

Package ‘effectR’

September 30, 2018

Title Predicts Oomycete Effectors

Version 1.0.2

Description Predicts cytoplasmic effector proteins using genomic data by searching for motifs of interest using regular expression searches and hidden Markov models (HMM) based in Haas et al. (2009) <doi:10.1038/nature08358>.

Depends R (>= 3.5.0)

Imports seqinr (>= 3.3.6), ggplot2 (>= 2.2.1), shiny, reshape2, utils, rmarkdown, viridis

License GPL-2 | GPL-3

Encoding UTF-8

LazyData true

RoxygenNote 6.1.0

Suggests knitr, testthat

VignetteBuilder knitr

NeedsCompilation no

Author Javier Tabima [aut, cre],
Niklaus J. Grunwald [ths]

Maintainer Javier Tabima <tabimaj@oregonstate.edu>

Repository CRAN

Date/Publication 2018-09-30 21:30:03 UTC

R topics documented:

candidate.rxlr	2
effector.summary	2
effectR	3
hmm.logo	3
hmm.search	4
regex.search	5
shiny.effectR	6

Index	8
--------------	----------

candidate.rxlr	<i>An example of 'effector.summary' output</i>
----------------	--

Description

'candidate.rxlr' is an example result of the [effector.summary](#) function for new users that have not installed the additional software used by the effectR package.

Author(s)

Javier F. Tabima <tabimaj@science.oregonstate.edu>

effector.summary	<i>Returns non-redundant sequences from hmm.search or regex.search and generates a motif table</i>
------------------	--

Description

This function summarize the results from [regex.search](#) or [hmm.search](#).

Usage

```
effector.summary(hmm.result, motif = "RxLR", reg.pat = NULL)
```

Arguments

hmm.result	A list of SeqFastadna objects obtained from regex.search or hmm.search
motif	A character string indicating the motif of interest. Motifs for two cytoplasmic effectors are added to the function: RxLR or CRN effectors. Each of these motifs are associated with a by-default REGEX (reg.pat). A third option, custom, allows for the search of custom motifs. The custom option requires the specification of the motif REGEX pattern in the reg.pat option, in a regex format. Default is RxLR
reg.pat	A character string indicating the REGEX pattern for the custom motif. The specification of the REGEX pattern in must be in regex format. Required for custom option of motif

Value

A list of two objects: Summary motif table and non-redundant sequences (only with results of [hmm.search](#))

Examples

```
## Not run:

fasta.file <- system.file("extdata", "test_infestans.fasta", package = "effectR")
ORF <- seqinr::read.fasta(fasta.file)
REGEX <- regex.search(ORF, motif='RxLR')
candidate.rxlr <- hmm.search(original.seq = fasta.file, regex.seq=REGEX, num.threads = 2)
effector.summary(candidate.rxlr)
# Custom motifs
reg.pat <- "^\\w{50,60}[w,v]"
REGEX <- regex.search(sequence = ORF, motif = "custom", reg.pat = reg.pat)
candidate.custom <- hmm.search(original.seq = fasta.file, regex.seq = REGEX)
effector.summary(candidate.custom, motif = "custom", reg.pat = reg.pat)

## End(Not run)
```

effectR

effectR: An R Package to Call Oomycete Effectors

Description

An R Package to Call Oomycete Effectors.

Author(s)

Javier F. Tabima <tabimaj@oregonstate.edu>

hmm.logo

Plots the relative frequencies of each position for hmmsearch table.

Description

This function plots the results from [hmm.search](#) as a barplot with amino acids in the x axis and the relative frequency of each amino acid in the y axis

Usage

```
hmm.logo(hmm.table)
```

Arguments

`hmm.table` The HMM profile table resulting from [hmm.search](#)

Examples

```
## Not run:

fasta.file <- system.file("extdata", "test_infestans.fasta", package = "effectR")
ORF <- seqinr::read.fasta(fasta.file)
REGEX <- regex.search(ORF, motif='RxLR')
candidate.rxlr <- hmm.search(original.seq = fasta.file, regex.seq=REGEX, num.threads = 2)
hmm.logo(candidate.rxlr$HMM_Table)

## End(Not run)
```

hmm.search

Searching for motifs using HMM searches

Description

This function uses MAFFT and HMMER to search for sequences with RxLR or CRN motifs using hidden markov models.

Usage

```
hmm.search(original.seq, regex.seq, alignment.file = NULL,
  save.alignment = FALSE, mafft.path = NULL, num.threads = 2,
  hmm.path = NULL, seed = 12345)
```

Arguments

original.seq	The absolute path for the original six-frame translation FASTA file
regex.seq	A list of SeqFastadna objects resulting from <code>regex.search</code> . The HMM profile will be constructed using these sequences
alignment.file	(Optional) The absolute path for an alignment file of the sequences to build the hmmer profile from. It's recommended that the alignment file contains the same sequences than the 'regex.seq' files. If the user provides the absolute path, *effectR* won't use MAFFT to align the sequences and will use the alignment for the HMMER searches. If no alignment file is provided, *effectR* will use MAFFT to align the sequences from 'regex.seq' and run HMMER.
save.alignment	(Optional) Save the alignment in the returning object. The MAFFT alignment will be saved as the first element of the returned object.
mafft.path	(Optional) Local path of folder containing the MAFFT binary executable file or the executable file itself. If not specified, then MAFFT must be in the program search path.
num.threads	(Optional) Number of threads to be used by MAFFT
hmm.path	(Optional) Local path of folder containing the HMMER binaries. If not specified, then HMMER executables must be in the program search path.
seed	(Optional) The seed to used with HMMER commands. Set this to get the same output each time

Details

hmm.search uses the results from [regex.search](#) to search for motifs of interest using Hidden Markov Models after aligning the sequences with MAFFT. After the multiple sequence alignment is complete, the function constructs a HMM profile using the alignment data. The HMM profile is in the original list of SeqFastadna objects to obtain the best HMM results with sequences with RxLR or CRN motifs.

