# Package 'Rphylip'

March 25, 2014

**Version** 0.1-23

**Date** 2014-3-25

Title An R interface for PHYLIP

Author Liam J. Revell, Scott A. Chamberlain
Maintainer Liam J. Revell <liam.revell@umb.edu></liam.revell@umb.edu>
<b>Depends</b> R (>= $2.10$ ), ape (>= $3.0-10$ )
<b>ZipData</b> no
<b>Description</b> Rphylip provides an R interface for the PHYLIP package. All users of Rphylip will thus first have to install the PHYLIP phylogeny methods program package (Felsenstein 2013). See http://www.phylip.com for more information about installing PHYLIP.
License GPL (>= 2)
URL http://www.phytools.org/Rphylip
Repository
Date/Publication 2014-3-25 12:00:00 EDT  R topics documented:
Rphylip-package       2         as.proseq       3         opt.Rdnaml       4         primates       5         print.proseq       6         Rclique       6         Rconsense       5         Rcontml       9         Rcontrast       10         Rdnacomp       12

2 Rphylip-package

Index	4	9
	setupOSX	8
	setPath	
	Rtreedist	_
	Rthreshml	4
	Rseqboot	3
	Rrestml	2
	Rrestdist	0
	Rprotpars	9
	Rprotdist	7
	Rproml	5
	Rpenny	4
	Rpars	2
	Rneighbor	1
	Rmix	
	Rkitsch	7
	Rgendist	_
	Rfitch	
	read.protein	_
	Rdolpenny	
	Rdollop	
	Rdnapenny	9

Rphylip-package

Rphylip: An R interface for PHYLIP

# **Description**

**Rphylip** provides an R interface for programs in the PHYLIP phylogeny methods package (Felsenstein 1989, 2013).

# **Details**

The complete list of functions can be displayed with library(help = Rphylip).

Obviously, before any of the functions of this package can be used, users must first install PHYLIP (Felsenstein 2013). More information about installing PHYLIP can be found on the PHYLIP webpage: http://evolution.genetics.washington.edu/phylip.html.

More information on **Rphylip** can be found at http://www.phytools.org/Rphylip/ or http://blog.phytools.org. The latest code for the development version of Rphylip can also be found on github at the following URL: http://github.com/liamrevell/Rphylip.

#### Author(s)

Liam J. Revell, Scott A. Chamberlain

Maintainer: Liam J. Revell < liam.revell@umb.edu>

as.proseq 3

## References

Felsenstein, J. (1989) PHYLIP-Phylogeny Inference Package (Version 3.2). Cladistics, 5, 164-166.

Felsenstein, J. (2013) PHYLIP (Phylogeny Inference Package) version 3.695. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Revell, L. J. (2013) Rphylip: An R interface for PHYLIP. R package version 0-1.09.

as.proseq

Converts objects to protein sequences, phylip.data, or rest.data object

## **Description**

Converts objects to class "proseq", "phylip.data", or "rest.data".

# Usage

```
as.proseq(x, ...)
as.phylip.data(x, ...)
as.rest.data(x, ...)
```

# **Arguments**

Х

an object containing amino sequences, a set of binary characters, an arbitrary user-defined multistate character, or presences ("+") and absences ("-") of restriction sites. as.proseq and as.phylip.data also converts objects of class "phyDat" from the phangorn package.

... optional arguments.

# Value

An object of class "proseq" containing protein sequences; an object of class "phylip.data"; or an object of class "rest.data".

# Author(s)

Liam J. Revell, Scott A. Chamberlain

Maintainer: Liam J. Revell < liam.revell@umb.edu>

# See Also

```
print.proseq, Rproml
```

4 opt.Rdnaml

opt.Rdnaml

Parameter optimizer for Rdnaml

## **Description**

This function is an wrapper for Rdnaml that attempts to optimize gamma (the alpha shape parameter of the gamma model of rate heterogeneity among sites), kappa (the transition:transversion ratio), and bf (the base frequencies).

# Usage

```
opt.Rdnaml(X, path=NULL, ...)
```

# **Arguments**

X an object of class "DNAbin".

path to the executable containing dnaml. If path = NULL, the R will search

several commonly used directories for the correct executable file.

... optional arguments. See details for more information.

#### **Details**

Optional arguments include the following: tree fixed tree to use in optimization - if not provided, it will be estimated using Rdnaml under the default conditions; bounds a list with bounds for optimization - for kappa and gamma this should be a two-element vector, whereas for bf this should be a 4 x 2 matrix with lower bounds in column 1 and upper bounds in column 2.

More information about the dnaml program in PHYLIP can be found here http://evolution.genetics.washington.edu/phylip/doc/dnaml.html.

Obviously, use of any of the functions of this package requires that PHYLIP (Felsenstein 1989, 2013) should first be installed. Instructions for installing PHYLIP can be found on the PHYLIP webpage: http://evolution.genetics.washington.edu/phylip.html.

## Value

This function returns a list with the following components: kappa, gamma, bf (see Details), and logLik (the log-likelihood of the fitted model).

# Author(s)

Liam J. Revell, Scott A. Chamberlain

Maintainer: Liam J. Revell < liam.revell@umb.edu>

# References

Felsenstein, J. (1981) Evolutionary trees from DNA sequences: A Maximum Likelihood approach. *Journal of Molecular Evolution*, 17, 368-376.

Felsenstein, J. (1989) PHYLIP–Phylogeny Inference Package (Version 3.2). *Cladistics*, 5, 164-166. Felsenstein, J. (2013) PHYLIP (Phylogeny Inference Package) version 3.695. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Felsenstein, J., Churchill, G. A. (1996) A Hidden Markov Model approach to variation among sites in rate of evolution. *Molecular Biology and Evolution*, 13, 93-104.

primates 5

#### See Also

Rdnam1

#### **Examples**

```
## Not run:
data(primates)
fit<-opt.Rdnaml(primates,bounds=list(kappa=c(0.1,40))
tree<-Rdnaml(primates,kappa=fit$kappa,gamma=fit$gamma,bf=fit$bf)
## End(Not run)</pre>
```

primates

Example datasets

# **Description**

primates is an object of class "DNAbin" containing nucleotide sequence data of mysterious origin for 12 species of primates. chloroplast is an object of class"proseq" containing a chloroplast alignment from the phangorn package (Schliep 2011). primates.bin is a an object of class "phylip.data" containing "0" and "1" character data for primates. cotton is a list with two parts: cotton\$tree is an object of class "phylo", containing a phylogenetic tree for 37 species of cotton (Gossypium); cotton\$data is an object of class "matrix" containing traits for each species. Most trait values are real, but a few values were not available and so random values were used so there were no missing values in the dataset. restriction.data is an object of class "rest.data" containing an example dataset of restriction site data.

## Usage

```
data(primates)
data(chloroplast)
data(primates.bin)
data(cotton)
data(restriction.data)
```

#### Format

The data are stored as an object of class "DNAbin", "proseq", "phylip.data", or "rest.data". cotton is a list containing an object "phylo" and a matrix of continuous trait data.

## Author(s)

Liam J. Revell, Scott A. Chamberlain

Maintainer: Liam J. Revell < liam.revell@umb.edu>

## Source

Chamberlain, S.A., & J.A. Rudgers. 2012. How do plants balance multiple mutualists? Correlations among traits for attracting protective bodyguards and pollinators in cotton (Gossypium). Evolutionary Ecology 26:65-77. DOI: 10.1007/s10682-011-9497-3.

6 Relique

print.proseq

Print method protein sequences, phylip.data, or rest.data object

# **Description**

Print method for an objects of class "proseq", "phylip.data", or "rest.data".

# Usage

```
## S3 method for class 'proseq'
print(x, printlen=6, digits=3, ...)
```

# **Arguments**

x an object of class "proseq", an object of class "phylip.data", or an object of class "rest.data".

printlen number of sequence names to print.

digits number of digits to print.
... optional arguments.

## Value

Prints to screen.

# Author(s)

Liam J. Revell, Scott A. Chamberlain

Maintainer: Liam J. Revell < liam.revell@umb.edu>

# See Also

```
as.proseq, print.DNAbin, Rproml
```

Rclique

R interface for clique

# Description

This function is an R interface for clique in the PHYLIP package (Felsenstein 2013). clique can be used for phylogeny inference using the compatibility method (Le Quesne 1969; Estabrook et al. 1976).

# Usage

```
Rclique(X, path=NULL, ...)
```

Rclique 7

## **Arguments**

Χ	an object of class "phylip.data" with a set of binary characters; or a matrix.
path	path to the executable containing clique. If path = NULL, the R will search several commonly used directories for the correct executable file.
	optional arguments to be passed to clique. See details for more information.

#### **Details**

Optional arguments include the following: quiet suppress some output to R console (defaults to quiet = FALSE); ancestral vector of ancestral states; weights vector of weights of length equal to the number of columns in X (defaults to unweighted); minimum.clique, if specified, will cause Rclique to print out all cliques (and corresponding trees) greater than minimum.clique; outgroup outgroup if outgroup rooting of the estimated tree is desired; and cleanup remove PHYLIP input & output files after the analysis is completed (defaults to cleanup = TRUE).

More information about the clique program in PHYLIP can be found here <a href="http://evolution.genetics.washington.edu/phylip/doc/clique.html">http://evolution.genetics.washington.edu/phylip/doc/clique.html</a>.

Obviously, use of any of the functions of this package requires that PHYLIP (Felsenstein 1989, 2013) should first be installed. Instructions for installing PHYLIP can be found on the PHYLIP webpage: http://evolution.genetics.washington.edu/phylip.html.

#### Value

This function returns an object of class "phylo" that is the optimized tree.

#### Author(s)

Liam J. Revell, Scott A. Chamberlain

Maintainer: Liam J. Revell < liam.revell@umb.edu>

#### References

Estabrook, G. F., Johnson, C. S. Jr., McMorris, F. R. (1976). A mathematical foundation for the analysis of character compatibility. *Mathematical Biosciences*, 23, 181-187.

Felsenstein, J. (1989) PHYLIP-Phylogeny Inference Package (Version 3.2). Cladistics, 5, 164-166.

Felsenstein, J. (2013) PHYLIP (Phylogeny Inference Package) version 3.695. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Le Quesne, W.J. (1969) A method of selection of characters in numerical taxonomy. *Systematic Zoology*, 18, 201-205.

## See Also

```
Rdnacomp, Rmix
```

## **Examples**

```
## Not run:
data(primates.bin)
tree<-Rclique(primates.bin)
## End(Not run)</pre>
```

8 Rconsense

Rconsense

R interface for consense

## **Description**

This function is an R interface for consense in the PHYLIP package (Felsenstein 2013). consense can be used to compute the consensus tree from a set of phylogenies.

# Usage

```
Rconsense(trees, path=NULL, ...)
```

## **Arguments**

trees an object of class "multiPhylo".

path path to the directory containing the executable consense. If path = NULL, the R

will search several commonly used directories for the correct executable file.

... optional arguments to be passed to consense. See details for more information.

#### **Details**

Optional arguments include the following: quiet suppress some output to R console (defaults to quiet = FALSE); method which can be "extended" (extended majority rule consensus, the default), "strict" (strict consensus), or regular majority rule consensus ("majority"); "outgroup" single taxon label or vector of taxa that should be used to root all trees before analysis; rooted logical value indicated whether to treat the input trees as rooted (defaults to rooted = FALSE); and cleanup remove PHYLIP input & output files after the analysis is completed (defaults to cleanup = TRUE).

More information about the consense program in PHYLIP can be found here http://evolution.genetics.washington.edu/phylip/doc/consense.html.

Obviously, use of any of the functions of this package requires that PHYLIP (Felsenstein 1989, 2013) should first be installed. Instructions for installing PHYLIP can be found on the PHYLIP webpage: http://evolution.genetics.washington.edu/phylip.html.

## Value

This function returns an object of class "phylo". For methods other than method = "strict", tree\$node.label contains the proportion of phylogenies in trees containing that subtree.

