

Package ‘crisprBwa’

October 18, 2022

Version 1.0.0

Date 2022-03-08

Title BWA-based alignment of CRISPR gRNA spacer sequences

Depends methods

Imports BiocGenerics, BSgenome, crisprBase (>= 0.99.15), GenomeInfoDb, Rbwa, readr, stats, stringr, utils

Suggests BiocStyle, BSgenome.Hsapiens.UCSC.hg38, knitr, rmarkdown, testthat

biocViews CRISPR, FunctionalGenomics, Alignment

Description Provides a user-friendly interface to map on-targets and off-targets of CRISPR gRNA spacer sequences using bwa. The alignment is fast, and can be performed using either commonly-used or custom CRISPR nucleases. The alignment can work with any reference or custom genomes. Currently not supported on Windows machines.

License MIT + file LICENSE

Encoding UTF-8

RoxygenNote 7.1.2

VignetteBuilder knitr

BugReports <https://github.com/Jfortin1/crisprBwa/issues>

URL <https://github.com/Jfortin1/crisprBwa>

OS_type unix

git_url <https://git.bioconductor.org/packages/crisprBwa>

git_branch RELEASE_3_15

git_last_commit ba7a773

git_last_commit_date 2022-04-27

Date/Publication 2022-10-18

Author Jean-Philippe Fortin [aut, cre]

Maintainer Jean-Philippe Fortin <fortin946@gmail.com>

R topics documented:

runBwa	2
runCrisprBwa	3

Index	5
--------------	----------

runBwa	<i>Run BWA short-read aligner</i>
--------	-----------------------------------

Description

Return BWA alignments for a list of short sequences for a prebuilt BWA index.

Usage

```
runBwa(sequences, bwa_index = NULL, n_mismatches = 3)
```

Arguments

sequences	Character vector of DNA sequences.
bwa_index	String specifying path to the BWA index.
n_mismatches	Integer specifying maximum number of mismatches allowed between the query sequences and the index sequences.

Details

runBwa can be used to map short DNA sequences to a reference genome. To search for sequences while imposing constraints on PAM sequences (such as gRNA spacer sequences), see runCrisprBwa instead.

Value

A data.frame of the alignments with the following columns:

- query — string specifying query DNA sequence
- chr - string specifying chromosome name
- pos - string specifying genomic coordinate of the start of the target DNA sequence
- strand - string specifying strand ("+" or "-")
- n_mismatches - integer specifying number of mismatches between query and target sequences

Author(s)

Jean-Philippe Fortin

See Also

link{runCrisprBwa} to map gRNA spacer sequences.

Examples

```

fasta <- system.file(package="crisprBwa", "example/chr12.fa")
outdir <- tempdir()
index <- file.path(outdir, "chr12")
Rbwa::bwa_build_index(fasta,
                      index_prefix=index)

seqs <- c("GGAAGTTG",
          "GTGGACAC",
          "GTGTGCAA")

aln <- runBwa(seqs,
              n_mismatches=1,
              bwa_index=index)

```

runCrisprBwa

Find gRNA spacer alignments with bwa

Description

Return bwa alignments for a list of gRNA spacer sequences.

Usage

```

runCrisprBwa(
  spacers,
  bwa_index = NULL,
  bsgenome = NULL,
  crisprNuclease = NULL,
  canonical = TRUE,
  ignore_pam = FALSE,
  n_mismatches = 0,
  force_spacer_length = FALSE,
  verbose = TRUE
)

```

Arguments

spacers	Character vector of DNA sequences corresponding to gRNA spacer sequences. Must all be of equal length.
bwa_index	Path to the bwa index to be used for alignment.
bsgenome	Bsgenome object.
crisprNuclease	CrisprNuclease object.
canonical	Should only canonical PAM sequences be considered? TRUE by default.
ignore_pam	If TRUE, will return all matches regardless of PAM sequence. FALSE by default.

<code>n_mismatches</code>	Integer specifying maximum number of mismatches allowed between spacer and protospacer sequences.
<code>force_spacer_length</code>	Should the spacer length be overwritten in the <code>crisprNuclease</code> object? FALSE by default.
<code>verbose</code>	Should messages be printed to the console? TRUE by default.

Details

`runCrisprBwa` is similar to `runBwa`, with the addition of imposing constraints on PAM sequences such that the query sequences are valid protospacer sequences in the searched genome.

Value

runBwa returns spacer alignment data, including genomic coordinates and sequence.

Author(s)

Jean-Philippe Fortin

See Also

`link{runBwa}` to map general DNA sequences.

Examples

```
# Building BWA index first:
fasta <- system.file(package="crisprBwa", "example/chr12.fa")
outdir <- tempdir()
index <- file.path(outdir, "chr12")
Rbwa::bwa_build_index(fasta,
                      index_prefix=index)

# Aligning Cas9 gRNA
library(BSgenome.Hsapiens.UCSC.hg38)
seqs <- c("AGCTGTCCGTGGGGTCCGC",
          "CCCCTGCTGCTGTGCCAGGC")
data(SpCas9, package="crisprBase")
bsgenome <- BSgenome.Hsapiens.UCSC.hg38
results <- runCrisprBwa(seqs,
                       bsgenome=bsgenome,
                       bwa_index=index,
                       n_mismatches=2,
                       crisprNuclease=SpCas9)
```

Index

runBwa, [2](#)
runCrisprBwa, [3](#)