Gene}

Description

mogene10stprobesetENTREZID 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 \texttt{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 concensus cannot be determined, we select the smallest identifier.

mogene10stprobesetENZYME2PROBE

Map between Enzyme Commission Numbers and Manufacturer Identifiers

mogene10stprobesetENZYME2PROBE

References


Examples

x <- mogene10stprobesetENTREZID
# 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]]
}

Description

mogene10stprobesetENZYME2PROBE is an R object that maps Enzyme Commission (EC) numbers to manufacturer identifiers.

Details

Each EC number maps to a named vector containing all of the manufacturer identifiers that correspond to the gene that produces that enzyme. The name of the vector corresponds to the EC number.

Enzyme Commission numbers are assigned by the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology http://www.chem.qmw.ac.uk/iubmb/enzyme/ to allow enzymes to be identified.

An Enzyme Commission number is of the format EC x.y.z.w, where x, y, z, and w are numeric numbers. In mogene10stprobesetENZYME2PROBE, EC is dropped from the Enzyme Commission numbers.

Enzyme Commission numbers have corresponding names that describe the functions of enzymes in such a way that EC x is a more general description than EC x.y that in turn is a more general description than EC x.y.z. The top level EC numbers and names are listed below:

EC 1 oxidoreductases
EC 2 transferases
EC 3 hydrolases
EC 4 lyases
EC 5 isomerases
EC 6 ligases

The EC name for a given EC number can be viewed at http://www.chem.qmul.ac.uk/iupac/jcbn/index.html#6
mogene10stprobesetENZYME

Mappings were based on data provided by: KEGG GENOME (ftp://ftp.genome.ad.jp/pub/kegg/genomes) on 2009-Mar10

References


Examples

# Convert to a list
xx <- as.list(mogene10stprobesetENZYME2PROBE)
if(length(xx) > 0){
    # Get the probe identifiers for the first five enzyme commission numbers
    xx[1:5]
    # Get the first one
    xx[[1]]
}

mogene10stprobesetENZYME

Map between Manufacturer IDs and Enzyme Commission (EC) Numbers

Description

mogene10stprobesetENZYME is an R object that provides mappings between manufacturer identifiers and EC numbers.

Details

Each manufacturer identifier maps to a named vector containing the EC number that corresponds to the enzyme produced by that gene. The name corresponds to the manufacturer identifier. If this information is unknown, the vector will contain an NA.

Enzyme Commission numbers are assigned by the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology http://www.chem.qmw.ac.uk/iubmb/enzyme/ to allow enzymes to be identified.

An Enzyme Commission number is of the format EC x.y.z.w, where x, y, z, and w are numeric numbers. In mogene10stprobesetENZYME2PROBE, EC is dropped from the Enzyme Commission numbers.

Enzyme Commission numbers have corresponding names that describe the functions of enzymes in such a way that EC x is a more general description than EC x.y that in turn is a more general description than EC x.y.z. The top level EC numbers and names are listed below:

EC 1 oxidoreductases
EC 2 transferases
EC 3 hydrolases
EC 4 lyases
EC 5 isomerases
EC 6 ligases
The EC name for a given EC number can be viewed at http://www.chem.qmul.ac.uk/iupac/jcbn/index.html#6

Mappings between probe identifiers and enzyme identifiers were obtained using files provided by: KEGG GENOME (ftp://ftp.genome.jp/pub/kegg/genomes) on 2009-Mar10

References

Examples

x <- mogene10stprobesetENZYME
# Get the probe identifiers that are mapped to an EC number
mapped_probes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_probes])
if(length(xx) > 0) {
    # Get the ENZYME for the first five probes
    xx[1:5]
    # Get the first one
    xx[[1]]
}

mogene10stprobesetGENENAME

Map between Manufacturer IDs and Genes

Description

mogene10stprobesetGENENAME 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.