## Author(s)

Liam J. Revell, Scott A. Chamberlain

Maintainer: Liam J. Revell < liam.revell@umb.edu>

# References

Margush, T., McMorris, F.R. (1981) Consensus n-trees. *Bulletin of Mathematical Biology*, 43, 239-244.

Felsenstein, J. (1989) PHYLIP–Phylogeny Inference Package (Version 3.2). *Cladistics*, 5, 164-166. Felsenstein, J. (2013) PHYLIP (Phylogeny Inference Package) version 3.695. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Rcontml 9

## **Examples**

```
## Not run:
trees<-rmtree(n=10,N=10)
tree<-Rconsense(trees)
## End(Not run)</pre>
```

Rcontml

R interface for contml

# Description

This function is an R interface for contml in the PHYLIP package (Felsenstein 2013). contml can be used for ML phylogeny estimation from gene frequency data or continuous characters. The continuous characters should be rotated so as to be uncorrelated before analysis.

# Usage

```
Rcontml(X, path=NULL, ...)
```

## Arguments

X	either (a) a <i>matrix</i> of continuous valued traits (in columns) with rownames containing species names; or (b) a list of matrices in which each row contains the relative frequency of alleles at a locus for a species. In the latter case the rownames of each matrix in the list should contain the species names.
path	path to the executable containing contml. If path = NULL, the R will search several commonly used directories for the correct executable file.
	optional arguments to be passed to contml. See details for more information.

# **Details**

Optional arguments include the following: quiet suppress some output to R console (defaults to quiet = FALSE); tree object of class "phylo" - if supplied, then the model will be optimized on a fixed input topology; global perform global search (defaults to global = TRUE); random.order add taxa to tree in random order (defaults to random.order = TRUE); random.addition number of random addition replicates for random.order = TRUE (defaults to random.addition = 10); outgroup outgroup if outgroup rooting of the estimated tree is desired; and cleanup remove PHYLIP input/output files after the analysis is completed (defaults to cleanup = TRUE).

More information about the contml program in PHYLIP can be found here http://evolution.genetics.washington.edu/phylip/doc/contml.html.

Obviously, use of any of the functions of this package requires that PHYLIP (Felsenstein 1989, 2013) should first be installed. Instructions for installing PHYLIP can be found on the PHYLIP webpage: http://evolution.genetics.washington.edu/phylip.html.

## Value

This function returns an object of class "phylo" that is the optimized tree.

10 Rcontrast

#### Author(s)

Liam J. Revell, Scott A. Chamberlain

Maintainer: Liam J. Revell < liam.revell@umb.edu>

## References

Felsenstein, J. (1981) Maximum likelihood estimation of evolutionary trees from continuous characters. *American Journal of Human Genetics*, 25, 471-492.

Felsenstein, J. (1981) Maximum likelihood estimation of evolutionary trees from continuous characters. *American Journal of Human Genetics*, 25, 471-492.

Felsenstein, J. (1989) PHYLIP-Phylogeny Inference Package (Version 3.2). Cladistics, 5, 164-166.

Felsenstein, J. (2013) PHYLIP (Phylogeny Inference Package) version 3.695. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

## See Also

Rdnaml, Rproml

Rcontrast

R interface for contrast

# **Description**

This function is an R interface for contrast in the PHYLIP package (Felsenstein 2013). contrast can be used to perform the among species phylogenetically independent contrasts method of Felsenstein (1985) and the within & among species method of Felsenstein (2008).

More information about the contrast program in PHYLIP can be found here http://evolution.genetics.washington.edu/phylip/doc/contrast.html.

Obviously, use of any of the functions of this package requires that PHYLIP (Felsenstein 1989, 2013) should first be installed. Instructions for installing PHYLIP can be found on the PHYLIP webpage: http://evolution.genetics.washington.edu/phylip.html.

#### Usage

```
Rcontrast(tree, X, path=NULL, ...)
```

# **Arguments**

tree	object of class "phylo".
X	a <i>matrix</i> of continuous valued traits (in columns) with rownames containing species names. For one trait, X can be a matrix with one column or a vector with names(X) containing species names matching tree\$tip.label. For within-species contrasts analysis, the matrix should contain repeating (identical) row names for conspecifics.
path	path to the executable containing contrast. If path = NULL, the R will search several commonly used directories for the correct executable file.
	optional arguments to be passed to contrast. See details for more information.

Rcontrast 11

#### **Details**

Optional arguments include the following: quiet suppress some output to R console (defaults to quiet = FALSE); and cleanup remove PHYLIP input/output files after the analysis is completed (defaults to cleanup = TRUE).

#### Value

If X contains one observation per species (say, the species mean), then Rcontrast returns a list with the following components: Contrasts, a matrix with all phylogenetically independent contrats; Covariance\_matrix, a matrix containing the evolutionary variances (on diagonals) and covariances; Regressions, a matrix containing the pair-wise bivariate regression coefficients (columns on rows); Correlations, a correlation matrix of contrasts.

If X contains more than one sample per species, then Rcontrast returns a list with the following elements: VarA, the estimated among-species variance-covariance matrix; VarE, the estimated within-species (i.e., 'environmental') variance-covariance matrix; VarA.Regression, a matrix containing the pair-wise bivariate among-species regression coefficients (columns on rows); VarA.Correlations, a matrix with the among-species evolutionary correlations; VarE.Regressions, the pair-wise bivariate within-species regression coefficients; VarE.Correlations, the within-species correlations; nonVa.VarE, nonVa.VarE.Regressions, and nonVa.VarA.Correlations, estimates obtained when VarA is not included in the model; logLik and nonVa.logLik, log-likelihood when VarA is included (or not) in the model; k and nonVa.k the number of parameters estimated in each model; and P the p-value of a likelihood-ratio test of VarA, in which df = k - nonVa.k.

#### Author(s)

Liam J. Revell, Scott A. Chamberlain

Maintainer: Liam J. Revell < liam.revell@umb.edu>

## References

Felsenstein, J. (1985) Phylogenies and the comparative method. American Naturalist, 125, 1-15.

Felsenstein, J. (1989) PHYLIP-Phylogeny Inference Package (Version 3.2). Cladistics, 5, 164-166.

Felsenstein, J. (2008) Comparative methods with sampling error and within-species variation: Contrasts revisited and revised. *American Naturalist*, 171, 713-725.

Felsenstein, J. (2013) PHYLIP (Phylogeny Inference Package) version 3.695. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

# See Also

```
pic, Rcontml, Rthreshml
```

# **Examples**

```
## Not run:
data(cotton)
contrasts<-Rcontrast(tree = cotton$tree, X = cotton$data)
## End(Not run)</pre>
```

12 Rdnacomp

Rdnacomp	R interface for dnacomp
----------	-------------------------

# **Description**

This function is an R interface for dnacomp in the PHYLIP package (Felsenstein 2013). dnacomp can be used for phylogeny inference from DNA sequences using the compatibility method (Le Quesne 1969; Fitch 1975).

## Usage

```
Rdnacomp(X, path=NULL, ...)
```

## **Arguments**

X an object of class "DNAbin".

path path to the executable containing dnacomp. If path = NULL, the R will search

several commonly used directories for the correct executable file.

... optional arguments to be passed to dnacomp. See details for more information.

#### **Details**

Optional arguments include the following: quiet suppress some output to R console (defaults to quiet = FALSE); tree object of class "phylo" - if supplied, then the number of compatible sites will be computed on a fixed input topology; random.order add taxa to tree in random order (defaults to random.order = TRUE); random.addition number of random addition replicates for random.order = TRUE (defaults to random.addition = 10); weights vector of weights of length equal to the number of columns in X (defaults to unweighted); outgroup outgroup if outgroup rooting of the estimated tree is desired; and cleanup remove PHYLIP input & output files after the analysis is completed (defaults to cleanup = TRUE).

More information about the dnacomp program in PHYLIP can be found here http://evolution.genetics.washington.edu/phylip/doc/dnacomp.html.

Obviously, use of any of the functions of this package requires that PHYLIP (Felsenstein 1989, 2013) should first be installed. Instructions for installing PHYLIP can be found on the PHYLIP webpage: http://evolution.genetics.washington.edu/phylip.html.

#### Value

This function returns an object of class "phylo" that is the optimized tree.

# Author(s)

Liam J. Revell, Scott A. Chamberlain

Maintainer: Liam J. Revell < liam.revell@umb.edu>

Rdnadist 13

#### References

Felsenstein, J. (1989) PHYLIP-Phylogeny Inference Package (Version 3.2). Cladistics, 5, 164-166.

Felsenstein, J. (2013) PHYLIP (Phylogeny Inference Package) version 3.695. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Fitch, W.M. (1975) Toward finding the tree of maximum parsimony. pp. 189-230 in *Proceedings of the Eighth International Conference on Numerical Taxonomy*, ed. G. F. Estabrook. W. H. Freeman, San Francisco.

Le Quesne, W.J. (1969) A method of selection of characters in numerical taxonomy. *Systematic Zoology*, 18, 201-205.

# See Also

Rdnapars, Rdnapenny

## **Examples**

```
## Not run:
data(primates)
tree<-Rdnacomp(primates)
## End(Not run)</pre>
```

Rdnadist

R interfaces for dnadist

# **Description**

This function is an R interface for dnadist in the PHYLIP package (Felsenstein 2013). dnadist can be used to estimate the evolutionary distances between DNA sequences under various models.

# Usage

```
Rdnadist(X, method=c("F84","K80","JC","LogDet"), path=NULL, ...)
```

# **Arguments**

. . .

Χ	an object of class "DNAbin".
method	method for calculating the distances. Can be "F84" (Kishino & Hasegawa 1989; Felsenstein & Churchill 1996), "K80" (Kimura 1980), "JC" (Jukes & Cantor 1969), or "LogDet" (Barry & Hartigan 1987; Lake 1994; Steel 1994; Lockhart et. al. 1994). Also method="similarity" computes the sequence similarity among the rows of X.
path	path to the executable containing dnadist. If path = NULL, the R will search several commonly used directories for the correct executable file.

optional arguments to be passed to dnadist. See details for more information.

14 Rdnadist

#### **Details**

Optional arguments include the following: quiet suppress some output to R console (defaults to quiet = FALSE); gamma alpha shape parameter of a gamma model of rate heterogeneity among sites (defaults to no gamma rate heterogeneity); kappa transition:transversion ratio (defaults to kappa = 2.0); rates vector of rates (defaults to single rate); rate.categories vector of rate categories corresponding to the order of rates; weights vector of weights of length equal to the number of columns in X (defaults to unweighted); bf vector of base frequencies in alphabetical order (i.e., A, C, G, & T) - if not provided, then defaults to empirical frequencies; and cleanup remove PHYLIP input & output files after the analysis is completed (defaults to cleanup = TRUE).

More information about the dnadist program in PHYLIP can be found here http://evolution.genetics.washington.edu/phylip/doc/dnadist.html.

Obviously, use of any of the functions of this package requires that PHYLIP (Felsenstein 1989, 2013) should first be installed. Instructions for installing PHYLIP can be found on the PHYLIP webpage: http://evolution.genetics.washington.edu/phylip.html.

#### Value

This function returns an object of class "dist".

#### Author(s)

Liam J. Revell, Scott A. Chamberlain

Maintainer: Liam J. Revell < liam.revell@umb.edu>

## References

Barry, D., Hartigan, J.A. (1987) Statistical analysis of hominoid molecular evolution. *Statistical Science*, 2, 191-200.

Felsenstein, J. (2013) PHYLIP (Phylogeny Inference Package) version 3.695. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Felsenstein, J., Churchill, G. A. (1996) A Hidden Markov Model approach to variation among sites in rate of evolution. *Molecular Biology and Evolution*, 13, 93-104.

Jukes, T.H., Cantor, C.R. (1969) Evolution of protein molecules. pp. 21-132 in *Mammalian Protein Metabolism Vol. III*, ed. M.N. Munro. Academic Press, New York.

Kimura, M. (1980) A simple model for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, 16, 111-120.

Kishino, H., Hasegawa, M. (1989) Evaluation of the maximum likelihood estimate of teh evolutionary tree topology from DNA sequence data, and the branching order in Hominoidea. *Journal of Molecular Evolutioon*, 29, 170-179.

