Package ‘convevol’

June 30, 2023

Type Package
Title Analysis of Convergent Evolution
Version 2.0.1
Author William H. Brightly, C. Tristan Stayton
Maintainer C. Tristan Stayton <tstayton@bucknell.edu>
Depends MASS, ape, cluster, geiger, phytools
Description Quantifies and assesses the significance of convergent evolution using multiple methods and measures as described in Stayton (2015) <DOI:10.1111/evo.12729> and Grossnickle et al. 2023. Also displays results in various ways.
License GPL-2
Encoding UTF-8
RoxygenNote 7.2.3
NeedsCompilation no
Repository CRAN
Date/Publication 2023-06-30 17:30:02 UTC

R topics documented:

calcConv .................................................. 2
calcConvCt .............................................. 3
calcCs .................................................. 5
calcCsCt .................................................. 6
convnum .................................................. 7
convnumsig ............................................. 9
convSig ................................................ 10
convSigCt .............................................. 11
multianc ................................................. 13
plotCt .................................................. 14
plotellipse ............................................. 15
pullNodeSeq ........................................... 16

Index 17
calcConv

**Description**

calcConv prepares arguments for the CalcCs function

**Usage**

calcConv(phy, traits, focaltaxa, VERBOSE = FALSE)

**Arguments**

- **phy**: The phylogeny of interest in phylo format
- **traits**: Phenotypic data for all tips
- **focaltaxa**: A list consisting of the names of all putatively convergent taxa
- **VERBOSE**: Whether or not to print progress

**Details**

calcConv is a wrapper function which formats data, performs ancestral state reconstructions, obtains distance matrices, and determines pairwise combinations of focal taxa, which are then used as arguments for the CalcCs function, which calculates values for C1-C4 of Stayton (2015) for each pair of putatively convergent taxa.

**Value**

C1-C4 convergence measures for all pairs of putatively convergent taxa.

**References**


**Examples**

```r
phy<-rtree(100)
traits<-fastBM(phy,nsim=3)
focaltaxa<-c("t1", "t50", "t100")
answer<-calcConv(phy,traits,focaltaxa,VERBOSE=FALSE)
```
calcConvCt

Computes Ct-metric scores for putatively convergent tips (or groups of tips) given a set of user provided phenotypic characters and a time calibrated phylogeny.

Description

calcConvCt Computes Ct-metric scores for putatively convergent tips (or groups of tips) given a set of user provided phenotypic characters and a time calibrated phylogeny.

Usage

calcConvCt(
  phy, 
  traits, 
  focaltaxa, 
  groups = NULL, 
  conservative = FALSE, 
  VERBOSE = FALSE 
)

Arguments

phy The time calibrated phylogeny of interest in phylo format

traits A matrix of numeric phenotypic traits with rownames matching tip labels of phy

focaltaxa A vector of tip labels for the putatively convergent taxa to be compared

groups An optional vector of groups with names matching focaltaxa. Indicates the group identity of all putatively convergent taxa and limits Ct measures to intergroup comparisons only

conservative Logical value indicating whether measurement of Dmax.t should be restricted to before the origin of the oldest lineage in each pairwise comparison of the focaltaxa. The origin of each convergent lineages is taken as the most recent common ancestor of tips in each user defined group. Where groups include a single tip, the parent node of the tip’s subtending branch is used. Requires group object to be provided by user.

VERBOSE Logical value indicating whether model information should be printed during computation

Details

Function incorporates the optimizations introduced by Zelditch et al. (2017), which significantly improve runtimes

Reconstructions part way along branches are obtained using equation [2] of Felsenstein (1985), following code modified from the phytools (Revell, 2012) function contMap
Value

A list of the following components:

mean A named vector of Ct-metrics averaged from all pairwise comparisons of focaltaxa. If user provided groups, this is based only on comparisons between taxa belonging to different groups.

Cmat A matrix of Ct-metrics for each pairwise comparison.

path_df A list of dataframes, one per pairwise comparison of the focal taxa, each containing information from all timepoint measurements of the two putatively convergent lineages. These provide the nodes at which comparisons were drawn, the evolutionary path along which that node fell (i.e., leading to one of two tips), the node height, reconstructed ancestral states at that node for each phenotypic trait, reconstructed ancestral values for each trait along the opposite path, and the phenotypic distance between the two lineages at that point.

grp.mean A matrix of Ct-metrics summarized for inter-group comparisons, returned only if user defined groups were specified. Provides overall results matching those reported in “mean”, results for each unique inter-group comparison, and results averaged with equal weight given to each unique inter-group comparison (i.e., ignoring differences in the number of tips in each group).

