Package ‘ginmappeR’

May 28, 2024

**Type** Package

**Title** Gene Identifier Mapper

**Version** 1.0.0


**License** GPL-3 + file LICENSE

**Encoding** UTF-8

**VignetteBuilder** knitr

**biocViews** Annotation, KEGG, Genetics, ThirdPartyClient, Software

**Imports** KEGGREST, UniProt.ws, XML, rentrez, httr, utils, memoise, cachem

**Suggests** RUnit, BiocGenerics, markdown, knitr

**git_url** https://git.bioconductor.org/packages/ginmappeR

**git_branch** RELEASE_3_19

**git_last_commit** f482726

**git_last_commit_date** 2024-04-30

**Repository** Bioconductor 3.19

**Date/Publication** 2024-05-28

**Author** Fernando Sola [aut, cre] (<https://orcid.org/0000-0003-1685-3270>),
  Daniel Ayala [aut] (<https://orcid.org/0000-0003-2095-1009>),
  Marina Pulido [aut] (<https://orcid.org/0000-0003-1633-4294>),
  Rafael Ayala [aut] (<https://orcid.org/0000-0002-9332-4623>),
  Lorena López-Cerero [aut] (<https://orcid.org/0000-0001-8950-4384>),
  Inma Hernández [aut] (<https://orcid.org/0000-0001-8920-6635>),
  David Ruiz [aut] (<https://orcid.org/0000-0003-4460-5493>)

**Maintainer** Fernando Sola <fsola@us.es>
Contents

changeCARDPath . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 2
getCARD2KEGG . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 3
getCARD2NCBIGene . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 4
getCARD2NCBInucleotide . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5
getCARD2NCBIProtein . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5
getCARD2UniProt . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6
getKEGG2CARD . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 7
getKEGG2NCBIGene . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 8
getKEGG2NCBInucleotide . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 9
getKEGG2NCBIProtein . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 10
getKEGG2UniProt . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 11
getNCBIGene2CARD . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 12
getNCBIGene2KEGG . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 12
getNCBIGene2NCBInucleotide . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 13
getNCBIGene2NCBIProtein . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 14
getNCBIGene2UniProt . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 15
getNCBIIdenticalProteins . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 16
getNCBInucleotide2CARD . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 17
getNCBInucleotide2KEGG . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 18
getNCBInucleotide2NCBIGene . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 19
getNCBInucleotide2NCBIProtein . . . . . . . . . . . . . . . . . . . . . . . . . . . . 20
getNCBInucleotide2UniProt . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 20
getNCBIProtein2CARD . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 21
getNCBIProtein2KEGG . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 22
getNCBIProtein2NCBIGene . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 23
getNCBIProtein2NCBInucleotide . . . . . . . . . . . . . . . . . . . . . . . . . . . . 24
getNCBIProtein2NCBIProtein . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 25
getUniProt2CARD . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 26
getUniProt2KEGG . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 27
getUniProt2NCBIGene . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 28
getUniProt2NCBInucleotide . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 29
getUniProt2NCBIProtein . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 30
getUniProtSimilarGenes . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 31
updateCARDDataBase . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 32

Index 33

changeCARDPath  Change path to download CARD database

Description

Defines a new path to the folder where to download CARD database. By default, this path is the temporary folder tempdir().
getCARD2KEGG

Usage

changeCARDPath(path=tempdir())

Arguments

path valid route to desired folder. Default value: tempdir()

Value

Updated path for CARD database to be downloaded.

getCARD2KEGG Translates a CARD database ID to KEGG database

Description

Retrieves the ID translation of an specific instance of CARD database ID to the KEGG Genes database.

Usage

getCARD2KEGG(cardId, exhaustiveMapping = FALSE, detailedMapping = FALSE,
byIdenticalProteins = TRUE, bySimilarGenes = TRUE)

Arguments

cardId valid CARD database ARO ID or a vector of them
exhaustiveMapping logical value (TRUE or FALSE) indicating to retrieve all possible translations or only the first and fastest one.
detailedMapping logical value (TRUE or FALSE) indicating to return a vector of KEGG IDs if FALSE, or a named vector if TRUE. This named vector details the origin of the translated IDs, dividing them in DT, if it is a Direct Translation, or 1.0, 0.9 or 0.5, if it has been retrieved through any of the identical gene clusters.
byIdenticalProteins logical value (TRUE or FALSE) indicating to translate also through NCBI Identical Proteins. TRUE value could mean higher translation time and FALSE, fewer chances to translate the ID
bySimilarGenes logical value (TRUE or FALSE) indicating to translate also through similar UniProt genes (100%, 90% or 50% of identity). Using TRUE value may take a long time, only advised if there is not a direct translation possible.

Value

A character vector of the same length as cardId. If parameter exhaustiveMapping is TRUE, it will return a list instead of a character vector to avoid mixing the identifiers.
getCARD2NCBIGene

Translates a CARD database ID to NCBI Gene database

Description

Retrieves the ID translation of an specific instance of CARD database ID to the NCBI Gene database.

Usage

getCARD2NCBIGene(cardId, exhaustiveMapping = FALSE)

Arguments

cardId valid CARD database ARO ID or a vector of them
exhaustiveMapping logical value (TRUE or FALSE) indicating to retrieve all possible translations or only the first and fastest one.

Value

A character vector of the same length as cardId. If parameter exhaustiveMapping is TRUE, it will return a list instead of a character vector to avoid mixing the identifiers.

Examples

# Translate CARD ID '3000938' to KEGG
gETCH2KEGG('3000938')

# Translate CARD IDs '3000938' and '3002557' to KEGG
gETCH2KEGG(c('3000938', '3002557'))

# Get all possible translations of IDs '3000938', '3002511' with detailed origin
gETCH2KEGG(c('3000938', '3002511'), exhaustiveMapping = TRUE, detailedMapping = TRUE)

# Translate CARD ID '3002525' to NCBI Gene
gETCH2NCBIGene('3002525')

# Translate CARD IDs '3002524' and '3002525' to NCBI Gene
gETCH2NCBIGene(c('3002524', '3002525'))

# Get all possible translations of IDs '3005061' and '3002524'
gETCH2NCBIGene(c('3005061', '3002524'), exhaustiveMapping = TRUE)
getCARD2NCBINucleotide

Translates a CARD database ID to NCBI Nucleotide database

Description

Retrieves the ID translation of an specific instance of CARD database ID to the NCBI Nucleotide database. Note that, due to the nature of NCBI Nucleotide database, a complete genome is often retrieved.