Lake, J.A. (1994) Reconstructing evolutionary trees from DNA and protein sequences: Paralinear distances. *Proceedings of the National Academy of Sciences*, 91, 1455-1459.

Lockhart, P.J., Steel, M.A., Hendy, M.D., Penny, D. (1994) Recovering evolutionary trees under a more realistic model of sequence evolution. *Molecular Biology and Evolution*, 11, 605-612.

Steel, M.A. (1994) Recovering a tree from the Markov leaf colourations it generates under a Markov model. *Applied Mathematics Letters*, 7, 19-23.

#### See Also

Rneighbor

Rdnainvar 15

#### **Examples**

```
## Not run:
data(primates)
D<-Rdnadist(primates,kappa=10)
tree<-Rneighbor(D)
## End(Not run)</pre>
```

Rdnainvar

R interface for dnainvar

# **Description**

This function is an R interface for dnainvar in the PHYLIP package (Felsenstein 2013). dnainvar reads DNA sequences and computest Lake's invariants (Cavender & Felsenstein 1987; Lake 1987).

## Usage

```
Rdnainvar(X, path=NULL, ...)
```

# **Arguments**

X an object of class "DNAbin". X cannot contain more than four species.

path path to the executable containing dnainvar. If path = NULL, the R will search

several commonly used directories for the correct executable file.

... optional arguments to be passed to dnainvar. See details for more information.

# **Details**

Optional arguments include the following: quiet suppress some output to R console (defaults to quiet = FALSE); weights vector of weights of length equal to the number of columns in X (defaults to unweighted); and cleanup remove PHYLIP input & output files after the analysis is completed (defaults to cleanup = TRUE).

More information about the dnainvar program in PHYLIP can be found here http://evolution.genetics.washington.edu/phylip/doc/dnainvar.html.

Obviously, use of any of the functions of this package requires that PHYLIP (Felsenstein 1989, 2013) should first be installed. Instructions for installing PHYLIP can be found on the PHYLIP webpage: http://evolution.genetics.washington.edu/phylip.html.

# Value

At present, this function merely prints the results of dnainvar to screen.

# Author(s)

Liam J. Revell, Scott A. Chamberlain

Maintainer: Liam J. Revell < liam.revell@umb.edu>

16 Rdnaml

#### References

Cavender, J.A., Felsenstein, J. (1987) Invariants of phylogenies in a simple case with discrete states. *Journal of Classification*, 4, 57-71.

Felsenstein, J. (1989) PHYLIP-Phylogeny Inference Package (Version 3.2). Cladistics, 5, 164-166.

Felsenstein, J. (2013) PHYLIP (Phylogeny Inference Package) version 3.695. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Lake, J.A. (1987) A rate-independent technique for analysis of nucleic acid sequences: evolutionary parsimony. *Molecular Biology and Evolution*, 4, 167-191.

## See Also

**Rdnapars** 

## **Examples**

```
## Not run:
data(primates)
primates<-primates[sample(nrow(primates),4),]
tree<-Rdnainvar(primates)
## End(Not run)</pre>
```

Rdnaml

R interfaces for dnaml and dnamlk

## **Description**

This function is an R interface for dnaml in the PHYLIP package (Felsenstein 2013). dnaml can be used for ML phylogeny estimation from DNA sequences (Felsenstein 1981; Felsenstein & Churchill 1996).

# Usage

```
Rdnaml(X, path=NULL, ...)
Rdnamlk(X, path=NULL, ...)
```

## **Arguments**

X an object of class "DNAbin".

path path to the executable containing dnaml. If path = NULL, the R will search

several commonly used directories for the correct executable file.

optional arguments to be passed to dnaml or dnamlk. See details for more infor-

mation.

Rdnaml 17

#### **Details**

Optional arguments include the following: quiet suppress some output to R console (defaults to quiet = FALSE); tree object of class "phylo" - if supplied, then the model will be optimized on a fixed input topology; kappa transition:transversion ratio (defaults to kappa = 2.0); bf vector of base frequencies in alphabetical order (i.e., A, C, G, & T) - if not provided, then defaults to empirical frequencies; rates vector of rates (defaults to single rate); rate.categories vector of rate categories corresponding to the order of rates; gamma alpha shape parameter of a gamma model of rate heterogeneity among sites (defaults to no gamma rate heterogeneity); ncat number of rate categories for the gamma model; inv proportion of invariant sites for the invariant sites model (defaults to inv = 0); weights vector of weights of length equal to the number of columns in X (defaults to unweighted); speedier speedier but rougher analysis (defaults to speedier = FALSE); global perform global search (defaults to global = TRUE); random.order add taxa to tree in random order (defaults to random.order = TRUE); random.addition number of random addition replicates for random.order = TRUE (defaults to random.addition = 10); outgroup outgroup if outgroup rooting of the estimated tree is desired; and cleanup remove PHYLIP input & output files after the analysis is completed (defaults to cleanup = TRUE).

Finally clock=TRUE enforces a molecular clock. The argument clock is only available for Rdnaml. If clock=TRUE then dnamlk is used internally. For Rdnamlk a molecular clock is assumed, thus Rdnaml(...,clock=TRUE) and Rdnamlk(...) are equivalent.

More information about the dnaml and dnamlk programs in PHYLIP can be found here http://evolution.genetics.washington.edu/phylip/doc/dnaml.html, and here http://evolution.genetics.washington.edu/phylip/doc/dnamlk.html.

Obviously, use of any of the functions of this package requires that PHYLIP (Felsenstein 1989, 2013) should first be installed. Instructions for installing PHYLIP can be found on the PHYLIP webpage: http://evolution.genetics.washington.edu/phylip.html.

#### Value

This function returns an object of class "phylo" that is the optimized tree.

## Author(s)

Liam J. Revell, Scott A. Chamberlain

Maintainer: Liam J. Revell < liam.revell@umb.edu>

#### References

Felsenstein, J. (1981) Evolutionary trees from DNA sequences: A Maximum Likelihood approach. *Journal of Molecular Evolution*, 17, 368-376.

Felsenstein, J. (1989) PHYLIP-Phylogeny Inference Package (Version 3.2). Cladistics, 5, 164-166.

Felsenstein, J. (2013) PHYLIP (Phylogeny Inference Package) version 3.695. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Felsenstein, J., Churchill, G. A. (1996) A Hidden Markov Model approach to variation among sites in rate of evolution. *Molecular Biology and Evolution*, 13, 93-104.

## See Also

opt.Rdnaml, Rcontml, Rproml

18 Rdnapars

#### **Examples**

```
## Not run:
data(primates)
tree<-Rdnaml(primates,kappa=10)
clockTree<-Rdnamlk(primates,kappa=10)
## End(Not run)</pre>
```

Rdnapars

R interface for dnapars

## **Description**

This function is an R interface for dnapars in the PHYLIP package (Felsenstein 2013). dnapars can be used for MP phylogeny estimation from DNA sequences (Eck & Dayhoff 1966; Kluge & Farris 1969; Fitch 1971).

# Usage

```
Rdnapars(X, path=NULL, ...)
```

## Arguments

X an object of class "DNAbin".

path path to the executable containing dnapars. If path = NULL, the R will search

several commonly used directories for the correct executable file.

... optional arguments to be passed to dnapars. See details for more information.

## **Details**

Optional arguments include the following: quiet suppress some output to R console (defaults to quiet = FALSE); tree object of class "phylo" - if supplied, then the parsimony score will be computed on a fixed input topology; thorough logical value indicating whether to conduct a more thorough search (defaults to thorough=TRUE); nsave number of trees to save if multiple equally parsimonious trees are found (defaults to nsave=10000); random.order add taxa to tree in random order (defaults to random.order = TRUE); random.addition number of random addition replicates for random.order = TRUE (defaults to random.addition = 10); threshold threshold value for threshold parsimony (defaults to ordinary parsimony); transversion logical value indicating whether to use transversion parsimony (defaults to transversion=FALSE); weights vector of weights of length equal to the number of columns in X (defaults to unweighted); outgroup outgroup if outgroup rooting of the estimated tree is desired; and cleanup remove PHYLIP input & output files after the analysis is completed (defaults to cleanup = TRUE).

More information about the dnapars program in PHYLIP can be found here http://evolution.genetics.washington.edu/phylip/doc/dnapars.html.

Obviously, use of any of the functions of this package requires that PHYLIP (Felsenstein 1989, 2013) should first be installed. Instructions for installing PHYLIP can be found on the PHYLIP webpage: http://evolution.genetics.washington.edu/phylip.html.

#### Value

This function returns an object of class "phylo" that is the optimized tree.

Rdnapenny 19

#### Author(s)

Liam J. Revell, Scott A. Chamberlain

Maintainer: Liam J. Revell < liam.revell@umb.edu>

#### References

Eck. R.V., Dayhoff, M.O. (1966) *Atlas of Protein Sequence and Structure 1966*. National Biomedical Research Foundation, Silver Spring, Maryland.

Felsenstein, J. (1989) PHYLIP-Phylogeny Inference Package (Version 3.2). Cladistics, 5, 164-166.

Felsenstein, J. (2013) PHYLIP (Phylogeny Inference Package) version 3.695. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Fitch, W.M. (1971) Toward defining the course of evolution: Minimu change for a specified tree topology. *Systematic Zoology*, 20, 406-416.

Kluge, A.G., Farris, J.S. (1969) Quantitative phyletics and the evolution of anurans. *Systematic Zoology*, 18, 1-32.

#### See Also

```
Rdnaml, Rdnapenny
```

## **Examples**

```
## Not run:
data(primates)
tree<-Rdnapars(primates)
## End(Not run)</pre>
```

Rdnapenny

R interface for dnapenny

# **Description**

This function is an R interface for dnapenny in the PHYLIP package (Felsenstein 2013). dnapenny performs branch & bound parsimony searching following Hendy & Penny (1982).

# Usage

```
Rdnapenny(X, path=NULL, ...)
```

# Arguments

X an object of class "DNAbin".

path path to the executable containing dnapenny. If path = NULL, the R will search

several commonly used directories for the correct executable file.

... optional arguments to be passed to dnapenny. See details for more information.

20 Rdnapenny

#### **Details**

Optional arguments include the following: quiet suppress some output to R console (defaults to quiet = FALSE); groups number of groups of 1,000 trees (defaults to groups = 10000); report reporting frequency, in numbers of trees (defaults to report = 1000); simple simple branch & bound (defaults to simple = TRUE); threshold threshold value for threshold parsimony (defaults to ordinary parsimony); weights vector of weights of length equal to the number of columns in X (defaults to unweighted); outgroup outgroup if outgroup rooting of the estimated tree is desired; and cleanup remove PHYLIP input & output files after the analysis is completed (defaults to cleanup = TRUE).

More information about the dnapenny program in PHYLIP can be found here http://evolution.genetics.washington.edu/phylip/doc/dnapenny.html.

Obviously, use of any of the functions of this package requires that PHYLIP (Felsenstein 1989, 2013) should first be installed. Instructions for installing PHYLIP can be found on the PHYLIP webpage: http://evolution.genetics.washington.edu/phylip.html.

#### Value

This function returns an object of class "phylo" or "multiPhylo" that is the tree or trees with the best parsimony score. tree\$score gives the parsimony score, for "phylo" object tree.

## Author(s)

Liam J. Revell, Scott A. Chamberlain

Maintainer: Liam J. Revell < liam.revell@umb.edu>

#### References

Felsenstein, J. (1989) PHYLIP-Phylogeny Inference Package (Version 3.2). Cladistics, 5, 164-166.

Felsenstein, J. (2013) PHYLIP (Phylogeny Inference Package) version 3.695. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Hendy, M.D., Penny, D. (1982) Branch and bound algorithms to determine minimal evolutionary trees. *Mathematical Biosciences*, 60, 133-142.

## See Also

Rdnapars

# **Examples**

```
## Not run:
data(primates)
tree<-Rdnapenny(primates)
## End(Not run)</pre>
```

Rdollop 21

Rdollop	R interface for dollop	

## **Description**

This function is an R interface for dollop in the PHYLIP package (Felsenstein 2013). dollop can be used for Dollo (Le Quesne 1974; Farris 1977) and polymorphism (Inger 1967; Farris 1978; Felsenstein 1979) parsimony tree inference.

