

Package ‘Rhisat2’

October 16, 2019

Type Package

Title R Wrapper for HISAT2 Aligner

Version 1.0.3

SystemRequirements GNU make

Description An R interface to the HISAT2 spliced short-read aligner by Kim et al. (2015). The package contains wrapper functions to create a genome index and to perform the read alignment to the generated index.

License GPL-3

Encoding UTF-8

RoxygenNote 6.1.1

Depends R (>= 3.6)

Suggests testthat, knitr, rmarkdown, BiocStyle

VignetteBuilder knitr

biocViews Alignment, Sequencing, SplicedAlignment

Imports GenomicFeatures, SGSeq, GenomicRanges, methods, utils

URL <https://github.com/fmicompbio/Rhisat2>

BugReports <https://github.com/fmicompbio/Rhisat2/issues>

Archs x64

git_url <https://git.bioconductor.org/packages/Rhisat2>

git_branch RELEASE_3_9

git_last_commit 1fb99f9

git_last_commit_date 2019-06-11

Date/Publication 2019-10-15

Author Charlotte Soneson [aut, cre] (<<https://orcid.org/0000-0003-3833-2169>>)

Maintainer Charlotte Soneson <charlottesoneson@gmail.com>

R topics documented:

| | |
|--------------------------------|---|
| Rhisat2-package | 2 |
| extract_splice_sites | 2 |
| hisat2 | 3 |
| hisat2_build | 4 |

| | |
|------------------------------|---|
| hisat2_build_usage | 6 |
| hisat2_usage | 6 |
| hisat2_version | 7 |

| | |
|--------------|----------|
| Index | 8 |
|--------------|----------|

| | |
|-----------------|-------------------------|
| Rhisat2-package | <i>R hisat2 wrapper</i> |
|-----------------|-------------------------|

Description

This package provides an R wrapper around the HISAT2 short read aligner. The package is used by the QuasR Bioconductor package. We recommend to use QuasR instead of calling Rhisat2 directly, since it covers the whole analysis workflow of typical high-throughput sequencing experiments.

Author(s)

Charlotte Soneson, based partly on code from Florian Hahne

References

Kim D, Langmead B and Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nature Methods* 12:357-360 (2015).

See Also

hisat2, hisat2_build, hisat2_build_usage, hisat2_usage

| | |
|----------------------|---|
| extract_splice_sites | <i>Extract splice sites from annotation</i> |
|----------------------|---|

Description

This function extracts splice sites from an annotation object (a gtf/gff3 file, a GRanges object or a TxDb object) and saves them in a text file formatted such that it can be directly used with HISAT2, by providing it as the argument `known-splicesite-infile`.

Usage

```
extract_splice_sites(features, outfile, min_length = 5)
```

Arguments

| | |
|------------|--|
| features | Either the path to a gtf/gff3 file containing the genomic features, a GRanges object or a TxDb object. |
| outfile | Character scalar. The path to a text file where the extracted splice sites will be written. |
| min_length | Integer scalar. Junctions corresponding to introns below this size will not be reported. The default setting in HISAT2 is 5. |

Value

Nothing is returned, but the splice junction coordinates are written to outfile.

Author(s)

Charlotte Soneson

References

Kim D, Langmead B and Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nature Methods 12:357-360 (2015).

Examples

```
tmp <- tempfile()
extract_splice_sites(features=system.file("extdata/refs/genes.gtf",
                                         package="Rhisat2"),
                    outfile=tmp, min_length=5)
```

hisat2

Align reads with HISAT2

Description

The function can be used to call the hisat2 binary.

Usage

```
hisat2(sequences, index, ..., type = c("single", "paired"), outfile,
      force = FALSE, strict = TRUE, execute = TRUE)
```

Arguments

| | |
|-----------|---|
| sequences | If type is single, a character vector of file names if the additional argument c is FALSE, otherwise a vector of read sequences. If type is paired, a length-2 list of file names or sequences, where the first list item corresponds to the first mate pair sequences, and the second list item to the second mate pair sequences. |
| index | Character scalar. The path+prefix of the HISAT2 index to align against (in the form <path/to/index>/<prefix>). |
| ... | Additional arguments passed to the binaries. |
| type | Character scalar, either "single" or "paired". If single, the input sequences are interpreted as single-end reads. If paired, they are supposed to be paired reads. |
| outfile | (optional) Character scalar. The path to the output file. If missing, the alignments will be returned as an R character vector. |
| force | Logical scalar. Whether to force overwriting of outdir. |
| strict | Logical scalar. Whether strict checking of input arguments should be enforced. |
| execute | Logical scalar. Whether to execute the assembled shell command. If FALSE, return a string with the command. |

Details

All additional arguments in ... are interpreted as additional arguments to the HISAT2 binaries. Any flags are supposed to be represented as logical values (e.g., `quiet=TRUE` will be translated into `--quiet`). Parameters with additional input are supposed to be character or numeric vectors, and the individual elements are collapsed into a single comma-separated string (e.g., `k=2` is translated into `-k 2`, `bmax=100` into `--bmax 100`). Some arguments to the HISAT2 binaries will be ignored if they are already handled as explicit function arguments. See the output of `hisat2_usage()` for details about available parameters.