Usage

getCARD2NCBINucleotide(cardId)

Arguments

cardId valid CARD database ARO ID or a vector of them

Value

A character vector of the same length as cardId.

Examples

# Translate CARD ID '3002535' to NCBI Nucleotide
getCARD2NCBINucleotide('3002535')

# Translate CARD ID 'ARO:3002535' to NCBI Nucleotide
getCARD2NCBINucleotide('ARO:3002535')

# Translate CARD IDs '3002535' and '3003988' to NCBI Nucleotide
getcARD2NCBINucleotide(c('3002535', '3003988'))

getCARD2NCBIProtein

Translates a CARD database ID to NCBI Protein database

Description

Retrieves the ID translation of an specific instance of CARD database ID to the NCBI Protein database.

Usage

getcARD2NCBIProtein(cardId)
getCARD2UniProt

Arguments

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>cardId</td>
<td>valid CARD database ARO ID or a vector of them</td>
</tr>
</tbody>
</table>

Value

A character vector of the same length as cardId.

Examples

```r
# Translate CARD ID '3002535' to NCBI Protein
getchCARD2NCBIProtein('3002535')

# Translate CARD ID 'ARO:3002535' to NCBI Protein
getCARD2NCBIProtein('ARO:3002535')

# Translate CARD IDs '3002535' and '3003988' to NCBI Protein
getCARD2NCBIProtein(c('3002535', '3003988'))
```

Description

Retrieves the ID translation of an specific instance of CARD database ID to the UniProt database.

Usage

```r
getchCARD2UniProt(cardId, exhaustiveMapping = FALSE, detailedMapping = FALSE)
```

Arguments

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>cardId</td>
<td>valid CARD database ARO ID or a vector of them</td>
</tr>
<tr>
<td>exhaustiveMapping</td>
<td>logical value (TRUE or FALSE) indicating to retrieve all possible translations or only the first and fastest one.</td>
</tr>
<tr>
<td>detailedMapping</td>
<td>logical value (TRUE or FALSE) indicating to return a vector of UniProt IDs if FALSE, or a named vector if TRUE. This named vector details the origin of the translated IDs, dividing them in DT, if it is a Direct Translation, or 1.0, 0.9 or 0.5, if it has been retrieved through any of the identical gene clusters.</td>
</tr>
</tbody>
</table>

Value

A character vector of the same length as cardId. If parameter exhaustiveMapping is TRUE, it will return a list instead of a character vector to avoid mixing the identifiers.
getKEGG2CARD

Examples

# Translate CARD ID '3002867' to UniProt
getCARD2UniProt('3002867')

# Translate CARD IDs '3002867' and '3002535' to UniProt
getCARD2UniProt(c('3002867', '3002535'))

# Get all possible translations of IDs '3002867', '3002535' with detailed origin
getCARD2UniProt(c('3002867', '3002535'), detailedMapping = TRUE, exhaustiveMapping = TRUE)

getKEGG2CARD

Translates a KEGG database ID to CARD database

Description

Retrieves the ID translation of an specific instance of KEGG Genes database ID to the CARD database.

Usage

getKEGG2CARD(keggId, exhaustiveMapping = FALSE, detailedMapping = FALSE, bySimilarGenes = TRUE)

Arguments

keggId valid KEGG Genes database ID or a vector of them
exhaustiveMapping logical value (TRUE or FALSE) indicating to retrieve all possible translations or only the first and fastest one.
detailedMapping logical value (TRUE or FALSE) indicating to return a vector of CARD IDs if FALSE, or a named vector if TRUE. This named vector details the origin of the translated IDs, dividing them in DT, if it is a Direct Translation, or 1.0, 0.9 or 0.5, if it has been retrieved through any of the identical gene clusters.
bySimilarGenes logical value (TRUE or FALSE) indicating to translate also through similar UniProt genes (100%, 90% or 50% of identity). Using TRUE value may take a long time, only advised if there is not a direct translation possible.

Value

A character vector of the same length as keggId. If parameter exhaustiveMapping is TRUE, it will return a list instead of a character vector to avoid mixing the identifiers.
getKEGG2NCBIGene

Translates a KEGG database ID to NCBI Gene database

Description

Retrieves the ID translation of an specific instance of KEGG Genes database ID to the NCBI Gene database.

Usage

getKEGG2NCBIGene(keggId, exhaustiveMapping = FALSE, detailedMapping = FALSE, bySimilarGenes = TRUE)

Arguments

keggId valid KEGG Genes database ID or a vector of them
exhaustiveMapping logical value (TRUE or FALSE) indicating to retrieve all possible translations or only the first and fastest one.
detailedMapping logical value (TRUE or FALSE) indicating to return a vector of NCBI Gene IDs if FALSE, or a named vector if TRUE. This named vector details the origin of the translated IDs, dividing them in DT, if it is a Direct Translation, or 1.0, 0.9 or 0.5, if it has been retrieved through any of the identical gene clusters.
bySimilarGenes logical value (TRUE or FALSE) indicating to translate also through similar UniProt genes (100%, 90% or 50% of identity). Using TRUE value may take a long time, only advised if there is not a direct translation possible.

Value

A character vector of the same length as keggId. If parameter exhaustiveMapping is TRUE, it will return a list instead of a character vector to avoid mixing the identifiers.
getKEGG2NCBINucleotide

Translates a KEGG database ID to NCBI Nucleotide database

Description

Retrieves the ID translation of an specific instance of KEGG Genes database ID to the NCBI Nucleotide database. Note that, due to the nature of NCBI Nucleotide database, a complete genome is often retrieved.

