

Package ‘grex’

October 13, 2022

Type Package

Title Gene ID Mapping for Genotype-Tissue Expression (GTEx) Data

Version 1.9

Maintainer Nan Xiao <me@nanx.me>

Description Convert 'Ensembl' gene identifiers from Genotype-Tissue Expression (GTEx) data to identifiers in other annotation systems, including 'Entrez', 'HGNC', and 'UniProt'.

License GPL-3 | file LICENSE

LazyData TRUE

VignetteBuilder knitr

URL <https://nanx.me/grex/>, <https://github.com/nanxstats/grex>

BugReports <https://github.com/nanxstats/grex/issues>

Depends R (>= 3.5.0)

Suggests knitr, rmarkdown

Encoding UTF-8

RoxygenNote 6.1.1

NeedsCompilation no

Author Nan Xiao [aut, cre] (<<https://orcid.org/0000-0002-0250-5673>>),
Gao Wang [aut],
Lei Sun [aut]

Repository CRAN

Date/Publication 2019-05-17 22:40:04 UTC

R topics documented:

cleanid	2
grex	2
gtexv6	3
gtexv6p	4
gtexv7	4

Index**5**

cleanid	<i>Remove Version Numbers in Raw GTEx (GENCODE) Gene IDs</i>
---------	--

Description

Remove the '.version' part in raw GTEx (GENCODE) gene IDs to produce Ensembl IDs.

Usage

```
cleanid(gtex_id)
```

Arguments

gtex_id Character vector of GTEx (GENCODE) gene IDs

Value

Character vector of Ensembl IDs

Examples

```
gtex_id <- c("ENSG00000227232.4", "ENSG00000223972.4", "ENSG00000268020.2")
cleanid(gtex_id)
```

grex	<i>Gene ID Mapping for Genotype-Tissue Expression (GTEx) Data</i>
------	---

Description

Map Ensembl IDs to Entrez Gene ID, HGNC symbol, and UniProt ID, with basic annotation information such as gene type.

Usage

```
grex(ensembl_id)
```

Arguments

ensembl_id Character vector of Ensembl IDs

Value

This function returns a data frame with the same number of rows as the length of input Ensembl IDs, containing:

- `ensembl_id` - Input Ensembl ID
- `entrez_id` - Entrez Gene ID
- `hgnc_symbol` - HGNC gene symbol
- `hgnc_name` - HGNC gene name
- `cyto_loc` - Cytogenetic location
- `uniprot_id` - UniProt ID
- `gene_biotype` - Gene type

The elements that cannot be mapped will be NA.

Examples

```
# Ensembl IDs in GTEx v6p gene count data
data("gtexv6p")
# select 100 IDs as example
id <- gtexv6p[101:200]
df <- grex(id)
# Rows that have a mapped Entrez ID
df[
  !is.na(df$"entrez_id"),
  c("ensembl_id", "entrez_id", "gene_biotype")
]
```

`gtexv6`*Ensembl IDs from GTEx V6 Gene Read Count Data*

Description

A dataset containing the Ensembl IDs from GTEx (V6) gene read count data.

Usage

```
gtexv6
```

Format

A character vector with 56,318 Ensembl IDs.

Source

<https://www.gtexportal.org>

gtexv6p

Ensembl IDs from GTEx V6p Gene Read Count Data

Description

A dataset containing the Ensembl IDs from GTEx (V6p) gene read count data.

Usage

gtexv6p

Format

A character vector with 56,238 Ensembl IDs.

Source

<https://www.gtexportal.org>

gtexv7

Ensembl IDs from GTEx V7 Gene Read Count Data

Description

A dataset containing the Ensembl IDs from GTEx (V7) gene read count data.

Usage

gtexv7

Format

A character vector with 56,202 Ensembl IDs.

Source

<https://www.gtexportal.org>

Index

* datasets

- gtexv6, 3
- gtexv6p, 4
- gtexv7, 4

cleanid, 2

greg, 2
gtexv6, 3
gtexv6p, 4
gtexv7, 4