Package 'phylotools'

February 20, 2015

Type Package

Title Phylogenetic tools for Eco-phylogenetics

Version 0.1.2

Date 2011-8-18

Author Jinlong Zhang, Nancai Pei, Xiangcheng Mi

Maintainer Jinlong Zhang <jinlongzhang01@gmail.com>

Description Building supermatrix for DNA barcodes using different genes, calculating the inequality among lineages and phylogenetic similarity for very large dataset using slicing methods by invoking Phylocom.

Depends seqRFLP, ape, picante, spaa, fields

Suggests vegan

License GPL-2

LazyLoad yes

Repository CRAN

Date/Publication 2012-10-29 08:59:27

NeedsCompilation no

R topics documented:

`phylotools-package` *Phylogenetic tools for ecologists*

Description

This package currently consists of a few functions for handling DNA-barcoding sequences to build supermatrix for further analysis with RAxML etc. Much more functions for conducting phylogenetic analysis would be added in the future, especially for community phylogentic analysis.

Details

Package:	phylotools
Type:	Package
Version:	0.1.2
Date:	2010-8-18
License:	GLP-2
LazyLoad:	yes

Author(s)

Jinlong Zhang, Nancai Pei, Xiangcheng Mi

Maintainer: Jinlong Zhang <jinlongzhang01@gmail.com>

References

Kress W., Erickson D., Jones F., Swenson N., Perez R., Sanjur O., Bermingham E., Plant DNA barcodes and community phylogeny of a tropical forest dynamics plot in Panama. Proceedings of the National Academy of Sciences of the United States of America. 2009 18621-18626

Examples

```
##############################################
## Example Part I###Building Supermatrix########
##############################################

### Build super matrix
dir <- system.file("extdata", package = "phylotools")
setwd(dir)

## Supermatrix with "rbcla","matk","trnH-psbA"
supermat <- supermat(rbcl = "rbcla.phy", matk = "matK.phy",
          trn = c("trn1.phy", "trn2.phy","trn3.phy","trn4.phy"))
## Save to file
write.mat(supermat, "result.phy")

## Delete file
unlink("result.phy")



##############################################
## Example Part II###Create a image plot#######
##############################################

## Create an Image with legend from named numerical vectors
x <- rnorm(600)
labmat <- expand.grid(paste("X",1:30, sep = ""),
                  paste("Y", 1:20, sep = ""))

lab <- paste(labmat[,1], labmat[,2], sep = "")
imagevect(x, labels = lab, col = cm.colors(10))



##############################################
## Example Part III###Handling fasta project###
```

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

## Handling Fasta files
library(seqRFLP)
## loading data
data(fil.fas)
## Get the names of the sequences
col1 <- gnames.fas(fil.fas)
## Generating new names
col2 <- paste("seq", 1:length(col1), sep = "")
reftable.rename <- data.frame(col1, col2)
renamed <- rename.fasta(fil.fas, reftable.rename)

##Generate split factor levels.
index <- rep(NA, length(col2))
level1 <- seq(1, length(col2), by = 2)
index[level1] <- 1
index[-level1] <- 2

## Reference table
reftable.split <- data.frame(col2, index)

## split the fasta object
fasta.split(renamed, reftable.split)



################################################
## Example Part IV###Phylosor for very large data
################################################
## Make sure the PhyloCom can be invoked by command line
## res <- phylocom.phylosor(sample.file = "sample",
##                                 phylo = "phylo")



################################################
## Example Part V###Rescale the FDP data #######
## to different scales ##########################
################################################
## 20m
## plotscale(inputdata = BCI, len = 1000, wid = 500,
##           scale = 20)
## 50m
## plotscale(inputdata = BCI, len = 1000, wid = 500,
##           scale = 50)
## 100m
## plotscale(inputdata = BCI, len = 1000, wid = 500,
##          scale = 100)



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

```
## Example Part VI###
## Lineages inequality and
## Mean Gini Coefficient amonge
## lineages caused by imperfect sampling
################################################
data(bird.orders)
rtr1 <- del.tree.tip(bird.orders,3)
inequality(bird.orders, rtr1[[1]], h = 25.25103)


data(bird.orders)
to.drop <- c("Craciformes", "Galliformes", "Gruiformes")
droped <- drop.tip(bird.orders, to.drop)
meangini(tree = bird.orders, subtree = droped,
         times = 10, plot = TRUE)
meangini(tree = bird.orders, subtree = droped,
         times = 50, plot = TRUE)
meangini(tree = bird.orders, subtree = droped,
          times = 100, plot = TRUE)
meangini(tree = bird.orders, subtree = droped,
          times = 200, plot = TRUE)
```

`add.mat` *Add matrix to current matrix*

Description

Concatenates second columns of matrix.

Usage

```
add.mat(mat1, mat2)
```

Arguments

`mat1`	A dataframe with 2 columns.
`mat2`	A dataframe with 2 columns.

Details

This function could be used to concatenate the second column of the second matrix with the first matrix's second column according to each matrix's first column (as the names of the sequences). The "?" will be added automatically.

Value

A matrix with first column the names of the sequences, the second column the concatenated sequences.

