GenomicFeatures

August 10, 2009

R topics documented:

`geneHuman` *UCSC Gene Predictions for hg18*

Description

A set of genomic coordinate data for gene transcripts, and related annotations, for H. sapiens. Coordinates are relative to the hg18 build. All coordinates are in nucleotides from the 5' end of the positive `"+"` strand.In the below, "gene" is synonymous with "spliced transcript", so each row corresponds to a transcript variant with a particular TSS, TES, and exon configuration. A single coding start (`cdsStart`) and end (`cdsEnd`) are predicted for each gene/transcript.

Usage

```
data(geneHuman)
```

Format

A data frame with 56722 observations on the following 12 variables.

`name` The name of the gene.

`chrom` The name of the chromosome the gene is located on.

`strand` The strand the gene is coded on, `"+"`, or `"-"`.

`txStart` Transcription start site.

`txEnd` Transcription stop site.

`cdsStart` Start position of the coding sequence.

`cdsEnd` End position of the coding sequence.

`exonCount` The number of exons.

`exonStarts` A comma separated list of the exon start positions.

exonEnds A comma separated list of exon stop positions.

proteinID An ID for the protein produced, missing values are coded as NA.

alignID Unique identifier of each gene and RNA alignment pair, apparently redundant with `name`.

Details

For genes coded on the negative strand the `txStart` is really the end, and similarly for the coding regions.

Source

This table was taken directly from the knownGene table in the UCSC database for hg18, see `http://genome.ucsc.edu/cgi-bin/hgTables` and Hsu F, Kent WJ, Clawson H, Kuhn RM, Diekhans M, Haussler D. The UCSC Known Genes. Bioinformatics. 2006 May 1;22(9):1036-46.

Examples

```
data(geneHuman)
str(geneHuman)
```

`geneMouse` *UCSC Gene Predictions for mm9*

Description

A set of genomic coordinate data for gene transcripts, and related annotations, for M. musculus. Coordinates are relative to the mm9 build. All coordinates are in nucleotides from the 5' end of the positive `"+"` strand. In the below, “gene” is synonymous with “spliced transcript”, so each row corresponds to a transcript variant with a particular TSS, TES, and exon configuration. A single coding start (`cdsStart`) and end (`cdsEnd`) are predicted for each gene/transcript.

Usage

```
data(geneMouse)
```

Format

A data frame with 49409 observations on the following 12 variables.

name The name of the gene.

chrom The name of the chromosome the gene is located on.

strand The strand the gene is coded on, `"+"`, or `"-"`.

txStart Transcription start site.

txEnd Transcription stop site.

cdsStart Start position of the coding sequence.

cdsEnd End position of the coding sequence.

exonCount The number of exons.

exonStarts A comma separated list of the exon start positions.

`exonEnds` A comma separated list of exon stop positions.

`proteinID` An ID for the protein produced, missing values are coded as NA.

`alignID` Unique identifier of each gene and RNA alignment pair, apparently redundant with `name`.

Details

For genes coded on the negative strand the `txStart` is really the end, and similarly for the coding regions.

Source

This table was taken directly from the knownGene table in the UCSC database for mm9, see `http://genome.ucsc.edu/cgi-bin/hgTables` and Hsu F, Kent WJ, Clawson H, Kuhn RM, Diekhans M, Haussler D. The UCSC Known Genes. Bioinformatics. 2006 May 1;22(9):1036-46.

Examples

```
data(geneMouse)
str(geneMouse)
```

`isochores.mm8` *Isochore boundaries for Mus musculus (build mm9).*

Description

Isochore boundaries for Mus musculus (build mm9). Isochores are large segments of the genome such that within-segment variability in GC content is substantially lower than between-segment variability. These isochores are computationally predicted by IsoFinder (see below).

Usage

```
data(isochores.mm8)
```

Format

A data frame with 32894 observations on the following 4 variables.

`Begin` isochore starts.

`End` isochore ends.

`GC` GC content in isochore.

`chromosome` chromosome identifier.

Source

`http://bioinfo2.ugr.es/isochores/`

Index