Usage

getKEGG2NCBINucleotide(keggId, exhaustiveMapping = FALSE, detailedMapping = FALSE, bySimilarGenes = TRUE)

Arguments

keggId
valid KEGG Genes database ID or a vector of them

exhaustiveMapping
logical value (TRUE or FALSE) indicating to retrieve all possible translations or only the first and fastest one.

detailedMapping
logical value (TRUE or FALSE) indicating to return a vector of NCBI Nucleotide IDs if FALSE, or a named vector if TRUE. This named vector details the origin of the translated IDs, dividing them in DT, if it is a Direct Translation, or 1.0, 0.9 or 0.5, if it has been retrieved through any of the identical gene clusters.

bySimilarGenes
logical value (TRUE or FALSE) indicating to translate also through similar UniProt genes (100%, 90% or 50% of identity). Using TRUE value may take a long time, only advised if there is not a direct translation possible.

Value

A character vector of the same length as keggId. If parameter exhaustiveMapping is TRUE, it will return a list instead of a character vector to avoid mixing the identifiers.
Examples

# Translate KEGG ID 'llo:LLO_2673' to NCBI Nucleotide
getKEGG2NCBINucleotide('llo:LLO_2673')

# Translate KEGG IDs 'llo:LLO_2673' and 'abh:M3Q_2831' to NCBI Nucleotide
getKEGG2NCBINucleotide(c('llo:LLO_2673', 'abh:M3Q_2831'))

# Get all possible translations of IDs 'aag:5579347', 'abh:M3Q_2831' with
# detailed origin
getKEGG2NCBINucleotide(c('aag:5579347', 'abh:M3Q_2831'),
  exhaustiveMapping = TRUE, detailedMapping = TRUE)

getKEGG2NCBIProtein  Translates a KEGG database ID to NCBI Protein database

Description

Retrieves the ID translation of an specific instance of KEGG Genes database ID to the NCBI Protein database.

Usage

getKEGG2NCBIProtein(keggId, exhaustiveMapping = FALSE, detailedMapping = FALSE, bySimilarGenes = TRUE)

Arguments

keggId  valid KEGG Genes database ID or a vector of them
exhaustiveMapping  logical value (TRUE or FALSE) indicating to retrieve all possible translations or only the first and fastest one.
detailedMapping  logical value (TRUE or FALSE) indicating to return a vector of NCBI Protein IDs if FALSE, or a named vector if TRUE. This named vector details the origin of the translated IDs, dividing them in DT, if it is a Direct Translation, or 1.0, 0.9 or 0.5, if it has been retrieved through any of the identical gene clusters.
bySimilarGenes  logical value (TRUE or FALSE) indicating to translate also through similar UniProt genes (100%, 90% or 50% of identity). Using TRUE value may take a long time, only advised if there is not a direct translation possible.

Value

A character vector of the same length as keggId. If parameter exhaustiveMapping is TRUE, it will return a list instead of a character vector to avoid mixing the identifiers.
Examples

# Translate KEGG ID 'llo:LLO_2673' to NCBI Protein
getKEGG2NCBIProtein('llo:LLO_2673')

# Translate KEGG IDs 'llo:LLO_2673' and 'abh:M3Q_2831' to NCBI Protein
getKEGG2NCBIProtein(c('llo:LLO_2673', 'abh:M3Q_2831'))

# Get all possible translations of IDs 'llo:LLO_2673', 'abh:M3Q_2831' with
detailed origin
getKEGG2NCBIProtein(c('llo:LLO_2673', 'abh:M3Q_2831'),
exhaustiveMapping = TRUE, detailedMapping = TRUE)

getKEGG2UniProt  Translates a KEGG Genes database ID to UniProt database

Description

Retrieves the ID translation of an specific instance of KEGG Genes database ID to the UniProt database.

Usage

getKEGG2UniProt(keggId, exhaustiveMapping = FALSE)

Arguments

keggId  valid KEGG Genes database ID or a vector of them
exhaustiveMapping  logical value (TRUE or FALSE) indicating to retrieve all possible translations or only the first and fastest one.

Value

A character vector of the same length as keggId. If parameter exhaustiveMapping is TRUE, it will return a list instead of a character vector to avoid mixing the identifiers.

Examples

# Translate KEGG Genes ID 'llo:LLO_2673' to UniProt
getKEGG2UniProt('llo:LLO_2673')

# Translate KEGG IDs 'llo:LLO_2673' and 'abh:M3Q_2831' to NCBI UniProt
getKEGG2UniProt(c('llo:LLO_2673', 'abh:M3Q_2831'))

# Get all possible translations of IDs 'llo:LLO_2673', 'abh:M3Q_2831'
getKEGG2UniProt(c('llo:LLO_2673', 'abh:M3Q_2831'), exhaustiveMapping = TRUE)
getNCBIGene2CARD

Translates a NCBI Gene database ID to CARD database

Description

Retrieves the ID translation of an specific instance of NCBI Gene database ID to the CARD database.

Usage

getNCBIGene2CARD(ncbiId, exhaustiveMapping = FALSE)

Arguments

ncbiId valid NCBI Gene database ID or a vector of them
exhaustiveMapping logical value (TRUE or FALSE) indicating to retrieve all possible translations or only the first and fastest one.

Value

A character vector of the same length as ncbiId. If parameter exhaustiveMapping is TRUE, it will return a list instead of a character vector to avoid mixing the identifiers.

Examples

# Translate NCBI Gene ID '3510143' to CARD
getNCBIGene2CARD('3510143')

# Translate NCBI Gene IDs '3510143' and '29426913' to CARD
getNCBIGene2CARD(c('3510143','29426913'))

# Get all possible translations of IDs '3510143' and '29426913'
getNCBIGene2CARD(c('3510143','29426913'), exhaustiveMapping = TRUE)

getNCBIGene2KEGG

Translates a NCBI Gene database ID to KEGG database

Description

Retrieves the ID translation of an specific instance of NCBI Gene database ID to the KEGG Genes database.

