

Package ‘D3GB’

April 11, 2017

Type Package

Version 1.1

Date 2017-04-10

Title Interactive Genome Browser with R

Description Creates interactive genome browser with 'R'. It joins the data analysis power of R and the visualization libraries of JavaScript in one package.

URL <http://d3gb.usal.es>

License CC BY-NC-SA 4.0

Depends R (>= 3.0.0)

Imports RSQLite (>= 1.0.0), DBI (>= 0.3.1), stats, utils

Suggests knitr, rmarkdown

VignetteBuilder knitr

NeedsCompilation no

Author Carlos Prieto <cprietos@usal.es>,
David Barrios <metal@usal.es>

Maintainer Carlos Prieto <cprietos@usal.es>

Repository CRAN

Date/Publication 2017-04-10 22:29:14 UTC

R topics documented:

createAssembly	2
D3GB	3
gbk2genomebrowser	3
genomebrowser	4
genomemap	5
genome_addGFF	6
genome_addSequence	7
genome_addTrack	8
genome_addVCF	9

getAssemblyFromFasta	10
GRCh37	11
GRCh37.bands	11
GRCh38	11
GRCh38.bands	12
NCBI36	12
segmentation	12

Index	14
--------------	-----------

createAssembly	<i>Creates a genome assembly for genomebrowser or genomemap.</i>
----------------	--

Description

createAssembly create assemblies for their use as parameter of genomebrowser or genomemap.

Usage

```
createAssembly(name, size)
```

Arguments

name	a vector with the chromosome or scaffold names.
size	a vector with the chromosome or scaffold sizes.

Value

a data frame in BED format to be used as an assembly in genomebrowser or genomemap.

Author(s)

David Barrios and Carlos Prieto. Bioinformatics, University of Salamanca. See <http://d3gb.usal.es/>

See Also

The 'D3GB' Website: <http://d3gb.usal.es>
[genomebrowser](#), [genomemap](#).

Examples

```
# A Micromonospora scaffolds' names and sizes
scaffolds <- c("NZ_HF570105.1", "NZ_HF570106.1", "NZ_HF570107.1", "NZ_HF570108.1")
sizes <- c(583, 1327, 241394, 7082520)

# Create an assembly
createAssembly(scaffolds, sizes)
```

D3GB

Interactive genome browsers with R

Description

Create interactive genome browsers with 'R'.

Author(s)

David Barrios and Carlos Prieto. Bioinformatics, University of Salamanca. See <http://d3gb.usal.es/>

gbk2genomebrowser

Generates an interactive genome browser.

Description

gbk2genomebrowser creates an interactive genome browser from a GenBank file.

Usage

```
gbk2genomebrowser(gbkfile, server = FALSE, dir = "GenomeBrowser")
```

Arguments

gbkfile	a "character" string representing the input GenBank file to be represented in the genome browser.
server	a logical value to enable (TRUE) or disable (FALSE) the server mode. Server mode: designed to be shared as a website. Resulting folder should be added to the Apache applications directory and enable writing permissions. In this way the genome browser will be working as a web site. Local mode: the genome browser will be functional on your local machine.
dir	a "character" string representing the directory where the graph will be saved.

Value

The function creates a folder in the working directory of your computer with an HTML document (index.html) which contains the graph. This file can be directly opened with Firefox. If you want to open this file stored in your local machine with other Web browser, please visit the help section on the D3GB Web site (<http://d3gb.usal.es>).

Author(s)

David Barrios and Carlos Prieto. Bioinformatics, University of Salamanca. See <http://d3gb.usal.es/>

Examples

```
## Not run:
# Download GenBank file
gbff <- tempfile()
download.file(paste0("ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/000/297/395/",
"GCF_000297395.2_ASM29739v2/GCF_000297395.2_ASM29739v2_genomic.gbff.gz"),gbff)

# Genome browser generation.
# It creates a genome browser ready to be viewed in Firefox.
# For a server version, ready to be shared with Apache as a Website, set the parameter server=True
gb <- gbk2genomebrowser(gbff, dir = "Micromonospora_gbk")

## End(Not run)
```

genomebrowser	<i>Generates an interactive genome browser.</i>
---------------	---

Description

genomebrowser creates an interactive genome browser.