Author(s)

Jinlong Zhang <jinlongzhang01@gmail.com>

References

Kress W., Erickson D., Jones F., Swenson N., Perez R., Sanjur O., Bermingham E., Plant DNA barcodes and community phylogeny of a tropical forest dynamics plot in Panama. Proceedings of the National Academy of Sciences of the United States of America. 2009 18621-18626

See Also

See Also supermat

Examples

```
## add.mat example ########
dir <- system.file("extdata", package = "phylotools")
setwd(dir)
rbcla <- read.phylip("rbcla.phy")
matk <- read.phylip("matK.phy" )
rbdat <- phy2dat(rbcla)
matdat <- phy2dat(matk)
add.mat(rbdat, matdat)
```

---

aln2dat *Convert ClustalX alignment data to dataframe*

Description

Convert ClustalX alignment data to dataframe

Usage

```
aln2dat(aln)
```

Arguments

aln	A character string indicating the the aln format file generated by Clustal

Details

None.

Value

Dataframe

Note

Kress W., Erickson D., Jones F., Swenson N., Perez R., Sanjur O., Bermingham E., Plant DNA barcodes and community phylogeny of a tropical forest dynamics plot in Panama. Proceedings of the National Academy of Sciences of the United States of America. 2009 18621-18626

Author(s)

Jinlong Zhang <jinlongzhang01@gmail.com>

References

None.

Examples

```
dir <- system.file("extdata", package = "phylotools")
setwd(dir)
##aln2dat examle ########
test <- readLines("aln")
aln2dat(test)
```

---

appendchar *Paste character according to the longest character.*

Description

Paste character for each elements at the second column which is shorter than the longest one.

Usage

```
appendchar(mat, pattern = "?")
```

Arguments

mat	data frame with two columns
pattern	Character to be added to the end of elements

Details

None.

Value

A data frame with specified character appended to the end of each shorter elements.

Note

None

Author(s)

Jinlong Zhang <jinlongzhang01@gmail.com>

References

Kress W., Erickson D., Jones F., Swenson N., Perez R., Sanjur O., Bermingham E., Plant DNA barcodes and community phylogeny of a tropical forest dynamics plot in Panama. Proceedings of the National Academy of Sciences of the United States of America. 2009 18621-18626

See Also

See Also framsub

Examples

```
##appendchar  example ########
aa <- c("1","2","3")
bb <- c("abc","abcd","abcdefg")
rrr <- data.frame(aa, bb)
rrr
appendchar(rrr)
```

`complement` *Get the complement sequences*

Description

This function could be used to convert the sequences to its complement sequence. The input data must be fasta format.

Usage

```
complement(fas)
```

Arguments

`fas`	Fasta format object

Details

This function could be used to obtain the complement sequence given a fasta file. For example, the complement sequence for `"5-TTGAACC-3"` is `"5-GGTTCAA-3"`.This may be used for converting the consensus sequence to its complement, when the user find that the sequence can not be aligned with other sequences.

Value

The complement sequences in fasta format.

Author(s)

Jinlong Zhang <jinlongzhang01@gmail.com>

References

None

See Also

revComp for more information.

Examples

```
cat(
">No305",
"NTTCGAAAAACACACCCACTACTAAAANTTATCAGTCACT",
file = "dna1.fas", sep = "\n")

sequence <- read.fasta("dna1.fas")
complement(sequence)
unlink("dna1.fas")
```

---

dat2phy *Convert dataframe to phylip format*

Description

Convert dataframe to phylip format

Usage

```
dat2phy(input, write = TRUE)
```

Arguments

input	dataframe with the first column the names of the sequence's and the second column the sequences
write	Write the file to disk.

Details

Convert dataframe to phylip format, whose first column are the names of the sequence's and the second column the sequences

Value

A vector of phylip format.

Note

None

Author(s)

Jinlong Zhang <jinlongzhang01@gmail.com>

References

Kress W., Erickson D., Jones F., Swenson N., Perez R., Sanjur O., Bermingham E., Plant DNA barcodes and community phylogeny of a tropical forest dynamics plot in Panama. Proceedings of the National Academy of Sciences of the United States of America. 2009 18621-18626

See Also

See Also phy2dat

Examples

```
##dat2phy example ########
##Convert dataframe to phylip object
dir <- system.file("extdata", package = "phylotools")
setwd(dir)
test <- read.phylip("matK.phy")
test2 <- phy2dat(test)
dat2phy(test2)
```

---

del.tree.tip *Randomly Delete tree tip labels*

Description

This function will randomly deleting tree tip labels, this is an internal function in phylotools

Usage

```
del.tree.tip(tree, n)
```

Arguments

tree	The phylogenetic tree
n	Number of tips to be deleted

Value

The phylogenetic tree in class 'phylo'

Author(s)

Jinlong Zhang < jinlongzhang01@gmail.com >

References

To be added

See Also

help

Examples

```
data(bird.orders)
del.tree.tip(bird.orders, 4)
```

---

dimension *Get the dimension of the matrix defined by "XnYm" labels*

Description

This is an internal function for handling "XnYm" labels, users do not need to call this function directly.

Usage

```
dimension(x, unique = FALSE, sort = FALSE)
```

Arguments

x	A vector contain the "XnYm" labels
unique	Logical, remove the duplicated value
sort	Logical, whether to sort the value in the result

Value

A list with the rows or columns

Author(s)

Jinlong Zhang < jinlongzhang01@gmail.com >

See Also

imagevect

Examples

```
dimension(c("X1Y4","X33Y56"))
```

---

edgesub *Substitute character for a sequence from both sides.*

Description

Substitute character for a sequence from both sides without changing the internal matched characters.

Usage

```
edgesub(x, pattern = "-", replacement = "?")
```

Arguments

x	The input DNA sequence.
pattern	The character that to be substitute.
replacement	The character that will replace the pattern.