References


Examples

```r
# create time calibrated tree
mytree<-rtree(100)
mycalibration <- makeChronosCalib(mytree, node="root", age.max=50)
phy <- chronos(mytree, calibration = mycalibration, control = chronos.control() )
class(phy)<-"phylo"

# create three normally distributed phenotypic traits
traits <- cbind(rnorm(Ntip(phy)),rnorm(Ntip(phy)),rnorm(Ntip(phy)))
rownames(traits) <- phy$tip.label
focaltaxa <- sample(phy$tip.label, 5)

system.time(run <- calcConvCt(phy, traits, focaltaxa))
```
The function `calcCs` calculates the C1-C4 measures of convergent evolution between two lineages as described in Stayton (2015). All measures quantify convergence by the ratio of current to maximum past phenotypic distance between lineages. Can be used as-is but more often will be used within the `calcConv` script. Code written by Jonathan S. Mitchell for Zelditch et al. (2017).

**Description**

`calcCs` calculates the C1-C4 measures of convergent evolution.

**Usage**

```r
calcCs(tips, ancList, allDists, phy, VERBOSE = FALSE)
```

**Arguments**

- `tips`: Two putatively convergent tips
- `ancList`: A list of ancestors of all tips. Most often obtained from `calcConv`
- `allDists`: A matrix of phenotypic distances between all nodes (tips and ancestors). Most often obtained from `calcConv`
- `phy`: The phylogeny of interest
- `VERBOSE`: Whether or not to print progress

**Details**

`calcCs` calculates values of C1-C4, all of which are fundamentally based on comparing the current phenotypic distance between two tips to the maximum past distances between the ancestors of those tips. Higher values indicate a greater amount of past phenotypic distance which has been "closed" by subsequent evolution, and thus greater convergence. C1 is the ratio of tip to maximum ancestral distance. C2 is the difference of those two values. C3 scales C2 by the total amount of evolution that has occurred in the two lineages. C4 scales C2 by the total amount of evolution that has occurred in the entire phylogeny. The arguments for this function will usually be obtained from the `calcConv` script in convevol - this allows certain computationally-intensive steps (e.g., calculating ancestral states) to only be performed once, thus saving a great deal of time. This script also corrects an error in the calculation of C4 of previous versions of convevol.

**Value**

- C1-C4 convergence measures for all pairs of putatively convergent taxa.
References

calcCsCt

Computes Ct values for a pair of tips. Internal, called in calcConv.

Description
calcCsCt Computes Ct values for a pair of tips. Internal, called in calcConv.

Usage
calcCsCt(
  tips,
  ancList,
  allDists,
  phy,
  VERBOSE = FALSE,
  allVals,
  edge,
  lim.height = NULL
)

Arguments
tips vector of two tips
ancList list of node paths for all tips in the user provided phylogeny
allDists matrix of phenotypic distances between all exterior and interior node pairs
phy The time calibrated phylogeny of interest in phylo format
VERBOSE logical value indicating whether model information should be printed during computation
allVals a matrix of observed and reconstructed phenotypes for all user supplied traits at interior and exterior nodes
edge a list of data frames, each including the edge matrix of user supplied phylogeny, along with node heights and reconstructed phenotype for each of the user supplied traits
lim.height an optional tree height used to limit Dmax.t, passed only if groups are defined and a conservative test is run (see calcConv)
Details

Function incorporates the optimizations introduced by Zelditch et al. (2017), which significantly improve runtimes.

Reconstructions part way along branches are obtained using equation [2] of Felsenstein (1985), following code modified from the phytools (Revell, 2012) function contMap.

Value

A list of the Ct values.

References


convnum

Quantify convergence by the number of convergent events

Description

This program takes in a set of taxa that are already suspected to be convergent in a particular area of morphospace. It then counts the number of times that a lineage has invaded that region of morphospace.

Usage

convnum(
    phy,
    phendata,
    convtips,
    plot = TRUE,
    ellipse = NULL,
    plotellipse = NULL
)
Arguments

- **phy**: The phylogeny of interest in phylo format
- **phendata**: Phenotypic data for all tips
- **convtips**: A list consisting of the names of all convergent taxa
- **plot**: Whether or not to plot a phylomorphospace with lineages that cross into the region of interest highlighted as red arrows. Default=TRUE
- **ellipse**: Optional. An ellipse defining the region of interest, into which groups may or may not converge.
- **plotellipse**: Optional. The ellipse defining the region of interest in the first two dimensions.

Details

This function will construct an ellipse around all convergent taxa. Then it will reconstruct ancestral states throughout the phylogeny, and use those to determine how many lineages have crossed into this ellipse from the outside.

Value

The number of lineages that have crossed into the region of trait space occupied by the convergent taxa.

