Package 'toprdata'

February 10, 2023

Title Gene and Exon Data from Ensembl

Version 1.0.2

Description Gene and exon information from Ensembl genome builds GRCh38.p13 (104) and GRCh37 (v40) to use with the 'topr' package.

License LGPL (>= 3)

Encoding UTF-8

LazyData true

LazyDataCompression xz

RoxygenNote 7.2.0

Suggests testthat (>= 3.0.0), knitr, rmarkdown, markdown

Config/testthat/edition 3

Imports

Depends R (>= 3.5.0)

NeedsCompilation no

Author Thorhildur Juliusdottir [cph, aut, cre],
Andri Stefansson [aut]

Maintainer Thorhildur Juliusdottir <totajuliusd@gmail.com>

Repository CRAN

Date/Publication 2023-02-10 12:00:02 UTC

R topics documented:

ENSEXONS   *Ensembl exons build HG38-104-5-2*

Description

https://www.ensembl.info/2021/05/05/ensembl-104-has-been-released/

Usage

ENSEXONS

Format

A data frame with 40,122 rows and 7 variables:

chrom Chromosome on build version 38 (GRCh38/hg38)
gene_start genetic position of gene start on build version 38
gene_end genetic position of gene end on build version 38
gene_symbol The name of the gene
exon_chromstart genetic positions of exon start
exon_chromend genetic position of exon end

Details

exons on chrY and chrM were excluded from the exon dataset

ENSEXONS_37   *Ensembl exons build HG37 (GRCh37) v40 from gencode*

Description

exons on chrY and chrM were excluded from the exon dataset

Usage

ENSEXONS_37

Format

A data frame with 62,739 rows and 7 variables:

chrom Chromosome on build version 38 (GRCh38/hg38)
gene_start genetic position of gene start on build version 38
gene_end genetic position of gene end on build version 38
gene_symbol The name of the gene
exon_chromstart genetic positions of exon start
exon_chromend genetic position of exon end

ENSGENES *Ensembl genes build HG38.104-5-2*

Description

https://www.ensembl.info/2021/05/05/ensembl-104-has-been-released/

Usage

ENSGENES

Format

A data frame with 40,122 rows and 5 variables:

chrom Chromosome on build version 38 (GRCh38/hg38)

gene_start genetic position of gene start on build version 38

gene_end genetic position of gene end on build version 38

gene_symbol The name of the gene

biotype the biotype of the gene

Details

genes on chrY and chrM were excluded

ENSGENES_37 *Ensembl genes build HG37 (GRCh37) v40 from gencode*

Description

genes on chrY and chrM were excluded

Usage

ENSGENES_37

Format

A data frame with 62,743 rows and 5 variables:

chrom Chromosome on build version 38 (GRCh38/hg38)

gene_start genetic position of gene start on build version 38

gene_end genetic position of gene end on build version 38

gene_symbol The name of the gene

biotype the biotype of the gene

toprdata	*toprdata*

Description

Gene and exon datasets to use with the 'topr' package

datasets

The two datasets are:

- ENSGENES Gene information from genome build GRCh38
- ENSEXONS Exon information from genome build GRCh38
- ENSGENES_37 Gene information from genome build GRCh37
- ENSEXONS_37 Exon information from genome build GRCh37

Examples

```
library(toprdata)
head(ENSGENES)
head(ENSEXONS)
```

Index