Details

None.

Value

This function will return to the substituted pattern of the input DNA sequence.

Note

None

Author(s)

Jinlong Zhang <jinlongzhang01@gmail.com>

References

Kress W., Erickson D., Jones F., Swenson N., Perez R., Sanjur O., Bermingham E., Plant DNA barcodes and community phylogeny of a tropical forest dynamics plot in Panama. Proceedings of the National Academy of Sciences of the United States of America. 2009 18621-18626

See Also

See Also framsub, for substitution of a dataframe, which is a general application of this function.

Examples

```
## edgesub example ########
#### Substitute each terminal of the sequence
dna = "---ATTGCCTTAS--TTTAAAAAACCGTTC-----"
edgesub(dna)
```

`fasta.split` *Split the fasta object to fasta objects*

Description

Split the fasta object to fasta objects according to the groups given

Usage

```
fasta.split(fasta, ref, save2disk = FALSE)
```

Arguments

`fasta`	The fasta object
`ref`	Data frame with first column of the names of the sequences , the second column which group it belongs to.
`save2disk`	Whether to save the groups of fasta files to separate files

Details

This function may be used to split the trnH-psbA sequences, given the order (usually APGIII) it belongs to.

Value

Returns a list, indicate the groups for the sequences

Author(s)

Jinlong Zhang <jinlongzhang01@gmail.com>

References

Non.

See Also

rename.fasta

Examples

```
library(seqRFLP)
## loading data
data(fil.fas)
## Get the names of the sequences
col1 <- gnames.fas(fil.fas)
## Generating new names
col2 <- paste("seq", 1:length(col1), sep = "")
reftable.rename <- data.frame(col1, col2)
renamed <- rename.fasta(fil.fas, reftable.rename)

##Generate split factor levels.
index <- rep(NA, length(col2))
level1 <- seq(1, length(col2), by = 2)
index[level1] <- 1
index[-level1] <- 2

## Reference table
reftable.split <- data.frame(col2, index)

## split the fasta object
fasta.split(renamed, reftable.split)
```

---

fmatch *Find the matching index*

Description

Find the matching index

Usage

```
fmatch(dna, pattern = "")
```

Arguments

dna	A character string including the DNA sequence.
pattern	A character to be used in matching the DNA sequence.

Details

None.

Value

The max number the input character could be matched from left to right.

Note

None

Author(s)

Jinlong Zhang <jinlongzhang01@gmail.com>

References

Kress W., Erickson D., Jones F., Swenson N., Perez R., Sanjur O., Bermingham E., Plant DNA barcodes and community phylogeny of a tropical forest dynamics plot in Panama. Proceedings of the National Academy of Sciences of the United States of America. 2009 18621-18626

See Also

See Also edgesub for a general application of this function.

Examples

```
### fmatch  example ########
dna = "---ATTGCCTTAS--TTTAAAAAACCGTTC-----"
fmatch(dna, "-")
```

---

formatXY *Formating XY label*

Description

Add "0" in characters where should be added, inorder to make the character labels sortable.

Usage

```
formatXY(x)
```

Arguments

x	A character vector containing the labels in "XnYm" format

Details

This function will add "0" before the number 'n' if it has fewer characters than the number that has more characters in the same position. Only then the elements "XnYm" in the label vector could be sorted.

Author(s)

Jinlong Zhang < jinlongzhang01@gmail.com >

See Also

imagevect

Examples

```
formatXY(c("X1Y1","X33Y567"))
```

---

framsub *Pattern of substitutions of a dataframe.*

Description

This function will replace the specified character in the egde of each element to the replacement character of the second column of the input dataframe.

Usage

```
framsub(x, pattern = "-", replacement = "?")
```

Arguments

x	The input dataframe.
pattern	A character to be matched.
replacement	A character to be replaced.

Details

This is a general application of the function edgesub.The only the matched characters at the egde of the input DNA string will be replaced.

Value

A dataframe with the matched characters at the two edges replaced.

Author(s)

Jinlong Zhang <jinlongzhang01@gmail.com>

References

Kress W., Erickson D., Jones F., Swenson N., Perez R., Sanjur O., Bermingham E., Plant DNA barcodes and community phylogeny of a tropical forest dynamics plot in Panama. Proceedings of the National Academy of Sciences of the United States of America. 2009 18621-18626

See Also

See Also edgesub

Examples

```
##framsub example ########
dir <- system.file("extdata", package = "phylotools")
setwd(dir)
testphy <- read.phylip("rbcla.phy")
testdat <- phy2dat(testphy)
framsub(testdat)
```

---

gini *Gini coefficient of inequality*

Description

The Gini coefficient of inequality

Usage

```
gini(x)
```

Arguments

x	A numeric vector

Details

Gini coefficient is a measure of inequality. Gini coefficient is based on Lorenz's curve, and the range of this coefficient is between 0 and 1. The more lower the value, the more equality the distribution is. The algorithm of this function is directly adopted from package "ineq" by Achim Zeileis.

Author(s)

Jinlong Zhang < jinlongzhang01@gmail.com >

References

F A Cowell: Measurement of Inequality, 2000, in A B Atkinson / F Bourguignon (Eds): Handbook of Income Distribution, Amsterdam

See Also

meangini

Examples

```
### 0.3076923
gini(c(1,3,2,1,1,1,4))
```

`hcreorder` *Reorder the height for hclust object*

Description

This function could be used to reorder the height values in hclust objects, especially when the hclust object was generated by `as.hclust.phylo()` in package ape. Since the method `rect.hclust()` or `cutree()` could not be applied to the hclust objected generated by ape before the height of the nodes be sorted in a increase order.

