

Package: TrIdent (via r-universe)

February 15, 2025

Type Package

BiocType Software

Title TrIdent - Transduction Identification

Version 0.99.3

Description The `TrIdent` R package automates the analysis of transductomics data by detecting, classifying, and characterizing read coverage patterns associated with potential transduction events. Transductomics is a DNA sequencing-based method for the detection and characterization of transduction events in pure cultures and complex communities. Transductomics relies on mapping sequencing reads from a viral-like particle (VLP)-fraction of a sample to contigs assembled from the metagenome (whole-community) of the same sample. Reads from bacterial DNA carried by VLPs will map back to the bacterial contigs of origin creating read coverage patterns indicative of ongoing transduction.

License GPL-2

Encoding UTF-8

LazyData false

Imports graphics, utils, stats, dplyr, ggplot2, patchwork, stringr, tidy, roll

biocViews Coverage, Metagenomics, PatternLogic, Classification, Sequencing

RoxygenNote 7.3.2

URL <https://github.com/jlmaier12/TrIdent>,
<https://jlmaier12.github.io/TrIdent/>

BugReports <https://github.com/jlmaier12/TrIdent/issues>

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

VignetteBuilder knitr

Depends R (>= 4.2.0)

Config/testthat/edition 3

Config/pak/sysreqs make libicu-dev

Repository https://bioc.r-universe.dev

RemoteUrl https://github.com/bioc/TrIdent

RemoteRef HEAD

RemoteSha bdb120b45c511f9590c8ee987461fd0dd8bf2cc9

Contents

plotTrIdentResults	2
specializedTransductionID	3
TrIdentClassifier	5
Index	7

plotTrIdentResults	<i>Plot read coverage graphs of contigs classified as Prophage-like, Sloping, or HighCovNoPattern</i>
--------------------	---

Description

Plot the read coverages of a contig and its associated pattern-match for Prophage-like, Sloping and HighCovNoPattern classifications. Returns a list of ggplot objects.

Usage

```
plotTrIdentResults(
  VLPpileup,
  WCPileup,
  TrIdentResults,
  matchScoreFilter,
  saveFilesTo
)
```

Arguments

VLPpileup	VLP-fraction pileup file generated by mapping sequencing reads from a sample's ultra-purified VLP-fraction mapped to the sample's whole-community metagenome assembly. The pileup file MUST have the following format: * V1: Contig accession * V2: Mapped read coverage values averaged over 100 bp windows * V3: Starting position (bp) of each 100 bp window. Restarts from 0 at the start of each new contig. * V4: Starting position (bp) of each 100 bp window. Does NOT restart at the start of each new contig.
-----------	---

WCpileup	A whole-community pileup file generated by mapping sequencing reads from a sample's whole-community mapped to the sample's whole-community metagenome assembly. The pileup file MUST have the following format: * V1: Contig accession * V2: Mapped read coverage values averaged over 100 bp windows * V3: Starting position (bp) of each 100 bp window. Restarts from 0 at the start of each new contig. * V4: Starting position (bp) of each 100 bp window. Does NOT restart at the start of each new contig.
TrIdentResults	Output from 'TrIdentClassifier()'.
matchScoreFilter	Optional, Filter plots using the normalized pattern match-scores. A suggested filtering threshold is provided by 'TrIdentClassifier()' if 'suggFiltThresh=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

Large list containing ggplot objects

Examples

```
data("VLPFractionSamplePileup")
data("WholeCommunitySamplePileup")
data("TrIdentSampleOutput")

patternMatches <- plotTrIdentResults(
  VLPpileup = VLPFractionSamplePileup,
  WCPileup = WholeCommunitySamplePileup,
  TrIdentResults = TrIdentSampleOutput
)
```

specializedTransductionID

Identify potential specialized transduction events on contigs classified as Prophage-like

Description

Search contigs classified as Prophage-like for dense read coverage outside of the pattern-match borders that may indicate specialized transduction. Returns a list with the first object containing a summary table and the second object containing a list of plots of with associated specialized transduction search results. If the plot is green, it has been identified as having potential specialized transduction.

