Package 'PACViR'

June 13, 2019

Version 0.4

Date 2019-06-13

Title Plastome Assembly Coverage Visualization in R

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

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

Depends R (>= 3.6.0)

biocViews

Imports RCircos (>= 1.2.0), optparse (>= 1.6.0), genbankr (>= 1.12.0),
BiocGenerics (>= 0.30.0)

SystemRequirements mosdepth (>= 0.2.2)

Description A user-friendly software tool to visualize the coverage depth of a complete plastid genome as well as the equality of its IR regions while simultaneously accounting for the circular, quadripartite structure of the genome, thus providing an aid to optimize the process of plastid genome assembly.

License BSD_3_clause + file LICENSE

R topics documented:

`PACViR-package` *Plastome Assembly Coverage Visualization in R*

Description

PACViR is a user-friendly software tool to visualize the coverage depth of a complete plastid genome as well as the equality of its IR regions while simultaneously accounting for the circular, quadripartite structure of the genome, thus providing an aid to optimize the process of plastid genome assembly.

Note

Software Requirements

Mandatory requirements for **PACViR**:

External installation of Mosdepth (>= 0.2.2) following https://github.com/brentp/mosdepth.

Dataset Requirements

The user-supplied data set must contain plastome data with information about the exact postion of the repeat regions. Those regions have to be named as 'IRa', 'IRb' or 'Inverted Repeat a', 'Inverted Repeat b'. The sequence length of the genome should be within 100kb-200kb; otherwise it may lead to an erroneous visualization.

Input File Requirements

In order to execute **PACViR**, a user must provide paths to two different types of input files: (a) a file in GenBank format and (b) a file in BAM format. The GenBank flat file format needs to comply with the GenBank record specifications of NCBI (https://www.ncbi.nlm.nih.gov/Sitemap/samplerecord.html). The BAM file format needs to comply with the specifications described in the Sequence Alignment/Map Format documentation (https://samtools.github.io/hts-specs/SAMv1.pdf). Please note that the accompanying BAM index file is mandatory for the use of **PACViR** and must, thus, be present.

Author(s)

Michael Gruenstaeudl, Nils Jenke Maintainer: Michael Gruenstaeudl <m.gruenstaeudl@fu-berlin.de>, Nils Jenke <nilsj24@zedat.fu-berlin.de>

References

Gruenstaeudl M., Jenke N. (2019) PACViR: Plastome Assembly Coverage Visualization in R. R package version 0.4.

`PACViR.complete` *Execute the complete PACViR pipeline via a single command*

Description

This function executes the complete **PACViR** pipeline.

Usage

```
PACViR.complete(gbk.file,
                bam.file,
                windowSize=250,
                mosdepthCmd='mosdepth',
                threshold=25,
                delete=TRUE,
                output='./PACViR_output.svg')
```

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 system
threshold	a numeric value that specifies the threshold for plotting coverage at different color
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

Details

This is the main function of **PACViR** to execute the full pipeline. It takes a Genbank file and a BAM file that are used to visualize plastome data in a circular way using the **RCircos** package. Those two input arguments are mandatory. The visualization contains gene names, gene locations and coverage data displayed as histrogram. The pipeline will produce a svg file format containing the circular visualization of the plastome data. Additionally, it will give the mosdepth output.

Value

invisible NULL.

Author(s)

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

References

Gruenstaeudl M., Jenke N. (2019) PACViR: Plastome Assembly Coverage Visualization in R. R package version 0.4.

Examples

```
# MH161174
gbkFile <- system.file("extdata", "MH161174/MH161174.gb", package="PACViR")
bamFile <- system.file("extdata", "MH161174/MH161174_PlastomeReadsOnly.sorted.bam",
                       package="PACViR")
outFile <- paste(getwd(), "/MH161174_AssemblyCoverage_viz.svg", sep="")
PACViR.complete(gbk.file=gbkFile, bam.file=bamFile, windowSize=250,
                mosdepthCmd='mosdepth', threshold=15, delete=FALSE, output=outFile)

# MH899017
gbkFile <- system.file("extdata", "MH899017/MH899017.gb", package="PACViR")
bamFile <- system.file("extdata", "MH899017/MH899017_PlastomeReadsOnly.sorted.bam",
                       package="PACViR")
outFile <- paste(getwd(), "/MH899017_AssemblyCoverage_viz.svg", sep="")
PACViR.complete(gbk.file=gbkFile, bam.file=bamFile, windowSize=250,
                mosdepthCmd='mosdepth', threshold=15, delete=FALSE, output=outFile)
```

Index