Package 'turboMEDUSA'

December 17, 2011

Type Package

Title MEDUSA: Modeling Evolutionary Diversification Using Stepwise AIC

Version 0.93

Date 2011-12-17

Author Joseph W. Brown, Richard G. FitzJohn, Michael E. Alfaro, and Luke J. Harmon.

Maintainer Joseph W. Brown <josephwb@uidaho.edu>

Description Piecewise birth-death models fit from a chronogram and complete extant species richness.

License GPL (>= 2)

LazyLoad yes

Depends ape, geiger, multicore (>= 0.1-3)

Enhances geiger

OS_type unix

R topics documented:

---

get.b.d   *Get Per-lineage Birth and Death Rates*

Description

Get per-lineage birth and death rates from inferred net diversification and relative extinction rates.

Usage

```
get.b.d(r, epsilon)
```

Arguments

r   Net diversification rate (b-d)

epsilon   Relative extinction rate (d/b)

Details

Useful if one is in interested in b (lambda) & d (mu) rather than r and epsilon.

Value

Returns a list containing:

$b   Per-lineage birth (speciation) rate

$d   Per-lineage death (extinction) rate

---

plotPrettyTree   *MEDUSA: Modeling Evolutionary Diversification Using Stepwise AIC*

Description

Plot a publication-quality colour-coded tree, where edge colours indicate piecewise diversification model membership from a runTurboMEDUSA analysis. A wrapper for plot.phylo in the APE package.

Usage

```
plotPrettyTree(treeParameters, time = TRUE, node.labels = FALSE, margin = FALSE, cex = 0.5,
label.offset = 0, font = 3, color.tip.label = FALSE, ...)
```

Arguments

treeParameters	The list produced from a summarizeTurboMEDUSA summary containing the tree and piecewise diversification model membership.
time	Whether a timescale should be plotted. Possibly not wanted for large trees. Default = true;
node.labels	Whether node labels should be plotted on the tree to indicate the position and order of diversification rate shifts.
margin	Whether plot should include margins. time (above) overrules this option, as a margin is requires for plotting a timescale.
cex	Specifies the font size for printing tip labels. Default is cex=0.5; may need to decrease for larger trees.
label.offset	Specifies how far tip labels are offset rightward from the default plot.phylo() settings.
font	Specifies the font to display for tip labels. The default font=3 plots labels in italic. Options are font=1 (plain text), font=2 (bold), font=3 (italic, the default), and font=4 (bold italic).
color.tip.label	
	Determines whether tip labels should be coloured according to their piecewise diversification model membership.
...	Potential additional arguments to be passed to plot.phylo.

Value

None.

Author(s)

Joseph W. Brown, Richard G. FitzJohn, Michael E. Alfaro, and Luke J. Harmon.

Maintainer: Joseph W. Brown <josephwb@uidaho.edu>

References

Alfaro, ME, F Santini, C Brock, H Alamillo, A Dornburg, DL Rabosky, G Carnevale, and LJ Harmon. 2009. Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proceedings of the National Academy of Sciences* **106**: 13410-13414.

See Also

runTurboMEDUSA summarizeTurboMEDUSA

---

runTurboMEDUSA	*MEDUSA: Modeling Evolutionary Diversification Using Stepwise AIC*

Description

Fits piecewise birth-death models to an ultrametric phylogenetic tree according to phylogenetic (edge-length) and taxonomic (richness) likelihoods. Optimal model size is determined via a stepwise AIC approach.

Usage

```
runTurboMEDUSA(phy, richness = NULL, model.limit = 20, stop = "model.limit", model="mixed",
criterion = "aicc", shiftCut = "both", initialR = 0.05, initialE = 0.5, plotFig = FALSE,
nexus = FALSE, verbose = TRUE, mc = FALSE, num.cores = NULL, ...)
```

