Package 'PIPETS'

November 12, 2026

Title Poisson Identification of PEaks from Term-Seq data

Version 1.8.0

Description PIPETS provides statistically robust analysis for 3'-seq/term-seq data. It utilizes a sliding window approach to apply a Poisson Distribution test to identify genomic positions with termination read coverage that is significantly higher than the surrounding signal. PIPETS then condenses proximal signal and produces strand specific results that contain all significant termination peaks.

License GPL-3

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.2

Depends R (>= 4.4.0)

Imports dplyr, utils, stats, GenomicRanges, BiocGenerics, methods

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

VignetteBuilder knitr

BugReports https://github.com/qfurumo/PIPETS/issues

URL https://github.com/qfurumo/PIPETS

biocViews Sequencing, Transcription, GeneRegulation, PeakDetection, Genetics, Transcriptomics, Coverage

Config/testthat/edition 3

git_url https://git.bioconductor.org/packages/PIPETS

git_branch RELEASE_3_23

git_last_commit ea66b69

git_last_commit_date 2026-04-28

Repository Bioconductor 3.23

Date/Publication 2026-11-11

Author Quinlan Furumo [aut, cre] (ORCID: <https://orcid.org/0000-0003-4486-302X>)

Maintainer Quinlan Furumo <furumo@bc.edu>

Contents

PIPETS_FullRun 2

Index **5**

`PIPETS_FullRun` *Analyze 3'-seq Data with PIPETS Poisson Identification of PEaks from Term-Seq data. This is the full run method that begins with input Bed file and returns the strand split results*

Description

PIPETS_FullRun

Usage

```
PIPETS_FullRun(
  inputData,
  readScoreMinimum,
  OutputFileID,
  OutputFileDir,
  slidingWindowSize = 25,
  slidingWindowMovementDistance = 25,
  threshAdjust = 0.75,
  threshAdjust_TopStrand = NA,
  threshAdjust_CompStrand = NA,
  user_pValue = 5e-04,
  highOutlierTrim = 0.01,
  highOutlierTrim_TopStrand = NA,
  highOutlierTrim_CompStrand = NA,
  adjacentPeakDistance = 2,
  peakCondensingDistance = 20,
  inputDataFormat = "bedFile"
)
```

Arguments

`inputData`	Either input Bed file or GRanges object. Either must have at least chromosome, start, stop, and strand information
`readScoreMinimum`	The user must input the minimum read score from the input bed files that is used to determine good quality reads. All values equal to and greater than the input are considered. In many modern sequencing runs, a score of 60 is used.
`OutputFileID`	User defined header for the output files of PIPETS. Will be the prefix for output bed and csv files.
`OutputFileDir`	User defined output file directory where all files generated by PIPETS will be placed

`slidingWindowSize`
: This parameter establishes the distance up and down stream of each position that a sliding window will be created around. The default value is 25, and this will result in a sliding window of total size 51 (25 upstream + position (1) + 25 downstream).

`slidingWindowMovementDistance`
: This parameter sets the distance that the sliding window will be moved. By default, it is set to move by half of the sliding window size in order to ensure that almost every position in the data is tested twice.

`threshAdjust`
: This parameter is used to establish a global cutoff threshold informed by the data. PIPETS sorts the genomic positions of each strand from highest to lowest, and starts with the highest read coverage position and subtracts that value from the total read coverage for that strand. By default, this continues until 75% of the total read coverage has been accounted for. Increasing the percentage (e.x. 0.9) will lower the strictness of the cutoff, thus increasing the total number of significant results.

`threshAdjust_TopStrand`
: Top strand specific threshAdjust value. If the user would like to run strand specific analysis, they should set threshAdjust to NA.

`threshAdjust_CompStrand`
: Comp strand specific threshAdjust value. If the user would like to run strand specific analysis, they should set threshAdjust to NA.

`user_pValue`
: Choose the minimum pValue that the Poisson distribution test must pass in order to be considered significant

`highOutlierTrim`
: This parameter is used along with threshAdjust to trim off the influence exerted by high read coverage outliers. By default, it removes the top 0.01 percent of the highest read coverage positions from the calculation of the global threshold (e.x. if there are 200 positions that make up 75% of the total reads, then this parameter will take the top 2 read coverage positions and remove them from the calculation of the global threshold). This parameter can be tuned to account for datasets with outliers that would otherwise severely skew the global threshold.

`highOutlierTrim_TopStrand`
: Top strand specific highOutlierTrim value. If the user would like to run strand specific analysis, they should set highOutlierTrim to NA.

`highOutlierTrim_CompStrand`
: Comp strand specific highOutlierTrim value. If the user would like to run strand specific analysis, they should set highOutlierTrim to NA.

`adjacentPeakDistance`
: During the peak condensing step, this parameter is used to define "adjacent" for significant genomic positions. This is used to identify initial peak structures in the data. By default this value is set to 2 to ensure that single instances of loss of signal are not sufficient to prevent otherwise contiguous peak signatures from being combined.

`peakCondensingDistance`
: Following the initial peak condensing step, this parameter is used to identify peak structures in the data that are close enough to be considered part of the

same termination signal. In testing, we have not identified cases in which two distinct termination signals so proximal that the default parameters incorrectly combine the signals together.

`inputDataFormat`

PIPETS currently supports "bedFile" (default) and "GRanges" as input formats

Value

PIPETS outputs strand specific results files as well as strand specific bed files to the directory that the R project is in.

Examples

```
## When run, the user will be prompted to provide a string for file names
## During the run, PIPETS will output the minumum read coverage cutoff for each strand
## After completion, the output files will be created in the R project directory

## For run with defualt strictness of analysis
PIPETS_FullRun(inputData = "PIPETS_TestData.bed", readScoreMinimum = 42,
OutputFileDir = "~/Desktop/", OutputFileID = "Antibiotic1")

## For a more strict run (can be run for files with high total read depth)
PIPETS_FullRun(inputData = "PIPETS_TestData.bed", readScoreMinimum = 42, threshAdjust = 0.6,
OutputFileDir = "~/Desktop/", OutputFileID = "Antibiotic1_Strict")

## For a less strict run (for data with low total read depth)
PIPETS_FullRun(inputData = "PIPETS_TestData.bed", readScoreMinimum = 42, threshAdjust = 0.9,
OutputFileDir = "~/Desktop/", OutputFileID = "Antibiotic1_Lax")
```

Index

PIPETS_FullRun, 2