Value

A list of three elements: REGEX candidate effectors, HMM candidate effectors, and HMM results table.

Note

If MAFFT/HMMER are not the program search path, the user has to specify the path for the MAFFT and the HMMER executable binaries and specify them in the `mafft.path` and `hmm.path`

Examples

```
## Not run:

fasta.file <- system.file("extdata", "test_infestans.fasta", package = "effectR")
ORF <- seqinr::read.fasta(fasta.file)
REGEX <- regex.search(ORF, motif="RxLR")
candidate.rxlr <- hmm.search(original.seq = fasta.file, regex.seq = REGEX,
  alignment.file=NULL, save.alignment=T)

# To save the alignment:
library(seqinr)
write.fasta(sequences = getSequence(candidate.rxlr$Alignment),
  names = getName(candidate.rxlr$Alignment),
  file.out = "Alin_infestans.fasta", nbchar = 10000)

# To use an alignment file
ALIN <- system.file("extdata", "Alin_infestans.fasta", package = "effectR")
candidate.rxlr <- hmm.search(original.seq = fasta.file, regex.seq = REGEX,
  alignment.file = ALIN)

## End(Not run)
```

regex.search

Searching for motifs using regular expressions (REGEX)

Description

This function searches a list of SeqFastadna objects for sequences with RxLR or CRN motifs.

Usage

```
regex.search(sequence, motif = "RxLR", reg.pat = NULL)
```

Arguments

sequence	A list of SeqFastadna objects from seqinr read.fasta . The SeqFastadna object must be comprised by amino acid sequences, not DNA sequences
motif	A character string indicating the motif to be searched. Motifs for two cytoplasmic effectors are added to the function: RxLR or CRN effectors. Each of these motifs are associated with a by-default REGEX (reg.pat). These motifs are adapted from Haas et al. (2009). A third option, custom, allows for the search of custom motifs. The custom option requires the specification of the motif REGEX pattern in the reg.pat option, in a regex format. Default motif is RxLR
reg.pat	A character string indicating the REGEX pattern for the custom motif. The specification of the REGEX pattern in must be in regex format. Required for custom option of motif

References

Haas, B.J., Kamoun, S., Zody, M.C., Jiang, R.H., Handsaker, R.E., Cano, L.M., Grabherr, M., Kodira, C.D., Raffaele, S., Torto-Alalibo, T. and Bozkurt, T.O., 2009. Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. *Nature*, 461(7262), p.393.

Examples

```
fasta.file <- system.file("extdata", "test_infestans.fasta", package = "effectR")
ORF <- seqinr::read.fasta(fasta.file)
rxlr.cand <- regex.search(ORF)
custom.cand <- regex.search(ORF, motif = "custom", reg.pat = "^\\w{12,60}r\\w1r\\w{6,10}eer")
```

shiny.effectR

Function to run the effectR package using a shiny app

Description

This function will launch an interactive web interface to run the effectR package functions to obtain effectors. It runs using the **shiny** R package.

Usage

```
shiny.effectR(mafft.path = NULL, hmm.path = NULL)
```

Arguments

<code>mafft.path</code>	Local path of folder containing the MAFFT binary executable file or the executable file itself. If not specified, then MAFFT must be in the program search path.
<code>hmm.path</code>	Local path of folder containing the HMMER binaries. If not specified, then HMMER executables must be in the program search path.

Details

To successfully run this function the user will need to set the

Index

- *Topic **data**
 - [candidate.rxlr](#), [2](#)
- *Topic **effector**
 - [effector.summary](#), [2](#)
 - [hmm.logo](#), [3](#)
 - [hmm.search](#), [4](#)
 - [regex.search](#), [5](#)
- *Topic **package**
 - [effectR](#), [3](#)
- *Topic **plot**
 - [hmm.logo](#), [3](#)
- *Topic **regex**
 - [effector.summary](#), [2](#)
 - [hmm.logo](#), [3](#)
 - [hmm.search](#), [4](#)
 - [regex.search](#), [5](#)

[candidate.rxlr](#), [2](#)

[effector.summary](#), [2](#), [2](#)
[effectR](#), [3](#)
[effectR-package \(effectR\)](#), [3](#)

[hmm.logo](#), [3](#)
[hmm.search](#), [2](#), [3](#), [4](#)

[read.fasta](#), [6](#)
[regex](#), [2](#), [6](#)
[regex.search](#), [2](#), [4](#), [5](#), [5](#)

[shiny.effectR](#), [6](#)