Arguments

phy	Ultrametric phylogenetic tree. Assumed to be of class phylo (i.e. ape format). Nexus trees can be passed in using nexus=TRUE (see below).
richness	Optional; only required if tree is NOT completely sampled. A dataframe with minimally two columns taxon and n.taxa. The identifier taxon must match exactly with a tip.label in the phylogeny phy; this is checked. May also include an exemplar column, used for renaming incompletely-sampled clades that require collapsing; in this case, exemplar must match exactly with a tip label and taxon will be the name for that clade (say, genus, family, etc.). If no richness information is provided then it is assumed tips represent single species with complete sampling.
model.limit	The number of piecewise models to be fit to the phylogeny. The program may overrule this value and fit fewer models is the tree is small (such that the denominator in the AIC correction factor becomes undefined).
model	The flavour of piecewise models to be fit to the phylogeny. Can be birth-death (model="bd"), pure-birth (model="yule"), or mixed (model="mixed") (the default). For mixed model analyses, both BD and Yule models are considered for all potential breaks.
stop	The criterion for when the analysis should be terminated. stop="model.limit" will terminate after model.limit piecewise models have been fit to the the phylogeny (see above). The alternative criterion is stop="threshold". Here threshold refers to the increase in AIC score that should be considered significant given the number of tips in the phylogeny; analysis will terminate when the addition of a piecewise BD model does not improve AIC score by threshold.
criterion	The statistical criterion used to determine whether larger models fit the data significantly better. Options are criterion="aicc" (default) or criterion="aic". As AIC and AICc will select the same breaks for any given model size, this option is only useful if stop="threshold". In this case, the default is strongly advised, as for some trees AIC can continue to improve well after AICc has deteriorated.
shiftCut	Determines where rate shifts are placed on the tree. Three options are available. For the first option shiftCut="stem", when a subtree is selected for submodel consideration, it is assumed that the shift in diversification occurred in the branch leading to the clade. The second option shiftCut="node" assumes the shift occurred at the node representing the most recent common ancestor of the clade (that is, omitting the leading branch). The final and default option shiftCut="both" considers both possibilities for shift placement, and carries the shift with the higher AICc score forward to piecewise model selection.
initialR	Starting guess for net diversification rate. May need to be decreased for the program to work on large trees.
initialE	Starting guess for relative extinction rate.
plotFig	Boolean value that indicates whether a plot of AIC and AICc against model size should be generated. Default is for no plot.

nexus	Boolean value that indicates whether the tree being passed in is in Nexus format. Default nexus=FALSE (i.e. the tree being passed in is in phylo format)
verbose	Print out additional information to screen, such as which taxa are pruned from the original tree given richness information. Mostly useful for debugging.
mc	Indicate whether multiple cores should be used (mc=TRUE). Works only on unix-like machines (i.e. not Windows), and only in non-GUI instantiations of R.
num.cores	The number of cores to use if mc=TRUE. By default all cores detected will be used. This option is only useful if one wants to limits the number of cares devoted to a particular R job.
...	Potential additional arguments for plotting etc.

Details

A series of diversification models is fit to a combination of phylogenetic and taxonomic data. The input is a phylogenetic tree with branch lengths proportional to time showing the relationship amongst clades, and the diversity for living species in all of those clades. All taxa missing from the tree have to be assigned to one of the tip clades in the richness matrix.

The algorithm first fits a single diversification model to the entire dataset. Then, a series of breaks are added, so that different parts of the tree evolve with different parameter values (per-lineage net diversification and relative extinction rates). The program first compares all single-breakpoint models with the overall model, and selects the best one. Then all possible two-breakpoint models are compared with the best single-breakpoint model, and so on.

Summarize results with summarizeTurboMEDUSA.

Value

Returned list has elements:

$z	A matrix listing branch times and richnesses; for summarizing results only.
$anc	A list of all ancestors; for summarizing results only.
$phy	The phylogeny that was analyzed. May have been pruned (from richness information) from originally tree passed in. Mainly for summarizing results.
$modelSummary	A dataframe containing break and fit values for optimal models at each size considered.
$threshold	The correct AIC threshold given the size of the tree analyzed. Used in summarizing results and (possibly) terminating the analysis.
$models	A list containing the optimal models at each successive size. For each model contains:

- $par A matrix containing estimates of r and epsilon for all piecewise models.
- $lnLik.part The likelihoods of each piecewise model.
- $lnLik The total likelihood for the piecewise model.
- $split.at The node(s) where individual BD models originate.
- $aic AIC for the piecewise model.
- $aicc AICc for the piecewise model.
- $num.par The number of parameters (total r + epsilon + breakpoints) in the piecewise model.

Author(s)

Joseph W. Brown, Richard G. FitzJohn, Michael E. Alfaro, and Luke J. Harmon.

Maintainer: Joseph W. Brown <josephwb@uidaho.edu>

References

Alfaro, ME, F Santini, C Brock, H Alamillo, A Dornburg, DL Rabosky, G Carnevale, and LJ Harmon. 2009. Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proceedings of the National Academy of Sciences* **106**: 13410-13414.

See Also

summarizeTurboMEDUSA plotPrettyTree

---

summarizeTurboMEDUSA   *MEDUSA: Modeling Evolutionary Diversification Using Stepwise AIC*

Description

Summarize results from a runTurboMEDUSA analysis.

