

Package ‘SanzCircos’

May 4, 2018

Type Package

Title Functions for Creating 'Circos' Files

Version 0.1.0

Author Matthew Woodruff, PhD

Maintainer ``Matthew Woodruff' <mcwoodr@emory.edu>

Description A series of functions designed to aid in the creation of input files for the 'Circos' visualization platform <<http://circos.ca/>>. Included are two main function types - `make_circos_x()` and `write_circos_x()`. The ``make" functions are designed to process an input data set, and return a data frame that can be manipulated and filtered to eliminate the need to write Perl scripts within the 'Circos' software platform for data filtering. The ``write" functions take processed data frames from the ``make" functions, and write .txt files compatible with the 'Circos' platform. In addition to these function types, there are other accessory functions that aid in the manipulation of the `make_circos_x()` returned data frames to enhance final visual appeal (such as `color_circos_links()`).

License GPL-3

Encoding UTF-8

LazyData true

Imports dplyr, tidyr, purrr, readr, tibble, pbapply, randomcoloR, parallel, grDevices

RoxygenNote 6.0.1.9000

NeedsCompilation no

Repository CRAN

Date/Publication 2018-05-04 10:52:54 UTC

R topics documented:

<code>color_circos_links</code>	2
<code>generate_searchable_links</code>	3
<code>make_circos_df_1</code>	4
<code>make_circos_links</code>	5

write_circos_histogram	6
write_circos_karyotype	7
write_circos_links	8

Index	9
--------------	----------

color_circos_links	<i>color_circos_links</i>
--------------------	---------------------------

Description

A function designed to assign random colors from the ‘randomcoloR’ package to a ‘make_circos_links()’ generated data frame.

Usage

```
color_circos_links(df, hue = c(" ", "random", "red", "orange", "yellow",
  "green", "blue", "purple", "pink", "monochrome"), luminosity = c(" ",
  "random", "light", "bright", "dark"))
```

Arguments

df	A data frame containing circos links information - as from ‘make_circos_links()’.
hue	Desired hue of the links.
luminosity	Desired luminosity of the links.

Value

Returns original links data frame with an added ‘\$color’ column in RGB format for Circos plotting.

Author(s)

Matthew Woodruff, Emory University

Examples

```
df <- data.frame(x = rnorm(50), y = rnorm(50))
str(df)
df <- color_circos_links(df)
str(df)
```

```
generate_searchable_links  
    generate_searchable_links
```

Description

A function serving as the core of the `make_circos_links` function. Takes a data frame containing a single lineage, and generates a links data frame containing all possible combinations of that lineage.

Usage

```
generate_searchable_links(df, chromosome_grouping, band_grouping, link_grouping,  
    start_position, end_position)
```

Arguments

<code>df</code>	A data frame containing all sub-lineages from the same lineageID
<code>chromosome_grouping</code>	A column defining the circos chromosomes. Passed as column index or "column name".
<code>band_grouping</code>	A column defining the circos chromosome bands. Passed as column index or "column name".
<code>link_grouping</code>	A column defining the circos linkage IDs. Passed as column index or "column name".
<code>start_position</code>	A column defining the start positions of the linkage ID.
<code>end_position</code>	A column defining the end positions of the linkage ID.

Value

Returns a data frame in a pre-"circos links" format for later use within the `'make_circos_links'` function.

Author(s)

Matthew Woodruff, Emory University

make_circos_df_1 *make_circos_df_1*

Description

First of a series of data-frame processing functions that take an initial data set, and processes it into a Circos-compatible format.

'make_circos_df_1()' groups sequences by population, isotype, and lineageID, and returns a summary data frame containing lineage start and stop positions, mutation frequencies, and scaled mutation frequencies. Can then be used to establish lineage highlighting, linkage assessment, or mutation plotting. The resulting data frame is finally arranged by population, isotype, mutation freq, and lineageID

Usage

```
make_circos_df_1(df)
```

Arguments

df Data frame to be transformed. Built to read data frames modified from the .db files that IgSeq generates. Must contain columns "population", "isotype", "lineageID", and "mut_freq"

Value

Returns a data frame suitable for circos plotting.

Author(s)

Matthew Woodruff, Emory University

Examples

```
df <- data.frame(population = c(rep("pop1", 10), rep("pop2", 10)),
  isotype = c(rep("M", 4), rep("G", 6), rep("M", 6), "G", "G", "A", "A"),
  lineageID = c(1, 2, 3, 4, 1, 1, 1, 2, 4, 5, 1, 1, 1, 5, 6, 7, 1, 3, 1, 1),
  mut_freq = rnorm(20))

head(df)

circos_df <- make_circos_df_1(df)

head(circos_df)
```

make_circos_links *make_circos_links*

Description

A function that takes a data frame in the format of those returned by the ‘make_circos_df_*’ functions, and returns a data frame containing all possible lineage links, and some additional metadata about each link for filtering purposes.

***NOTE: This function is optimized for parallelized computing. Setting the "status" argument to true will eliminate this feature, and instead provide a status bar to track progress.