References


Examples

```r
phy<-rtree(10)
phendata<-fastBM(phy,nsim=2)
convtips<-c("t1","t2","t3")
answer<-convnum(phy,phendata,convtips,plot=TRUE,ellipse=NULL,plotellipse=NULL)
```
Assess the significance of convergent evolution using simulations and the convnum metric

Description
Simulates evolution along a given phylogeny, using parameters derived from observed data, and calculates the convnum metric for each simulation for a set of user-defined taxa. Then compares the observed convnum value to the simulated values to assess the significance of the observed levels of convergent evolution.

Usage
```
convnumsig(
  phy, phendata, convtips, nsim, ellipse = NULL, plot = FALSE, plotellipse = NULL
)
```

Arguments
- `phy`: The phylogeny of interest in phylo format
- `phendata`: Phenotypic data for all tips
- `convtips`: A list consisting of the names of all convergent taxa
- `nsim`: The number of simulations to conduct
- `ellipse`: Optional. An ellipse defining the region of interest, into which groups may or may not converge.
- `plot`: Optional. Describes whether or not to show phylomorphospaces for all of the simulations.
- `plotellipse`: Optional. The ellipse defining the region of interest in the first two dimensions.

Details
None

Value
A list, consisting first of the p-value for the observed convnum, and second of a vector containing all of the simulated convnum values. Also displays a histogram of all of the simulated convnum values.
References


Examples

phy<-rtree(10)
phendata<-fastBM(phy,nsim=2)
convtips<-c("t1","t2","t3")
answer<-convnumsig(phy,phendata,convtips,10,plot=FALSE,ellipse=NULL,plotellipse=NULL)

convSig


Description

convSig calculates the significance of measures of convergent evolution.

Usage

convSig(phy, traits, focaltaxa, nsim = 1000)

Arguments

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>phy</td>
<td>The phylogeny of interest in phylo format</td>
</tr>
<tr>
<td>traits</td>
<td>Phenotypic data for all tips</td>
</tr>
<tr>
<td>focaltaxa</td>
<td>A list consisting of the names of all putatively convergent taxa</td>
</tr>
<tr>
<td>nsim</td>
<td>The number of simulations to use to assess significance</td>
</tr>
</tbody>
</table>

Details

This script simulates data according to a Brownian motion model of evolution, and then assesses convergene on that simulated data. The number of times that the simulated data produces greater convergence than that observed in the empirical data is used to calculate a p-value.

Value

C1-C4 convergence measures for all pairs of putatively convergent taxa and their associated p-values.
convSigCt

References

Examples
phy<-rtree(100)
traits<-fastBM(phy,nsim=3)
focaltaxa<-c("t1","t50","t100")
answer<-convSig(phy,traits,focaltaxa,nsim=10)

convSigCt
Computes and conducts significance tests on Ct-metric scores for putatively convergent tips (or groups of tips) given a set of user provided phenotypic characters and a time calibrated phylogeny.

Description
convSigCt Computes and conducts significance tests on Ct-metric scores for putatively convergent tips (or groups of tips) given a set of user provided phenotypic characters and a time calibrated phylogeny.

Usage
convSigCt(phy, traits, focaltaxa, groups = NULL, nsim = 1000, ...)

Arguments
phy
The time calibrated phylogeny of interest in phylo format
traits
a matrix of numeric phenotypic traits with rownames matching tip labels of phy
focaltaxa
a vector of tip labels for the putatively convergent taxa to be compared
groups
an optional vector of groups with names matching focaltaxa. Indicates the group identity of all putatively convergent taxa and limits Ct measures to intergroup comparisons only
nsim
number of simulated (Brownian motion) datasets used to build the null distribution
...
optional arguments to be passed to calcConv. If conservsative is TRUE, Dmax.t is restricted to before the origin of the oldest lineage in each pairwise comparison of the focal taxa. The origin of convergent lineages is taken as the most recent common ancestors of each user defined group. Where these groups include a single tip, the parent node of its subtending branch is used. Requires user
Function incorporates the optimizations introduced by Zelditch et al. (2017), which significantly improve runtimes.

Reconstructions part way along branches are obtained using equation [2] of Felsenstein (1985), following code modified from the phytools (Revell, 2012) function contMap.

