

Package ‘sveval’

December 12, 2018

Title SV evaluation

Version 0.0.0.9000

Description Evaluate SV in a call set against a truth set using overlap-based approaches and sequence comparison for insertions.

Depends R (\geq 3.4.4)

License MIT + file LICENSE

Encoding UTF-8

LazyData true

RoxygenNote 6.1.0

Imports VariantAnnotation,
GenomicRanges,
magrittr,
dplyr,
rlang,
DelayedArray,
Biostrings,
parallel

R topics documented:

filterSVs	1
readSVvcf	2
svevalOI	3

Index	4
--------------	-------------------

filterSVs	<i>Filter SVs for regions of interest</i>
-----------	---

Description

Filter SVs for regions of interest

Usage

```
filterSVs(sv.gr, regions.gr, ol.prop = 0.5)
```

Arguments

<code>sv.gr</code>	the input SVs (e.g. read from <code>readSVvcf</code>)
<code>regions.gr</code>	the regions of interest
<code>ol.prop</code>	minimum proportion of <code>sv.gr</code> that must overlap <code>regions.gr</code> . Default is 0.5

Value

a subset of `sv.gr` that overlaps `regions.gr`

Author(s)

Jean Monlong

<code>readSVvcf</code>	<i>Read SVs from a VCF file</i>
------------------------	---------------------------------

Description

Read SVs from a VCF file

Usage

```
readSVvcf(vcf.file, keep.ins.seq = FALSE, sample.name = NULL)
```

Arguments

<code>vcf.file</code>	the path to the VCF file
<code>keep.ins.seq</code>	should it keep the inserted sequence? Default is FALSE.
<code>sample.name</code>	the name of the sample to use. If NULL (default), use first sample.

Value

a GRanges object with relevant information.

Author(s)

Jean Monlong

sveval0l*SV evaluation based on overlap and variant size*

Description

SV evaluation based on overlap and variant size

Usage

```
sveval0l(calls.gr, truth.gr, max.ins.dist = 20, min.cov = 0.5,  
min.del.rol = 0.1, ins.seq.comp = FALSE, nb.cores = 1,  
min.size = 0, max.size = Inf, bed.regions = NULL,  
sample.name = NULL, outfile = NULL, out.bed.prefix = NULL)
```

Arguments

<code>calls.gr</code>	call set. A GRanges or the path to a VCF file.
<code>truth.gr</code>	truth set. A GRanges or the path to a VCF file.
<code>max.ins.dist</code>	maximum distance for insertions to be clustered. Default is 20.
<code>min.cov</code>	the minimum coverage to be considered a match. Default is 0.5
<code>min.del.rol</code>	minimum reciprocal overlap for deletions. Default is 0.1
<code>ins.seq.comp</code>	compare sequence instead of insertion sizes. Default is FALSE.
<code>nb.cores</code>	number of processors to use. Default is 1.
<code>min.size</code>	the minimum SV size to be considered. Default 0.
<code>max.size</code>	the maximum SV size to be considered. Default is Inf.
<code>bed.regions</code>	If non-NULL, a GRanges object or path to a BED file (no headers) with regions of interest.
<code>sample.name</code>	the name of the sample to use if VCF files given as input. If NULL (default), use first sample.
<code>outfile</code>	the TSV file to output the results. If NULL (default), returns a data.frame.
<code>out.bed.prefix</code>	prefix for the output BED files. If NULL (default), no BED output.

Value

a data.frame with TP, FP and FN for each SV type.

Author(s)

Jean Monlong

Index

`filterSVs`, [1](#)

`readSVvcf`, [2](#)

`sveval01`, [3](#)