Usage

```
genomebrowser(assembly, tracks = NA, types = NA, colors = NA,
mapTrack = NA, server = FALSE, dir = "GenomeBrowser")
```

Arguments

tracks	a list of data frames with track information that will be viewed in the genome browser. These data frames should have a BED format with the following columns: chromosome/scaffold, start, end, name, score, strand.
types	a "character" vector with type of tracks that will be represented. It should be one of: "gene", "exons", "domain", "value" or "score".
colors	a "character" vector with color used for the track representation.
mapTrack	a data frame with values to represent on the genome map in BED format: chromosome, start, end, name, value.
assembly	a genome assembly data frame. D3GB provides human assemblies (NCBI36, GRCh37, GRCh38), human assemblies with cytobands (GRCh37.bands or GRCh38.bands), or methods for creating your own assembly from a FASTA file getAssemblyFromFasta .
server	a logical value to enable (TRUE) or disable (FALSE) the server mode. Server mode: designed to be shared as a website. Resulting folder should be added to the Apache applications directory and enable writing permissions. In this way the genome browser will be working as a web site. Local mode: the genome browser will be functional on your local machine.
dir	a "character" string representing the directory where the graph will be saved.

Value

The function creates a folder in the working directory of your computer with an HTML document (index.html) which contains the graph. This file can be directly opened with Firefox. If you want to open this file stored in your local machine with other Web browser, please visit the help section on the D3GB Web site (<http://d3gb.usal.es>).

Author(s)

David Barrios and Carlos Prieto. Bioinformatics, University of Salamanca. See <http://d3gb.usal.es/>

See Also

The 'D3GB' Website: <http://d3gb.usal.es>
[createAssembly](#), [getAssemblyFromFasta](#), [GRCh37.bands](#), [GRCh38.bands](#), [genome_addGFF](#), [genome_addSequence](#), [genome_addTrack](#), [genome_addVCF](#).

Examples

```
# Genome browser generation.
# It creates a genome browser ready to be viewed in Firefox.
# For a server version, ready to be shared with Apache as a Website, set the parameter server=True
gb <- genomebrowser(GRCh38.bands, dir="GRCh38_browser")
# Now you can add sequences, tracks, gff or vcf
```

genomemap

Create an interactive genome map.

Description

genomemap creates an interactive genome map.

Usage

```
genomemap(assembly, mapTrack = NA, dir = "GenomeMap")
```

Arguments

assembly	a genome assembly data frame. D3GB provides human assemblies (NCBI36, GRCh37, GRCh38), human assemblies with cytobands (GRCh37.bands or GRCh38.bands), or methods for creating your own assembly from a FASTA file getAssemblyFromFasta .
mapTrack	a data frame with values to represent on the genome map in BED format: chromosome, start, end, name, value.
dir	a "character" string representing the directory where the graph will be saved.

Value

The function creates a folder in your computer with an HTML document named index.html which contains the graph. This file can be directly opened with Firefox. If you want to open this file stored in your local machine with other Web browser, please visit the help section on the D3GB Web site (<http://d3gb.usal.es>).

Author(s)

David Barrios and Carlos Prieto. Bioinformatics, University of Salamanca. See <http://d3gb.usal.es/>

See Also

The 'D3GB' Website: <http://d3gb.usal.es>
[createAssembly](#), [getAssemblyFromFasta](#), [GRCh37.bands](#), [GRCh38.bands](#).

Examples

```
genomemap(GRCh38.bands, dir="GRCh38_map")
```

genome_addGFF	<i>Add track in a gff file to genomebrowser.</i>
---------------	--

Description

genome_addGFF add track in a gff file to genomebrowser.

Usage

```
genome_addGFF(gb, gfffile)
```

Arguments

gb	a genome browser object produced by the function genomebrowser.
gfffile	a "character" string representing the input gff file to be represented in the genome browser.

Value

The function add track in a gff file to genomebrowser.

Author(s)

David Barrios and Carlos Prieto. Bioinformatics, University of Salamanca. See <http://d3gb.usal.es/>

See Also

The 'D3GB' Website: <http://d3gb.usal.es>
[genomebrowser](#), [genome_addSequence](#), [genome_addTrack](#), [genome_addVCF](#).