Usage

getNCBIGene2KEGG(ncbiId, exhaustiveMapping = FALSE, detailedMapping = FALSE, byIdenticalProteins = TRUE, bySimilarGenes = TRUE)
Arguments

ncbiId valid NCBI Gene database ID or a vector of them

exhaustiveMapping

logical value (TRUE or FALSE) indicating to retrieve all possible translations or only the first and fastest one.

detailedMapping

logical value (TRUE or FALSE) indicating to return a vector of KEGG IDs if FALSE, or a named vector if TRUE. This named vector details the origin of the translated IDs, dividing them in DT, if it is a Direct Translation, or 1.0, 0.9 or 0.5, if it has been retrieved through any of the identical gene clusters.

byIdenticalProteins

logical value (TRUE or FALSE) indicating to translate also through NCBI Identical Proteins. TRUE value could mean higher translation time and FALSE, fewer chances to translate the ID

bySimilarGenes

logical value (TRUE or FALSE) indicating to translate also through similar UniProt genes (100%, 90% or 50% of identity). Using TRUE value may take a long time, only advised if there is not a direct translation possible.

Value

A character vector of the same length as ncbiId. If parameter exhaustiveMapping is TRUE, it will return a list instead of a character vector to avoid mixing the identifiers.

Examples

# Translate NCBI Gene ID '76524190' to KEGG
getNCBIGene2KEGG('76524190')

# Translate NCBI Gene IDs '76524190' and 'AFU38919' to KEGG
getNCBIGene2KEGG(c('76524190', 'AFU38919'))

# Get all possible translations of IDs '76524190' and 'AFU38919' with # detailed origin
getNCBIGene2KEGG(c('76524190', 'AFU38919'), exhaustiveMapping = TRUE, detailedMapping = TRUE)
getNCBIGene2NCBIProtein

Usage

getNCBIGene2NCBINucleotide(id, exhaustiveMapping=FALSE)

Arguments

id
valid NCBI Gene database ID or a vector of them

exhaustiveMapping
logical value (TRUE or FALSE) indicating to retrieve all possible translations or only the first and fastest one.

Value

A character vector of the same length as id. If parameter exhaustiveMapping is TRUE, it will return a list instead of a character vector to avoid mixing the identifiers.

Examples

# Translate NCBI Gene ID '76524190' to NCBI Nucleotide
getNCBIGene2NCBINucleotide('76524190')

# Translate NCBI Gene IDs '76524190' and '1272' to NCBI Nucleotide
getNCBIGene2NCBINucleotide(c('76524190', '1272'))

# Get all possible translations of IDs '76524190' and '1272'
getNCBIGene2NCBINucleotide(c('76524190', '1272'), exhaustiveMapping = TRUE)

getNCBIGene2NCBIProtein

Translates a NCBI Gene database ID to NCBI Protein database

Description

Retrieves the ID translation of an specific instance of NCBI Gene database ID to the NCBI Protein database.

Usage

getNCBIGene2NCBIProtein(id, exhaustiveMapping=FALSE)

Arguments

id
valid NCBI Gene database ID or a vector of them

exhaustiveMapping
logical value (TRUE or FALSE) indicating to retrieve all possible translations or only the first and fastest one.
getNCBIGene2UniProt

Value

A character vector of the same length as id. If parameter exhaustiveMapping is TRUE, it will return a list instead of a character vector to avoid mixing the identifiers.

Examples

# Translate NCBI Gene ID '76524190' to NCBI Protein
getNCBIGene2NCBIProtein('76524190')

# Translate NCBI Gene IDs '76524190' and '1272' to NCBI Protein
getNCBIGene2NCBIProtein(c('76524190', '1272'))

# Get all possible translations of IDs '76524190', '1272'
getNCBIGene2NCBIProtein(c('76524190', '1272'), exhaustiveMapping = TRUE)

getNCBIGene2UniProt

Translates a NCBI Gene database ID to UniProt database

Description

Retrieves the ID translation of an specific instance of NCBI Gene database ID to the UniProt database.

Usage

getNCBIGene2UniProt(ncbiId, exhaustiveMapping = FALSE, detailedMapping = FALSE, byIdenticalProteins = TRUE)

Arguments

ncbiId valid NCBI Gene database ID or a vector of them

exhaustiveMapping logical value (TRUE or FALSE) indicating to retrieve all possible translations or only the first and fastest one.

detailedMapping logical value (TRUE or FALSE) indicating to return a vector of UniProt IDs if FALSE, or a named vector if TRUE. This named vector details the origin of the translated IDs, dividing them in DT, if it is a Direct Translation, or 1.0, 0.9 or 0.5, if it has been retrieved through any of the identical gene clusters.

byIdenticalProteins logical value (TRUE or FALSE) indicating to translate also through NCBI Identical Proteins. TRUE value could mean higher translation time and FALSE, fewer chances to translate the ID
getNCBIIdenticalProteins

Search NCBI Identical Protein database by a given NCBI ID

Description

Retrieves identical NCBI proteins to that of a given ID.

Usage

getNCBIIdenticalProteins(ncbiId, format='ids')

Arguments

ncbiId 
valid NCBI ID or a vector of them

format 
return format of the response. Default value ids returns a vector of the identical proteins IDs; dataframe value returns a dataframe with different information about the found identical proteins

Value

Depending on parameter format, it will return a list of same legth as ncbiId with identical NCBI Protein IDs or with dataframes containing different information about the identical proteins found: Id, Source, Nucleotide.Accession, Start, Stop, Strand, Protein, Protein.Name, Organism, Strain and Assembly.

