Package 'PhyloWGA'

May 7, 2020

Title PhyloWGA: a user-friendly framework for chromosome-aware phylogenetic interrogation of whole ge-nome alignments

Version 0.0.0.9000

Description PhyloWGA Contains a suite of functions of phylogenetic analyses and interrogation of WGAs

License What license it uses

Encoding UTF-8

LazyData true

RoxygenNote 7.1.0

R topics documented:

`Chromo.Crawl` *Chromo.Crawl: function to crawl over an alignment of chromosomes while attempting to concatenate adjacent genomic windows*

Description

This function writes the results to an output file. Additionally, this function provides a directory where the resulting supergenes are placed.

Usage

```
Chromo.Crawl(
  string.PathParentDir,
  numeric.WindowSize,
  numeric.StepSize,
  string.PathToFastaFile,
  numeric.NumberOfCores
)
```

Arguments

string.PathParentDir	Path to parent dir for all analyses
numeric.WindowSize	Size of each chromosomal window
numeric.StepSize	Spacing of each chromosomal window
string.PathToFastaFile	Path to fasta file
numeric.NumberOfCores	Number of cores

Value

Results Results written to outputfiles

Examples

```
####################################
# Read example chromosome alignment #
####################################
String.Path_ExampleChromosomeAlignment <- system.file("extdata", "Example_Chr10.fasta", package="PhyloWGA")

########################
# Conduct.ChromoPhylome #
########################
Chromo.Crawl(string.PathParentDir = '~/Desktop/',
                        numeric.WindowSize = 1000,
                        numeric.StepSize = 1000,
                       string.PathToFastaFile = String.Path_ExampleChromosomeAlignment)
```

Chromo.Phylome — *Chromo.Phylome: function to infer a chromosome-specific set of phylogenetic trees*

Description

This function returns a list of the phylogenetic tree models, one for each window set by the user-defined window and step size input parameters

Usage

```
Chromo.Phylome(
  string.PathParentDir,
  string.PathToFastaFile,
  numeric.WindowSize,
  numeric.StepSize,
  string.Commands_iqtree
)
```

Arguments

`string.PathParentDir`	Path to parent dir for all analyses
`string.PathToFastaFile`	Path to fasta file
`numeric.WindowSize`	Size of each chromosomal window
`numeric.StepSize`	Spacing of each chromosomal window
`string.Commands_iqtree`	String used for nucleotide substitution model for IqTree. Leave as "" for model selection

Value

Results Results written to output files

Examples

```
####################################
# Read example chromosome alignment #
####################################
String.Path_ExampleChromosomeAlignment <- system.file("extdata", "Example_Chr10.fasta", package="PhyloWGA")


########################
# Conduct.ChromoPhylome #
########################
handle.RESULTS <- Chromo.Phylome(string.PathParentDir = '~/Desktop/',
                           string.PathToFastaFile = String.Path_ExampleChromosomeAlignment,
                                        numeric.WindowSize = 1000,
                                        numeric.StepSize = 1000,
                                        string.Commands_iqtree = "-m GTR+G")
```

`Chromo.Phylome.Custom` *Chromo.Phylome.Custom: function to infer a chromosome-specific set of phylogenetic trees*

Description

This function returns a list of the phylogenetic tree models, one for each window set by the user-defined window and step size input parameters

Usage

```
Chromo.Phylome.Custom(
  string.PathParentDir,
  string.PathToFastaFile,
  matrix.WindowCoordinates,
  string.Commands_iqtree
)
```

Arguments

`string.PathParentDir`
: Path to parent dir for all analyses

`string.PathToFastaFile`
: Path to fasta file

`string.Commands_iqtree`
: String used for nucleotide substitution model for IqTree. Leave as "" for model selection

`numeric.WindowSize`
: Size of each chromosomal window

`numeric.StepSize`
: Spacing of each chromosomal window

Value

Results Results written to output files

`Define.PhyloWGA_Experiment` *Define.PhyloWGA_Experiment: function to generate a matrix of window coordinates for a given alignment*

Description

This function returns a matrix of the window coordinates for an alignment

Usage

```
Define.PhyloWGA_Experiment(
  numeric.WindowSize,
  numeric.StepSize,
  numeric.TotalLength
)
```

Arguments

`numeric.WindowSize`
: Number of bp for each window, window length

`numeric.StepSize`
: Number of bp seperating each window

`numeric.TotalLength`
: Total number of bp for the entire alignment

Value

matrix.WindowCoordinates Matrix of k x 2 dimensions, each row indicating the start and end coordinates for the respective window