Usage

```
hcreorder(x)
```

Arguments

`x`	object in class hclust.

Details

It should be note that if the height of the nodes are not sorted in increasing order, the `rect.hclust()` or `cutree()` will give wrong results. If the hclust object is generated by `hclust()` on distance matrix, there is no need to call this function. However, as hclust object was specified manually, or if generated using ape from a phylogenetic tree using as.hclust.phylo, you may have to call hcreorder first to reorder the heights of nodes. It should be noted that the merge matrix in hclust has also been regenerated according the definition of hclust object.

Value

hclust object with heights been reordered

Author(s)

Jinlong Zhang <jinlongzhang01@gmail.com>

References

None.

See Also

`hclust`,`cutree`,`rect.hclust`

Examples

```
### help file for hreorder()
### load ape library

library(ape)
data(bird.orders)
### convert the bird.orders to class hclust
bird.hclust <- as.hclust(multi2di(bird.orders))

### ready to draw tree/dendrograms
par(mfrow = c(1,3))

### plot phylogenetic tress
plot(bird.orders, direction = "downwards", cex = 1.5)


### the wrong results, because of the unordered height in hclust
### plot the bird.hclust
plot(bird.hclust, hang = -1 )

### add rectangles to the plot
rect.hclust(bird.hclust, k = 4)


### the proper results.

### reorder the hclust object
bird.hclust.ordered <- hcreorder(bird.hclust)
### plot the reordered hclust object
plot(bird.hclust.ordered, hang = -1 )

### add rectangles to the plot
rect.hclust(bird.hclust.ordered, k = 4)


### call cutree to divide the taxon to groups/clades
cutree(bird.hclust.ordered, k = 4)

## Not run
```

`imagevect` *Image plot of a numeric vector*

Description

An image plot of a vector given the labels in "XnYn" in format.

Usage

```
imagevect(x, labels, contour = FALSE, gridsize = 20, axes = TRUE, nlabx = 5, nlaby = 5, ...)
```

Arguments

x	A numeric vector representing the value for each quadrat.
labels	A character vector representing the row and column for each value in x. Should be in format "X10Y5", which is account for column 10 (X = 10), and row 5(Y = 5).
contour	Whether to add contour lines to the existing image plot.
gridsize	Size for each quadrat or grid
axes	Whether to draw axes for the plot
nlabx	Number of labels to be added to the x axis
nlaby	Number of labels to be added to the y axis
...	Further arguments to passed on for plotting

Details

This function will automatically sort the values in vector according to the information of rows and columns that extracted from labels, and subsequently build a matrix for invoking the `image.plot` function in "fields" package. The labels must be in "XnYn" format, and must be in the same length with vec.

Value

Image plot for the vector.

Author(s)

Jinlong Zhang < jinlongzhang01@gmail.com >

References

None

Examples

```
x <- rnorm(600)
labmat <- expand.grid(paste("X",1:30, sep = ""), paste("Y", 1:20, sep = ""))
lab <- paste(labmat[,1], labmat[,2], sep = "")
imagevect(x, labels = lab, col = cm.colors(10))
```

inequality *The inequality of a cutting time*

Description

This function could be applied to calculate the inequality of a given time.

Usage

```
inequality(tree, subtree, h, detail = FALSE)
```

Arguments

tree	The phylogenetic tree, in class 'phylo' as defined in ape
subtree	The phylogenetic tree, usually the original 'tree' with some of the tips removed.
h	Numeric: The given time
detail	Logical,the time

Details

The inequality of the insufficient sampling is evaluated by Gini coefficient with a range of c(0, 1). If the value equals to zero, a ideally equality observed; if the values equals to 1, a extremely inequality observed.

Author(s)

Jinlong Zhang < jinlongzhang01@gmail.com >

References

To be added

Examples

```
data(bird.orders)
rtr1 <- del.tree.tip(bird.orders,3)
inequality(bird.orders, rtr1[[1]], h = 25.25103)
```

meangini *Mean Gini coefficient*

Description

Calculating the inequality of clade insufficient sampling Gini coefficient using Monte Carlo simulations

Usage

```
meangini(tree, subtree, times = 1000, plot = FALSE)
```

Arguments

tree	A phylogenetic tree in class 'phylo' as in package 'ape'
subtree	A phylogenetic tree in class 'phylo', but have fewer tips than the input 'tree'.
times	Times for calculating the Gini coefficient
plot	Whether to plot the cutting pattern and the histogram

Details

This function could be used to calculating the inequality of clade insufficient sampling. The inequality can be measured by Gini coefficient, which is frequently used in economy. The Gini coefficient is based on Lorenz's curve, and the range of this coefficient is between 0 and 1. The more lower the value, the more equality the distribution is. A random time between the oldest node and the latest node in the phylogenetic tree was generated, and the phylogenetic tree was divided into two to n clades according to the time. For each clade in each "cutting" pattern, number of tips that can not be found in the perfect phylogenetic tree was recorded as number of "missing" tips. The number of missing tips was further divided by the number of expectation in the clade in perfect sampling tree. This ratio was calculated for each clade.The Gini coefficient was calculated for these values to represent the inequality of species that not included among clades. This procedure will be replicated for 1000 times in default and require intensively computationally power, this may last for minutes, even in handling small phylogenetic trees. Therefore, please be patient on the time consuming task.