## Usage

```
Rdollop(X, path=NULL, ...)
```

#### **Arguments**

X	an object of class "phylip.data" with a set of binary characters; or a matrix. Data should be "0" and "1". The states "P" (or "B"), for polymorphism, and "?" are also permitted.
path	path to the executable containing dollop. If path = NULL, the R will search several commonly used directories for the correct executable file.
	optional arguments to be passed to dollop. See details for more information.

#### **Details**

Optional arguments include the following: quiet suppress some output to R console (defaults to quiet = FALSE); tree object of class "phylo" - if supplied, then the Dollo or polymorphism parsimony score will be computed on a fixed input topology; method indicating whether Dollo (method="dollo", the default) or polymorphism (method="polymorphism") should be used; random.order add taxa to tree in random order (defaults to random.order = TRUE); random.addition number of random addition replicates for random.order = TRUE (defaults to random.addition = 10); threshold threshold value for threshold parsimony (defaults to ordinary parsimony); ancestral vector of ancestral states, which allows for the unordered Dollo parsimony method of Felsenstein (1984); weights vector of weights of length equal to the number of columns in X (defaults to unweighted); and cleanup remove PHYLIP input & output files after the analysis is completed (defaults to cleanup = TRUE).

More information about the dollop program in PHYLIP can be found here http://evolution.genetics.washington.edu/phylip/doc/dollop.html.

Obviously, use of any of the functions of this package requires that PHYLIP (Felsenstein 1989, 2013) should first be installed. Instructions for installing PHYLIP can be found on the PHYLIP webpage: http://evolution.genetics.washington.edu/phylip.html.

## Value

This function returns an object of class "phylo" that is the optimized tree.

## Author(s)

Liam J. Revell, Scott A. Chamberlain

Maintainer: Liam J. Revell < liam.revell@umb.edu>

22 Rdolpenny

#### References

Farris, J.S. (1977) Phylogenetic analysis under Dollo's Law. Systematic Zoology, 26, 77-88.

Farris, J.S. (1978) Inferring phylogenetic trees from chromosome inversion data. *Systematic Zoology*, 27, 275-284.

Felsenstein, J. (1979) Alternative methods of phylogenetic inference and their interrelationship. *Systematic Zoology*, 28, 49-62.

Felsenstein, J. (1984) The statistical approach to inferring evolutionary trees and what it tells us about parsimony and compatibility. pp. 169-191 in: *Cladistics: Perspectives in the Reconstruction of Evolutionary History*, edited by T. Duncan and T. F. Stuessy. Columbia University Press, New York.

Felsenstein, J. (1989) PHYLIP-Phylogeny Inference Package (Version 3.2). Cladistics, 5, 164-166.

Felsenstein, J. (2013) PHYLIP (Phylogeny Inference Package) version 3.695. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Inger, R.F. (1967) The development of a phylogeny of frogs. Evolution, 21, 369-384.

Le Quesne, W.J. (1974) The uniquely evolved character concept and its cladistic application. *Systematic Zoology*, 23, 513-517.

## See Also

Rpars, Rdnapars

## **Examples**

```
## Not run:
data(primates.bin)
tree<-Rdollop(primates.bin)
## End(Not run)</pre>
```

Rdolpenny

R interface for dolpenny

## **Description**

This function is an R interface for dolpenny in the PHYLIP package (Felsenstein 2013). dolpenny performs Dollo (Le Quesne 1974; Farris 1977) or polymorphism (Inger 1967; Farris 1978; Felsenstein 1979) parsimony using the branch and bound algorithm of Hendy & Penny (1982).

# Usage

```
Rdolpenny(X, path=NULL, ...)
```

# Arguments

X	an object of class "phylip.data" with a binary character; or a matrix. Data should be "0" and "1". The states "P" (or "B"), for polymorphism, and "?" are also permitted.
path	path to the executable containing dolpenny. If path $=$ NULL, the R will search several commonly used directories for the correct executable file.
• • •	optional arguments to be passed to dolpenny. See details for more information.

Rdolpenny 23

#### **Details**

Optional arguments include the following: quiet suppress some output to R console (defaults to quiet = FALSE); method indicating whether Dollo (method="dollo", the default) or polymorphism (method="polymorphism") should be used; groups number of groups of 1,000 trees (defaults to groups = 100); report reporting frequency, in numbers of trees (defaults to report = 1000); simple simple branch & bound (defaults to simple = TRUE); threshold threshold value for threshold parsimony (defaults to ordinary parsimony); ancestors vector of ancestral states, which allows for the unordered Dollo parsimony method of Felsenstein (1984); weights vector of weights of length equal to the number of columns in X (defaults to unweighted); outgroup outgroup if outgroup rooting of the estimated tree is desired; and cleanup remove PHYLIP input & output files after the analysis is completed (defaults to cleanup = TRUE).

More information about the dolpenny program in PHYLIP can be found here http://evolution.genetics.washington.edu/phylip/doc/dolpenny.html.

Obviously, use of any of the functions of this package requires that PHYLIP (Felsenstein 1989, 2013) should first be installed. Instructions for installing PHYLIP can be found on the PHYLIP webpage: http://evolution.genetics.washington.edu/phylip.html.

#### Value

This function returns an object of class "phylo" or "multiPhylo" that is the tree or trees with the best parsimony score. tree\$score gives the parsimony score, for "phylo" object tree.

## Author(s)

Liam J. Revell, Scott A. Chamberlain

Maintainer: Liam J. Revell < liam.revell@umb.edu>

#### References

Farris, J.S. (1977) Phylogenetic analysis under Dollo's Law. Systematic Zoology, 26, 77-88.

Farris, J.S. (1978) Inferring phylogenetic trees from chromosome inversion data. *Systematic Zoology*, 27, 275-284.

Felsenstein, J. (1979) Alternative methods of phylogenetic inference and their interrelationship. *Systematic Zoology*, 28, 49-62.

Felsenstein, J. (1984) The statistical approach to inferring evolutionary trees and what it tells us about parsimony and compatibility. pp. 169-191 in: *Cladistics: Perspectives in the Reconstruction of Evolutionary History*, edited by T. Duncan and T. F. Stuessy. Columbia University Press, New York.

Felsenstein, J. (1989) PHYLIP-Phylogeny Inference Package (Version 3.2). Cladistics, 5, 164-166.

Felsenstein, J. (2013) PHYLIP (Phylogeny Inference Package) version 3.695. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Hendy, M.D., Penny, D. (1982) Branch and bound algorithms to determine minimal evolutionary trees. *Mathematical Biosciences*, 60, 133-142.

Inger, R.F. (1967) The development of a phylogeny of frogs. Evolution, 21, 369-384.

Le Quesne, W.J. (1974) The uniquely evolved character concept and its cladistic application. *Systematic Zoology*, 23, 513-517.

## See Also

**Rdnapars** 

24 read.protein

# **Examples**

```
## Not run:
data(primates.bin)
tree<-Rdolpenny(primates.bin)
## End(Not run)</pre>
```

read.protein

Reads protein sequences from file in multiple formats

# Description

Reads protein sequences from a file.

# Usage

```
read.protein(file, format="fasta", ...)
```

# Arguments

file file name for file containing protein sequences.

format of input file. Permitted formats are "fasta" and "sequential". See

read. dna for more information.

... optional arguments.

# Value

An object of class "proseq" containing protein sequences.

# Author(s)

```
Liam J. Revell, Scott A. Chamberlain
```

Maintainer: Liam J. Revell < liam.revell@umb.edu>

## See Also

```
as.proseq, print.proseq, Rproml
```

Rfitch 25

Rfitch R interface for fitch

## **Description**

This function is an R interface for fitch in the PHYLIP package (Felsenstein 2013). fitch can be used for tree inference using the Fitch-Margoliash method (Fitch & Margoliash 1967), the Cavalli-Sforza & Edwards leasts-squares method (Cavalli-Sforza & Edwards 1967), and the minimum evolution criterion (Kidd & Sgaramella-Zonta 1971; Nei & Rzhetsky 1993).

# Usage

```
Rfitch(D, path=NULL , ...)
```

## **Arguments**

D	a distance matrix as an object of class "matrix" or "dist". If a matrix, then D should be symmetrical with a diagonal of zeros.
path	path to the executable containing fitch. If path = NULL, the R will search several commonly used directories for the correct executable file.
	optional arguments to be passed to fitch. See details for more information.

#### **Details**

Optional arguments include the following: quiet suppress some output to R console (defaults to quiet = FALSE); method - can be "FM" or "fm" (for the Fitch-Margoliash criterion; Fitch & Margoliash 1967), "LS" or "1s" for the Cavalli-Sforza & Edwards (1967) least-squares criterion, or "ME" or "me" (for Minimum Evolution); tree object of class "phylo" - if supplied, then branch lengths are optimized & the optimality criterion is computed on the input tree; negative a logical value indicating whether negative branch lengths should be permitted (defaults to negative = TRUE; global perform global search (defaults to global = TRUE); random.order add taxa to tree in random order (defaults to random.order = TRUE); random.addition number of random addition replicates for random.order = TRUE (defaults to random.addition = 10); outgroup outgroup if outgroup rooting of the estimated tree is desired; and cleanup remove PHYLIP input & output files after the analysis is completed (defaults to cleanup = TRUE).

More information about the fitch program in PHYLIP can be found here http://evolution.genetics.washington.edu/phylip/doc/fitch.html.

Obviously, use of any of the functions of this package requires that PHYLIP (Felsenstein 1989, 2013) should first be installed. More information about installing PHYLIP can be found on the PHYLIP webpage: http://evolution.genetics.washington.edu/phylip.html.

#### Value

This function returns an object of class "phylo" that is the Fitch-Margoliash, LS, or ME tree.

## Author(s)

Liam J. Revell, Scott A. Chamberlain

Maintainer: Liam J. Revell < liam.revell@umb.edu>

26 Rgendist

#### References

Cavalli-Sforza, L.L., Edwards, A.W.F. (1967) Phylogenetic analysis: models and estimation procedures. *Evolution*, 32, 550-570.

Fitch, W.M., Margoliash, E. (1967) Construction of phylogenetic trees. Science, 155, 279-284.

Kidd, K.K., Sgaramella-Zonta, L.A. (1971) Phylogenetic analysis: concepts and methods. *American Journal of Human Genetics*, 23, 235-252.

Felsenstein, J. (1989) PHYLIP-Phylogeny Inference Package (Version 3.2). Cladistics, 5, 164-166.

Felsenstein, J. (2013) PHYLIP (Phylogeny Inference Package) version 3.695. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

#### See Also

Rdnadist

## **Examples**

```
## Not run:
data(primates)
D<-dist.dna(data(primates),model="JC")
tree<-Rfitch(D)
## End(Not run)</pre>
```

Rgendist

R interface for gendist

# Description

This function is an R interface for gendist in the PHYLIP package (Felsenstein 2013). gendist can be used to estimate the evolutionary distances between populations based on gene frequency data. Three methods can be used for computing these distances. These are (1) Nei's genetic distance (Nei 1972), (2) Cavalli-Sforza's chord measure (Cavalli-Sforza & Edwards 1967), and (3) Reynolds et al.'s (1983) genetic distance.

## Usage

```
Rgendist(X, path=NULL, ...)
```

## **Arguments**

Χ

either: (a) A *matrix* of allele frequencies (in columns) with rownames containing species names. In this case it is assumed that m - 1 allele frequencies are reported for a locus with m segregating alleles. (The mth frequency is just one minus the sum of the other m-1.) Or (b) a list of matrices in which each row contains the relative frequency of alleles at a locus for a species. In this case, the rownames of each matrix in the list should contain the species names. Here the number of alleles at each locus can be inferred from the number of columns in the matrix, and no frequencies should be omitted. (I.e., rows should add to 1.0.) path to the executable containing gendist. If path = NULL, the R will search several paths for gendist. See details for more information.

path

optional arguments to be passed to gendist. See details for more information.