Examples

```
###################################
# Specify experimental parameters #
###################################
matrix.ExperimentalParams <- Define.PhyloWGA_Experiment(numeric.WindowSize = 1000, numeric.StepSize = 500, nu
```

`Execute.PhylogeneticCongruence_concatepillar`

Execute.PhylogeneticCongruence_concatepillar: function to execute concatepillar on a pair of alignments

---

Description

This function returns a list of results from concatepillar

Usage

```
Execute.PhylogeneticCongruence_concatepillar(
  matrix.WindowAlignment_01,
  matrix.WindowAlignment_02,
  string.PathParentDir,
  numeric.NumberOfCores
)
```

Arguments

`matrix.WindowAlignment_01`
: Matrix of alignment for the first window

`matrix.WindowAlignment_02`
: Matrix of alignment for the second window

`string.PathParentDir`
: Path to parent dir for all analyses

`numeric.NumberOfCores`
: Number of cores

Value

List List of results from concatepillar

Examples

```
################
# Load depends #
################
library(PhyloWGA)
library(ape)

####################################
# Read example chromosome alignment #
####################################
String.Path_ExampleChromosomeAlignment <-  system.file("extdata", "Example_Chr10.fasta", package="PhyloWGA")

########################
# Get alignment length #
########################
numeric.AlignmentLength <- Get.AlignmentLength(string.PathToFastaFile = String.Path_ExampleChromosomeAlignme

######################
# Get sequence names #
######################
vector.SequenceNames <- Get.SequenceNames(string.PathToFastaFile = String.Path_ExampleChromosomeAlignment)

#################################
# Define experimental parameters #
#################################
matrix.ExperimentalParams <- Define.PhyloWGA_Experiment(numeric.WindowSize = 1000, numeric.StepSize = 1000, n

############################
# Extract window alignment #
############################
matrix.WindowAlignment_01 <- Extract.WindowAlignment(string.PathToFastaFile = String.Path_ExampleChromosomeA
                                          vector.SequenceNames = vector.SequenceNames,
                                 vector.Coordinates_ExperimentalLocus = matrix.ExperimentalParams[5,])


#####################################################
# Prune taxa that are only missin data from matrix #
#####################################################
matrix.Pruned_WindowAlignment_01 <- Prune.MissingDataTaxa(matrix.WindowAlignment = matrix.WindowAlignment_01

############################
# Extract window alignment #
############################
matrix.WindowAlignment_02 <- Extract.WindowAlignment(string.PathToFastaFile = String.Path_ExampleChromosomeA
                                          vector.SequenceNames = vector.SequenceNames,
                                 vector.Coordinates_ExperimentalLocus = matrix.ExperimentalParams[6,])


#####################################################
# Prune taxa that are only missin data from matrix #
#####################################################
```

```
matrix.Pruned_WindowAlignment_02 <- Prune.MissingDataTaxa(matrix.WindowAlignment = matrix.WindowAlignment_02

##############################
# Estimate model with iqtree #
##############################
Execute.PhylogeneticCongruence_concatepillar(matrix.Pruned_WindowAlignment_01, matrix.Pruned_WindowAlignmen
```

`Execute.PhylogeneticInference_iqtree`

Execute.PhylogeneticInference_iqtree: function to estimate a phylogenetic tree model for a window with iqtree

---

Description

This function returns a list containing parameter estimates and a phylogenetic tree estimate

Usage

```
Execute.PhylogeneticInference_iqtree(
  matrix.WindowAlignment,
  string.PathParentDir,
  string.Commands_iqtree
)
```

Arguments

`matrix.WindowAlignment`
: Matrix of the alignment for a window

`string.PathParentDir`
: Path used for estimating trees with iqtree

Value

List List containing (1) list of parameter estimates and (2) phylogenetic tree estimate

Examples

```
################
# Load depends #
################
library(PhyloWGA)
library(ape)

####################################
# Read example chromosome alignment #
####################################
String.Path_ExampleChromosomeAlignment <- system.file("extdata", "Example_Chr10.fasta", package="PhyloWGA")

########################
```

```
# Get alignment length #
########################
numeric.AlignmentLength <- Get.AlignmentLength(string.PathToFastaFile = String.Path_ExampleChromosomeAlignme

######################
# Get sequence names #
######################
vector.SequenceNames <- Get.SequenceNames(string.PathToFastaFile = String.Path_ExampleChromosomeAlignment)

#################################
# Define experimental parameters #
#################################
matrix.ExperimentalParams <- Define.PhyloWGA_Experiment(numeric.WindowSize = 1000, numeric.StepSize = 1000, n

############################
# Extract window alignment #
############################
matrix.WindowAlignment <- Extract.WindowAlignment(string.PathToFastaFile = String.Path_ExampleChromosomeAlig
                                                  vector.SequenceNames = vector.SequenceNames,
                                    vector.Coordinates_ExperimentalLocus = matrix.ExperimentalParams[5,])


#####################################################
# Prune taxa that are only missin data from matrix #
#####################################################
matrix.Pruned_WindowAlignment <- Prune.MissingDataTaxa(matrix.WindowAlignment = matrix.WindowAlignment, nume

############################
# Estimate model with iqtree #
############################
Execute.PhylogeneticInference_iqtree(matrix.WindowAlignment = matrix.Pruned_WindowAlignment, string.PathPare
```

