

ecoli2.db

April 29, 2025

ecoli2ACCNUM

Map Manufacturer identifiers to Accession Numbers

Description

ecoli2ACCNUM 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: 2021-Apr14

See Also

- [AnnotationDb-class](#) for use of the `select()` interface.

Examples

```
## select() interface:
## Objects in this package can be accessed using the select() interface
## from the AnnotationDbi package. See ?select for details.

## Bimap interface:
x <- ecoli2ACCNUM
# Get the probe identifiers that are mapped to an ACCNUM
mapped_probes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_probes][1:300])
if(length(xx) > 0) {
  # Get the ACCNUM for the first five probes
  xx[1:5]
  # Get the first one
  xx[[1]]
}
```

ecoli2ALIAS2PROBE	<i>Map between Common Gene Symbol Identifiers and Manufacturer Identifiers</i>
-------------------	--

Description

ecoli2ALIAS 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: 2021-Apr14

See Also

- [AnnotationDb-class](#) for use of the `select()` interface.

Examples

```
## select() interface:
## Objects in this package can be accessed using the select() interface
## from the AnnotationDbi package. See ?select for details.

## Bimap interface:
# Convert the object to a list
xx <- as.list(ecoli2ALIAS2PROBE)
if(length(xx) > 0){
  # Get the probe identifiers for the first two aliases
  xx[1:2]
  # Get the first one
  xx[[1]]
}
```

ecoli2.db

Bioconductor annotation data package

Description

Welcome to the ecoli2.db annotation Package. The purpose of this package is to provide detailed information about the ecoli2 platform. This package is updated biannually.

Objects in this package are accessed using the `select()` interface. See `?select` in the AnnotationDbi package for details.

See Also

- [AnnotationDb-class](#) for use of keys(), columns() and select().

Examples

```
## select() interface:
## Objects in this package can be accessed using the select() interface
## from the AnnotationDbi package. See ?select for details.
columns(ecoli2.db)

## Bimap interface:
## The 'old style' of interacting with these objects is manipulation as
## bimap. While this approach is still available we strongly encourage the
## use of select().
ls("package:ecoli2.db")
```

ecoli2ENTREZID

Map between Manufacturer Identifiers and Entrez Gene

Description

ecoli2ENTREZID 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: 2021-Apr14

References

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

See Also

- [AnnotationDb-class](#) for use of the select() interface.

Examples

```
## select() interface:
## Objects in this package can be accessed using the select() interface
## from the AnnotationDbi package. See ?select for details.

## Bimap interface:
x <- ecoli2ENTREZID
# 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][1:300])
if(length(xx) > 0) {
  # Get the ENTREZID for the first five probes
  xx[1:5]
  # Get the first one
  xx[[1]]
}
```

ecoli2ENZYME

Maps between Manufacturer IDs and Enzyme Commission (EC) Numbers

Description

ecoli2ENZYME is an R object that provides mappings between manufacturer identifiers and EC numbers. ecoli2ENZYME2PROBE is an R object that maps Enzyme Commission (EC) numbers to manufacturer identifiers.

Details

When the ecoli2ENZYME mapping is viewed as a list, each manufacturer identifier maps to a named vector containing the EC number that corresponds to the enzyme produced by that gene. The names correspond to the manufacturer identifiers. If this information is unknown, the vector will contain an NA.

For the ecoli2ENZYME2PROBE, 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 ecoli2ENZYME2PROBE, 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> With a date stamp from the source of: 2011-Mar15

References

<ftp://ftp.genome.ad.jp/pub/kegg/pathways>

See Also

- [AnnotationDb-class](#) for use of the `select()` interface.