Value

If `execute` is `TRUE`, the output generated by calling the `hisat2` binary. If `execute` is `FALSE`, the `hisat2` command.

Author(s)

Charlotte Sonesson, based on code from Florian Hahne.

References

Kim D, Langmead B and Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nature Methods* 12:357-360 (2015).

Examples

```
tmp <- tempdir()
refs <- list.files(system.file("extdata/refs", package="Rhisat2"),
                  full.names=TRUE, pattern="\\.fa$")
hisat2_build(references=refs, outdir=file.path(tmp, "index"),
             force=TRUE, prefix="index")
reads <- list.files(system.file("extdata/reads", package="Rhisat2"),
                   full.names=TRUE, pattern="\\.fastq$")
hisat2(sequences=as.list(reads), index=file.path(tmp, "index/index"),
      type="paired", outfile=file.path(tmp, "out.sam"), force=TRUE)
```

hisat2_build

Generate HISAT2 index

Description

This function can be used to call the `hisat2-build` binary.

Usage

```
hisat2_build(references, outdir, ..., prefix = "index", force = FALSE,
             strict = TRUE, execute = TRUE)
```

Arguments

| | |
|------------|--|
| references | Character vector. The path to the files containing the reference sequences from which to build the HISAT2 index. |
| outdir | Character scalar. The path to the output directory in which to store the HISAT2 index. If the directory already exists, the function will throw an error, unless <code>force=TRUE</code> . |
| ... | Additional arguments passed to the binaries. |
| prefix | Character scalar. The prefix to use for the HISAT2 index files. |
| force | Logical scalar. Whether to force overwriting of <code>outdir</code> . |
| strict | Logical scalar. Whether strict checking of input arguments should be enforced. |
| execute | Logical scalar. Whether to execute the assembled shell command. If <code>FALSE</code> , return a string with the command. |

Details

All additional arguments in `...` are interpreted as additional arguments to the HISAT2 binaries. Any flags are supposed to be represented as logical values (e.g., `quiet=TRUE` will be translated into `--quiet`). Parameters with additional input are supposed to be character or numeric vectors, and the individual elements are collapsed into a single comma-separated string (e.g., `k=2` is translated into `-k 2`, `bmax=100` into `--bmax 100`). Some arguments to the HISAT2 binaries will be ignored if they are already handled as explicit function arguments. See the output of `hisat2_build_usage()` for details about available parameters.

Value

If `execute` is `TRUE`, the output generated by calling the `hisat2-build` binary. If `execute` is `FALSE`, the `hisat2-build` command.

Author(s)

Charlotte Sonesson, based on code from Florian Hahne.

References

Kim D, Langmead B and Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nature Methods* 12:357-360 (2015).

Examples

```
tmp <- tempdir()
refs <- list.files(system.file(package="Rhisat2", "extdata/refs"),
                  full.names=TRUE, pattern="\\.fa$")
x <- hisat2_build(references=refs, outdir=file.path(tmp, "index"),
                 force=TRUE)
head(x)
list.files(file.path(tmp, "index"))
```

hisat2_build_usage *Print usage of hisat2-build*

Description

Print usage of hisat2-build

Usage

```
hisat2_build_usage()
```

Value

No value is returned, the usage of hisat2_build is printed to the console.

Author(s)

Charlotte Soneson

Examples

```
hisat2_build_usage()
```

hisat2_usage *Print usage of hisat2*

Description

Print usage of hisat2

Usage

```
hisat2_usage()
```

Value

No value is returned, the usage of hisat2 is printed to the console.

Author(s)

Charlotte Soneson

Examples

```
hisat2_usage()
```

| | |
|----------------|-----------------------------|
| hisat2_version | <i>Print HISAT2 version</i> |
|----------------|-----------------------------|

Description

Print HISAT2 version

Usage

hisat2_version()

Value

No value is returned, the version information for hisat2 is printed to the console.

Author(s)

Charlotte Soneson

Examples

hisat2_version()

Index

[extract_splice_sites](#), [2](#)

[hisat2](#), [3](#)

[hisat2_build](#), [4](#)

[hisat2_build_usage](#), [6](#)

[hisat2_usage](#), [6](#)

[hisat2_version](#), [7](#)

[Rhisat2-package](#), [2](#)