Examples

# Translate NCBI Gene ID '76524190' to UniProt
getNCBIGene2UniProt('76524190')

# Translate NCBI Gene IDs '76524190' and '1272' to UniProt
getNCBIGene2UniProt(c('76524190', '1272'))

# Get all possible translations of IDs '76524190', '1272' with detailed origin
getNCBIGene2UniProt(c('76524190', '1272'), detailedMapping = TRUE,
exhaustiveMapping = TRUE)
getNCBIIdenticalProteins

Examples

# Get IDs on identical proteins to the one with ID 'AHA80958'
getNCBIIdenticalProteins('AHA80958', format = 'ids')

# Get IDs on identical proteins to the ones with IDs 'AHA80958' and 'CAA76794'
getNCBIIdenticalProteins(c('AHA80958', 'CAA76794'), format = 'ids')

getNCBINucleotide2CARD

 Translates a NCBI Nucleotide database ID to CARD database

Description

Retrieves the ID translation of an specific instance of NCBI Nucleotide database ID to the CARD database.

Usage

getNCBINucleotide2CARD(ncbiId, exhaustiveMapping = FALSE)

Arguments

ncbiId valid NCBI Nucleotide database ID or a vector of them
exhaustiveMapping logical value (TRUE or FALSE) indicating to retrieve all possible translations or only the first and fastest one.

Value

A character vector of the same length as ncbiId. If parameter exhaustiveMapping is TRUE, it will return a list instead of a character vector to avoid mixing the identifiers.

Examples

# Translate NCBI Nucleotide ID 'AY034138' to CARD
getNCBINucleotide2CARD('AY034138')

# Translate NCBI Nucleotide IDs 'AY034138' and 'X54723.1' to CARD
getNCBINucleotide2CARD(c('AY034138', 'X54723.1'))

# Get all possible translations of IDs 'AY034138', 'X54723.1'
getNCBINucleotide2CARD(c('AY034138', 'X54723.1'), exhaustiveMapping = TRUE)
getNCBINucleotide2KEGG

**Description**

Retrieves the ID translation of an specific instance of NCBI Nucleotide database ID to the KEGG Genes database.

**Usage**

```r
getNCBINucleotide2KEGG(ncbiId, exhaustiveMapping = FALSE, detailedMapping = FALSE, byIdenticalProteins = TRUE, bySimilarGenes = TRUE)
```

**Arguments**

- `ncbiId`: valid NCBI Nucleotide database ID or a vector of them
- `exhaustiveMapping`: logical value (TRUE or FALSE) indicating to retrieve all possible translations or only the first and fastest one.
- `detailedMapping`: logical value (TRUE or FALSE) indicating to return a vector of KEGG IDs if FALSE, or a named vector if TRUE. This named vector details the origin of the translated IDs, dividing them in DT, if it is a Direct Translation, or 1.0, 0.9 or 0.5, if it has been retrieved through any of the identical gene clusters.
- `byIdenticalProteins`: logical value (TRUE or FALSE) indicating to translate also through NCBI Identical Proteins. TRUE value could mean higher translation time and FALSE, fewer chances to translate the ID
- `bySimilarGenes`: logical value (TRUE or FALSE) indicating to translate also through similar UniProt genes (100%, 90% or 50% of identity). Using TRUE value may take a long time, only advised if there is not a direct translation possible.

**Value**

A character vector of the same length as ncbiId. If parameter exhaustiveMapping is TRUE, it will return a list instead of a character vector to avoid mixing the identifiers.

**Examples**

```r
# Translate NCBI Nucleotide ID 'AY536519' to KEGG
getNCBINucleotide2KEGG('AY536519')

# Translate NCBI Nucleotide IDs 'AY536519' and 'NC_013861.1' to KEGG
getNCBINucleotide2KEGG(c('AY536519', 'NC_013861.1'))

# Get all possible translations of IDs 'AY536519', 'NC_013861.1'
```
getNCBINucleotide2NCBIGene

Translates a NCBI Nucleotide database ID to NCBI Gene database

Description

Retrieves the ID translation of an specific instance of NCBI Nucleotide database ID to the NCBI Gene database.

Usage

getNCBINucleotide2NCBIGene(id, exhaustiveMapping=FALSE)

Arguments

id valid NCBI Nucleotide database ID or a vector of them
exhaustiveMapping logical value (TRUE or FALSE) indicating to retrieve all possible translations or only the first and fastest one.

Value

A character vector of the same length as id. If parameter exhaustiveMapping is TRUE, it will return a list instead of a character vector to avoid mixing the identifiers.

Examples

# Translate NCBI Nucleotide ID 'Z21488' to NCBI Gene
getNCBINucleotide2NCBIGene('Z21488')

# Get all possible translations of ID 'Z21488'
getNCBINucleotide2NCBIGene(c('Z21488'), exhaustiveMapping = TRUE)
getNCBINucleotide2NCBIProtein

Translates a NCBI Nucleotide database ID to NCBI Protein database

Description

Retrieves the ID translation of an specific instance of NCBI Nucleotide database ID to the NCBI Protein database.

Usage

getNCBINucleotide2NCBIProtein(id, exhaustiveMapping=FALSE)

Arguments

id valid NCBI Nucleotide database ID or a vector of them
exhaustiveMapping logical value (TRUE or FALSE) indicating to retrieve all possible translations or only the first and fastest one.

Value

A character vector of the same length as id. If parameter exhaustiveMapping is TRUE, it will return a list instead of a character vector to avoid mixing the identifiers.