Usage

```
specializedTransductionID(
  VLPpfileup,
  TrIdentResults,
  specificContig,
  noReadCov = 500,
  specTransLength = 2000,
  matchScoreFilter,
  logScale = FALSE,
  verbose = TRUE,
  SaveFilesTo
)
```

Arguments

VLPpfileup	VLP-fraction pileup file generated by mapping sequencing reads from a sample's ultra-purified VLP-fraction mapped to the sample's whole-community metagenome assembly. The pileup file MUST have the following format: * V1: Contig accession * V2: Mapped read coverage values averaged over 100 bp windows * V3: Starting position (bp) of each 100 bp window. Restarts from 0 at the start of each new contig. * V4: Starting position (bp) of each 100 bp window. Does NOT restart at the start of each new contig.
TrIdentResults	Output from 'TrIdentClassifier()'
specificContig	Optional, Search a specific contig classified as Prophage-like ("NODE_1").
noReadCov	Number of basepairs of zero read coverage encountered before specialized transduction searching stops. Default is 500. Must be at least 100.
specTransLength	Number of basepairs of non-zero read coverage needed for specialized transduction to be considered. Default is 2000. Must be at least 100.
matchScoreFilter	Optional, Filter plots using the normalized pattern match-scores. A suggested filtering threshold is provided by 'TrIdentClassifier()' if 'suggFiltThresh=TRUE'.
logScale	TRUE or FALSE, display VLP-fraction read coverage in log10 scale. Default is FALSE.
verbose	TRUE or FALSE. Print progress messages to console. Default is TRUE.
SaveFilesTo	Provide a path to the directory you wish to save output to. 'specializedTransductionID()' will make a folder within the provided directory to store results.

Value

Large list containing two objects

Examples

```
data("VLPFractionSamplePileup")
data("TrIdentSampleOutput")
```

```

specTransduction <- specializedTransductionID(
  VLPpfileup = VLPFractionSamplePileup,
  TrIdentResults = TrIdentSampleOutput
)

specTransductionNODE62 <- specializedTransductionID(
  VLPpfileup = VLPFractionSamplePileup,
  TrIdentResults = TrIdentSampleOutput,
  specificContig = "NODE_62"
)

```

TrIdentClassifier	<i>Classify contigs as Prophage-like, Sloping, HighCovNoPattern, and NoPattern</i>
-------------------	--

Description

Performs all the pattern-matching and summarizes the results into a list. The first item in the list is a table consisting of the summary information of all the contigs that passed through pattern-matching (i.e. were not filtered out). The second item in the list is a table consisting of the summary information of all contigs that were classified via pattern-matching. The third item in the list contains the pattern-match information associated with each contig in the previous table. The fourth object in the list is a table containing the contigs that were filtered out prior to pattern-matching. The fifth item is the windowSize used for the search.

Usage

```

TrIdentClassifier(
  VLPpfileup,
  WCPfileup,
  windowSize = 1000,
  minBlockSize = 10000,
  maxBlockSize = Inf,
  minContigLength = 30000,
  minSlope = 0.001,
  suggFiltThresh = FALSE,
  verbose = TRUE,
  SaveFilesTo
)

```

Arguments

VLPpfileup	VLP-fraction pileup file generated by mapping sequencing reads from a sample's ultra-purified VLP-fraction mapped to the sample's whole-community metagenome assembly. The pileup file MUST have the following format: * V1: Contig accession * V2: Mapped read coverage values averaged over 100 bp windows * V3: Starting position (bp) of each 100 bp window. Restarts from 0 at the start of each new contig. * V4: Starting position (bp) of each 100 bp window. Does NOT restart at the start of each new contig.
------------	---

WCpileup	A whole-community pileup file generated by mapping sequencing reads from a sample's whole-community mapped to the sample's whole-community metagenome assembly. The pileup file MUST have the following format: * V1: Contig accession * V2: Mapped read coverage values averaged over 100 bp windows * V3: Starting position (bp) of each 100 bp window. Restarts from 0 at the start of each new contig. * V4: Starting position (bp) of each 100 bp window. Does NOT restart at the start of each new contig.
windowSize	The number of basepairs to average read coverage values over. Options are 100, 200, 500, 1000 ONLY. Default is 1000.
minBlockSize	The minimum size (in bp) of the Prophage-like block pattern. Default is 10000. Must be at least 1000.
maxBlockSize	The maximum size (in bp) of the Prophage-like block pattern. Default is NA (no maximum).
minContigLength	The minimum contig size (in bp) to perform pattern-matching on. Must be at least 25000. Default is 30000.
minSlope	The minimum slope value to test for sloping patterns. Default is 0.001 (i.e minimum change of 10x read coverage over 100,000 bp).
suggFiltThresh	TRUE or FALSE, Suggest a filtering threshold for TrIdent classifications based on the normalized pattern-match scores. Default is FALSE.
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

Large list containing 5 objects

Examples

```
data("VLPFractionSamplePileup")
data("WholeCommunitySamplePileup")

TrIdent_results <- TrIdentClassifier(
  VLPpileup = VLPFractionSamplePileup,
  WCpileup = WholeCommunitySamplePileup
)
```

Index

`plotTrIdentResults`, [2](#)

`specializedTransductionID`, [3](#)

`TrIdentClassifier`, [5](#)