Rkitsch 27

#### **Details**

Optional arguments include the following: quiet suppress some output to R console (defaults to quiet = FALSE); method, the method for calculating the genetic distance - options are "Nei", "Cavalli-Sforza", and "Reynolds" (defaults to method="Nei"); nalleles a vector containing the number of alleles per locus in X (not used if X is a list, and assumed to be a vector of 2s if ommitted); and cleanup remove PHYLIP input & output files after the analysis is completed (defaults to cleanup = TRUE).

More information about the gendist program in PHYLIP can be found here http://evolution.genetics.washington.edu/phylip/doc/gendist.html.

Obviously, use of any of the functions of this package requires that PHYLIP (Felsenstein 1989, 2013) should first be installed. Instructions for installing PHYLIP can be found on the PHYLIP webpage: http://evolution.genetics.washington.edu/phylip.html.

## Value

This function returns an object of class "dist".

#### Author(s)

Liam J. Revell, Scott A. Chamberlain

Maintainer: Liam J. Revell < liam.revell@umb.edu>

#### References

Cavalli-Sforza, L.L., Edwards, A.W.F. (1967) Phylogenetic analysis: models and estimation procedures. *Evolution*, 32, 550-570.

Felsenstein, J. (1989) PHYLIP-Phylogeny Inference Package (Version 3.2). Cladistics, 5, 164-166.

Felsenstein, J. (2013) PHYLIP (Phylogeny Inference Package) version 3.695. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Nei, M. (1972) Genetic distance between populations. American Naturalist, 106, 283-292.

Reynolds, J.B., Weir, B.S., Cockerham, C.C. (1983) Estimation of the coancestry coefficient: basis for a short-term genetic distance. *Genetics*, 105, 767-779.

# See Also

Rdnadist, Rprotdist

Rkitsch

R interface for kitsch

# Description

This function is an R interface for kitsch in the PHYLIP package (Felsenstein 2013). kitsch can be used for tree inference using the Fitch-Margoliash method (Fitch & Margoliash 1967), the Cavalli-Sforza & Edwards leasts-squares method (Cavalli-Sforza & Edwards 1967), and the minimum evolution criterion (Kidd & Sgaramella-Zonta 1971; Nei & Rzhetsky 1993), but contraining the branch lengths of the tree to be clock-like (i.e., all paths from the root to any tip are equal in length).

28 Rkitsch

## Usage

```
Rkitsch(D, path=NULL , ...)
```

#### **Arguments**

D a distance matrix as an object of class "matrix", "data.frame", or "dist". If

a matrix, then D should be symmetrical with a diagonal of zeros.

path path to the executable containing kitsch. If path = NULL, the R will search

several commonly used directories for the correct executable file.

... optional arguments to be passed to kitsch. See details for more information.

#### **Details**

Optional arguments include the following: quiet suppress some output to R console (defaults to quiet = FALSE); method - can be "FM" or "fm" (for the Fitch-Margoliash criterion; Fitch & Margoliash 1967), "LS" or "1s" for the Cavalli-Sforza & Edwards (1967) least-squares criterion, or "ME" or "me" (for Minimum Evolution); tree object of class "phylo" - if supplied, then branch lengths are optimized & the optimality criterion is computed on the input tree; negative a logical value indicating whether negative branch lengths should be permitted (defaults to negative = TRUE; random.order add taxa to tree in random order (defaults to random.order = TRUE); random.addition number of random addition replicates for random.order = TRUE (defaults to random.addition = 10); outgroup outgroup if outgroup rooting of the estimated tree is desired; subreplicate, a logical value indicating whether or not to perform the subreplication test of Felsenstein (1986); and cleanup remove PHYLIP input & output files after the analysis is completed (defaults to cleanup = TRUE). Note that for some reason subreplicate=TRUE does not appear to work in the current version of kitsch.

More information about the kitsch program in PHYLIP can be found here http://evolution.genetics.washington.edu/phylip/doc/kitsch.html.

Obviously, use of any of the functions of this package requires that PHYLIP (Felsenstein 1989, 2013) should first be installed. More information about installing PHYLIP can be found on the PHYLIP webpage: http://evolution.genetics.washington.edu/phylip.html.

#### Value

This function returns an object of class "phylo" that is the NJ or UPGMA tree.

#### Author(s)

Liam J. Revell, Scott A. Chamberlain

Maintainer: Liam J. Revell < liam.revell@umb.edu>

#### References

Cavalli-Sforza, L.L., Edwards, A.W.F. (1967) Phylogenetic analysis: models and estimation procedures. *Evolution*, 32, 550-570.

Felsenstein, J. (1986) Distance methods: a reply to Farris. Cladistics, 2, 130-144.

Fitch, W.M., Margoliash, E. (1967) Construction of phylogenetic trees. Science, 155, 279-284.

Kidd, K.K., Sgaramella-Zonta, L.A. (1971) Phylogenetic analysis: concepts and methods. *American Journal of Human Genetics*, 23, 235-252.

Felsenstein, J. (1989) PHYLIP-Phylogeny Inference Package (Version 3.2). Cladistics, 5, 164-166.

Rmix 29

Felsenstein, J. (2013) PHYLIP (Phylogeny Inference Package) version 3.695. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

#### See Also

Rfitch

# **Examples**

```
## Not run:
data(primates)
D<-dist.dna(data(primates),model="JC")
tree<-Rkitsch(D)
## End(Not run)</pre>
```

Rmix

R interface for mix

## **Description**

This function is an R interface for mix in the PHYLIP package (Felsenstein 2013). mix performs branch & bound parsimony searching following Hendy & Penny (1982).

#### Usage

```
Rmix(X, path=NULL, ...)
```

## **Arguments**

x an object of class "phylip.data" with a binary character; or a matrix. Data should be "0", "1", and "?".

path path to the executable containing mix. If path = NULL, the R will search several commonly used directories for the correct executable file.

... optional arguments to be passed to mix. See details for more information.

# **Details**

Optional arguments include the following: quiet suppress some output to R console (defaults to quiet = FALSE); tree object of class "phylo" - if supplied, then the parsimony score will be computed on a fixed input topology; mixture, a vector containing "W" and "C" in upper or lower case indicating which characters should be evaluated using Wagner (Eck & Dayhoff 1966; Kluge & Farris 1969) or Camin-Sokal (Camin & Sokal 1965) parsimony methods, respectively - this argument takes precedence over method (see below); method either "Wagner" or "Camin-Sokal", in upper or lower case, indicating whether Wagner (Eck & Dayhoff 1966; Kluge & Farris 1969) or Camin-Sokal (Camin & Sokal 1965) parsimony should be used; random.order add taxa to tree in random order (defaults to random.order = TRUE); random.addition number of random addition replicates for random.order = TRUE (defaults to random.addition = 10); threshold threshold value for threshold parsimony (defaults to ordinary parsimony); ancestral a vector of length equal to the number of columns in X containing the the ancestral state at the root of the tree for each character, these should be "0", "1", and "?"; weights vector of weights of length equal to the number of columns in X (defaults to unweighted); outgroup outgroup if outgroup rooting of the

30 Rmix

estimated tree is desired; and cleanup remove PHYLIP input & output files after the analysis is completed (defaults to cleanup = TRUE).

More information about the mix program in PHYLIP can be found here http://evolution.genetics.washington.edu/phylip/doc/mix.html.

Obviously, use of any of the functions of this package requires that PHYLIP (Felsenstein 1989, 2013) should first be installed. Instructions for installing PHYLIP can be found on the PHYLIP webpage: http://evolution.genetics.washington.edu/phylip.html.

## Value

This function returns an object of class "phylo" or "multiPhylo" that is the tree or trees with the best parsimony score. tree\$score gives the parsimony score, for "phylo" object tree.

# Author(s)

Liam J. Revell, Scott A. Chamberlain

Maintainer: Liam J. Revell < liam.revell@umb.edu>

#### References

Camin, J.H., Sokal, R.R. (1965) A method for deducing branching sequences in phylogeny. *Evolution*, 19, 311-326.

Eck, R.V., Dayhoff, M.O. (1966) *Atlas of Protein Sequence and Structure 1966*. National Biomedical Research Foundation, Silver Spring, Maryland.

Felsenstein, J. (1989) PHYLIP-Phylogeny Inference Package (Version 3.2). Cladistics, 5, 164-166.

Felsenstein, J. (2013) PHYLIP (Phylogeny Inference Package) version 3.695. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Kluge, A.G., Farris, J.S. (1969) Quantitative phyletics and the evolution of anurans. *Systematic Zoology*, 18, 1-32.

Hendy, M.D., Penny, D. (1982) Branch and bound algorithms to determine minimal evolutionary trees. *Mathematical Biosciences*, 60, 133-142.

# See Also

Rdnapars

## **Examples**

```
## Not run:
data(primates.bin)
tree<-Rmix(primates.bin)
## End(Not run)</pre>
```

Rneighbor 31

Rneighbor	R interface for neighbor	

# **Description**

This function is an R interface for neighbor in the PHYLIP package (Felsenstein 2013). neighbor can be used for neighbor-joining (Saitou & Nei 1987) and UPGMA (Sokal & Michener 1958) phylogeny inference.

# Usage

```
Rneighbor(D, path=NULL , ...)
```

## **Arguments**

D	a distance matrix as an object of class "matrix" or "dist". If a matrix, then D should be symmetrical with a diagonal of zeros.
path	path to the executable containing neighbor. If path $=$ NULL, the R will search several commonly used directories for the correct executable file.
	optional arguments to be passed to neighbor. See details for more information.

#### **Details**

Optional arguments include the following: quiet suppress some output to R console (defaults to quiet = FALSE); method - can be "NJ" or "nj" (for neighbor-joining), or "UPGMA" or "UPGMA" (for UPGMA); random.order add taxa to tree in random order (defaults to random.order = TRUE); outgroup outgroup if outgroup rooting of the estimated tree is desired (only works for method = "NJ", UPGMA trees are already rooted); and cleanup remove PHYLIP input & output files after the analysis is completed (defaults to cleanup = TRUE).

More information about the neighbor program in PHYLIP can be found here http://evolution.genetics.washington.edu/phylip/doc/neighbor.html.

Obviously, use of any of the functions of this package requires that PHYLIP (Felsenstein 1989, 2013) should first be installed. More information about installing PHYLIP can be found on the PHYLIP webpage: http://evolution.genetics.washington.edu/phylip.html.

## Value

This function returns an object of class "phylo" that is the NJ or UPGMA tree.

## Author(s)

Liam J. Revell, Scott A. Chamberlain

Maintainer: Liam J. Revell < liam.revell@umb.edu>

32 **Rpars** 

#### References

Saitou, N., Nei, M. (1987) The neighbor-joining method: A new method for reconstructing phylogenetic trees. Molecular Biology and Evolution, 4, 406-425.

Sokal, R., Michener, C. (1958) A statistical method for evaluating systematic relationships. University of Kansas Science Bulletin, 38, 1409-1438.

Felsenstein, J. (1989) PHYLIP-Phylogeny Inference Package (Version 3.2). Cladistics, 5, 164-166.

Felsenstein, J. (2013) PHYLIP (Phylogeny Inference Package) version 3.695. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

#### See Also

Rdnadist

## **Examples**

```
## Not run:
data(primates)
D<-dist.dna(data(primates), model="JC")
tree<-Rneighbor(D)</pre>
## End(Not run)
```

**Rpars** 

R interface for pars

# **Description**

This function is an R interface for pars in the PHYLIP package (Felsenstein 2013). pars can be used for MP phylogeny estimation from DNA sequences (Eck & Dayhoff 1966; Kluge & Farris 1969; Fitch 1971).

## Usage

```
Rpars(X, path=NULL, ...)
```

# **Arguments**

X	an object of class "phylip.data" with a set of binary or multistate characters; or a matrix.
path	path to the executable containing pars. If path = NULL, the R will search several commonly used directories for the correct executable file.
	optional arguments to be passed to pars. See details for more information.

