Package 'revert'

December 18, 2025

Type Package

Title Reversion Mutation Detector for DNA Sequencing Data

Version 1.0.0

Date 2025-12-16

Author Hui Xiao [aut, cre],
Stephen Pettitt [aut],
Syed Haider [aut]

Maintainer Hui Xiao <hui.xiao@icr.ac.uk>

Description
Reversion mutations are secondary mutations that reverse the deleterious effects of an original pathogenic mutation, partially or fully restoring the gene's function. The revert package detects reversion mutations for a specific pathogenic mutation from DNA-seq bam files.

Depends R (>= 4.4.0)

License GPL-2

Imports BSgenome, BiocGenerics, Biostrings, GenomeInfoDb,
GenomicRanges, IRanges, Rsamtools, S4Vectors, ensembldb,
pwalign, stats, utils

Suggests knitr, rmarkdown, futile.logger, BSgenome.Hsapiens.UCSC.hg19,
EnsDb.Hsapiens.v75, BSgenome.Hsapiens.UCSC.hg38,
EnsDb.Hsapiens.v86

RoxygenNote 7.3.2

Encoding UTF-8

VignetteBuilder knitr

NeedsCompilation no

Repository CRAN

Date/Publication 2025-12-17 23:30:02 UTC

Contents

`getReversions` *Detect reversion mutations*

Description

getReversions() detects reversion mutations for a given pathogenic mutation from a BAM file of DNA sequencing data.

Usage

```
getReversions(
  bam.file,
  out.dir,
  reference,
  pathog.mut,
  gene.name = NULL,
  transcript.id = NULL,
  detection.window = 100,
  splice.region = 8,
  check.soft.clipping = TRUE,
  softClippedReads.realign.window = 1000,
  softClippedReads.realign.match = 1,
  softClippedReads.realign.mismatch = 4,
  softClippedReads.realign.gapOpening = 6,
  softClippedReads.realign.gapExtension = 0,
  check.wildtype.reads = FALSE,
  is.paired.end = TRUE,
  keep.duplicate.reads = TRUE,
  keep.secondary.alignment = TRUE,
  keep.supplementary.alignment = TRUE,
  minimum.mapping.quality = 0,
  verbose = TRUE,
  out.failed.reads = FALSE
)
```

Arguments

`bam.file`	A character file name of the BAM file containing aligned reads to be processed.
`out.dir`	A character file path to write output files.
`reference`	A character variable specifying the reference genome version (hg19, hg38, mm10) or a FASTA file containing the open reading frames of reference sequences.
`pathog.mut`	A character variable specifying the genomic position of pathogenic mutation following the HGVS-like syntax for substitution, deletion, insertion, deletion-insertion (delins), or duplication.
`gene.name`	A character gene name for the pathogenic mutation.
`transcript.id`	A character Ensembl Transcript ID for the pathogenic mutation.

`detection.window`
: A non-negative integer specifying the length of flanking regions to be added to both ends of pathogenic mutation locus for detecting reversion mutations. Default is 100.

`splice.region`
: A positive integer specifying the length of splicing junction region to be considered in introns. Default is 8.

`check.soft.clipping`
: A logical value indicating whether soft-clipped reads to be realigned. Default is TRUE.

`softClippedReads.realign.window`
: A non-negative integer specifying the length of flanking regions to be added to both ends of pathogenic mutation locus for realigning soft-clipped reads. Default is 1000.

`softClippedReads.realign.match`
: A non-negative integer specifying the scoring for a nucleotide match for realigning soft-clipped reads. Default is 1.

`softClippedReads.realign.mismatch`
: A non-negative integer specifying the scoring for a nucleotide mismatch for realigning soft-clipped reads. Default is 4.

`softClippedReads.realign.gapOpening`
: A non-negative integer specifying the cost for opening a gap in the realignment of soft-clipped reads. Default is 6.

`softClippedReads.realign.gapExtension`
: A non-negative integer specifying the incremental cost incurred along the length of the gap in the realignment of soft-clipped reads. Default is 0.

`check.wildtype.reads`
: A logical value indicating whether wild type reads to be processed as revertant-to-wildtype reads. Default is FALSE.

`is.paired.end`
: A logical value indicating whether reads in BAM file are paired-end (TRUE) or single-end (FALSE). Default is TRUE.

`keep.duplicate.reads`
: A logical value indicating whether duplicated reads in the BAM file to be processed (TRUE) or discarded (FALSE). Default is TRUE.

`keep.secondary.alignment`
: A logical value indicating whether secondary alignment reads in the BAM file to be processed (TRUE) or discarded (FALSE). Default is TRUE.

`keep.supplementary.alignment`
: A logical value indicating whether supplementary alignment reads in the BAM file to be processed (TRUE) or discarded (FALSE). Default is TRUE.

`minimum.mapping.quality`
: A non-negative integer specifying the minimum mapping quality of reads in the BAM file to be processed. Default is 0.

`verbose`
: A logical value indicating whether progress logging to be printed to stdout. Default is TRUE.

`out.failed.reads`
: A logical value indicating whether the name of failed reads to be written to the '.failed_reads.txt' file. Default is FALSE.

Value

Results written into output directory:

- ".reversions.txt" contains all reversions identified for the pathogenic mutation from the BAM file.
- ".split_mutations.txt" contains information of each single mutation in a reversion.
- ".revert_assembly.bam" contains all reads realigned to the pathogenic mutation.
- ".revert_assembly.bam.bai" is the index file for '.revert_assembly.bam'.
- ".revert_settings.txt" contains the summary of running parameters and processed reads.
- ".failed_reads.txt" (optional) contains the names of reads failed for reversion detection.

For more details of the output files see the help vignette

Examples

```
getReversions(
    bam.file = system.file("extdata", "toy_data_1.bam", package="revert"),
    out.dir = tempdir(),
    reference = "hg19",
    pathog.mut = "chr13:g.32913319_32913320delTG",
    gene.name = "BRCA2",
    transcript.id = "ENST00000544455")

# For more examples see the help vignette
```

Index