Details

Value

A list of the following components:

pvals: a matrix containing Ct1 - Ct4 and p-values from significance tests for each
meas.Cmat: a matrix of Ct values for each pairwise comparison of focaltaxa
meas.path: a list of dataframes, one per pairwise comparison of focaltaxa, each containing information from all timepoint measurements of the two putatively convergent lineages. These provide the nodes at which comparisons were drawn, the evolutionary path along which that node fell (i.e., leading to one of two tips), the node height, reconstructed ancestral states at that node for each phenotypic trait, reconstructed ancestral values for each trait along the opposite path, and the phenotypic distance between the two lineages at that point.
sim.avg: average Ct values from all pairwise comparisons between focaltaxa using simulated Brownian motion traits, number of columns corresponds to the user provided number of simulations
sim.path: a list of dataframes as in meas.path, but obtained using simulated data. Length of object determined by number of pairwise comparisons multiplied by the number of simulated datasets.
grp.mean: a matrix of Ct-metrics summarized for inter-group comparisons, returned only if user defined groups were specified. Provides overall results matching those reported in "mean", results for each unique inter-group comparison, and results averaged with equal weight given to each unique inter-group comparison (i.e., ignoring differences in the number of tips in each group).
grp.pvals: a matrix of p-values associated with Ct values in grp.mean object. Returned only if user defined groups were specified.

References

Examples

# create time calibrated tree
mytree<-rtree(50)
mycalibration <- makeChronsCalib(mytree, node="root", age.max=50)
phy <- chronos(mytree, calibration = mycalibration, control = chronos.control() )
class(phy)<="phylo"

# create three normally distributed phenotypic traits
traits <- cbind(rnorm(Ntip(phy)),rnorm(Ntip(phy)),rnorm(Ntip(phy)))
rownames(traits) <- phy$tip.label
focaltaxa <- sample(phy$tip.label, 5)

system.time(run2 <- convSigCt(phy, traits, focaltaxa, nsim=10))

---

multianc

Reconstructs ancestral states for multiple characters

Description

Uses fastAnc to reconstruct ancestral states for multiple phenotypic characters

Usage

multianc(phy, phendata)

Arguments

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>phy</td>
<td>The phylogeny of interest in phylo format</td>
</tr>
<tr>
<td>phendata</td>
<td>Phenotypic data for all tips</td>
</tr>
</tbody>
</table>

Details

None

Value

A matrix with the tips data in the first n rows and the ancestral data in the remaining n-1 rows.

References


plotCt

### Examples

```r
phy<-rtree(10)
phendata<-fastBM(phy,nsim=2)
ancs<-multianc(phy,phendata)
```

### Description

plotCt Plots calcConv or convSig output.

### Usage

```r
plotCt(
  output,  # object containing calcConv or convSig output
  phy,    # The time calibrated phylogeny of interest in phylo format
  focaltaxa,  # a vector of tip labels for the putatively convergent taxa to be compared
  nsim = 25,  # number of null simulations to plot
  col = c("black", "forest green", "dodgerblue2", "firebrick1", "purple", "orange",  
          "salmon", "goldenrod", "springgreen2", "plum1"),  
  groups = NULL,  
  ...  
)
```

### Arguments

- **output**: object containing calcConv or convSig output
- **phy**: The time calibrated phylogeny of interest in phylo format
- **focaltaxa**: a vector of tip labels for the putatively convergent taxa to be compared
- **nsim**: number of null simulations to plot
- **col**: vector of colors to use for all unique intergroup comparisons a default option is given usable with up to five groups. If number of groups is 1 or less than length of col, not all colors will be used
- **groups**: an optional vector of groups with names matching focaltaxa, indicating the group identity of all focaltaxa
- **...**: optional arguments to be passed to tiplabels

### Details

Creates a plot that shows the phenotypic distances between pairs of putatively convergent lineages over time. When these distances decrease, convergence has occurred. When more than two putatively convergent taxa are analyzed, all pairs are plotted.
*plotellipse*

**Value**

A plot identifying putatively convergent taxa in the provided phylogeny and tracking the change in phenotypic distance between taxa since their most recent common ancestor.

**References**


---

**plotellipse**

*Plots an ellipse*

**Description**

Plots a minimum ellipse around a set of data.

**Usage**

`plotellipse(ellipse)`

**Arguments**

- `ellipse`: Gives the parameters of the ellipse - output from the ellipsoidhull function in cluster.

**Details**

Routine adapted from a suggestion made on CrossValidated: http://stats.stackexchange.com/questions/9898/how-to-plot-an-ellipse-from-eigenvalues-and-eigenvectors-in-r

**Value**

Nothing - just plots the ellipse.

**References**

pullNodeSeq


Description

pullNodeSeq Extracts a vector of all ancestors of a given taxon in a phylogeny.

Usage

pullNodeSeq(phy, tip)

Arguments

- **phy**
  - The phylogeny of interest in phylo format
- **tip**
  - The tip of interest

Value

A vector of ancestors

References


Examples

```r
phylogeny<-rtree(100)
answer<-pullNodeSeq(phy=phylogeny, tip="t1")
```
Index

calcConv, 2
calcConvCt, 3
calcCs, 5
calcCsCt, 6
convnum, 7
convnumSig, 9
convSig, 10
convSigCt, 11
multianc, 13
plotCt, 14
plotellipse, 15
pullNodeSeq, 16