Examples

```
## select() interface:
## Objects in this package can be accessed using the select() interface
## from the AnnotationDbi package. See ?select for details.

## Bimap interface:
x <- ecoli2ENZYME
# 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][1:3])
if(length(xx) > 0) {
  # Get the ENZYME for the first five probes
  xx[1:5]
  # Get the first one
  xx[[1]]
}

# Now convert ecoli2ENZYME2PROBE to a list to see inside
x <- ecoli2ENZYME2PROBE
mapped_probes <- mappedkeys(x)
## convert to a list
xx <- as.list(x[mapped_probes][1:3])
if(length(xx) > 0){
  # Get the probe identifiers for the first five enzyme
  #commission numbers
  xx[1:5]
  # Get the first one
  xx[[1]]
}
```

ecoli2GENENAME

Map between Manufacturer IDs and Genes

Description

ecoli2GENENAME 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: 2021-Apr14

See Also

- [AnnotationDb-class](#) for use of the `select()` interface.

Examples

```
## select() interface:
## Objects in this package can be accessed using the select() interface
## from the AnnotationDbi package. See ?select for details.

## Bimap interface:
x <- ecoli2GENENAME
# 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][1:300])
if(length(xx) > 0) {
  # Get the GENENAME for the first five probes
  xx[1:5]
  # Get the first one
  xx[[1]]
}
```

ecoli2GO

Maps between manufacturer IDs and Gene Ontology (GO) IDs

Description

ecoli2GO is an R object that provides mappings between manufacturer identifiers and the GO identifiers that they are directly associated with. This mapping and its reverse mapping (ecoli2GO2PROBE) do NOT associate the child terms from the GO ontology with the gene. Only the directly evidenced terms are represented here.

ecoli2GO2ALLPROBES is an R object that provides mappings between a given GO identifier and all of the manufacturer identifiers annotated at that GO term OR TO ONE OF IT'S CHILD NODES in the GO ontology. Thus, this mapping is much larger and more inclusive than ecoli2GO2PROBE.

Details

If ecoli2GO is cast as a list, each manufacturer 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 manufacturer id. Some of 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

A more complete listing of evidence codes can be found at:

<http://www.geneontology.org/GO.evidence.shtml>

If `ecoli2GO2ALLPROBES` or `ecoli2GO2PROBE` is cast as a list, each GO term maps to a named vector of manufacturer identifiers and evidence codes. 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`.

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 <http://current.geneontology.org/ontology/go-basic.obo> With a date stamp from the source of: 2021-02-01

References

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

See Also

- [ecoli2GO2ALLPROBES](#)
- [AnnotationDb-class](#) for use of the `select()` interface.

Examples

```
## select() interface:
## Objects in this package can be accessed using the select() interface
## from the AnnotationDbi package. See ?select for details.

## Bimap interface:
x <- ecoli2GO
# 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][1:300])
if(length(xx) > 0) {
  # Try the first one
  got <- xx[[1]]
  got[[1]][["GOID"]]
  got[[1]][["Ontology"]]
  got[[1]][["Evidence"]]
}
# For the reverse map:
x <- ecoli2G02PROBE
mapped_genes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_genes][1:300])
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]])
}

x <- ecoli2G02ALLPROBES
mapped_genes <- mappedkeys(x)
# Convert ecoli2G02ALLPROBES to a list
xx <- as.list(x[mapped_genes][1:300])
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]])
}

```

ecoli2MAPCOUNTS

Number of mapped keys for the maps in package ecoli2.db

Description

DEPRECATED. Counts in the MAPCOUNT table are out of sync and should not be used.

ecoli2MAPCOUNTS provides the "map count" (i.e. the count of mapped keys) for each map in package ecoli2.db.

Details

DEPRECATED. Counts in the MAPCOUNT table are out of sync and should not be used.

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.

ecoli2ORGANISM

The Organism information for ecoli2

Description

ecoli2ORGANISM is an R object that contains a single item: a character string that names the organism for which ecoli2 was built. ecoli2ORGPKG is an R object that contains a character vector with the name of the organism package that a chip package depends on for its gene-centric annotation.