Rpars 33

#### **Details**

Optional arguments include the following: quiet suppress some output to R console (defaults to quiet = FALSE); tree object of class "phylo" - if supplied, then the parsimony score will be computed on a fixed input topology; thorough logical value indicating whether to conduct a more thorough search (defaults to thorough=TRUE); nsave number of trees to save if multiple equally parsimonious trees are found (defaults to nsave=10000); random.order add taxa to tree in random order (defaults to random.order = TRUE); random.addition number of random addition replicates for random.order = TRUE (defaults to random.addition = 10); threshold threshold value for threshold parsimony (defaults to ordinary parsimony); weights vector of weights of length equal to the number of columns in X (defaults to unweighted); outgroup outgroup if outgroup rooting of the estimated tree is desired; and cleanup remove PHYLIP input & output files after the analysis is completed (defaults to cleanup = TRUE).

More information about the pars program in PHYLIP can be found here http://evolution.genetics.washington.edu/phylip/doc/pars.html.

Obviously, use of any of the functions of this package requires that PHYLIP (Felsenstein 1989, 2013) should first be installed. Instructions for installing PHYLIP can be found on the PHYLIP webpage: http://evolution.genetics.washington.edu/phylip.html.

#### Value

This function returns an object of class "phylo" that is the optimized tree.

## Author(s)

Liam J. Revell, Scott A. Chamberlain

Maintainer: Liam J. Revell < liam.revell@umb.edu>

# References

Eck. R.V., Dayhoff, M.O. (1966) *Atlas of Protein Sequence and Structure 1966*. National Biomedical Research Foundation, Silver Spring, Maryland.

Felsenstein, J. (1989) PHYLIP-Phylogeny Inference Package (Version 3.2). Cladistics, 5, 164-166.

Felsenstein, J. (2013) PHYLIP (Phylogeny Inference Package) version 3.695. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Fitch, W.M. (1971) Toward defining the course of evolution: Minimu change for a specified tree topology. *Systematic Zoology*, 20, 406-416.

Kluge, A.G., Farris, J.S. (1969) Quantitative phyletics and the evolution of anurans. *Systematic Zoology*, 18, 1-32.

#### See Also

Rdnaml, Rdnapenny

#### **Examples**

```
## Not run:
data(primates.bin)
tree<-Rpars(primates.bin)
## End(Not run)</pre>
```

34 Rpenny

Rpenny	R interface for penny	

#### **Description**

This function is an R interface for penny in the PHYLIP package (Felsenstein 2013). penny performs branch & bound parsimony searching following Hendy & Penny (1982).

## Usage

```
Rpenny(X, path=NULL, ...)
```

#### Arguments

X	an object of class "phylip.data" with a binary character; or a matrix. Data should be "0", "1", and "?".
path	path to the executable containing penny. If path = NULL, the R will search several commonly used directories for the correct executable file.
	optional arguments to be passed to penny. See details for more information.

#### **Details**

Optional arguments include the following: quiet suppress some output to R console (defaults to quiet = FALSE); mixture, a vector containing "W" and "C" in upper or lower case indicating which characters should be evaluated using Wagner (Eck & Dayhoff 1966; Kluge & Farris 1969) or Camin-Sokal (Camin & Sokal 1965) parsimony methods, respectively - this argument takes precedence over method (see below); method either "Wagner" or "Camin-Sokal", in upper or lower case, indicating whether Wagner (Eck & Dayhoff 1966; Kluge & Farris 1969) or Camin-Sokal (Camin & Sokal 1965) parsimony should be used; groups number of groups of 1,000 trees (defaults to groups = 100); report reporting frequency, in numbers of trees (defaults to report = 1000); simple simple branch & bound (defaults to simple = TRUE); threshold threshold value for threshold parsimony (defaults to ordinary parsimony); ancestral a vector of length equal to the number of columns in X containing the the ancestral state at the root of the tree for each character, these should be "0", "1", and "?"; weights vector of weights of length equal to the number of columns in X (defaults to unweighted); outgroup outgroup if outgroup rooting of the estimated tree is desired; and cleanup remove PHYLIP input & output files after the analysis is completed (defaults to cleanup = TRUE).

More information about the penny program in PHYLIP can be found here http://evolution.genetics.washington.edu/phylip/doc/penny.html.

Obviously, use of any of the functions of this package requires that PHYLIP (Felsenstein 1989, 2013) should first be installed. Instructions for installing PHYLIP can be found on the PHYLIP webpage: http://evolution.genetics.washington.edu/phylip.html.

#### Value

This function returns an object of class "phylo" or "multiPhylo" that is the tree or trees with the best parsimony score. tree\$score gives the parsimony score, for "phylo" object tree.

Rproml 35

#### Author(s)

Liam J. Revell, Scott A. Chamberlain

Maintainer: Liam J. Revell < liam.revell@umb.edu>

#### References

Camin, J.H., Sokal, R.R. (1965) A method for deducing branching sequences in phylogeny. *Evolution*, 19, 311-326.

Eck, R.V., Dayhoff, M.O. (1966) *Atlas of Protein Sequence and Structure 1966*. National Biomedical Research Foundation, Silver Spring, Maryland.

Felsenstein, J. (1989) PHYLIP-Phylogeny Inference Package (Version 3.2). Cladistics, 5, 164-166.

Felsenstein, J. (2013) PHYLIP (Phylogeny Inference Package) version 3.695. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Kluge, A.G., Farris, J.S. (1969) Quantitative phyletics and the evolution of anurans. *Systematic Zoology*, 18, 1-32.

Hendy, M.D., Penny, D. (1982) Branch and bound algorithms to determine minimal evolutionary trees. *Mathematical Biosciences*, 60, 133-142.

## See Also

Rdnapars

## **Examples**

```
## Not run:
data(primates.bin)
tree<-Rpenny(primates.bin)
## End(Not run)</pre>
```

Rproml

R interfaces for proml and promlk

## **Description**

This function is an R interface for proml in the PHYLIP package (Felsenstein 1989, 2013). proml can be used for ML phylogeny estimation from amino acid sequences.

# Usage

```
Rproml(X, path=NULL, ...)
Rpromlk(X, path=NULL, ...)
```

# **Arguments**

X an object of class "proseq".

path path to the executable containing proml. If path = NULL, the R will search

several commonly used directories for the correct executable file.

... optional arguments to be passed to proml or promlk. See details for more infor-

mation.

36 Rproml

#### **Details**

Optional arguments include the following: quiet suppress some output to R console (defaults to quiet = FALSE); tree object of class "phylo" - if supplied, then the model will be optimized on a fixed input topology; model amino acid model - could be "JTT" (Jones et al. 1992), "PMB" (Veerassamy et al. 2003), or "PAM" (Dayhoff & Eck 1968; Dayhoff et al. 1979; Koisol & Goldman 2005); rates vector of rates (defaults to single rate); rate.categories vector of rate categories corresponding to the order of rates; gamma alpha shape parameter of a gamma model of rate heterogeneity among sites (defaults to no gamma rate heterogeneity); ncat number of rate categories for the gamma model; inv proportion of invariant sites for the invariant sites model (defaults to inv = 0); weights vector of weights of length equal to the number of columns in X (defaults to unweighted); speedier speedier but rougher analysis (defaults to speedier = FALSE); global perform global search (defaults to global = TRUE); random.order add taxa to tree in random order (defaults to random.order = TRUE); random.addition number of random addition replicates for random.order = TRUE (defaults to random.addition = 10); outgroup outgroup if outgroup rooting of the estimated tree is desired; and cleanup remove PHYLIP input & output files after the analysis is completed (defaults to cleanup = TRUE).

Finally clock=TRUE enforces a molecular clock. The argument clock is only available for Rproml. If clock=TRUE then promlk is used internally. For Rpromlk a molecular clock is assumed, thus Rproml(...,clock=TRUE) and Rpromlk(...) are equivalent. Note that in PHYLIP 3.695 my tests of promlk yielded peculiar results (all branch lengths zero length, random topology), so I'm not sure what to make of that.

More information about the proml and promlk programs in PHYLIP can be found here http://evolution.genetics.washington.edu/phylip/doc/proml.html, and here http://evolution.genetics.washington.edu/phylip/doc/promlk.html.

Obviously, use of any of the functions of this package requires that PHYLIP (Felsenstein 1989, 2013) should first be installed. Instructions for installing PHYLIP can be found on the PHYLIP webpage: http://evolution.genetics.washington.edu/phylip.html.

# Value

This function returns an object of class "phylo" that is the optimized tree.

# Author(s)

Liam J. Revell, Scott A. Chamberlain

Maintainer: Liam J. Revell < liam.revell@umb.edu>

## References

Dayhoff, M.O., Eck, R.V. (1968) *Atlas of Protein Sequence and Structure 1967-1968*. National Biomedical Research Foundation, Silver Spring, Maryland.

Dayhoff, M.O., Schwartz, R.M., Orcutt, B.C. (1979) A model of evolutionary change in proteins. pp. 345-352 in *Atlas of Protein Sequence and Structure, Volume 5, Supplement 3, 1978*, ed. M.O. Dayhoff. National Biomedical Research Foundataion, Silver Spring, Maryland.

Felsenstein, J. (1989) PHYLIP-Phylogeny Inference Package (Version 3.2). Cladistics, 5, 164-166.

Felsenstein, J. (2013) PHYLIP (Phylogeny Inference Package) version 3.695. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Jones, D.T., Taylor, W.R., Thornton, J.M. (1992) The rapid generation of mutation data matrices from protein sequences. *Computer Applications in the Biosciences (CABIOS)*, 8, 275-282.

Rprotdist 37

Koisol, C., Goldman, N. (2005) Different versions of the Dayhoff rate matrix. *Molecular Biology and Evolution*, 22, 193-199.

Veerassamy, S., Smith, A., Tillier, E.R. (2003) A transition probability model for amino acid substitutions from blocks. *Journal of Computational Biology*, 10, 997-1010.

#### See Also

```
as.proseq, Rdnaml, read.protein
```

#### **Examples**

```
## Not run:
data(chloroplast)
tree<-Rproml(chloroplast)
## End(Not run)</pre>
```

Rprotdist

R interfaces for protdist

#### **Description**

This function is an R interface for protdist in the PHYLIP package (Felsenstein 2013). protdist can be used to estimate the evolutionary distances between amino acid sequences under various models.

#### Usage

```
Rprotdist(X, path=NULL, ...)
```

#### **Arguments**

path

X an object of class "proseq" containing aligned amino acid sequences.

path to the executable containing protdist. If path = NULL, the R will search

sev arguments to be passed to protdist. See details for more information.

... optional arguments to be passed to protdist. See details for more information.

## **Details**

Optional arguments include the following: quiet suppress some output to R console (defaults to quiet = FALSE); model can be "JTT" (Jones et al. 1992), "PMB" (Veerassamy et al. 2003), "PAM" (Dayhoff & Eck 1968; Dayhoff et al. 1979; Koisol & Goldman 2005), "Kimura" (a simple model based on Kimura 1980), "similarity" which gives the similarity between sequences, and "categories" which is due to Felsenstein; gamma alpha shape parameter of a gamma model of rate heterogeneity among sites (defaults to no gamma rate heterogeneity) - note that gamma rate heterogeneity does not apply to model = "Kimura" or model = "similarity"; kappa transition:transversion ratio (defaults to kappa = 2.0), genetic.code, type of genetic code to assume (options are "universal", the default, "mitochondrial", "vertebrate.mitochondrial", "fly.mitochondrial", and "yeast.mitochondrial"), categorization, categorization scheme for amino acids (options are "GHB", the George et al. 1988 classification, "Hall", a classification scheme provided by Ben Hall, and "chemical", a scheme based on Conn & Stumpf 1963); and, finally, ease, a numerical parameter that indicates the facility of getting between amino

38 Rprotdist

acids of different categories in which 0 is nearly impossible, and 1 is no difficulty (defaults to ease = 0.457) - note that kappa, bf, genetic.code, categorization, and ease apply only to model = "categories"; rates vector of rates (defaults to single rate); rate.categories vector of rate categories corresponding to the order of rates; weights vector of weights of length equal to the number of columns in X (defaults to unweighted); and cleanup remove PHYLIP input & output files after the analysis is completed (defaults to cleanup = TRUE).