Examples

```
## Not run:
# Download fasta file
fasta <- tempfile()
download.file(paste0("ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/000/297/395/",
"GCF_000297395.2_ASM29739v2/GCF_000297395.2_ASM29739v2_genomic.fna.gz"), fasta)

# Genome browser generation.
# It creates a genome browser ready to be viewed in Firefox.
# For a server version, ready to be shared with Apache as a Website, set the parameter server=True
gb <- genomebrowser(getAssemblyFromFasta(fasta), dir = "Micromonospora_gff")

genome_addSequence(gb, fasta)

# Download gff file and add to the genome browser
gff <- tempfile()
download.file(paste0("ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/000/297/395/",
"GCF_000297395.2_ASM29739v2/GCF_000297395.2_ASM29739v2_genomic.gff.gz"), gff)

genome_addGFF(gb, gff)

## End(Not run)
```

genome_addSequence *Add genome sequence on Fasta format to genomebrowser.*

Description

genome_addSequence add FASTA sequences to genomeviewer.

Usage

```
genome_addSequence(gb, fastafilename)
```

Arguments

gb a genome browser object produced by the function genomebrowser.
fastafilename a "character" string representing the input Fasta file to be added in the genome browser.

Value

The function add genome sequence on Fasta format to genomebrowser.

Author(s)

David Barrios and Carlos Prieto. Bioinformatics, University of Salamanca. See <http://d3gb.usal.es/>

See Also

The 'D3GB' Website: <http://d3gb.usal.es>
[genomebrowser](#), [genome_addGFF](#), [genome_addTrack](#), [genome_addVCF](#).

Examples

```
## Not run:
# Download fasta file
fasta <- tempfile()
download.file(paste0("ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/000/297/395/",
"GCF_000297395.2_ASM29739v2/GCF_000297395.2_ASM29739v2_genomic.fna.gz"), fasta)

# Genome browser generation.
# It creates a genome browser ready to be viewed in Firefox.
# For a server version, ready to be shared with Apache as a Website, set the parameter server=True
gb <- genomebrowser(getAssemblyFromFasta(fasta), dir = "Micromonospora_sequence")

genome_addSequence(gb, fasta)

## End(Not run)
```

genome_addTrack	<i>Add tracks to genomebrowser.</i>
-----------------	-------------------------------------

Description

genome_addTrack add tracks to genomebrowser.

Usage

```
genome_addTrack(gb, track, trackname = NULL, type = "gene", color = "#000", scale = NA)
```

Arguments

gb	a genome browser object produced by the function genomebrowser.
track	a "character" string giving the input BED file or a data frame with track information to be represented in the genome browser.
trackname	a "character" string giving a name for the track.
type	a "character" string with the type of track should be drawn. Possible types are: "gene", "exons", "domain", "value" or "score".
color	a "character" string giving the color of the track.
scale	a vector with two values which specifies the minimum and maximum limits in the representation of the "score" or "value" tracks. By default maximum and minimum scores are taken as the limits.

Value

The function add an input track to genomebrowser.

Author(s)

David Barrios and Carlos Prieto. Bioinformatics, University of Salamanca. See <http://d3gb.usal.es/>

Examples

```
# Segmentate tracks.  
  
genome_addTrack(gb, track)
```

genome_addVCF	<i>Add VCF tracks to genomebrowser.</i>
---------------	---

Description

genome_addVCF add VCF tracks to genomebrowser.

Usage

```
genome_addVCF(gb, vcffile, trackname = NULL, show = NULL)
```

Arguments

gb	a genome browser object produced by the function genomebrowser.
vcffile	a "character" string representing the input VCF file to be represented in the genome browser.
trackname	a "character" string giving a name for the track.
show	a "character" vector giving the info features to display.

Value

The function add a VCF track to genomebrowser.

Author(s)

David Barrios and Carlos Prieto. Bioinformatics, University of Salamanca. See <http://d3gb.usal.es/>

getAssemblyFromFasta *Create an assembly from a FASTA file for its use as parameter of genomebrowser or genomemap.*

Description

getAssemblyFromFasta create assemblies from a FASTA file for their use as parameter of genomebrowser or genomemap.

Usage

```
getAssemblyFromFasta(fasta)
```

Arguments

fasta a "character" string representing the input Fasta file to be added in the genome browser.