Examples

x <- mogene10stprobesetGENENAME
# 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]]
}
mogene10stprobesetGO2ALLPROBES

Map between Gene Ontology (GO) Identifiers and all Manufacturer Identifiers in the subtree

Description

mogene10stprobesetGO2ALLPROBES is an R object that provides mappings between a given GO identifier and all manufacturer identifiers annotated at that GO term or one of its children in the GO ontology.

Details

GO consists of three ontologies—molecular function (MF), biological process (BP), and cellular component (CC). All ontologies are structured as directed acyclic graphs (DAGs). Each node in each DAG (tree) is a GO term (id) associated with a named vector of manufacturer identifiers. The name associated with each manufacturer id corresponds to the evidence code for that GO identifier. This object mogene10stprobesetGO2ALLPROBES maps between a given GO identifier and all manufacturer identifiers annotated at that GO term or one of its children in the GO ontology.

The evidence code indicates what kind of evidence supports the association between the GO and Entrez Gene identifiers. Evidence codes currently 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

A GO identifier may be 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.

Mappings were based on data provided by:


References

Examples

```r
# Convert to a list
xx <- as.list(mogene10stprobesetGO2ALLPROBES)
if(length(xx) > 0){
  # Get the probe identifiers for the top 2nd and 3rd GO identifiers
goids <- xx[2:3]
  # Get all the probe identifiers for the first element of goids
goids[[1]]
  # Evidence code for the mappings
  names(goids[[1]])
}
```

mogene10stprobesetGO2PROBE

*Map between Gene Ontology (GO) and Manufacturer Identifiers*

Description

mogene10stprobesetGO2PROBE is an R object that provides mappings between GO identifiers and manufacturer identifiers.

Details

Each GO term maps to a named vector of manufacturer identifiers. The name associated with each manufacturer identifier corresponds to the evidence code for that GO identifier. The evidence code indicates what kind of evidence supports the association between the GO and Entrez Gene identifiers. Evidence codes currently 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

A GO identifier may be mapped to the same probe 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.

Mappings were based on data provided by:


References

mogene10stprobesetGO

Map between Manufacturer IDs and Gene Ontology (GO)

Examples

# Convert to a list
xx <- as.list(mogene10stprobesetGO2PROBE)
if(length(xx) > 0){
  # Get the probe identifiers for the top 2nd and 3rd GO identifiers
goids <- xx[2:3]
  # Get the probe identifiers for the first element of goids
goids[[1]]
  # Evidence code for the mappings
names(goids[[1]])
}

mogene10stprobesetGO

Map between Manufacturer IDs and Gene Ontology (GO)

Description

mogene10stprobesetGO is an R object that provides mappings between manufacturer identifiers and the GO identifiers that they are directly associated with.

Details

Each Entrez Gene identifier is mapped to a list of lists. The names on the outer list are GO identifiers. Each inner list consists of three named elements: GOID, Ontology, and Evidence.

The GOID element matches the GO identifier named in the outer list and is included for convenience when processing the data using 'lapply'.

The Ontology element indicates which of the three Gene Ontology categories this identifier belongs to. The categories are biological process (BP), cellular component (CC), and molecular function (MF).

The Evidence element contains a code indicating what kind of evidence supports the association of the GO identifier to the Entrez Gene 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 between probe identifiers and GO information were obtained through their mappings to Entrez Gene identifiers. NAs are assigned to probe 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.
Mappings were based on data provided by: Gene Ontology (ftp://ftp.geneontology.org/pub/go/godatabase/archive/latest) on 200903

Examples

```r
x <- mogene10stprobesetGO
# Get the probe identifiers that are mapped to a GO ID
mapped_probes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_probes])
if(length(xx) > 0) {
    # Try the first one
    got <- xx[[1]]
    got[[1]]"GOID"
    got[[1]]"Ontology"
    got[[1]]"Evidence"
}
```

mogene10stprobesetMAPCOUNTS

**Number of mapped keys for the maps in package mogene10stprobeset.db**

Description

mogene10stprobesetMAPCOUNTS provides the "map count" (i.e. the count of mapped keys) for each map in package mogene10stprobeset.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

```r
mogene10stprobesetMAPCOUNTS
mapnames <- names(mogene10stprobesetMAPCOUNTS)
mogene10stprobesetMAPCOUNTS[mapnames[1]]
x <- get(mapnames[1])
sum(!is.na(as.list(x)))
count.mappedkeys(x)  # much faster!
```
mogene10stprobesetMAP

Map between Manufacturer Identifiers and cytogenetic maps/bands

Description

mogene10stprobesetMAP 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.