More information about the protdist program in PHYLIP can be found here http://evolution.genetics.washington.edu/phylip/doc/protdist.html.

Obviously, use of any of the functions of this package requires that PHYLIP (Felsenstein 1989, 2013) should first be installed. Instructions for installing PHYLIP can be found on the PHYLIP webpage: http://evolution.genetics.washington.edu/phylip.html.

#### Value

This function returns an object of class "dist".

#### Author(s)

Liam J. Revell, Scott A. Chamberlain

Maintainer: Liam J. Revell < liam.revell@umb.edu>

#### References

Conn, E.E., Stumpf, P.K. (1963) Outlines of Biochemistry. John Wiley and Sons, New York.

Dayhoff, M.O., Eck, R.V. (1968) *Atlas of Protein Sequence and Structure 1967-1968*. National Biomedical Research Foundation, Silver Spring, Maryland.

Dayhoff, M.O., Schwartz, R.M., Orcutt, B.C. (1979) A model of evolutionary change in proteins. pp. 345-352 in *Atlas of Protein Sequence and Structure, Volume 5, Supplement 3, 1978*, ed. M.O. Dayhoff. National Biomedical Research Foundataion, Silver Spring, Maryland.

Felsenstein, J. (1989) PHYLIP-Phylogeny Inference Package (Version 3.2). Cladistics, 5, 164-166.

Felsenstein, J. (2013) PHYLIP (Phylogeny Inference Package) version 3.695. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

George, D.G., Hunt, L.T., Barker., W.C. (1988) *Current methods in sequence comparison and analysis*. pp. 127-149 in Macromolecular Sequencing and Synthesis, ed. D. H. Schlesinger. Alan R. Liss, New York.

Jones, D.T., Taylor, W.R., Thornton, J.M. (1992) The rapid generation of mutation data matrices from protein sequences. *Computer Applications in the Biosciences (CABIOS)*, 8, 275-282.

Kimura, M. (1980) A simple model for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, 16, 111-120.

Koisol, C., Goldman, N. (2005) Different versions of the Dayhoff rate matrix. *Molecular Biology and Evolution*, 22, 193-199.

Veerassamy, S., Smith, A., Tillier, E.R. (2003) A transition probability model for amino acid substitutions from blocks. *Journal of Computational Biology*, 10, 997-1010.

## See Also

Rneighbor

Rprotpars 39

## **Examples**

```
## Not run:
data(chloroplast)
D<-Rprotdist(chloroplast,model="PAM")
tree<-Rneighbor(D)
## End(Not run)</pre>
```

**Rprotpars** 

R interface for protpars

## **Description**

This function is an R interface for protpars in the PHYLIP package (Felsenstein 2013). protpars can be used for MP phylogeny estimation from protein sequences (Eck & Dayhoff 1966; Fitch 1971).

## Usage

```
Rprotpars(X, path=NULL, ...)
```

#### **Arguments**

X an object of class "proseq" containing aligned amino acid sequences.

path path to the executable containing protpars. If path = NULL, the R will search

several commonly used directories for the correct executable file.

... optional arguments to be passed to protpars. See details for more information.

# Details

Optional arguments include the following: quiet suppress some output to R console (defaults to quiet = FALSE); tree object of class "phylo" - if supplied, then the parsimony score will be computed on a fixed input topology; random.order add taxa to tree in random order (defaults to random.order = TRUE); random.addition number of random addition replicates for random.order = TRUE (defaults to random.addition = 10); threshold threshold value for threshold parsimony (defaults to ordinary parsimony); genetic.code, type of genetic code to assume (options are "universal", the default, "mitochondrial", "vertebrate.mitochondrial", "fly.mitochondrial", and "yeast.mitochondrial"); weights vector of weights of length equal to the number of columns in X (defaults to unweighted); outgroup outgroup if outgroup rooting of the estimated tree is desired; and cleanup remove PHYLIP input & output files after the analysis is completed (defaults to cleanup = TRUE).

More information about the protpars program in PHYLIP can be found here http://evolution.genetics.washington.edu/phylip/doc/protpars.html.

Obviously, use of any of the functions of this package requires that PHYLIP (Felsenstein 1989, 2013) should first be installed. Instructions for installing PHYLIP can be found on the PHYLIP webpage: http://evolution.genetics.washington.edu/phylip.html.

## Value

This function returns an object of class "phylo" that is the optimized tree.

40 Rrestdist

#### Author(s)

Liam J. Revell, Scott A. Chamberlain

Maintainer: Liam J. Revell < liam.revell@umb.edu>

#### References

Eck. R.V., Dayhoff, M.O. (1966) *Atlas of Protein Sequence and Structure 1966*. National Biomedical Research Foundation, Silver Spring, Maryland.

Felsenstein, J. (1989) PHYLIP-Phylogeny Inference Package (Version 3.2). Cladistics, 5, 164-166.

Felsenstein, J. (2013) PHYLIP (Phylogeny Inference Package) version 3.695. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Fitch, W.M. (1971) Toward defining the course of evolution: Minimu change for a specified tree topology. *Systematic Zoology*, 20, 406-416.

## See Also

```
as.proseq, Rdnapars, read.protein
```

# **Examples**

```
## Not run:
data(chloroplast)
tree<-Rprotpars(chloroplast)
## End(Not run)</pre>
```

Rrestdist

R interface for restdist

## **Description**

This function is an R interface for restdist in the PHYLIP package (Felsenstein 2013). restdist can be used to estimate the evolutionary distances between populations based on restriction site or fragment data.

# Usage

```
Rrestdist(X, path=NULL, ...)
```

# **Arguments**

X an object of class "rest.data" containing data for restriction sites.

path to the executable containing restdist. If path = NULL, the R will search

several paths for restdist. See details for more information.

... optional arguments to be passed to restdist. See details for more information.

Rrestdist 41

#### **Details**

Optional arguments include the following: quiet suppress some output to R console (defaults to quiet = FALSE); method, the method for calculating the genetic distance - options are "Modified" (the default), and "Nei/Li" (Nei & Li 1979); data, the type of data - either "sites" or "fragments"; gamma alpha shape parameter of a gamma model of rate heterogeneity among sites (defaults to no gamma rate heterogeneity); kappa transition:transversion ratio (defaults to kappa = 2.0); site.length the restriction site length (defaults to site.length=6); and cleanup remove PHYLIP input & output files after the analysis is completed (defaults to cleanup = TRUE).

More information about the restdist program in PHYLIP can be found here http://evolution.genetics.washington.edu/phylip/doc/restdist.html.

Obviously, use of any of the functions of this package requires that PHYLIP (Felsenstein 1989, 2013) should first be installed. Instructions for installing PHYLIP can be found on the PHYLIP webpage: http://evolution.genetics.washington.edu/phylip.html.

#### Value

This function returns an object of class "dist".

#### Author(s)

Liam J. Revell, Scott A. Chamberlain

Maintainer: Liam J. Revell < liam.revell@umb.edu>

# References

Felsenstein, J. (1989) PHYLIP-Phylogeny Inference Package (Version 3.2). Cladistics, 5, 164-166.

Felsenstein, J. (2013) PHYLIP (Phylogeny Inference Package) version 3.695. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Nei, M., Li, W.-H. (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences*, 76, 5269-5273.

# See Also

Rdnadist, Rgendist, Rprotdist

# **Examples**

```
## Not run:
data(restriction.data)
D<-Rrestdist(restriction.data)
## End(Not run)</pre>
```

42 Rrestml

Rrestml

R interface for restml

## **Description**

This function is an R interface for restml in the PHYLIP package (Felsenstein 2013). restml can be used to conduct ML phylogeny inference from restriction site or fragment data (Nei & Li 1979; Smouse & Li 1987; Felsenstein 1992).

## Usage

```
Rrestml(X, path=NULL, ...)
```

## **Arguments**

X an object of class "rest.data" containing data for restriction sites.

path path to the executable containing restml. If path = NULL, the R will search

several paths for restml. See details for more information.

... optional arguments to be passed to restml. See details for more information.

#### **Details**

Optional arguments include the following: quiet suppress some output to R console (defaults to quiet = FALSE); tree object of class "phylo" - if supplied, then the model will be optimized on a fixed input topology; speedier speedier but rougher analysis (defaults to speedier = FALSE); global perform global search (defaults to global = TRUE); random.order add taxa to tree in random order (defaults to random.order = TRUE); random.addition number of random addition replicates for random.order = TRUE (defaults to random.addition = 10); site.length the restriction site length (defaults to site.length=6); outgroup outgroup if outgroup rooting of the estimated tree is desired; and cleanup remove PHYLIP input & output files after the analysis is completed (defaults to cleanup = TRUE).

More information about the restml program in PHYLIP can be found here http://evolution.genetics.washington.edu/phylip/doc/restml.html.

Obviously, use of any of the functions of this package requires that PHYLIP (Felsenstein 1989, 2013) should first be installed. Instructions for installing PHYLIP can be found on the PHYLIP webpage: http://evolution.genetics.washington.edu/phylip.html.

## Value

This function returns an object of class "dist".

## Author(s)

Liam J. Revell, Scott A. Chamberlain

Maintainer: Liam J. Revell < liam.revell@umb.edu>

Rseqboot 43

#### References

Felsenstein, J. (1989) PHYLIP-Phylogeny Inference Package (Version 3.2). Cladistics, 5, 164-166.

Felsenstein, J. (1992) Phylogenies from restriction sites, a maximum likelihood approach. *Evolution*, 46, 159-173.

Felsenstein, J. (2013) PHYLIP (Phylogeny Inference Package) version 3.695. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Nei, M., Li, W.-H. (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences*, 76, 5269-5273.

Smouse, P. E., Li, W.-H. (1987) Likelihood analysis of mitochondrial restriction-cleavage patterns for the human-chimpanzee-gorilla trichotomy. *Evolution*, 41, 1162-1176.

#### See Also

Rrestdist

# **Examples**

```
## Not run:
data(restriction.data)
tree<-Rrestml(restriction.data)
## End(Not run)</pre>
```

Rseqboot

R interface for seqboot

# Description

This function is an R interface for seqboot in the PHYLIP package (Felsenstein 2013). seqboot can be to perform the non-parametric bootstrap following Felsenstein (1985).

## Usage

```
Rseqboot(X, path=NULL, ...)
```

## **Arguments**

X	er: (a) an object of class "DNAbin" containing DNA sequences or an ob- of class "proseq" containing amino acid sequences; (b) an object of class ylip.data" containing data for a binary or multi-state discrete character; (c) bject of class "rest.data" containing restriction site data; or (d) a matrix	
path	with gene frequency data (this option does not presently work properly).  path to the executable containing seqboot. If path = NULL, the R will search several paths for seqboot. See details for more information.	

... optional arguments to be passed to seqboot. See details for more information.

44 Rthreshml

#### **Details**

Optional arguments include the following: quiet suppress some output to R console (defaults to quiet = FALSE); type the data type (options are "sequence", "morph", "rest", or "gene.freq"); method analysis method (options are "bootstrap", "jacknife", and "permute"); percentage a bootstrap or jacknife percentage (can be >100); block.size block size in number of characters for the bootstrap, jacknife, or permutation (defaults to 1, the regular bootstrap); replicates number of replicates (defaults to 100); weights a vector of weights that must be 0 or 1 - characters with weight of 0 are excluded from the resampling procedure; rate.categories vector of integers between 1 and 9 indicating the rate category for each character - this is supplied so that resampled rate categories can be output by the method; mixture vector of "W" and "C" (see Rmix); and cleanup remove PHYLIP input & output files after the analysis is completed (defaults to cleanup = TRUE).

More information about the seqboot program in PHYLIP can be found here <a href="http://evolution.genetics.washington.edu/phylip/doc/seqboot.html">http://evolution.genetics.washington.edu/phylip/doc/seqboot.html</a>.

Obviously, use of any of the functions of this package requires that PHYLIP (Felsenstein 1989, 2013) should first be installed. Instructions for installing PHYLIP can be found on the PHYLIP webpage: http://evolution.genetics.washington.edu/phylip.html.