Usage

```
summarizeTurboMEDUSA(results, modelNum = NULL, cutoff = "threshold", criterion = "aicc",
plotTree = TRUE, time = TRUE, node.labels = TRUE, cex = 0.5, plotSurface = FALSE,
printTitle = TRUE, n.points = 100, ...)
```

Arguments

results	The list produced from a runTurboMEDUSA analysis.
modelNum	Optional. An integer indicating which model in results should be summarized. This is only useful if one wants to summarize a suboptimal model (i.e. not the best criterion or cutoff model; see below).
cutoff	Determines the increase in criterion score (see below) that is considered significant i.e. is used in determining the optimal size of model. The default is cutoff="threshold"; this is a predetermined threshold that is a function of the number of tips in the phylogeny. This relationship was determined via simulation. Alternatively, a numeric value can be passed in as the cutoff. In either case, the program selects the smallest model where criterion(model_i - model_i-1) > threshold
criterion	The statistical criterion used to determine whether larger models fit the data significantly better. Options are criterion="aicc" (default) or criterion="aic". The default is strongly advised, as for some trees AIC can continue to improve well after AICc has deteriorated.
plotTree	Boolean value that determines whether the (possibly pruned) tree is plotted. If plotTree=TRUE (the default) the tree is plotted with edges coloured according to which piecewise model they belong to, and positions (nodes) of break points are potentially indicated (see node.labels below).

time	Boolean value that indicates whether a timescale is printed below the tree. Only has meaning if `plotTree=TRUE`. For large trees it might be desirable to set `time=FALSE` to make tip labels legible.
node.labels	Boolean value that indicates whether node breaks are plotted on the tree. Only has meaning if `plotTree=TRUE`.
cex	Specifies the font size for printing tip labels. Default is `cex=0.5`; may need to decrease for larger trees. Only has meaning if `plotTree=TRUE`.
plotSurface	Boolean value that indicates whether likelihood surface(s) for the selected piecewise model is plotted. If `plotSurface=TRUE` the optimal model is evaluated from 1e-10 to 1 at `n.points` (see below), and contours are plotted. Default is no plot.
printTitle	Whether to print titles on surface plots.
n.points	The number of values between 1e-10 and 1 for each of r and epsilon to evaluate the selected piecewise model (hence, a total of `n.points` x `n.points` parameter combinations considered). A greater value of `n.points` gives better resolution of the surface, but comes with a computational burden. Default is `n.points=100`. Only has meaning if `plotSurface=T`.
...	Potential additional arguments for plotting etc.

Value

Returned list has elements:

$z	A matrix listing branch times and richnesses, coded by model membership.
$mm	Colour-coding for edges.
$break.point	The position in the tree where shifts in models occur.
$phy	The tree.

Author(s)

Joseph W. Brown, Richard G. FitzJohn, Michael E. Alfaro, and Luke J. Harmon.

Maintainer: Joseph W. Brown <josephwb@uidaho.edu>

References

Alfaro, ME, F Santini, C Brock, H Alamillo, A Dornburg, DL Rabosky, G Carnevale, and LJ Harmon. 2009. Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proceedings of the National Academy of Sciences* **106**: 13410-13414.

See Also

runTurboMEDUSA plotPrettyTree

---

turboMEDUSA	*MEDUSA: Modeling Evolutionary Diversification Using Stepwise AIC*

Description

Fits piecewise birth-death models to an ultrametric phylogenetic tree according to phylogenetic (edge-length) and taxonomic (richness) likelihoods. A reimplementation of the original **geiger** MEDUSA method, meant especially for the analysis of large trees. Optimal model size is determined via a stepwise AIC approach. Run with `runTurboMEDUSA` and summarize results with `summarizeTurboMEDUSA`. Trees with coloured branches (depending on model membership) can be plotted using `plotPrettyTree`.

Details

Package:	turboMEDUSA
Type:	Package
Version:	0.19
Date:	2011-08-03
License:	GPL version 2 or greater
LazyLoad:	yes

Author(s)

Joseph W. Brown, Richard G. FitzJohn, Michael E. Alfaro, and Luke J. Harmon.

Maintainer: Joseph W. Brown <josephwb@uidaho.edu>

References

Alfaro, ME, F Santini, C Brock, H Alamillo, A Dornburg, DL Rabosky, G Carnevale, and LJ Harmon. 2009. Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proceedings of the National Academy of Sciences* **106**: 13410-13414.

---

turboMEDUSA-internal	*Turbo-MEDUSA Internal Functions*

Description

`turboMEDUSA-internal` are either not typically called by the user or are currently undocumented

Details

This is an internal **turboMEDUSA** function, not intended to be called directly by the user.

Author(s)

Joseph W. Brown, Richard G. FitzJohn, Michael E. Alfaro, and Luke J. Harmon.

Maintainer: Joseph W. Brown <josephwb@uidaho.edu>

References

Alfaro, ME, F Santini, C Brock, H Alamillo, A Dornburg, DL Rabosky, G Carnevale, and LJ Harmon. 2009. Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proceedings of the National Academy of Sciences* **106**: 13410-13414.

---

whales   *Whale Phylogeny*

Description

Tree, species richnesses.

Usage

```
data(whales)
```

Format

The data are stored in RData (binary) format.

Source

Graham Slater, PaleoBioDB.

Examples

```
data(whales)
str(whales)
```

Index