The physical location of each band on a chromosome can be obtained from another object named "organism"CYTOLOC in a separate data package for human(humanCHRLOC), mouse(mouseCHRLOC), and rat(ratCHRLOC).


References


Examples

x <- mogene10stprobesetMAP
# 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]]
}
mogene10stprobesetMGI

Map MGI gene accession numbers with manufacturer identifiers

Description

mogene10stprobesetMGI is an R object that contains mappings between manufacturer identifiers and Jackson Laboratory MGI gene accession numbers.

Details


Examples

```r
x <- mogene10stprobesetMGI
# Get the manufacturer IDs that are mapped to an MGI ID
mapped_genes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_genes])
if(length(xx) > 0) {
  # Get the MGI IDs for the first five genes
  xx[1:5]
  # Get the first one
  xx[[1]]
}
# For the reverse map MGI2EG:
# Convert to a list
xx <- as.list(mogene10stprobesetMGI2PROBE)
if(length(xx) > 0){
  # Gets the manufacturer IDs for the first five MGI IDs
  xx[1:5]
  # Get the first one
  xx[[1]]
}
```

mogene10stprobesetORGANISM

The Organism for mogene10stprobeset

Description

mogene10stprobesetORGANISM is an R object that contains a single item: a character string that names the organism for which mogene10stprobeset was built.

Details

Although the package name is suggestive of the organism for which it was built, mogene10stprobesetORGANISM provides a simple way to programmatically extract the organism name.
**mogene10stprobesetPATH2PROBE**

*Map between Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway identifiers and Manufacturer Identifiers*

**Description**

*mogene10stprobesetPATH2PROBE* is an R object that provides mappings between KEGG identifiers and manufacturer identifiers.

**Details**

Each KEGG identifier is mapped to a named vector of manufacturer identifiers. The name represents the KEGG identifier and the vector contains all manufacturer identifiers that are found in that particular pathway. An `NA` is reported for any KEGG identifier that cannot be mapped to any manufacturer identifiers.

Pathway name for a given pathway identifier can be obtained using the KEGG data package that can either be built using AnnBuilder or downloaded from Bioconductor [http://www.bioconductor.org](http://www.bioconductor.org).


**References**


**Examples**

```r
# Convert the object to a list
xx <- as.list(mogene10stprobesetPATH2PROBE)
if(length(xx) > 0){
    # Get the probe identifiers for the first two pathway identifiers
    xx[1:2]
    # Get the first one
    xx[[1]]
}
```
mogene10stprobesetPATH

*Mappings between probe identifiers and KEGG pathway identifiers*

**Description**

KEGG (Kyoto Encyclopedia of Genes and Genomes) maintains pathway data for various organisms. `mogene10stprobesetPATH` maps probe identifiers to the identifiers used by KEGG for pathways in which the genes represented by the probe identifiers are involved.

**Details**

Each KEGG pathway has a name and identifier. Pathway name for a given pathway identifier can be obtained using the KEGG data package that can either be built using AnnBuilder or downloaded from Bioconductor [http://www.bioconductor.org](http://www.bioconductor.org).