Value

returns a list with the first the mean for all value, and the second a vector representing the all Gini coefficient for each randomly cutting pattern.

Note

This function require intensive computationally power, and can be very slow during handling large phylogenies.

Author(s)

Jinlong Zhang < jinlongzhang01@gmail.com >

References

Jinlong Zhang, Nathan Swenson, Xiangcheng Mi, Jihong Huang, Keping Ma, The imperfect sampling and the accuracy of phylogenetic patterns in communities.

See Also

gini

Examples

```
data(bird.orders)
to.drop <- c("Craciformes", "Galliformes", "Gruiformes")
dropped <- drop.tip(bird.orders, to.drop)
meangini(tree = bird.orders, subtree = dropped, times = 10, plot = TRUE)
meangini(tree = bird.orders, subtree = dropped, times = 50, plot = TRUE)
meangini(tree = bird.orders, subtree = dropped, times = 100, plot = TRUE)
meangini(tree = bird.orders, subtree = dropped, times = 200, plot = TRUE)
```

---

phy2dat *Convert phylip file to dataframe*

Description

Convert phylip file to dataframe

Usage

```
phy2dat(x)
```

Arguments

x	Vector of phylip format generated by Clustal software.

Details

Convert phylip file to dataframe. The phylip format vector can be read from local files via `read.phylip`, or even `readLines`.

Value

data frame with the first column the names of the input sequences, and the sequences in the second column.

Note

None

Author(s)

Jinlong Zhang <jinlongzhang01@gmail.com>

References

Kress W., Erickson D., Jones F., Swenson N., Perez R., Sanjur O., Bermingham E., Plant DNA barcodes and community phylogeny of a tropical forest dynamics plot in Panama. Proceedings of the National Academy of Sciences of the United States of America. 2009 18621-18626

See Also

See Also read.phylip

Examples

```
##phy2dat example ########
dir <- system.file("extdata", package = "phylotools")
setwd(dir)
test <- read.phylip("rbcla.phy")
phy2dat(test)
```

---

phylocom.comdist *Calculating phylogenetic comdist using Phylocom*

Description

This is a wrapper function for phylocom.

Usage

```
phylocom.comdist(sample = "sample", phylo = "phylo", method = "comdist", aw = TRUE, file = "comdist.tx
```

Arguments

sample	Character string, the name of the sample file used in Phylocom
phylo	Character string, the name of the "phylo" file used in Phylocom
method	character, only "comdist" or "comdistn"
aw	Logical, if TRUE, the comdist will be computed in accounting of abundance. Otherwise, for presence - absence.
file	Result File that to be saved to .

Author(s)

Jinlong Zhang < jinlongzhang01@gmail.com >

References

Campbell Webb, David Ackerly, Steve Kembel (2010) Phylocom software for the analysis of phylogenetic community structure and character evolution (with phylomatic and ecovolve) User S Manual Version 4.1

See Also

phylocom.pd,phylocom.comstruct

Examples

```
### Assume the file sample1 and phylo1 have been saved
### to the R working directory. The phylocom executable
### file has been copied to the same working directory
### Or could be invoked in console by command lines.
### in order to run phylocom in R, simply copy the
### following command:
#   i = 1
#   phylocom.comdist(sample = paste("sample",i, sep = ""),
#                   phylo = paste("phylo",i, sep = ""),
#                   method = "comdist", aw = TRUE,
#                   file =  paste("comdist.aw",  i,
#                   ".txt", sep = ""))
#
```

`phylocom.comstruct` *Calculating community phylogenetic structure using Phylocom*

Description

This is a wrapper function for phylocom.

Usage

```
phylocom.comstruct(sample = "sample", phylo = "phylo",
    swapmethod = 2, runs = 999, swaps = 1000,
    aw = TRUE, file = "comstruct.txt")
```

Arguments

`sample`	Character string, the name of the sample file used in Phylocom
`phylo`	Character string, the name of the "phylo" file used in Phylocom
`swapmethod`	Swap method used in phylocom
`runs`	Times to generating null models
`swaps`	Times to swap
`aw`	Logical, if TRUE, the comdist will be computed in accounting of abundance. Otherwise, for presence - absence.
`file`	Result File that to be saved to .

Author(s)

Jinlong Zhang < jinlongzhang01@gmail.com >

References

Campbell Webb, David Ackerly, Steve Kembel (2010) Phylocom software for the analysis of phylogenetic community structure and character evolution (with phylomatic and ecovolve) User S Manual Version 4.1

See Also

phylocom.pd,phylocom.comdist

Examples

```
### Assume the file sample1 and phylo1 have been
### saved to the R working directory.
### The phylocom executable file has been copied
### to the same working directory
### Or could be invoked in console by command lines.
### in order to run phylocom in R, simply copy the
### following command:

#   i = 1
#   phylocom.comstruct(sample = paste("sample",i, sep = ""),
#            phylo = paste("phylo",i, sep = ""),
#            swapmethod = 2, runs = 999, swaps = 1000, aw = TRUE,
#            file = paste("comstruct.aw",   i,".txt", sep = ""))
```

---

phylocom.pd *Calculate phylogenetic diversity using Phylocom*

Description

This is a wrapper function of phylocom for Calculate phylogenetic diversity using Phylocom

Usage

```
phylocom.pd(sample = "sample", phylo = "phylo")
```

Arguments

sample	Character string, the name of the sample file used in Phylocom
phylo	Character string, the name of the "phylo" file used in Phylocom