`Extract.WindowAlignment`

Extract.WindowAlignment: function to extract an alignment matrix given an input vector of sequences, coordinates, and a path to a fasta file

---

Description

This function returns matrix consisting of an extracted alignment

Usage

```
Extract.WindowAlignment(
  string.PathToFastaFile,
  vector.SequenceNames,
  vector.Coordinates_ExperimentalLocus
)
```

Arguments

`string.PathToFastaFile`
: String defining the path to the input fasta file

`vector.SequenceNames`
: Vector of sequence names to be extracted

`vector.Coordinates_ExperimentalLocus`
: Vector containing two coordinates for the window (start and end)

Value

matrix.Window_Alignment Matrix containing the alignment for the input window coordinates from the fasta file

Examples

```
################
# Load depends #
################
library(PhyloWGA)
library(ape)

####################################
# Read example chromosome alignment #
####################################
String.Path_ExampleChromosomeAlignment <-  system.file("extdata", "Example_Chr10.fasta", package="PhyloWGA")

########################
# Get alignment length #
########################
numeric.AlignmentLength <- Get.AlignmentLength(string.PathToFastaFile = String.Path_ExampleChromosomeAlignme

######################
# Get sequence names #
######################
vector.SequenceNames <- Get.SequenceNames(string.PathToFastaFile = String.Path_ExampleChromosomeAlignment)

#################################
# Define experimental parameters #
#################################
matrix.ExperimentalParams <- Define.PhyloWGA_Experiment(numeric.WindowSize = 1000, numeric.StepSize = 1000, n

############################
# Extract window alignment #
############################
matrix.WindowAlignment <- Extract.WindowAlignment(string.PathToFastaFile = String.Path_ExampleChromosomeAlig
                                                  vector.SequenceNames = vector.SequenceNames,
                                    vector.Coordinates_ExperimentalLocus = matrix.ExperimentalParams[5,])


write.dna(x = matrix.WindowAlignment, file = '~/Desktop/ExampleWindow.fasta', colsep = "", format = "fasta")
```

Get.AlignmentLength — *Get.AlignmentLength: get the total length of an alignment in the input fasta file*

Description

This function returns the number of bp found in the input fasta file

Usage

```
Get.AlignmentLength(string.PathToFastaFile)
```

Arguments

`string.PathToFastaFile`
String defining the path to the input fasta file

Value

numeric.Total_AlignmentLength Number of base pairs in the input fasta file

Examples

```
####################################
# Read example chromosome alignment #
####################################
String.Path_ExampleChromosomeAlignment <- system.file("extdata", "Example_Chr10.fasta", package="PhyloWGA")

########################
# Get alignment length #
########################
numeric.AlignmentLength <- Get.AlignmentLength(string.PathToFastaFile = String.Path_ExampleChromosomeAlignme
```

Get.SequenceNames — *Get.SequenceNames: get a vector of the sequence names provided in a fasta file*

Description

This function returns a vector of sequence names

Usage

```
Get.SequenceNames(string.PathToFastaFile)
```

Arguments

`string.PathToFastaFile`
String defining the path to the input fasta file

Value

vector.SequenceNames Vector containing the set of sequence names obtained from fasta file

Examples

```
####################################
# Read example chromosome alignment #
####################################
String.Path_ExampleChromosomeAlignment <- system.file("extdata", "Example_Chr10.fasta", package="PhyloWGA")

########################
# Get alignment length #
########################
vector.SequenceNames <- Get.SequenceNames(string.PathToFastaFile = String.Path_ExampleChromosomeAlignment)
```

Organize.ParallelPhyloWGA

Organize.ParallelPhyloWGA: function to partition and organize subdirectories for parallel WGA analyses

---

Description

This function returns a XXX

Usage

```
Organize.ParallelPhyloWGA(
  numeric.NumberSubsets,
  string.PathToFastaFile,
  string.PathParentDir,
  string.Analysis,
  numeric.WindowSize,
  string.Commands_iqtree,
  numeric.StepSize,
  numeric.NumberOfCores
)
```