**References**


**Examples**

```r
x <- mogene10stprobesetPATH
# Get the probe identifiers that are mapped to a KEGG pathway ID
mapped_probes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_probes])
if(length(xx) > 0) {
  # Get the PATH for the first five probes
  xx[1:5]
  # Get the first one
  xx[[1]]
}
```

---

mogene10stprobesetPFAM

*Map Manufacturer IDs to Pfam IDs*

**Description**

`mogene10stprobesetPFAM` 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: The International Protein Index (ftp://ftp.ebi.ac.uk/pub/databases/IPI/current) on 2009-Mar03

References


Examples

```r
x <- mogene10stprobesetPFAM
# 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)
```

mogene10stprobesetPMID2PROBE

Map between PubMed Identifiers and Manufacturer Identifiers

Description

mogene10stprobesetPMID2PROBE is an R object that provides mappings between PubMed identifiers and manufacturer identifiers.

Details

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.

Examples

```r
xx <- as.list(mogene10stprobesetPMID2PROBE)
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")
  }
}
```

mogene10stprobesetPMID

*Map between Manufacturer Identifiers and PubMed Identifiers*

Description

mogene10stprobesetPMID is an R object that provides mappings between manufacturer identifiers and PubMed identifiers.

Details

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.

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.


References


Examples

```r
x <- mogene10stprobesetPMID
# 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
```
mogene10stprobesetPROSITE

Map Manufacturer IDs to PROSITE ID

Description

mogene10stprobesetPROSITE 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 numbe 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: The International Protein Index (ftp://ftp.ebi.ac.uk/pub/databases/IPI/current) on 2009-Mar03

References


Examples

```r
x <- mogene10stprobesetPROSITE
# 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)]
```
mogene10stprobesetREFSEQ

*Map between Manufacturer Identifiers and RefSeq Identifiers*

**Description**

mogene10stprobesetREFSEQ is an R object that provides mappings between manufacturer identifiers and RefSeq identifiers.

**Details**

Each manufacturer identifier is mapped to a named vector of RefSeq identifiers. The name represents the manufacturer identifier and the vector contains all RefSeq identifiers that can be mapped to that manufacturer identifier. The length of the vector may be one or greater, depending on how many RefSeq identifiers a given manufacturer identifier can be mapped to. An NA is reported for any manufacturer identifier that cannot be mapped to a RefSeq identifier at this time.

RefSeq identifiers differ in format according to the type of record the identifiers are for as shown below:

- NG_XXXXX: RefSeq accessions for genomic region (nucleotide) records
- NM_XXXXX: RefSeq accessions for mRNA records
- NC_XXXXX: RefSeq accessions for chromosome records
- NP_XXXXX: RefSeq accessions for protein records
- XR_XXXXX: RefSeq accessions for model RNAs that are not associated with protein products
- XM_XXXXX: RefSeq accessions for model mRNA records
- XP_XXXXX: RefSeq accessions for model protein records

Where XXXXX is a sequence of integers.


**References**


**Examples**

```r
x <- mogene10stprobesetREFSEQ  
# Get the probe identifiers that are mapped to any RefSeq ID
mapped_probes <- mappedkeys(x)  
# Convert to a list
xx <- as.list(x[mapped_probes])
if(length(xx) > 0) {
  # Get the REFSEQ for the first five probes
  xx[1:5]
  # Get the first one
  xx[[1]]
}
```
mogene10stprobesetSYMBOL

Map between Manufacturer Identifiers and Gene Symbols

Description

mogene10stprobesetSYMBOL 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.


References


Examples

```r
x <- mogene10stprobesetSYMBOL
# 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]]
}
```

mogene10stprobesetUNIGENE

Map between Manufacturer Identifiers and UniGene cluster identifiers

Description

mogene10stprobesetUNIGENE 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.


References


Examples

```r
x <- mogene10stprobesetUNIGENE
# 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]]
}
```

mogene10stprobesetUNIPROT

Map Uniprot accession numbers with Entrez Gene identifiers

Description

mogene10stprobesetUNIPROT is an R object that contains mappings between Entrez Gene identifiers and Uniprot accession numbers.