Details

Although the package name is suggestive of the organism for which it was built, ecoli2ORGANISM provides a simple way to programmatically extract the organism name. ecoli2ORGPKG provides a simple way to programmatically extract the name of the parent organism package. The parent organism package is a strict dependency for chip packages as this is where the gene centric information is ultimately extracted from. The full package name will always be this string plus the extension ".db". But most programatic acces will not require this extension, so its more convenient to leave it out.

See Also

- [AnnotationDb-class](#) for use of the `select()` interface.

Examples

```
## select() interface:
## Objects in this package can be accessed using the select() interface
## from the AnnotationDbi package. See ?select for details.

## Bimap interface:
ecoli2ORGANISM
ecoli2ORGPKG
```

ecoli2PATH

Mappings between probe identifiers and KEGG pathway identifiers

Description

KEGG (Kyoto Encyclopedia of Genes and Genomes) maintains pathway data for various organisms.

ecoli2PATH maps probe identifiers to the identifiers used by KEGG for pathways in which the genes represented by the probe identifiers are involved

ecoli2PATH2PROBE is an R object that provides mappings between KEGG identifiers and manufacturer identifiers.

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>.

Graphic presentations of pathways are searchable at url <http://www.genome.ad.jp/kegg/pathway.html> by using pathway identifiers as keys.

Mappings were based on data provided by: KEGG GENOME <ftp://ftp.genome.jp/pub/kegg/genomes>
With a date stamp from the source of: 2011-Mar15

References

<http://www.genome.ad.jp/kegg/>

See Also

- [AnnotationDb-class](#) for use of the `select()` interface.

Examples

```
## select() interface:
## Objects in this package can be accessed using the select() interface
## from the AnnotationDbi package. See ?select for details.

## Bimap interface:
x <- ecoli2PATH
# 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][1:300])
if(length(xx) > 0) {
  # Get the PATH for the first five probes
  xx[1:5]
  # Get the first one
  xx[[1]]
}

x <- ecoli2PATH
mapped_probes <- mappedkeys(x)
# Now convert the ecoli2PATH2PROBE object to a list
xx <- as.list(x[mapped_probes][1:300])
if(length(xx) > 0){
  # Get the probe identifiers for the first two pathway identifiers
  xx[1:2]
  # Get the first one
  xx[[1]]
}
```

Description

ecoli2PMID is an R object that provides mappings between manufacturer identifiers and PubMed identifiers. ecoli2PMID2PROBE is an R object that provides mappings between PubMed identifiers and manufacturer identifiers.

Details

When ecoli2PMID 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 ecoli2PMID2PROBE 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: 2021-Apr14

References

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

See Also

- [AnnotationDb-class](#) for use of the select() interface.

Examples

```
## select() interface:
## Objects in this package can be accessed using the select() interface
## from the AnnotationDbi package. See ?select for details.

## Bimap interface:
x <- ecoli2PMID
# 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][1:300])
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")
  }
}
```


References

<https://www.ncbi.nlm.nih.gov> <https://www.ncbi.nlm.nih.gov/RefSeq/>

See Also

- [AnnotationDb-class](#) for use of the `select()` interface.

Examples

```
## select() interface:
## Objects in this package can be accessed using the select() interface
## from the AnnotationDbi package. See ?select for details.

## Bimap interface:
x <- ecoli2REFSEQ
# 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][1:300])
if(length(xx) > 0) {
  # Get the REFSEQ for the first five probes
  xx[1:5]
  # Get the first one
  xx[[1]]
}
```

ecoli2SYMBOL

Map between Manufacturer Identifiers and Gene Symbols

Description

ecoli2SYMBOL 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: 2021-Apr14

References

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

See Also

- [AnnotationDb-class](#) for use of the `select()` interface.

