Package 'ColeOR'

August 18, 2022

Title A comprehensive R package for finding and validating olfactory receptor gene models in beetle genomes

Version 0.0.0.9000

Description ColeOR includes a suite of functions for finding, testing, and validating olfactory receptor genes in beetle genomes.

License `use_mit_license()`, `use_gpl3_license()` or friends to pick a license

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.2.1

R topics documented:

Function.ReadORFasta   *Function.ReadORFasta: function to read and format a fasta file containing multiple OR amino acid sequences*

Description

This function returns a data frame with the first column representing the OR sequence names, and the second column containing a string for each OR amino acid sequence

Usage

```
Function.ReadORFasta(string.PathFastaFile)
```

Arguments

string.PathFastaFile

> String containing system path to fasta file. Make sure: no slash characters in the sequence names

Value

dataframe.ORSequences Returns a data.frame with four columns: Species (genome), Sequence names (fasta names), sequence scaffolds (taken from fasta file), and sequence strings (amino acid strings for each OR gene)

Examples

```
################
# load depends #
################
library(seqinr)

###########################
# read input fasta file #
###########################
handle.InputFastaFile <- Function.ReadORFasta(string.PathFastaFile = '~/Desktop/Beetle ORs Protein Translatio
```

Function.ReadResults_DeepTMHMM

Function.ReadResults_DeepTMHMM: function to read the results from DeepTMTHMM

Description

This function returns a vector with results from DeepTMHMM

Usage

```
Function.ReadResults_DeepTMHMM(
  string.Path_RESULTS,
  string.SequenceName,
  string.SpeciesName,
  string.SequenceScaffold
)
```

Arguments

string.Path_RESULTS

                  String containing path to DeepTMHMM results

string.SequenceName

                  String containing the name of the sequence (used to name the directory as well)

string.SpeciesName

                  String containing the name of species (genome) taken from the original fasta file

string.SequenceScaffold

                  String containing the name of the scaffold with putative BLAST hits

Value

vector.RESULTS Vector containing results that summerize the TMD inferences

Examples

```
###############
# load depends #
###############
library(seqinr)


##########################
# read input fasta file #
##########################
handle.InputFastaFile <- Function.ReadORFasta(string.PathFastaFile = '~/Desktop/Beetle ORs Protein Translatio


###################################
## run Deep TMHMM on single gene #
###################################
#Function.RunDeepTMHMM(string.SequenceName = handle.InputFastaFile$SequenceName[1],
#                       string.SequenceData = handle.InputFastaFile$SequenceData[1],
#                       string.SpeciesName = handle.InputFastaFile$SequenceSpecies[1],
#                 string.SequenceScaffold = handle.InputFastaFile$SequenceScaffolds[1],
#                       string.PathDir = '~/Desktop/')


################
# read results #
################
Function.ReadResults_DeepTMHMM(string.SequenceName = handle.InputFastaFile$SequenceName[1],
                         string.SpeciesName = handle.InputFastaFile$SequenceSpecies[1],
                   string.SequenceScaffold = handle.InputFastaFile$SequenceScaffolds[1],
                   string.Path_RESULTS = '~/Desktop/Conduct.DeepTMHMM_AglaOr1_Orco_2022-08-02/biolib_resul
```

Function.RunDeepTMHMM   *Function.RunDeepTMHMM: function to run DeepTMHMM on a single amino acid sequence*

Description

This function returns a subdirectory containing the results from DeepTMHMM

Usage

```
Function.RunDeepTMHMM(
  string.SequenceName,
  string.SpeciesName,
  string.SequenceScaffold,
  string.SequenceData,
  string.PathDir
)
```

Arguments

string.SequenceName

                String containing the name of the sequence (used to name the directory as well)

string.SpeciesName

                String containing the name of species (genome) taken from the original fasta file

string.SequenceScaffold
                    String containing the name of the scaffold with putative BLAST hits
string.SequenceData
                    String containing the sequence data for a single sequence to run DeepTMHMM
string.PathDir   String containing the directory path to where the analyses will be conducted

Value

Directory External directory with results from DeepTMHMM

Examples

```
################
# load depends #
################
library(seqinr)

##########################
# read input fasta file #
##########################
handle.InputFastaFile <- Function.ReadORFasta(string.PathFastaFile = '~/Desktop/ChemoGenes_ProteinSet1.fasta

###################################
# run Deep TMHMM on single gene #
###################################
Function.RunDeepTMHMM(string.SequenceName = handle.InputFastaFile$SequenceName[1],
                      string.SequenceData = handle.InputFastaFile$SequenceData[2],
                      string.PathDir = '~/Desktop/')
```

Function.RunDeepTMHMM_Pipeline

Function.RunDeepTMHMM_Pipeline: function to run DeepTMHMM on each sequence of a AA fast file

---

Description

This function returns a data.frame containing results of DeepTMHMM running on each sequence in fasta file

Usage

```
Function.RunDeepTMHMM_Pipeline(string.PathDir, string.FastaFile)
```

Arguments

string.PathDir   String containing the directory path to where the analyses will be conducted
string.FastaFile
                    String containing the path to the input AA fasta file

Value

handle.DF_Results Data.frame containing results from fasta file

Examples

```
################
# load depends #
################
library(seqinr)


###############################
# run on complete fasta file #
###############################
handle.RESULTS <- Function.RunDeepTMHMM_Pipeline(string.PathDir = '~/Desktop/', string.FastaFile = '~/Desktop

ANOTHER EXAMPLE:
################
# Load depends #
################
library(ColeOR)
library(seqinr)


#######################################
# Read example chromosome alignment #
#######################################
String.Path_ExampleModels <- system.file("extdata", "ExampleModelsORs.fasta", package="ColeOR")


#########################
# run TMD predictions #
#########################
handle.RESULTS <- Function.RunDeepTMHMM_Pipeline(string.PathDir = '~/Desktop/',
                                    string.FastaFile = String.Path_ExampleModels)
```

Function.WriteORFasta   *Function.WriteORFasta: function to create directory and write file for amino acid sequence for OR gene*

Description

This function returns a directory with fasta file containing amino acid sequence

Usage

```
Function.WriteORFasta(
  string.SpeciesName,
  string.SequenceName,
  string.SequenceScaffold,
  string.SequenceData,
  string.PathDir
)
```

Arguments

string.SpeciesName
           String containing the name of species (genome) taken from the original fasta file

string.SequenceName
           String containing the name of the OR sequence

string.SequenceScaffold
                    String containing the name of the scaffold with putative BLAST hits
string.SequenceData
                    String containing the sequence data (amino acid sequence from Function.ReadORFasta)
string.PathDir   String containing the directory path to where the analyses will be conducted

Value

File Function creates a directory and writes OR amino acid sequence to file in directory

Examples

```
################
# load depends #
################
library(seqinr)

##########################
# read input fasta file #
##########################
handle.InputFastaFile <- Function.ReadORFasta(string.PathFastaFile = '~/Desktop/Beetle ORs Protein Translatio

##################################
# run Deep TMHMM on single gene #
##################################
Function.WriteORFasta(string.SequenceName = handle.InputFastaFile$SequenceName[1],
                      string.SpeciesName = handle.InputFastaFile$SequenceSpecies[1],
                  string.SequenceScaffold = handle.InputFastaFile$SequenceScaffolds[1],
                      string.SequenceData = handle.InputFastaFile$SequenceData[1],
                      string.PathDir = '~/Desktop/')
```