Examples

# Translate NCBI Nucleotide ID 'AY536519' to NCBI Protein
getNCBINucleotide2NCBIProtein('AY536519')

# Translate NCBI Nucleotide IDs 'AY536519' and 'X54723.1' to NCBI Protein
getNCBINucleotide2NCBIProtein(c('AY536519', 'X54723.1'))

# Get all possible translations of IDs 'AY536519', 'X54723.1'
getNCBINucleotide2NCBIProtein(c('AY536519', 'X54723.1'), exhaustiveMapping = TRUE)

getNCBINucleotide2UniProt

Translates a NCBI Nucleotide database ID to UniProt database

Description

Retrieves the ID translation of an specific instance of NCBI Nucleotide database ID to the UniProt database.
Usage

getNCBINucleotide2UniProt(ncbiId, exhaustiveMapping = FALSE, detailedMapping = FALSE, byIdenticalProteins = TRUE)

Arguments

ncbiId valid NCBI Nucleotide database ID or a vector of them
exhaustiveMapping logical value (TRUE or FALSE) indicating to retrieve all possible translations or only the first and fastest one.
detailedMapping logical value (TRUE or FALSE) indicating to return a vector of UniProt IDs if FALSE, or a named vector if TRUE. This named vector details the origin of the translated IDs, dividing them in DT, if it is a Direct Translation, or 1.0, 0.9 or 0.5, if it has been retrieved through any of the identical gene clusters.
byIdenticalProteins logical value (TRUE or FALSE) indicating to translate also through NCBI Identical Proteins. TRUE value could mean higher translation time and FALSE, fewer chances to translate the ID

Value

A character vector of the same length as ncbiId. If parameter exhaustiveMapping is TRUE, it will return a list instead of a character vector to avoid mixing the identifiers.

Examples

# Translate NCBI Nucleotide ID 'AY536519' to UniProt
getNCBINucleotide2UniProt('AY536519')

# Translate NCBI Nucleotide IDs 'AY536519' and 'X54723.1' to UniProt
getNCBINucleotide2UniProt(c('AY536519', 'X54723.1'))

# Get all possible translations of IDs 'AY536519', 'X54723.1'
# with detailed origin
getNCBINucleotide2UniProt(c('AY536519', 'X54723.1'), detailedMapping = TRUE, exhaustiveMapping = TRUE)
Usage

getNCBIProtein2CARD(ncbiId, exhaustiveMapping = FALSE)

Arguments

ncbiId valid NCBI Protein database ID or a vector of them
exhaustiveMapping logical value (TRUE or FALSE) indicating to retrieve all possible translations or only the first and fastest one.

Value

A character vector of the same length as ncbiId. If parameter exhaustiveMapping is TRUE, it will return a list instead of a character vector to avoid mixing the identifiers.

Examples

# Translate NCBI Protein ID 'AAK64581' to CARD
getNCBIProtein2CARD('AAK64581')

# Translate NCBI Protein IDs 'AAK64581' and 'WP_010896559.1' to CARD
getNCBIProtein2CARD(c('AAK64581', 'WP_010896559.1'))

# Get all possible translations of IDs 'AAK64581', 'WP_010896559.1'
getNCBIProtein2CARD(c('AAK64581', 'WP_010896559.1'), exhaustiveMapping = TRUE)

getNCBIProtein2KEGG Translates a NCBI Protein database ID to KEGG database

Description

Retrieves the ID translation of an specific instance of NCBI Protein database ID to the KEGG Genes database.

Usage

getNCBIProtein2KEGG(ncbiId, exhaustiveMapping = FALSE, detailedMapping = FALSE, byIdenticalProteins = TRUE, bySimilarGenes = TRUE)

Arguments

ncbiId valid NCBI Protein database ID or a vector of them
exhaustiveMapping logical value (TRUE or FALSE) indicating to retrieve all possible translations or only the first and fastest one.
getNCBIProtein2NCBIGene

DetailedMapping
Logical value (TRUE or FALSE) indicating to return a vector of KEGG IDs if FALSE, or a named vector if TRUE. This named vector details the origin of the translated IDs, dividing them in DT, if it is a Direct Translation, or 1.0, 0.9 or 0.5, if it has been retrieved through any of the identical gene clusters.

byIdenticalProteins
Logical value (TRUE or FALSE) indicating to translate also through NCBI Identical Proteins. TRUE value could mean higher translation time and FALSE, fewer chances to translate the ID.

bySimilarGenes
Logical value (TRUE or FALSE) indicating to translate also through similar UniProt genes (100%, 90% or 50% of identity). Using TRUE value may take a long time, only advised if there is not a direct translation possible.

Value
A character vector of the same length as ncbiId. If parameter exhaustiveMapping is TRUE, it will return a list instead of a character vector to avoid mixing the identifiers.

Examples

# Translate NCBI Protein ID 'WP_010896559.1' to KEGG
getNCBIProtein2KEGG('WP_010896559.1', exhaustiveMapping = FALSE, detailedMapping = FALSE, byIdenticalProteins = FALSE)

# Translate NCBI Protein IDs 'AFU38919' and 'WP_010896559.1' to KEGG
getNCBIProtein2KEGG(c('AFU38919', 'WP_010896559.1'), byIdenticalProteins = FALSE)

# Get all possible translations of ID 'AFU38919', 'WP_010896559.1'
# with detailed origin
getNCBIProtein2KEGG(c('AFU38919', 'WP_010896559.1'), exhaustiveMapping = TRUE, detailedMapping = TRUE)
getNCBIProtein2NCBINucleotide

Arguments

id  valid NCBI Protein database ID or a vector of them
exhaustiveMapping
    logical value (TRUE or FALSE) indicating to retrieve all possible translations or
    only the first and fastest one.

Value

A character vector of the same length as id. If parameter exhaustiveMapping is TRUE, it will return
a list instead of a character vector to avoid mixing the identifiers.

Examples

# Translate NCBI Protein ID 'CAA79696' to NCBI Gene
getNCBIProtein2NCBIGene('CAA79696')

# Translate NCBI Protein IDs 'CAA79696' and 'WP_001082319' to NCBI Gene
getNCBIProtein2NCBIGene(c('CAA79696', 'WP_001082319'))

# Get all possible translations of IDs 'CAA79696', 'WP_001082319'
getNCBIProtein2NCBIGene(c('CAA79696', 'WP_001082319'), exhaustiveMapping=TRUE)

getNCBIProtein2NCBINucleotide

Translates a NCBI Protein database ID to NCBI Nucleotide database

Description

Retrieves the ID translation of a specific instance of NCBI Protein database ID to the NCBI Nu-
cleotide database. Note that, due to the nature of NCBI Nucleotide database, a complete genome is
often retrieved.

