Package 'PACVr'

February 21, 2020

Version 0.9.2

Date 2020-02-09

Title Plastome Assembly Coverage Visualization

Author Michael Gruenstaeudl [aut, cre],
Nils Jenke [ctb]

Maintainer Michael Gruenstaeudl <m.gruenstaeudl@fu-berlin.de>

Depends R (>= 3.3.0)

Imports RCircos (>= 1.2.0), optparse (>= 1.6.0), genbankr (>= 1.2.1), BiocGenerics (>= 0.20.0), Biostrings (>= 2.48.0), GenomicAlignments (>= 1.18.1)

SystemRequirements mosdepth

Description Visualizes the coverage depth of a complete plastid genome as well as the equality of its inverted repeat regions in relation to the circular, quadripartite genome structure and the location of individual genes. For more information, please see Gruenstaeudl and Jenke (2019) <doi:10.1101/697821>.

License BSD 3-clause License + file LICENSE

OS_type unix

NeedsCompilation no

Repository CRAN

Date/Publication 2020-02-21 00:10:07 UTC

R topics documented:

PACVr-package *Plastome Assembly Coverage Visualization in R*

Description

PACVr visualizes the coverage depth of a complete plastid genome as well as the equality of its inverted repeat regions in relation to the circular, quadripartite genome structure and the location of individual genes.

Note

Software Dependencies

For full functionality, **PACVr** requires the installation of mosdepth (https://github.com/brentp/mosdepth) on the system.

Input Requirements

The input to **PACVr** consists of two input files that contain information on (a) genome sequence and structure, and (b) on coverage depth. Specifically, users must provide (a) a file in GenBank flat file format that complies with the GenBank record specifications (https://www.ncbi.nlm.nih.gov/Sitemap/samplerecord.html), and (b) a file in BAM format that complies with the specifications described in the Sequence Alignment/Map Format documentation (https://samtools.github.io/hts-specs/SAMv1.pdf) and is accompanied by an ancillary index file.

Data Requirements

The user-supplied GenBank flat file must contain a sequence record of a complete plastid genome. For effective visualizations, the sequence record should contain a sequence feature for each of the inverted repeat regions, whereby the regions should be named 'IRa' and 'IRb' or 'Inverted Repeat a' and 'Inverted Repeat b', respectively. For effective visualizations, the total sequence length of the complete genome should be between 50 kb and 250 kb.

Author(s)

Michael Gruenstaeudl, Nils Jenke

Maintainer: Michael Gruenstaeudl <m.gruenstaeudl@fu-berlin.de>

References

Gruenstaeudl, M. and Jenke, N. (2019). PACVr: Plastome Assembly Coverage Visualization in R. bioRxiv 697821; doi: https://doi.org/10.1101/697821

PACVr.complete *Execute the complete pipeline of* **PACVr**

Description

This function executes the complete pipeline of **PACVr** via a single command.

Usage

```
PACVr.complete(gbk.file,
               bam.file,
               windowSize = 250,
               mosdepthCmd = 'mosdepth',
               logScale = FALSE,
               threshold = 0.5,
               syntenyLineType = 1,
               relative = TRUE,
               textSize = 0.5,
               delete = TRUE,
               output = NA)
```

Arguments

gbk.file	a character vector that specifies the name of, and path to, the GenBank input file
bam.file	a character vector that specifies the name of, and path to, the BAM input file
windowSize	a numeric value that specifies window size in which the coverage is calculated
mosdepthCmd	a character vector that specifies the command to execute mosdepth on the system
logScale	a boolean that specifies if the coverage depth is to be log-transformed before visualizing it
threshold	a numeric value that specifies the threshold for plotting coverage depth bars in red as opposed to the default black
syntenyLineType	a numeric value of 1, 2 or 3 that specifies the line type for visualizing IR gene synteny; 1 = ribbon lines, 2 = solid lines, 3 = no line
relative	a boolean that specifies whether the threshold is a relative value of the average coverage instead of an absolute value
textSize	a numeric value that specifies the relative font size of the text element in the visualization
delete	the decision to delete temporary files upon program execution
output	a character vector that specifies the name of, and path to, the output file

Examples

```
gbkFile <- system.file("extdata", "NC_045072/NC_045072.gb", package="PACVr")
bamFile <- system.file("extdata", "NC_045072/NC_045072_PlastomeReadsOnly.sorted.bam",
                       package="PACVr")
outFile <- paste(tempdir(), "/NC_045072__all_reads.pdf", sep="")

PACVr.complete(gbk.file=gbkFile, bam.file=bamFile, windowSize=250,
               mosdepthCmd='mosdepth', threshold=0.5, syntenyLineType=1,
               relative=TRUE, textSize=0.5, delete=TRUE, output=outFile)
```

`RCircos.Env` *Export the custom environment 'RCircos.Env'*

Description

This function exports the custom environment 'RCircos.Env' used by **RCircos**.

Details

PACVr employs **RCircos** as its visualization engine. In its operation, **RCircos** defines a custom environment (called 'RCircos.Env') and reads/writes variables to this environment from various of its functions. In order to make this environment accessible to **RCircos** within **PACVr** and, simultaneously, fulfil the requirements of CRAN, this export command was created. For more information, please see the Stackoverflow post at https://stackoverflow.com/questions/56875962/r-package-transferring-environment-from-imported-package.

Index