Package: colocalized (via r-universe)

October 31, 2024

Type Package

Title Clusters of Colocalized Sequences

Version 0.2.0

Description Also abbreviates to ``CCSeq''. Finds clusters of colocalized
sequences in .bed annotation files up to a specified cut-off
distance. Two sequences are colocalized if they are within the
cut-off distance of each other, and clusters are sets of
sequences where each sequence is colocalized to at least one
other sequence in the cluster. For a set of .bed annotation
tables provided in a list along with a cut-off distance, the
program will output a file containing the locations of each
cluster. Annotated .bed files are from the 'pwmscan'
application at <https://ccg.epfl.ch/pwmtools/pwmscan.php>.
Personal machines might crash or take excessively long
depending on the number of annotated sequences in each file and
whether chromsearch() or gensearch() is used.

License MIT + file LICENSE

Encoding UTF-8

LazyData true

RoxygenNote 6.1.1.9000

Depends R (>= 3.0.0)

Imports foreach, doParallel, purrr, utils

NeedsCompilation no

Author Stefan Golas [cre, aut]

Maintainer Stefan Golas <stefanmgolas@gmail.com>

Date/Publication 2019-10-18 21:20:12 UTC

Repository https://stefangolas.r-universe.dev

RemoteUrl https://github.com/cran/colocalized

RemoteRef HEAD

RemoteSha 496527ec2f80edd1b0599924bb14e623ed8965b3

Contents

---

chromsearch   *Search one chromsome*

Description

Search a single chromosome for clusters of TF binding sequences. Example produces a null result, test the same complex on "chr9" for a positive reading.

Usage

```
chromsearch(choose, n, chrom)
```

Arguments

choose	List of .bed tables
n	Cut-off distance between colocalized sequences
chrom	Chromosome to be searched given as e.g. "chr19"

Value

A table containing the addresses (as one dimensional intervals) of the members of every cluster, with some annotation data.

Examples

```
complex<-list(nfkb1,nfkb2,relb)
chromsearch(complex,150,"chrY")
```

---

colocalized *Colocalized cluster search.*

Description

Searches for clusters of colocalized transcription factor (TF) binding sequences. `colocalized(choose,chr,n)` searches for instances where the sequences from each table element in choose are colocalized to within a cut-off distance.

Usage

```
colocalized(choose, chr, n, cores)
```

Arguments

choose	List of .bed tables
chr	Chromosome
n	The cut-off distance
cores	Number of cores for parallel processing. Leaving this blank causes the program to use default (series) processing

Value

Table of clusters found in `chr`

Examples

```
complex<-list(nfkb1,nfkb2,relb)
colocalized(complex,"chrY",150)
```

ColocalizedFullSearch *Colocalized full search.*

Description

Wrapper for `colocalized` that searches every chromosome shared between the given .bed files.

Usage

```
ColocalizedFullSearch(choose, n, cores)
```

Arguments

choose	List of .bed tables
n	The cut-off distance
cores	Number of cores for parallel processing. Leaving this blank causes the program to use default (series) processing

Value

List of lists of each cluster found in each chromosome.

```
colocalized_sequential
```
Sequential cluster search

Description

Search one chromosome for clusters using default non-parallel processing.

Usage

```
colocalized_sequential(choose, chr, n)
```

Arguments

choose	List of .bed tables
chr	Chromosome to be searched given as e.g. "chr19"
n	Cut-off distance between colocalized sequences

Value

Table containing the addresses (as one dimensional intervals) of the members of every cluster, with some annotation data.

Examples

```
complex<-list(nfkb1,nfkb2,relb)
colocalized_sequential(complex,"chrY",150)
```

```
gensearch
```
Whole genome search.

Description

Search the whole genome for clusters of colocalized TF binding sequences.

Usage

```
gensearch(choose, n, cores)
```

Arguments

`choose`	List of .bed tables
`n`	The cut-off distance
`cores`	Number of cores for parallel processing. Leaving this blank causes the program to use default (series) processing

Value

Table containing the addresses (as one dimensional intervals) of the members of every cluster, with some annotation data.

`nfkb1`   *NFKB1 bed file*

Description

nfkb1

Usage

```
nfkb1

nfkb1
```

Format

A dataframe with 230505 rows and 8 columns

Author(s)

Stefan Golas

Source

<https://ccg.epfl.ch/pwmtools/pwmscan.php>

<https://ccg.epfl.ch/pwmtools/pwmscan.php>

---

nfkb2   *NFKB2 bed file*

Description

nfkb2

Usage

nfkb2

nfkb2

Format

A dataframe with 1901 rows and 8 columns

Author(s)

Stefan Golas

Source

<https://ccg.epfl.ch/pwmtools/pwmscan.php>

<https://ccg.epfl.ch/pwmtools/pwmscan.php>

---

onedim_dist   *Create a colocalization matrix*

Description

Create a colocalization matrix

Usage

onedim_dist(bed1, bed2, n)

Arguments

bed1	A .bed table
bed2	A .bed table
n	The cut-off distance

Value

A colocalization matrix whose dimensions are the number of rows in bed1 by the number of rows in bed2. Entry i,j is a 1 if the ith sequence in bed1 is within the cut-off distance of the jth sequence in bed2, and 0 otherwise.

Examples

```
chr<-"chrY"
onedim_dist(nfkb1[which(nfkb1[,1]==chr),],nfkb2[which(nfkb2[,1]==chr),],150)
```

---

relb	*RELB bed file*

Description

relb

Usage

```
relb

relb
```

Format

A dataframe with 1448 rows and 8 columns

Author(s)

Stefan Golas

Source

<https://ccg.epfl.ch/pwmtools/pwmscan.php>

<https://ccg.epfl.ch/pwmtools/pwmscan.php>

Index