Usage

getNCBIProtein2NCBINucleotide(id, exhaustiveMapping=FALSE)

Arguments

id  valid NCBI Protein database ID or a vector of them
exhaustiveMapping
    logical value (TRUE or FALSE) indicating to retrieve all possible translations or
    only the first and fastest one.

Value

A character vector of the same length as id. If parameter exhaustiveMapping is TRUE, it will return
a list instead of a character vector to avoid mixing the identifiers.
getNCBIProtein2UniProt

Examples

# Translate NCBI Protein ID 'AAS48620' to NCBI Nucleotide
getNCBIProtein2NCBINucleotide('AAS48620')

# Translate NCBI Protein IDs 'AAS48620' and 'CAA38525' to NCBI Nucleotide
getNCBIProtein2NCBINucleotide(c('AAS48620', 'CAA38525'))

# Get all possible translations of IDs 'AAS48620', 'CAA38525'
getNCBIProtein2NCBINucleotide(c('AAS48620', 'CAA38525'), exhaustiveMapping=TRUE)

getNCBIProtein2UniProt

Translates a NCBI Protein database ID to UniProt database

Description

Retrieves the ID translation of an specific instance of NCBI Protein database ID to the UniProt database.

Usage

getNCBIProtein2UniProt(ncbiId, exhaustiveMapping = FALSE, detailedMapping = FALSE, byIdenticalProteins = TRUE)

Arguments

ncbiId valid NCBI Protein database ID or a vector of them
exhaustiveMapping logical value (TRUE or FALSE) indicating to retrieve all possible translations or only the first and fastest one.
detailedMapping logical value (TRUE or FALSE) indicating to return a vector of UniProt IDs if FALSE, or a named vector if TRUE. This named vector details the origin of the translated IDs, dividing them in DT, if it is a Direct Translation, or 1.0, 0.9 or 0.5, if it has been retrieved through any of the identical gene clusters.
byIdenticalProteins logical value (TRUE or FALSE) indicating to translate also through NCBI Identical Proteins. TRUE value could mean higher translation time and FALSE, fewer chances to translate the ID

Value

A character vector of the same length as ncbiId. If parameter exhaustiveMapping is TRUE, it will return a list instead of a character vector to avoid mixing the identifiers.
getUniProt2CARD

Translates an UniProt database ID to CARD database

Description

Retrieves the ID translation of an specific instance of UniProt database ID to the CARD database.

Usage

getUniProt2CARD(upId, exhaustiveMapping = FALSE, detailedMapping = FALSE, bySimilarGenes = TRUE)

Arguments

upId valid UniProt database ID or a vector of them
exhaustiveMapping logical value (TRUE or FALSE) indicating to retrieve all possible translations or only the first and fastest one.
detailedMapping logical value (TRUE or FALSE) indicating to return a vector of CARD IDs if FALSE, or a named vector if TRUE. This named vector details the origin of the translated IDs, dividing them in DT, if it is a Direct Translation, or 1.0, 0.9 or 0.5, if it has been retrieved through any of the identical gene clusters.
bySimilarGenes logical value (TRUE or FALSE) indicating to translate also through similar UniProt genes (100%, 90% or 50% of identity). Using TRUE value may take a long time, only advised if there is not a direct translation possible.

Value

A character vector of the same length as upId. If parameter exhaustiveMapping is TRUE, it will return a list instead of a character vector to avoid mixing the identifiers.
getUniProt2KEGG

Translates an UniProt database ID to KEGG database

Description

Retrieves the ID translation of an specific instance of UniProt database ID to the KEGG Genes database.

Usage

getUniProt2KEGG(upId, exhaustiveMapping = FALSE, bySimilarGenes = TRUE, detailedMapping = FALSE)

Arguments

upId valid UniProt database ID or a vector of them.

exhaustiveMapping logical value (TRUE or FALSE) indicating to retrieve all possible translations or only the first and fastest one.

bySimilarGenes logical value (TRUE or FALSE) indicating to translate also through similar UniProt genes (100%, 90% or 50% of identity). Using TRUE value may take a long time, only advised if there is not a direct translation possible.

detailedMapping logical value (TRUE or FALSE) indicating to return a vector of KEGG IDs if FALSE, or a named vector if TRUE. This named vector details the origin of the translated IDs, dividing them in DT, if it is a Direct Translation, or 1.0, 0.9 or 0.5, if it has been retrieved through any of the identical gene clusters.

Value

A character vector of the same length as upId. If parameter exhaustiveMapping is TRUE, it will return a list instead of a character vector to avoid mixing the identifiers.

Examples

# Translate UniProt ID 'A0A1S7BGS4' to CARD
getUniProt2CARD('A0A1S7BGS4')

# Translate UniProt IDs 'A0A1S7BGS4' and 'A0A6H2TXZ6' to CARD
getUniProt2CARD(c('A0A1S7BGS4', 'A0A6H2TXZ6'))

# Get all possible translations of IDs 'A0A1S7BGS4', 'A0A6H2TXZ6'
# with detailed origin
getUniProt2CARD(c('A0A1S7BGS4', 'A0A6H2TXZ6'), detailedMapping = TRUE, exhaustiveMapping = TRUE)
getUniProt2NCBIGene

Translates an UniProt database ID to NCBI Gene database

Description

Retrieves the ID translation of an specific instance of UniProt database ID to the NCBI Gene database.

Usage

getUniProt2NCBIGene(upId, exhaustiveMapping = FALSE, detailedMapping = FALSE, bySimilarGenes = TRUE)

Arguments

upId valid UniProt database ID or a vector of them
exhaustiveMapping logical value (TRUE or FALSE) indicating to retrieve all possible translations or only the first and fastest one.
detailedMapping logical value (TRUE or FALSE) indicating to return a vector of NCBI Gene IDs if FALSE, or a named vector if TRUE. This named vector details the origin of the translated IDs, dividing them in DT, if it is a Direct Translation, or 1.0, 0.9 or 0.5, if it has been retrieved through any of the identical gene clusters.
bySimilarGenes logical value (TRUE or FALSE) indicating to translate also through similar UniProt genes (100%, 90% or 50% of identity). Using TRUE value may take a long time, only advised if there is not a direct translation possible.