Usage

```
make_circos_links(df, chromosome_grouping, band_grouping, link_grouping,  
start_position, end_position, status = FALSE)
```

Arguments

df	A data frame in the format of those returned by the ‘make_circos_df_*’ functions
chromosome_grouping	A column defining the circos chromosomes. Passed as column index or "column name".
band_grouping	A column defining the circos chromosome bands. Passed as column index or "column name".
link_grouping	A column defining the circos linkage IDs. Passed as column index or "column name".
start_position	A column defining the start positions of the linkage ID.
end_position	A column defining the end positions of the linkage ID.
status	Defaults to FALSE. Removes parallelized computing, and provides a status bar to track progress.

Value

Returns a links data frame suitable for filtering and exporting using the ‘write_circos_links()’ function.

Author(s)

Matthew Woodruff, Emory University

Examples

```
links_df <- data.frame(chrom = c(rep("chr1", 5), rep("chr2", 5)),
  band = c(rep("band1", 3), rep("band2", 2), "band1", rep("band2", 4)),
  link = c(1, 2, 3, 1, 2, 1, 1, 3, 4, 5),
  start = c(1, 3, 5, 10, 35, 1, 5, 8, 13, 15),
  end = c(3, 5, 10, 35, 39, 5, 8, 13, 15, 21))

links <- make_circos_links(links_df, "chrom", "band", "link", "start", "end", status = TRUE)

print(links)
```

```
write_circos_histogram
      write_circos_histogram
```

Description

A function that takes a data.frame, and creates a histogram.txt file for input into circos

Usage

```
write_circos_histogram(df, chromosome_grouping, start_position, end_position,
  values, file_name = "histogram.txt", file_path = NULL)
```

Arguments

df	A database to draw Circos data from
chromosome_grouping	A column defining the circos chromosomes. Passed as column index or "column name"
start_position	A column defining the start positions of the desired value
end_position	A column defining the end positions of the desired value
values	The value to be plotted
file_name	The desired file name. Defaults to histogram.txt
file_path	The desired file path destination folder. Defaults to NULL

Value

Writes a Circos-compatible histogram file to the desired directory

Author(s)

Matthew Woodruff, Emory University

Examples

```
df <- data.frame(chrom = c(1,1,2,2), start = c(1, 5, 1, 8),
end = c(5, 10, 8, 13), plotting_value = c(5, 78, 9, 2))

write_circos_histogram(df = df, chromosome_grouping = "chrom",
start_position = "start", end_position = "end",
value = "plotting_value", file_name = "histogram.txt", file_path = tempdir())
```

```
write_circos_karyotype
      write_circos_karyotype
```

Description

A function that takes a data frame in the format of those returned by the ‘make_circos_df_*’ functions, and writes a "karyotype" file for Circos plotting

Usage

```
write_circos_karyotype(df, chromosome_grouping = FALSE,
band_grouping = FALSE, start_position, end_position,
file_name = "karyotype.txt", file_path = NULL)
```

Arguments

df	A database to draw Circos data from
chromosome_grouping	A column defining the circos chromosomes. Passed as column index or "column name"
band_grouping	A column defining the circos bands. Passed as column index or "column name"
start_position	A column defining the lineage start positions. Passed as column index or "column name"
end_position	A column defining the lineage start positions. Passed as column index or "column name"
file_name	The desired file name. Defaults to karyotype.txt in the current working directory
file_path	The desired file path destination folder. Defaults to NULL

Value

Writes a Circos-compatible karyotype file to the desired directory

Author(s)

Matthew Woodruff, Emory University

Examples

```
df <- data.frame(chrom = c(1, 1, 2, 2), band = c(1, 2, 1, 2), start = c(1, 5, 1, 8),
end = c(5, 10, 8, 13), n = c(5, 5, 8, 5))

write_circos_karyotype(df = df, chromosome_grouping = "chrom", band_grouping = "band",
start_position = "start", end_position = "end",
file_name = "karyotype.txt", file_path = tempdir())
```

write_circos_links *write_circos_links*

Description

A function that takes a data frame in the format of those returned by the 'make_circos_links' function, and writes a "links" file for Circos plotting

Usage

```
write_circos_links(df, include_colors = FALSE, file_name = "links.txt",
file_path = NULL)
```

Arguments

df	A data frame in the format of those returned by the 'make_circos_links' function
include_colors	Include colors generated by the 'color_circos_links' function in the write file
file_name	The desired file name. Defaults to links.txt in the current working directory
file_path	The desired file path destination folder. Defaults to NULL

Value

Writes a Circos-compatible links file to the desired directory

Author(s)

Matthew Woodruff, Emory University

Examples

```
df <- data.frame(lin_id = c(1,2), chr1 = c(1,1), band1 = c(1,1),
chr1_start = c(1,5), chr1_end = c(5,8),
n1 = c(5,3), chr2 = c(1,2), band2 = c(2,1),
chr2_start = c(8,1), chr2_end = c(13,5), n2 = c(5,5))

write_circos_links(df = df, file_name = "links.txt", file_path = tempdir())
```


Index

`color_circos_links`, 2

`generate_searchable_links`, 3

`make_circos_df_1`, 4

`make_circos_links`, 5

`write_circos_histogram`, 6

`write_circos_karyotype`, 7

`write_circos_links`, 8