Value

a data frame in BED format ready to use as an assembly in genomebrowser or genomemap.

Note

FASTA format is a text-based format for representing either nucleotide sequences or peptide sequences, in which nucleotides or amino acids are represented using single-letter codes.

Author(s)

David Barrios and Carlos Prieto. Bioinformatics, University of Salamanca. See <http://d3gb.usal.es/>

See Also

The 'D3GB' Website: <http://d3gb.usal.es>
[genomebrowser](#), [genomemap](#).

Examples

```
# Download fasta file
fasta <- tempfile()
download.file(paste0("ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/000/297/395/",
"GCF_000297395.2_ASM29739v2/GCF_000297395.2_ASM29739v2_genomic.fna.gz"), fasta)

# Assembly generation
getAssemblyFromFasta(fasta)
```

GRCh37	<i>Length of human chromosomes based on GRCh37 assembly</i>
--------	---

Description

Length of human chromosomes based on GRCh37 assembly.

Source

See http://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.13/#/st.

Examples

```
head(GRCh37)
```

GRCh37.bands	<i>Cytoband information based on GRCh37</i>
--------------	---

Description

A Data Frame containing cytobands of GRCh37 assembly.

Source

See http://grch37.ensembl.org/Homo_sapiens/Location/Genome.

Examples

```
head(GRCh37.bands)
```

GRCh38	<i>Length of human chromosomes based on GRCh38 assembly</i>
--------	---

Description

Length of human chromosomes based on GRCh38 assembly.

Source

See <http://www.ncbi.nlm.nih.gov/assembly/883148/#/st>.

Examples

```
head(GRCh38)
```

`GRCh38.bands`*Cytoband information based on GRCh38*

Description

A Data Frame containing cytobands of GRCh38 assembly.

Source

See http://www.ensembl.org/Homo_sapiens/Location/Genome.

Examples

```
head(GRCh38.bands)
```

`NCBI36`*Length of human chromosomes based on NCBI36 assembly*

Description

Length of human chromosomes based on NCBI36 assembly.

Source

See http://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.12/#/st.

Examples

```
head(NCBI36)
```

`segmentation`*Function for segmentate tracks in BED format.*

Description

`segmentation` performs a segmentation of track data in bed format in order to enable its effective representation.

Usage

```
segmentation(track, cell)
```

Arguments

track a Data Frame with genomic features (in BED format).
cell a numeric value with the size of each segment.

Value

a segmented Data Frame in BED format.

Note

Input file is provided in BED format which is an universal format for genomic features.

Author(s)

David Barrios and Carlos Prieto. Bioinformatics, University of Salamanca. See <http://d3gb.usal.es/>

See Also

The 'D3GB' Website: <http://d3gb.usal.es>
[genomebrowser](#), [genomemap](#), [genome_addTrack](#).

Examples

```
# Segmentate tracks.  
# Create test data  
chr <- character()  
pos <- numeric()  
  
for(i in 1:nrow(GRCh38)){  
  chr <- c(chr,as.character(rep(GRCh38[i,"chr"],100)))  
  pos <- c(pos,sample(GRCh38[i,"start"]:GRCh38[i,"end"],100))  
}  
  
value <- round(rexp(length(pos)),2)  
  
track <- data.frame(chr,pos,pos+1,NA,value)  
segments <- segmentation(track, 50000)
```

Index

`createAssembly`, [2](#), [5](#), [6](#)

D3GB, [3](#)

D3GB-package (D3GB), [3](#)

`gbk2genomebrowser`, [3](#)

`genome_addGFF`, [5](#), [6](#), [8](#)

`genome_addSequence`, [5](#), [7](#), [7](#)

`genome_addTrack`, [5](#), [7](#), [8](#), [8](#), [13](#)

`genome_addVCF`, [5](#), [7](#), [8](#), [9](#)

`genomebrowser`, [2](#), [4](#), [7](#), [8](#), [10](#), [13](#)

`genomemap`, [2](#), [5](#), [10](#), [13](#)

`getAssemblyFromFasta`, [4–6](#), [10](#)

GRCh37, [11](#)

GRCh37.bands, [5](#), [6](#), [11](#)

GRCh38, [11](#)

GRCh38.bands, [5](#), [6](#), [12](#)

NCBI36, [12](#)

`segmentation`, [12](#)