Package: ProActive (via r-universe)

January 14, 2025

Title Detect Elevations and Gaps in Mapped Sequencing Read Coverage

Version 0.0.2

Maintainer Jessie Maier <jlmaier@ncsu.edu>

Description Automate the detection of gaps and elevations in mapped sequencing read coverage using a 2D pattern-matching algorithm. 'ProActive' detects, characterizes and visualizes read coverage patterns in both genomes and metagenomes. Optionally, users may provide gene predictions associated with their genome or metagenome in the form of a .gff file. In this case, 'ProActive' will generate an additional output table containing the gene predictions found within the detected regions of gapped and elevated read coverage.

License GPL-2

Encoding UTF-8

LazyData true

RoxygenNote 7.3.2

URL https://github.com/jlmaier12/ProActive, https://jlmaier12.github.io/ProActive/

BugReports https://github.com/jlmaier12/ProActive/issues

Imports utils, stats, dplyr, ggplot2, stringr

Suggests knitr, rmarkdown, testthat (>= 3.0.0), kableExtra

VignetteBuilder knitr

Depends R (>= 4.2.0)

Config/testthat/edition 3

NeedsCompilation no

Author Jessie Maier [aut, cre, cph] (<https://orcid.org/0009-0001-8575-5386>), Manuel Kleiner [aut, ths] (<https://orcid.org/0000-0001-6904-0287>)

Repository CRAN

Date/Publication 2025-01-14 09:50:10 UTC

Config/pak/sysreqs libicu-dev

Contents

`plotProActiveResults` *Plot results of 'ProActive()' pattern-matching*

Description

Plot read coverage of contigs/chunks with detected gaps and elevations and their associated pattern-match.

Usage

```
plotProActiveResults(pileup, ProActiveResults, elevFilter, saveFilesTo)
```

Arguments

`pileup`	A .txt file containing mapped sequencing read coverages averaged over 100 bp windows/bins.
`ProActiveResults`	The output from 'ProActive()'.
`elevFilter`	Optional, only plot results with pattern-matches that achieved an elevation ratio (max/min) greater than the specified values. Default is no filter.
`saveFilesTo`	Optional, Provide a path to the directory you wish to save output to. A folder will be made within the provided directory to store results.

Value

A list containing ggplot objects

Examples

```
ProActivePlots <- plotProActiveResults(sampleMetagenomePileup,
                                       sampleMetagenomeResults)
```

`ProActive` *Detect elevations and gaps in mapped read coverage patterns.*

Description

Performs read coverage pattern-matching and summarizes the results into a list. The first list item summarizes the pattern-matching results. The second list item is the 'cleaned' version of the summary table with all the 'noPattern' classifications removed. (i.e were not filtered out). The third list item contains the pattern-match information needed for pattern-match visualization with 'plotProActiveResults()'. The fourth list item is a table containing all the contigs that were filtered out prior to pattern-matching. The fifth list item contains arguments used during pattern-matching (windowSize, mode, chunkSize, chunkContigs). If the user provides a gffTSV files, then the last list is a table consisting of ORFs found within the detected gaps and elevations in read coverage.

Usage

```
ProActive(
  pileup,
  mode,
  gffTSV,
  windowSize = 1000,
  chunkContigs = FALSE,
  minSize = 10000,
  maxSize = Inf,
  minContigLength = 30000,
  chunkSize = 1e+05,
  IncludeNoPatterns = FALSE,
  verbose = TRUE,
  saveFilesTo
)
```

Arguments

`pileup`	A .txt file containing mapped sequencing read coverages averaged over 100 bp windows/bins.
`mode`	Either "genome" or "metagenome"
`gffTSV`	Optional, a .gff file (TSV) containing gene predictions associated with the .fasta file used to generate the pileup.
`windowSize`	The number of basepairs to average read coverage values over. Options are 100, 200, 500, 1000 ONLY. Default is 1000.
`chunkContigs`	TRUE or FALSE, If TRUE and 'mode'="metagenome", contigs longer than the 'chunkSize' will be 'chunked' into smaller subsets and pattern-matching will be performed on each subset. Default is FALSE.
`minSize`	The minimum size (in bp) of elevation or gap patterns. Default is 10000.
`maxSize`	The maximum size (in bp) of elevation or gap patterns. Default is NA (i.e. no maximum).

`minContigLength`
: The minimum contig/chunk size (in bp) to perform pattern-matching on. Default is 25000.

`chunkSize`
: If ‘mode‘="genome" OR if ‘mode‘="metagenome" and ‘chunkContigs‘=TRUE, chunk the genome or contigs, respectively, into smaller subsets for pattern-matching. ‘chunkSize‘ determines the size (in bp) of each ’chunk’. Default is 100000.

`IncludeNoPatterns`
: TRUE or FALSE, If TRUE the noPattern pattern-matches will be included in the ProActive PatternMatches output list. If you would like to visualize the noPattern pattern-matches in ‘plotProActiveResults()‘, this should be set to TRUE.

`verbose`
: TRUE or FALSE. Print progress messages to console. Default is TRUE.

`saveFilesTo`
: Optional, Provide a path to the directory you wish to save output to. A folder will be made within the provided directory to store results.

Value

A list containing 6 objects described in the function description.

Examples

```
metagenome_results <- ProActive(
  pileup = sampleMetagenomePileup,
  mode = "metagenome",
  gffTSV = sampleMetagenomegffTSV
)
```

Index