Value

PD data frame

Author(s)

Jinlong Zhang < jinlongzhang01@gmail.com >

References

Campbell Webb, David Ackerly, Steve Kembel (2010) Phylocom software for the analysis of phylogenetic community structure and character evolution (with phylomatic and ecovolve) User S Manual Version 4.1

See Also

phylocom.comstruct,phylocom.comdist

Examples

```
### Assume the file sample1 and phylo1 have been saved to the R working directory.
### The phylocom executable file has been copied to the same working directory
### Or could be invoked in console by command lines.
### inorder to run phylocom in R, simply copy the following command:
#   i = 1
#   PD.res <- phylocom.pd(sample = paste("sample",i, sep = ""),
#          phylo = paste("phylo",i, sep = ""),
#          )
#   PD.res
#
```

---

phylocom.phylosor *Calculating Phylosor similarity*

Description

This function can be used to calculating phylosor (Phylogenetic Sorenson) similarity between each pair of places

Usage

```
phylocom.phylosor(sample.file = "sample", phylo = "phylo")
```

Arguments

sample.file	Phylocom "Sample" file in the working directory .
phylo	The phylo file in the working directory, must be in newick format.

Details

This function has been tested for calculating phylosor for very large phylogenies (more than 17000 tips that generated by Phylomatic)

Value

a distance matrix

Author(s)

Jinlong Zhang < jinlongzhang01@gmail.com >

References

Bryant, J.A., Lamanna, C., Morlon, H., Kerkhoff, A.J., Enquist, B.J. & Green, J.L. (2008) Microbes on mountainsides: Contrasting elevational patterns of bacterial and plant diversity. Proceedings of the National Academy of Sciences of the United States of America, 105, 11505-11511.

Morlon, H., Schwilk, D.W., Bryant, J.A., Marquet, P.A., Rebelo, A.G., Tauss, C., Bohannan, B.J.M. & Green, J.L. (2011) Spatial patterns of phylogenetic diversity. Ecology Letters, 14, 141-149.

See Also

phylocom.pd

Examples

```
#  Please make sure the Phylocom executable file can
#  be invoked by console by adding the system path for it.
#  res <- phylocom.phylosor(sample.file = "sample", phylo = "phylo")
#
```

---

phyloshuffle *Tip label shuffling of a phylogenetic tree*

Description

Internal function used in phylotools.

Usage

```
phyloshuffle(tree)
```

Arguments

tree	Phylogenetic tree

Details

Shuffle the phylogenetic tips

Value

Randomized phylogenetic tree

Author(s)

Jinlong Zhang < jinlongzhang01@gmail.com >

Examples

```
data(bird.orders)
par(mfrow = c(1,2))
plot(bird.orders)
rand <- phyloshuffle(bird.orders)
plot(rand)
```

---

plotcut *Cutting the phylogenetic cutting pattern by a given time or times*

Description

Add a line to the plot phylogenetic tree

Usage

```
plotcut(tree, n)
```

Arguments

tree	The phylogenetic tree in class 'phylo' defined in ape
n	Number of randomly placed lines to add to.

Details

This function add lines to a phylogenetic tree randomly. There for it is only a demonstration for cutting patters of the phylogenetic tree. This cutting pattern could be used to evaluate the inequality of clade insufficient sampling as implemented in meangini

Author(s)

Jinlong Zhang < jinlongzhang01@gmail.com >

See Also

meangini

Examples

```
data(bird.orders)
plotcut(bird.orders, 10)
```

`plotscale` *Rescaling the individuals occurred in a Forestry Dynamic Plot*

Description

Rescaling the individuals occurred in a Forestry Dynamic Plot. It will add a column called "gridnames". This function could make the analysis based on different scales easier.

Usage

```
plotscale(inputdata, len = NULL, wid = NULL, scale = NULL)
```

Arguments

`inputdata`	A data matrix inputdata that has "tag", "x" and "y".
`len`	Length of the FDP
`wid`	Width of the FDP
`scale`	Size of the quadrat

Details

The input data must he following columns, "Tag" : the tag number for each individual in the FDP. "x", "y" are the coordinats of the individuals in the FDP.

Value

A column will be added to the input matrix , to represent the position for each individual. For example, "X5Y20" means the corresponding individual lies at cell X = 5,y = 20. The labels could be further used to represent the species matrix.

Author(s)

Jinlong Zhang < `jinlongzhang01@gmail.com` >

References

Nathan J. B. Kraft, Renato Valencia, David D. Ackerly (2008) Functional Traits and Niche-Based Tree Community Assembly in an Amazonian Forest Science 322, 580 ;

Examples

```
## 20m
## plotscale(inputdata = BCI, len = 1000, wid = 500,
##           scale = 20)
## 50m
## plotscale(inputdata = BCI, len = 1000, wid = 500,
##           scale = 50)
## 100m
## plotscale(inputdata = BCI, len = 1000, wid = 500,
##          scale = 100)
```

`read.phylip` *Read phylip file to memory.*

Description

Read phylip file to memory.

Usage

```
read.phylip(fil = NULL)
```

Arguments

`fil`	file path for the input phylip file.

Details

The phylip format should be generated by the Clustal X software.

Value

Phylip objected to be used in further analysis.

Note

None.