Arguments

`numeric.NumberSubsets`
: Number of subsets for the original alignment and analyses

`string.PathToFastaFile`
: String defining the path to the input fasta file

`string.PathParentDir`
: Path to parent dir for all analyses

`numeric.WindowSize`
: Size of each chromosomal window

`numeric.StepSize`
: Spacing of each chromosomal window

numeric.NumberOfCores
: Number of cores

Value

XXX XXX

Prune.MissingDataTaxa *Prune.MissingDataTaxa: function to pruned matrix (high-missing data sequences removed)*

Description

This function returns a matrix that has been pruned of sequences with missing data above a given threshold

Usage

```
Prune.MissingDataTaxa(matrix.WindowAlignment, numeric.MissingDataThreshold)
```

Arguments

matrix.WindowAlignment
: Matrix of input chromosomal window alignment

numeric.MissingDataThreshold
: Numeric (proportion) of missing data permited of any one sequence

Value

matrix.Pruned_WindowAlignment Matrix of window alignment with sequences with

Examples

```
################
# Load depends #
################
library(PhyloWGA)
library(ape)

####################################
# Read example chromosome alignment #
####################################
String.Path_ExampleChromosomeAlignment <- system.file("extdata", "Example_Chr10.fasta", package="PhyloWGA")

########################
# Get alignment length #
########################
numeric.AlignmentLength <- Get.AlignmentLength(string.PathToFastaFile = String.Path_ExampleChromosomeAlignme

######################
# Get sequence names #
######################
vector.SequenceNames <- Get.SequenceNames(string.PathToFastaFile = String.Path_ExampleChromosomeAlignment)
```

```
#################################
# Define experimental parameters #
#################################
matrix.ExperimentalParams <- Define.PhyloWGA_Experiment(numeric.WindowSize = 1000, numeric.StepSize = 1000, n

############################
# Extract window alignment #
############################
matrix.WindowAlignment <- Extract.WindowAlignment(string.PathToFastaFile = String.Path_ExampleChromosomeAlig
                                                  vector.SequenceNames = vector.SequenceNames,
                                    vector.Coordinates_ExperimentalLocus = matrix.ExperimentalParams[5,])


#################################################
# Prune taxa that are only missin data from matrix #
#################################################
matrix.Pruned_WindowAlignment <- Prune.MissingDataTaxa(matrix.WindowAlignment = matrix.WindowAlignment, nume
```

`Read.OutputFile_iqtree`

Read.OutputFile_iqtree: function to extract a list of parameter estimates from an iqtree outputfile

---

Description

This function returns a list of the parameter estimates

Usage

```
Read.OutputFile_iqtree(string.Path_OutputFile_iqtree)
```

Arguments

`string.Path_OutputFile_iqtree`
: String defining path to the output iqtree file

Value

List List of containing estimates of the GTR+G model

Examples

```
################
# Load depends #
################
library(PhyloWGA)
library(ape)

####################################
# Read example chromosome alignment #
```

```
####################################
String.Path_ExampleChromosomeAlignment <-  system.file("extdata", "Example_Chr10.fasta", package="PhyloWGA")

########################
# Get alignment length #
########################
numeric.AlignmentLength <- Get.AlignmentLength(string.PathToFastaFile = String.Path_ExampleChromosomeAlignme

######################
# Get sequence names #
######################
vector.SequenceNames <- Get.SequenceNames(string.PathToFastaFile = String.Path_ExampleChromosomeAlignment)

#################################
# Define experimental parameters #
#################################
matrix.ExperimentalParams <- Define.PhyloWGA_Experiment(numeric.WindowSize = 1000, numeric.StepSize = 1000, n

############################
# Extract window alignment #
############################
matrix.WindowAlignment <- Extract.WindowAlignment(string.PathToFastaFile = String.Path_ExampleChromosomeAlig
                                                  vector.SequenceNames = vector.SequenceNames,
                                    vector.Coordinates_ExperimentalLocus = matrix.ExperimentalParams[5,])


#####################################################
# Prune taxa that are only missin data from matrix #
#####################################################
matrix.Pruned_WindowAlignment <- Prune.MissingDataTaxa(matrix.WindowAlignment = matrix.WindowAlignment, nume

##############################
# Estimate model with iqtree #
##############################
Execute.PhylogeneticInference_iqtree(matrix.WindowAlignment = matrix.Pruned_WindowAlignment, string.PathPare

####################
# Read output file #
####################
Read.OutputFile_iqtree(string.Path_OutputFile_iqtree = '~/Desktop/PhylogeneticInference_iqtree/WindowAlignm
```

Index