Package 'RcisTarget.hg19.motifDBs.cisbpOnly.500bp'

April 24, 2025

Type Package

Title RcisTarget motif databases for human (hg19) - Subset of 4.6k motifs

Version 1.28.0

Date 2018-05-03

Author Sara Aibar, Gert Hulselmans, Stein Aerts. Laboratory of Computational Biology, KU Leuven Center for Human Genetics. Leuven, Belgium.

Maintainer Sara Aibar <sara.aibar@kuleuven.vib.be>

Description RcisTarget databases: Gene-based motif rankings and annotation to transcription factors. This package contains a subset of 4.6k motifs (cisbp motifs), scored only within 500bp upstream and the TSS.
See RcisTarget tutorial to download the full databases, containing 20k motifs and search space up to 10kbp around the TSS.

URL http://scenic.aertslab.org

Depends R (>= 3.3)

Imports data.table

License GPL-3

biocViews Homo_sapiens_Data

LazyData FALSE

git_url https://git.bioconductor.org/packages/RcisTarget.hg19.motifDBs.cisbpOnly.500bp

git_branch RELEASE_3_21

git_last_commit 6030af7

git_last_commit_date 2025-04-15

Repository Bioconductor 3.21

Date/Publication 2025-04-24

Contents

RcisTarget.hg19.motifDBs.cisbpOnly.500bp 2

Index 4

`RcisTarget.hg19.motifDBs.cisbpOnly.500bp`

RcisTarget motif databases for human (hg19) - Subset of 4.6k motifs

---

Description

RcisTarget databases: gene-based motif rankings and their annotation to transcription factors for Human (hg19).

For other organisms or the **full databases** (containing ~20k motifs and search space up to 10kbp around the TSS) see RcisTarget tutorial.

This package contains a *subset* of 4.6k motifs (cisbp motifs), scored only within 500bp upstream and the TSS (source database: `hg19-500bp-upstream-7species.mc9nr`). Only rankings under 5050 are kept, this implies that this is the maximum value to use as `aucMaxRank` in `calcAUC`, and as `maxRank` in `addSignificantGenes` or `getSignificantGenes`.

Format

This dataset contains two data objects which are meant to be used with package `RcisTarget` (http://bioconductor.org/packages/RcisTarget):

`hg19_500bpUpstream_motifRanking_cispbOnly`: **Motif rankings** (4.6k motifs scored within the 500 bp upstream of the TSS and the TSS of the gene).

- This object is provided to RcisTarget *as is* (e.g. not meant to be modified by the user). Internally, the ranking is a matrix of motifs (rows) by genes (columns). The value in the matrix contains the ranking of the gene for each motif (i.e. a ranking of genes per motif).

`hg19_motifAnnotation_cisbpOnly`: **Motif annotations** (data.table containing the annotation of the motif to transcription factors).

This object is also meant to be provided to RcisTarget without modification, but it can also be explored by the user to obtain further information about the motifs.

Columns:

- **motif**: Motif ID.
- **TF**: Transcription factor (or inferred gene).
- **directAnnotation**, **inferred_Orthology**, **inferred_MotifSimil**: Boolean values indicating whether the motif is annotated to the TF in the source database ("directAnnotation"), or whether it was inferred by orthology ("inferred_Orthology") or motif similarity ("inferred_MotifSimil").
- **Description**: Description of the source of the annotation.
- **annotationSource**: Source of the annotation formatted as factor (e.g. for subsetting). Levels: directAnnotation, inferredBy_Orthology, inferredBy_MotifSimilarity, inferredBy_MotifSimilarity_n_Orthology.

Author(s)

Sara Aibar, Gert Hulselmans, Stein Aerts. Laboratory of Computational Biology, KU Leuven Center for Human Genetics. Leuven, Belgium.

Maintainer: Sara Aibar <sara.aibar@kuleuven.vib.be>

Source

This is a subset of the databases (20k motif collections) used in iRegulon (http://iregulon.aertslab.org/collections.html#motifcolldesc).

See *iRegulon* paper and documentation for **details** on how the rankings and annotations were built.

References

Janky R, Verfaillie A, Imrichova H, Van de Sande B, Standaert L, Christiaens V, Hulselmans G, Herten K, Naval Sanchez M, Potier D, Svetlichnyy D, Kalender Atak Z, Fiers M, Marine JM, and Aerts S. **iRegulon: From a Gene List to a Gene Regulatory Network Using Large Motif and Track Collections**. PLoS Comput Biol. 2014 Jul 24;10(7):e1003731.

Index

∗ **datasets**
RcisTarget.hg19.motifDBs.cisbpOnly.500bp, 2

hg19_500bpUpstream_motifRanking_cispbOnly *(*RcisTarget.hg19.motifDBs.cisbpOnly.500bp*)*, 2
hg19_motifAnnotation_cisbpOnly *(*RcisTarget.hg19.motifDBs.cisbpOnly.500bp*)*, 2

RcisTarget.hg19.motifDBs.cisbpOnly.500bp, 2