Author(s)

Jinlong Zhang <jinlongzhang01@gmail.com>

References

Kress W., Erickson D., Jones F., Swenson N., Perez R., Sanjur O., Bermingham E., Plant DNA barcodes and community phylogeny of a tropical forest dynamics plot in Panama. Proceedings of the National Academy of Sciences of the United States of America. 2009 18621-18626

See Also

See Also as readLines

Examples

```
## read.phylip() examle ########

dir <- system.file("extdata", package = "phylotools")
setwd(dir)
read.phylip("rbcla.phy")
```

---

rename.fasta *Renaming a fasta object given a reference table*

Description

this function could be used to rename the sequences in the fasta object, given a reference dataframe.

Usage

```
rename.fasta(fas, ref, fil = NULL, prefix = NULL)
```

Arguments

fas	the fasta object.
ref	Dataframe with first column of the names of the sequences , the second column which name to change into.
fil	A string indicating the file name, the result written to.
prefix	prefix of which the names of the sequences.

Details

Prefix providing the flexibility of modifying the input names of sequences, for example, if the sequences names begin with "P_seq1", and unfortunately the names in reference table are lack of "P_". You may set the prefix as "P_", and without changing the reference table.

Value

fasta object

Author(s)

Jinlong Zhang <jinlongzhang01@gmail.com>

References

None

See Also

gnames.fas

Examples

```
library(seqRFLP)
## loading data
data(fil.fas)
## Get the names of the sequences
col1 <- gnames.fas(fil.fas)
## Generating new names
col2 <- paste("seq", 1:length(col1), sep = "")
reftable.rename <- data.frame(col1, col2)
##
rename.fasta(fil.fas, reftable.rename)
```

---

resid.tree *The undeleted terminals of a phylogenetic tree*

Description

This is a internal function, and only handling the object that generated by del.tree.tip.

Usage

```
resid.tree(tree, deltree)
```

Arguments

tree	A phylogenetic tree
deltree	An object that generated by del.tree.tip

Value

The phylogenetic tree of the deleted tipes

Author(s)

Jinlong Zhang < jinlongzhang01@gmail.com >

See Also

del.tree.tip

Examples

```
data(bird.orders)
del <- del.tree.tip(bird.orders,3)
del.tree <- resid.tree(bird.orders, del)
plot(del.tree)
```

`reverse` *Get the reverse sequence.*

Description

This function could be used to obtain the reverse sequence of the input sequence.

Usage

```
reverse(dna)
```

Arguments

`dna`	A character string of input DNA sequence

Details

This function could be used to obtain the reverse sequence of the input sequence.

Value

A string contains the reverse sequence.

Note

None

Author(s)

Jinlong Zhang <jinlongzhang01@gmail.com>

References

Kress W., Erickson D., Jones F., Swenson N., Perez R., Sanjur O., Bermingham E., Plant DNA barcodes and community phylogeny of a tropical forest dynamics plot in Panama. Proceedings of the National Academy of Sciences of the United States of America. 2009 18621-18626

See Also

See Also `fmatch`

Examples

```
## reverse example ########
dna = "---ATTGCCTTAS--TTTAAAAAACCGTTC-----"
reverse(dna)
```

---

rh *Generating random time between the oldest node and the latest node*

Description

Random time between the oldest node and the latest node. This function only handle the ultrametric tree.

Usage

```
rh(tree, times = 1)
```

Arguments

tree	The phylogenetic tree
times	Number of random time values to be generated

Value

A vector representing the time, that randomly generated

Note

This function only handle the ultrametric tree.

Author(s)

Jinlong Zhang < jinlongzhang01@gmail.com >

References

To be added

Examples

```
data(bird.orders)
rh(bird.orders, 10)
```

rm.col *Delete specified columns for dataframe or matrix*

Description

Delete specified columns for dataframe or matrix.

Usage

```
rm.col(mat, to.rm = NULL)
```

Arguments

mat	a dataframe or matrix, must have column names.
to.rm	a character vector representing the column names that to be removed.

Value

A dataframe or matrix (according to the input data)with certain column deleted.

Author(s)

Jinlong Zhang < jinlongzhang01@gmail.com >

References

None

Examples

```
x <- rnorm(600)
dim(x) <- c(20,30)
colnames(x) <- paste("col",1:30,sep = "")
rm.col(x, to.rm = c("col1", "col3", "col8", "col22", "col30"))
```

RMPD *Ratio of mean phylogenetic distance of two phylogenetic trees*

Description

Ratio of the mean phylogenetic distance between the tips in subtree to the mean phylogenetic distance between the tips in whole tree.

Usage

```
RMPD(subtree, tree)
```

Arguments

subtree	A subset of phylogenetic tree
tree	A phylogenetic tree

Details

RMPD > 1 indicate the tips in the "subtree" are less phylogenetic related than in "tree". This ratio could be used as a representative of phylogenetic relationships of species that occurred in subtree in comparison with in "tree". RMPD < 1 indicate the tips in "subtree" are more close related

Author(s)

Jinlong Zhang < jinlongzhang01@gmail.com >

References

To be added

See Also

resid.tree

Examples

```
data(bird.orders)
del <- del.tree.tip(bird.orders,3)
del.tree <- resid.tree(bird.orders, del)
RMPD(subtree = del.tree, tree = bird.orders)
```

`seq2fasta` *Convert Seq file to fasta format*

Description

This function convert .seq file generated by dnaMAN to fasta format.

Usage

```
seq2fasta(file)
```

Arguments

`file`	A character string indicate the ".seq" file name.

Details

Seq file is a DNA format generated by dnaMAN, often the exported sequences assembly results. Sometimes these files will have to be converted to fasta format in order to carry on further analysis.

Value

fasta file.