Examples

```
## select() interface:
## Objects in this package can be accessed using the select() interface
## from the AnnotationDbi package. See ?select for details.

## Bimap interface:
x <- ecoli2SYMBOL
# 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][1:300])
if(length(xx) > 0) {
  # Get the SYMBOL for the first five probes
  xx[1:5]
  # Get the first one
  xx[[1]]
}
```

ecoli2_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

```
ecoli2_dbconn()
ecoli2_dbfile()
ecoli2_dbschema(file="", show.indices=FALSE)
ecoli2_dbInfo()
```

Arguments

<code>file</code>	A connection, or a character string naming the file to print to (see the <code>file</code> argument of the <code>cat</code> function for the details).
<code>show.indices</code>	The CREATE INDEX statements are not shown by default. Use <code>show.indices=TRUE</code> to get them.

Details

`ecoli2_dbconn` returns a connection object to the package annotation DB. **IMPORTANT:** Don't call `dbDisconnect` on the connection object returned by `ecoli2_dbconn` or you will break all the `AnnDbObj` objects defined in this package!

`ecoli2_dbfile` returns the path (character string) to the package annotation DB (this is an SQLite file).

`ecoli2_dbschema` prints the schema definition of the package annotation DB.

`ecoli2_dbInfo` prints other information about the package annotation DB.

Value

ecoli2_dbconn: a DBIConnection object representing an open connection to the package annotation DB.

ecoli2_dbfile: a character string with the path to the package annotation DB.

ecoli2_dbschema: none (invisible NULL).

ecoli2_dbInfo: none (invisible NULL).

See Also

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

Examples

```
library(DBI)
## Count the number of rows in the "probes" table:
dbGetQuery(ecoli2_dbconn(), "SELECT COUNT(*) FROM probes")

ecoli2_dbschema()

ecoli2_dbInfo()
```

Index

* datasets

ecoli2.db, [2](#)
ecoli2_dbconn, [14](#)
ecoli2ACCNUM, [1](#)
ecoli2ALIAS2PROBE, [2](#)
ecoli2ENTREZID, [3](#)
ecoli2ENZYME, [4](#)
ecoli2GENENAME, [5](#)
ecoli2G0, [6](#)
ecoli2MAPCOUNTS, [8](#)
ecoli2ORGANISM, [9](#)
ecoli2PATH, [9](#)
ecoli2PMID, [10](#)
ecoli2REFSEQ, [12](#)
ecoli2SYMBOL, [13](#)

* utilities

ecoli2_dbconn, [14](#)

AnnDbObj, [14](#)

cat, [14](#)

checkMAPCOUNTS, [8](#)

dbconn, [15](#)

dbConnect, [15](#)

dbDisconnect, [14](#)

dbfile, [15](#)

dbGetQuery, [15](#)

dbInfo, [15](#)

dbschema, [15](#)

ecoli2 (ecoli2.db), [2](#)

ecoli2.db, [2](#)

ecoli2_dbconn, [14](#)

ecoli2_dbfile (ecoli2_dbconn), [14](#)

ecoli2_dbInfo (ecoli2_dbconn), [14](#)

ecoli2_dbschema (ecoli2_dbconn), [14](#)

ecoli2ACCNUM, [1](#)

ecoli2ALIAS2PROBE, [2](#)

ecoli2ENTREZID, [3](#)

ecoli2ENZYME, [4](#)

ecoli2ENZYME2PROBE (ecoli2ENZYME), [4](#)

ecoli2GENENAME, [5](#)

ecoli2G0, [6](#)

ecoli2G02ALLPROBES, [7](#)

ecoli2G02ALLPROBES (ecoli2G0), [6](#)

ecoli2G02PROBE (ecoli2G0), [6](#)

ecoli2LOCUSID (ecoli2ENTREZID), [3](#)

ecoli2MAPCOUNTS, [8](#)

ecoli2ORGANISM, [9](#)

ecoli2ORGPKG (ecoli2ORGANISM), [9](#)

ecoli2PATH, [9](#)

ecoli2PATH2PROBE (ecoli2PATH), [9](#)

ecoli2PMID, [10](#)

ecoli2PMID2PROBE (ecoli2PMID), [10](#)

ecoli2REFSEQ, [12](#)

ecoli2SYMBOL, [13](#)