Value

A character vector of the same length as upId. If parameter exhaustiveMapping is TRUE, it will return a list instead of a character vector to avoid mixing the identifiers.
getUniProt2NCBINucleotide

Description
Retrieves the ID translation of an specific instance of UniProt database ID to the NCBI Nucleotide database. Note that, due to the nature of NCBI Nucleotide database, a complete genome is often retrieved.

Usage
getUniProt2NCBINucleotide(upId, exhaustiveMapping = FALSE, detailedMapping = FALSE, bySimilarGenes = TRUE)

Arguments
upId valid UniProt database ID or a vector of them
exhaustiveMapping logical value (TRUE or FALSE) indicating to retrieve all possible translations or only the first and fastest one.
detailedMapping logical value (TRUE or FALSE) indicating to return a vector of NCBI Nucleotide IDs if FALSE, or a named vector if TRUE. This named vector details the origin of the translated IDs, dividing them in DT, if it is a Direct Translation, or 1.0, 0.9 or 0.5, if it has been retrieved through any of the identical gene clusters.
bySimilarGenes logical value (TRUE or FALSE) indicating to translate also through similar UniProt genes (100%, 90% or 50% of identity). Using TRUE value may take a long time, only advised if there is not a direct translation possible.

Value
A character vector of the same length as upId. If parameter exhaustiveMapping is TRUE, it will return a list instead of a character vector to avoid mixing the identifiers.
Examples

# Translate UniProt ID 'G9JVE6' to NCBI Nucleotide
getUniProt2NCBINucleotide('G9JVE6')

# Translate UniProt IDs 'A0SNL9' and 'G9JVE6' to NCBI Nucleotide
getUniProt2NCBINucleotide(c('A0SNL9', 'G9JVE6'))

# Get all possible translations of IDs 'A0SNL9', 'G9JVE6' with detailed origin
getUniProt2NCBINucleotide(c('A0SNL9', 'G9JVE6'), detailedMapping = TRUE, exhaustiveMapping = TRUE)

getUniProt2NCBIProtein

Translates an UniProt database ID to NCBI Protein database

Description

Retrieves the ID translation of a specific instance of UniProt database ID to the NCBI Protein database.

Usage

getiUniProt2NCBIProtein(upId, exhaustiveMapping = FALSE, detailedMapping = FALSE, bySimilarGenes = TRUE)

Arguments

upId valid UniProt database ID or a vector of them

exhaustiveMapping logical value (TRUE or FALSE) indicating to retrieve all possible translations or only the first and fastest one.

detailedMapping logical value (TRUE or FALSE) indicating to return a vector of NCBI Protein IDs if FALSE, or a named vector if TRUE. This named vector details the origin of the translated IDs, dividing them in DT, if it is a Direct Translation, or 1.0, 0.9 or 0.5, if it has been retrieved through any of the identical gene clusters.

bySimilarGenes logical value (TRUE or FALSE) indicating to translate also through similar UniProt genes (100%, 90% or 50% of identity). Using TRUE value may take a long time, only advised if there is not a direct translation possible.

Value

A character vector of the same length as upId. If parameter exhaustiveMapping is TRUE, it will return a list instead of a character vector to avoid mixing the identifiers.
getUniProtSimilarGenes

Get similar UniProt genes by a given identity percentage

Description

Retrieves the cluster of genes with a certain percentage of identity where a given UniProt gene belongs.

Usage

getUniProtSimilarGenes(upId, clusterIdentity = '1.0', clusterNames = FALSE)

Arguments

upId valid UniProt ID or a vector of them
clusterIdentity identity of the cluster to be retrieved. Default value 1.0 returns the 100% identical genes cluster, values 0.9 and 0.5, the 90% and 50% identical genes clusters, respectively
clusterNames logical value (TRUE or FALSE) indicating if cluster names are retrieved

Value

A list of the same length as upId. If parameter clusterNames is TRUE it will return a named list.

Examples

# Get UniProt IDs of 50% identical genes of UniProt gene 'G0L217'
getUniProtSimilarGenes('G0L217', clusterIdentity = '0.5')

# Get UniProt IDs of 50% identical genes of UniProt genes 'G0L217' and 'A0SNL9'
geuniProtSimilarGenes(c('G0L217', 'A0SNL9'), clusterIdentity = '0.5')
updateCARDDataBase

*Downloads the most up-to-date version of CARD database*

**Description**

Cleans previous CARD database versions if there were, and downloads its latest version.

**Usage**

```r
updateCARDDataBase()
```

**Value**

Downloaded CARD database.
Index

changeCARDPath, 2
getCARD2KEGG, 3
getCARD2NCBIGene, 4
getCARD2NCBINucleotide, 5
getCARD2NCBIProtein, 5
getCARD2UniProt, 6
getKEGG2CARD, 7
getKEGG2NCBIGene, 8
getKEGG2NCBINucleotide, 9
getKEGG2NCBIProtein, 10
getKEGG2UniProt, 11
getNCBIGene2CARD, 12
getNCBIGene2KEGG, 12
getNCBIGene2NCBINucleotide, 13
getNCBIGene2NCBIProtein, 14
getNCBIGene2UniProt, 15
getNCBIIdenticalProteins, 16
getNCBINucleotide2CARD, 17
getNCBINucleotide2KEGG, 18
getNCBINucleotide2NCBIGene, 19
getNCBINucleotide2NCBIProtein, 20
getNCBINucleotide2UniProt, 20
getNCBIProtein2CARD, 21
getNCBIProtein2KEGG, 22
getNCBIProtein2NCBIGene, 23
getNCBIProtein2NCBINucleotide, 24
getNCBIProtein2UniProt, 25
getUniProt2CARD, 26
getUniProt2KEGG, 27
getUniProt2NCBIGene, 28
getUniProt2NCBINucleotide, 29
getUniProt2NCBIProtein, 30
getUniProtSimilarGenes, 31
updateCARDDataBase, 32