#### Value

This function returns a list of objects of class "DNAbin", "proseq", "phylip.data", or "rest.data" depending on the input data type. If ancestors, mixture, or rate.categories are supplied, then each element of the list contains an object data (containing the data), and vectors ancestors, mixture, and/or categories, containing the resampled input vectors.

#### Author(s)

Liam J. Revell, Scott A. Chamberlain

Maintainer: Liam J. Revell < liam.revell@umb.edu>

#### References

Felsenstein, J. (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39, 783-791.

Felsenstein, J. (1989) PHYLIP–Phylogeny Inference Package (Version 3.2). *Cladistics*, 5, 164-166. Felsenstein, J. (2013) PHYLIP (Phylogeny Inference Package) version 3.695. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Rthreshml

R interface for threshml

# **Description**

This function is an R interface for threshml in the PHYLIP package (Felsenstein 1989, 2013). threshml fits the threshold model of Felsenstein (2005; 2012). Note that threshml is new & not in the currently released version of PHYLIP (as of December 2013). It can be downloaded from its webpage here: http://evolution.gs.washington.edu/phylip/download/threshml/. If not specifying path, the executable file for threshml (e.g., threshml.exe in Windows) should be placed in the folder containing all other executable files for PHYLIP (e.g., C:/Program Files/phylip=3.695/exe in Windows).

Rthreshml 45

Obviously, use of any of the functions of this package requires that PHYLIP (Felsenstein 2013) should first be installed. More information about installing PHYLIP can be found on the PHYLIP webpage: http://evolution.genetics.washington.edu/phylip.html.

# Usage

```
Rthreshml(tree, X, types=NULL, path=NULL, ...)
```

## **Arguments**

tree	object of class "phylo".
Х	a <i>data.frame</i> of continuous valued or discrete character traits with rownames containing species names. Discrete & continuous characters can be supplied in any order. All discrete character traits must be two-state, but can be coded using any convention (i.e., $\emptyset$ , 1, "A", "B", etc.).
types	character vector containing the types (e.g., "discrete", "continuous"). If types are not supplied, Rthreshml will try to figure out which columns via is.numeric.
path	path to the executable containing threshml. If path = NULL, the R will search several commonly used directories for the correct executable file.
	optional arguments to be passed to threshml. See details for more information.

## **Details**

Optional arguments include the following: quiet suppress some output to R console (defaults to quiet = FALSE); burnin burnin generations for the MCMC; nchain number of chains of the MCMC; ngen number of generations in each chain; proposal variance on the proposal distribution for the MCMC; 1rtest logical value indicating whether to conduct a likelihood-ratio test of the hypothesis that some correlations are zero (does not appear to work in the present version); and cleanup remove PHYLIP input/output files after the analysis is completed (defaults to cleanup = TRUE).

# Value

This function returns a list containing the results from threshml.

## Author(s)

Liam J. Revell, Scott A. Chamberlain

Maintainer: Liam J. Revell < liam.revell@umb.edu>

#### References

Felsenstein, J. (1989) PHYLIP-Phylogeny Inference Package (Version 3.2). Cladistics, 5, 164-166.

Felsenstein, J. (2005) Using the quantitative genetic threshold model for inferences between and within species *Philosophical Transactions of the Royal Society London B*, 360, 1427-1434.

Felsenstein, J. (2012) A comparative method for both discrete and continuous characters using the threshold model. *American Naturalist*, 179, 145-156.

Felsenstein, J. (2013) PHYLIP (Phylogeny Inference Package) version 3.695. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

46 Rtreedist

#### See Also

Rcontrast

Rtreedist	R interface for treedist	

# **Description**

This function is an R interface for treedist in the PHYLIP package (Felsenstein 2013). treedist can be used to compute the distance between trees by two different methods.

# Usage

```
Rtreedist(trees, method=c("branch.score","symmetric"), path=NULL, ...)
```

## **Arguments**

trees	an object of class "multiPhylo". (Or, under rare circumstances, an object of class "phylo". See below.)
method	method to compute the distance between trees. method="branch.score" is the branch score method of Kuhner & Felsenstein (1994). method="symmetric" is the symmetric distance or Robinson-Foulds distance (Bourque 1978; Robinson & Foulds 1981).
path	path to the directory containing the executable treedist. If path = NULL, the R will search several commonly used directories for the correct executable file.
	optional arguments to be passed to treedist. See details for more information.

#### **Details**

Optional arguments include the following: quiet suppress some output to R console (defaults to quiet = FALSE); trees2 object of class "multiPhylo" or "phylo" - if two sets of trees are to be compared; rooted logical value indicating whether trees should be treated as rooted (defaults to rooted = FALSE); distances argument telling treedist which distances to compute - could be "all" (all pairwise in trees), "all.1to2" (all in trees by all in trees2), "adjacent" (adjacent species in trees only), and "corresponding" (all corresponding trees in trees and trees2); and cleanup remove PHYLIP input & output files after the analysis is completed (defaults to cleanup = TRUE).

More information about the treedist program in PHYLIP can be found here http://evolution.genetics.washington.edu/phylip/doc/treedist.html.

Obviously, use of any of the functions of this package requires that PHYLIP (Felsenstein 1989, 2013) should first be installed. Instructions for installing PHYLIP can be found on the PHYLIP webpage: http://evolution.genetics.washington.edu/phylip.html.

# Value

This function returns a matrix of pairwise distances for distances = "all" and distances = "all.1to2", or a named vector for distances = "adjacent" and distances = "corresponding".

setPath 47

## Author(s)

Liam J. Revell, Scott A. Chamberlain

Maintainer: Liam J. Revell < liam.revell@umb.edu>

#### References

Bourque, M. (1978) *Arbres de Steiner et reseaux dont certains sommets sont a localisation variable*. Ph.D. Dissertation, Universite de Montreal, Montreal, Quebec.

Felsenstein, J. (1989) PHYLIP-Phylogeny Inference Package (Version 3.2). Cladistics, 5, 164-166.

Felsenstein, J. (2013) PHYLIP (Phylogeny Inference Package) version 3.695. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Kuhner, M.K., Felsenstein, J. (1994) A simulation comparison of phylogeny algorithms under equal and unequal evolutionary rates. *Molecular Biology and Evolution*, 11, 459-468.

Robinson, D.F., Foulds, L.R. (1981) Comparison of phylogenetic trees. *Mathematical Biosciences*, 53, 131-147.

# **Examples**

```
## Not run:
trees<-rmtree(n=10,N=10)
D<-Rtreedist(trees,method="symmetric")
## End(Not run)</pre>
```

setPath

Set path to the folder containing PHYLIP executables for the current R session

## **Description**

Sets the path to the folder containing PHYLIP executables for the current R session. Once setPath has been used, Rphylip functions will no longer search for the path to the PHYLIP executable, using path instead.

clearPath clears the path.

## Usage

```
setPath(path)
clearPath()
```

# **Arguments**

path

path to the folder contains PHYLIP executables.

# Author(s)

Liam J. Revell, Scott A. Chamberlain

Maintainer: Liam J. Revell < liam.revell@umb.edu>

48 setupOSX

setupOSX

Help set up PHYLIP in Mac OS X

# **Description**

This function attempts to help set up PHYLIP on a Mac OS X machine.

## Usage

```
setupOSX(path=NULL)
```

## **Arguments**

path

path to the folder containing the PHYLIP package. If path = NULL, the R will search several commonly used directories.

#### **Details**

This function can be used to help set up PHYLIP (http://evolution.genetics.washington.edu/phylip.html) following the special instructions found here: http://evolution.genetics.washington.edu/phylip/install.html. setupOSX should only be run once - when PHYLIP is first installed.

## Author(s)

Liam J. Revell, Scott A. Chamberlain

Maintainer: Liam J. Revell < liam.revell@umb.edu>

# References

Felsenstein, J. (2013) PHYLIP (Phylogeny Inference Package) version 3.695. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

# **Examples**

```
## Not run:
setupOSX()
## End(Not run)
```

# Index

*Topic <b>DNA</b>	Rfitch, 25
opt.Rdnam1,4	Rkitsch, 27
*Topic <b>amino acid</b>	Rmix, 29
as.proseq, 3	Rneighbor, 31
print.proseq,6	Rpars, 32
read.protein,24	Rpenny, 34
Rproml, 35	Rproml, 35
Rprotdist, 37	Rprotpars, 39
Rprotpars, 39	*Topic maximum likelihood
*Topic binary character	opt.Rdnaml,4
as.proseq, 3	Rcontml, 9
print.proseq, 6	Rcontrast, 10
*Topic comparative method	Rdnadist, 13
Rcontrast, 10	Rdnaml, 16
Rthreshml, 44	Rproml, 35
*Topic <b>consensus</b>	Rprotdist, 37
Rconsense, 8	*Topic <b>package</b>
*Topic continuous characters	Rphylip-package, 2
Rcontml, 9	*Topic parsimony
Rcontrast, 10	Rclique, 6
*Topic datasets	Rdnacomp, 12
primates, 5	Rdnapars, 18
*Topic distance method	Rdnapenny, 19
Rdnadist, 13	Rdollop, 21
Rfitch, 25	Rdolpenny, 22
Rgendist, 26	Rmix, 29
Rkitsch, 27	Rpars, 32
Rneighbor, 31	Rpenny, 34
Rprotdist, 37	Rprotpars, 39
Rrestdist, 40	*Topic phylogenetics
Rrestml, 42	opt.Rdnaml,4
*Topic inference	Rclique, 6
opt.Rdnaml,4	Rconsense, 8
Rclique, 6	Rcontml, 9
Rcontml, 9	Rcontrast, 10
Rdnacomp, 12	Rdnacomp, 12
Rdnadist, 13	Rdnadist, 13
Rdnainvar, 15	Rdnainvar, 15
Rdnaml, 16	Rdnaml, 16
Rdnapars, 18	Rdnapars, 18
Rdnapenny, 19	Rdnapenny, 19
Rdollop, 21	Rdollop, 21
Rdolpenny, 22	Rdolpenny, 22

50 INDEX

read.protein, 24 Rfitch, 25 Rgendist, 26 Rkitsch, 27 Rmix, 29 Rneighbor, 31 Rpars, 32 Rpenny, 34 Rproml, 35 Rprotdist, 37 Rprotpars, 39 Rrestdist, 40 Rrestml, 42 Rseqboot, 43 Rthreshml, 44 Rtreedist, 46 setupOSX, 48 *Topic restriction sites as.proseq, 3 print.proseq, 6 *Topic utilities as.proseq, 3 print.proseq, 6 read.protein, 24 setPath, 47 setupOSX, 48  as.phylip.data (as.proseq), 3 as.proseq, 3, 6, 24, 37, 40 as.rest.data (as.proseq), 3 chloroplast (primates), 5 clearPath (setPath), 47 cotton (primates), 5 opt.Rdnaml, 4, 17 pic, 11 primates, 5 print.DNAbin, 6 print.phylip.data (print.proseq), 6 Print.proseq, 3, 6, 24 print.rest.data (print.proseq), 6 Rclique, 6 Rconsense, 8 Rcontml, 9, 17 Rcontrast, 10, 46 Rdnacomp, 7, 12 Podedict, 12, 26, 37, 23, 41	Rdnapars, 13, 16, 18, 20, 22, 23, 30, 35, 40 Rdnapenny, 13, 19, 19, 33 Rdollop, 21 Rdolpenny, 22 read.dna, 24 read.protein, 24, 37, 40 restriction.data (primates), 5 Rfitch, 25, 29 Rgendist, 26, 41 Rkitsch, 27 Rmix, 7, 29, 44 Rneighbor, 14, 31, 38 Rpars, 22, 32 Rpenny, 34 Rphylip (Rphylip-package), 2 Rphylip-package, 2 Rproml, 3, 6, 10, 17, 24, 35 Rprotdist, 27, 37, 41 Rprotpars, 39 Rrestdist, 40, 43 Rrestml, 42 Rseqboot, 43 Rthreshml, 44 Rtreedist, 46 setPath, 47 setupOSX, 48
Rconsense, $8$ Rcontml, $9$ , $17$ Rcontrast, $10$ , $46$	