Details


Examples

```r
x <- mogene10stprobesetUNIPROT
# 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]]
Index

*Topic datasets
mogene10stprobeset.db, 2
mogene10stprobeset_dbconn, 5
mogene10stprobesetACCNUM, 1
mogene10stprobesetALIAS2PROBE, 2
mogene10stprobesetCHR, 4
mogene10stprobesetCHRENGTHS, 3
mogene10stprobesetCHRLOC, 3
mogene10stprobesetENSEMBL, 6
mogene10stprobesetENTREZID, 7
mogene10stprobesetENZYME, 9
mogene10stprobesetENZYME2PROBE, 8
mogene10stprobesetGENENAME, 10
mogene10stprobesetGO, 13
mogene10stprobesetGO2ALLPROBES, 11
mogene10stprobesetGO2PROBE, 12
mogene10stprobesetMAP, 15
mogene10stprobesetMAPCOUNTS, 14
mogene10stprobesetMGI, 16
mogene10stprobesetORGANISM, 16
mogene10stprobesetPATH, 18
mogene10stprobesetPATH2PROBE, 17
mogene10stprobesetPFAM, 18
mogene10stprobesetPMID, 20
mogene10stprobesetPMID2PROBE, 19
mogene10stprobesetPROSITE, 21
mogene10stprobesetREFSEQ, 22
mogene10stprobesetSYMBOL, 23
mogene10stprobesetUNIGENE, 23
mogene10stprobesetUNIPROT, 24

*Topic utilities
mogene10stprobeset_dbconn, 5
AnnDbObj, 5

cat, 5
cHECKMAPCOUNTS, 14
count.mappedkeys, 14
dbconn, 5
dbConnect, 5
dbDisconnect, 5
dbfile, 5
dbGetQuery, 5
dbInfo, 5
dbschema, 5

mappedkeys, 14
mogene10stprobeset
  (mogene10stprobeset.db), 2
mogene10stprobeset.db, 2
mogene10stprobeset_dbconn, 5
mogene10stprobeset_dbfile
  (mogene10stprobeset_dbconn), 5
mogene10stprobeset_dbinfo
  (mogene10stprobeset_dbconn), 5
mogene10stprobeset_dbschema
  (mogene10stprobeset_dbconn), 5
mogene10stprobeset_ACCNUM, 1
mogene10stprobesetALIAS2PROBE, 2
mogene10stprobesetCHR, 4
mogene10stprobesetCHRENGTHS, 3
mogene10stprobesetCHRLOC, 3
mogene10stprobesetCHRLOCEND
  (mogene10stprobesetCHRLOC), 3
mogene10stprobesetENSEMBL, 6
mogene10stprobesetENSEMBL2PROBE
  (mogene10stprobesetENSEMBL), 6
mogene10stprobesetENTREZID, 7
mogene10stprobesetENSEMBL, 6
mogene10stprobesetENZYME2PROBE, 8
mogene10stprobesetGENENAME, 10
mogene10stprobesetGO, 13
INDEX

mogene10stprobesetGO2ALLPROBES, 11
mogene10stprobesetGO2PROBE, 12
mogene10stprobesetLOCUSID
(mogene10stprobesetENTREZID), 7
mogene10stprobesetMAP, 15
mogene10stprobesetMAPCOUNTS, 14
mogene10stprobesetMGI, 16
mogene10stprobesetMGI2PROBE
(mogene10stprobesetMGI), 16
mogene10stprobesetORGANISM, 16
mogene10stprobesetPATH, 18
mogene10stprobesetPATH2PROBE, 17
mogene10stprobesetPFAM, 18
mogene10stprobesetPMID, 20
mogene10stprobesetPMID2PROBE, 19
mogene10stprobesetPROSITE, 21
mogene10stprobesetREFSEQ, 22
mogene10stprobesetSYMBOL, 23
mogene10stprobesetUNIGENE, 23
mogene10stprobesetUNIPROT, 24
mogene10stprobesetUNIPROT2PROBE
(mogene10stprobesetUNIPROT), 24