Author(s)

Jinlong Zhang <jinlongzhang01@gmail.com>

References

To be added

Examples

```
dir <- system.file("extdata", package = "phylotools")
seq2fasta(paste(dir,"/DNAman.seq", sep = ""))
```

sub.tip.label *Substitute the tip labels of a phylogenetic tree*

Description

This function may be used in changing the tip labels of a phylogenetic tree according to a reference data table.

Usage

```
sub.tip.label(tree, dat)
```

Arguments

tree	Phylogenetic tree
dat	A dataframe with the first the tip labels and the second the substitutions.

Details

To be added

Value

A Phylogenetic tree with the tip labels substituted

Author(s)

Jinlong Zhang <jinlongzhang01@gmail.com>

References

None

Examples

```
library(ape)
data(bird.families)
tips <- bird.families$tip.label
abr <- paste("fam",1:length(tips), sep = "")
dat <- data.frame(tips, abr)
ntree <- sub.tip.label(bird.families, dat)
```

supermat *Build super matrix using PHYLIP files*

Description

This function could be used to build the super dataframe described by Kress et al.(2009). The out put dataframe could be converted to Phylip format file for phylogenetic analysis.

Usage

```
supermat(rbcl = NULL, matk = NULL, trn = NULL)
```

Arguments

rbcl	A character string indicting the name of rbcl phylip file.
matk	A character string indicting the name of matK phylip file.
trn	A vector of character strings indicting the names of trnH-psbA phylip files.

Details

The phylip files should be first put into a directory. Users are encouraged to copy the phylip files which generated by software Clustal to a directory, and then set it as the working directory. The function will search the input phylip file names from the working directory and build the "super" dataframe. The dataframe then could be converted to phylip format by dat2phy and be written to the working directory by writeLines so that to generate a super matrix.

Value

The super data frame to be converted to supermatrix.

Note

None.

Author(s)

Jinlong Zhang <jinlongzhang01@gmail.com>

References

Kress W., Erickson D., Jones F., Swenson N., Perez R., Sanjur O., Bermingham E., Plant DNA barcodes and community phylogeny of a tropical forest dynamics plot in Panama. Proceedings of the National Academy of Sciences of the United States of America. 2009 18621-18626

See Also

See Also dat2phy, writeLines

Examples

```
dir <- system.file("extdata", package = "phylotools")
setwd(dir)
## Supermatrix with "rbcla","matk","trnH-psbA"
supermat1 <- supermat(rbcl = "rbcla.phy", matk = "matK.phy",
           trn = c("trn1.phy", "trn2.phy","trn3.phy","trn4.phy"))
write.mat(supermat1, "result1.phy")


## Supermatrix with"matk","trnH-psbA"
supermat2 <- supermat(matk = "matK.phy",
           trn = c("trn1.phy", "trn2.phy","trn3.phy","trn4.phy"))
write.mat(supermat2, "result2.phy")


## Supermatrix with "rbcla","matk"
supermat3 <- supermat(matk = "matK.phy", rbcl = "rbcla.phy")
write.mat(supermat3, "result3.phy")


## Supermatrix with "rbcla","trnH-psbA"
supermat4 <- supermat(rbcl = "rbcla.phy",
           trn = c("trn1.phy", "trn2.phy","trn3.phy","trn4.phy"))
write.mat(supermat4, "result4.phy")

## Delete the results
unlink(c("result1.phy", "result2.phy", "result3.phy", "result4.phy"))
```

`uniquefasta` *Deleting duplicated sequences in fasta object*

Description

This function may be used to deleting the duplicated sequences in the fasta object. The sequences with its name same to any sequences will be ignored. Only the first sequence will be retained.

Usage

```
uniquefasta(fasta)
```

Arguments

`fasta`	The fasta object

Value

This function returns to the fasta object with unique sequence.

Note

This function will remove the duplicated sequences and retain the first sequence only, all according to the names of the sequences.

Author(s)

Jinlong Zhang <jinlongzhang01@mail.com>

References

None

See Also

unique

Examples

```
library(seqRFLP)
## loading data
data(fil.fas)
## Get the names of the sequences
fasta.names <- gnames.fas(fil.fas)
new.names0 <- substring(fasta.names,1,2 )
new.names <- paste(new.names0, sep = "")
reftable.rename <- data.frame(fasta.names, new.names)
renamed <- rename.fasta(fil.fas, reftable.rename)
uniquefasta(renamed)
```

---

write.mat *Save supermatrix*

Description

Save supermatrix to a phylip file

Usage

```
write.mat(supermat, file = NULL)
```

Arguments

supermat	Supermatrix as build by function supermat
file	File name to be specified

Details

None

Value

Save supermatrix to phylip file.

Author(s)

Jinlong Zhang <jinlongzhang01@gmail.com>

References

Kress W., Erickson D., Jones F., Swenson N., Perez R., Sanjur O., Bermingham E., Plant DNA barcodes and community phylogeny of a tropical forest dynamics plot in Panama. Proceedings of the National Academy of Sciences of the United States of America. 2009 18621-18626

See Also

supermat

Examples

```
dir <- system.file("extdata", package = "phylotools")
file.copy(from = dir, to = getwd())
## Supermatrix with "rbcla","matk","trnH-psbA"
supermat1 <- supermat(rbcl = "rbcla.phy", matk = "matK.phy",
           trn = c("trn1.phy", "trn2.phy","trn3.phy","trn4.phy"))
write.mat(supermat1, "result1.phy")
```

Index