

Package: closesamples (via r-universe)

August 18, 2024

Type Package

Title Gets feasible taxon samples on a tree

Version 1.0

Date 2020-02-25

Author Brian C. O'Meara

Maintainer Brian C. O'Meara <omeara.brian@gmail.com>

Description Often you want to pick species at random, but don't have access to all the possible species, and your group might be a nonrandom set. However, you might have a taxonomy, perhaps a bad one, for all the species. This will let you pick species at random on that full taxonomy and then pick a species out of the possible set closest to this. You could still have a biased sample, but way less biased.

License GPL 3.0

Imports ape, data.table, phangorn, taxize, Rfast2

Suggests anRpackage, testthat (>= 2.1.0)

Encoding UTF-8

RoxygenNote 7.1.0

Repository <https://phylotastic.r-universe.dev>

RemoteUrl <https://github.com/bomeara/closesamples>

RemoteRef HEAD

RemoteSha 1beb4a721edb00887d2fbcd05c7995bb35dc6c30

Contents

GetClosest	2
GetClosestSamples	2
GetEnclosingTaxonomy	3
GetSimilarity	4
LabelNodesWithChosenDescendants	4
LabelNodesWithFeasibleDescendants	5
SubsampleTree	5

Index[7](#)

GetClosest	<i>Get the closest taxon to the focal taxon</i>
------------	---

Description

Get the closest taxon to the focal taxon

Usage

```
GetClosest(focal_taxon, similarity_matrix)
```

Arguments

focal_taxon	String with taxon name to find closest match to
similarity_matrix	Matrix from GetSimilarity function

Value

The taxon closest to it

GetClosestSamples	<i>Get n samples</i>
-------------------	----------------------

Description

Get n samples

Usage

```
GetClosestSamples(
  n,
  phy_full,
  taxa_feasible,
  replace_full = TRUE,
  replace_feasible = FALSE,
  truncate_full_to_mrca = FALSE,
  less_memory = FALSE,
  descendant_labeled = FALSE,
  fast_ultrametric = FALSE,
  verbose = TRUE
)
```

Arguments

n	How many taxa to sample
phy_full	A phylo object with all feasible taxa to sample from
taxa_feasible	A vector of taxon names that are feasible to study (often on another tree)
replace_full	If TRUE, will allow selecting the same taxon from the full tree more than once. If FALSE, forbids this. Both cases return n taxa
replace_feasible	If TRUE, will allow selecting the same taxon from the set of feasible taxa more than once (returning a smaller than desired tree). If FALSE, forbids this.
truncate_full_to_mrca	If TRUE, prune the full tree to the node that is the MRCA of the
less_memory	If TRUE, uses a much slower approach that will not create giant matrices
descendant_labeled	If TRUE, assumes the phy_full has been labeled with LabelNodesWithFeasibleDescendants
fast_ultrametric	If TRUE, uses a fast algorithm for ultrametric trees
verbose	If TRUE, all the output will print to the screen

Value

A data.frame of chosen taxa, closest feasible match, and distance between them

GetEnclosingTaxonomy *Get enclosing taxonomy for a feasible tree*

Description

If you don't have a taxonomy tree for the feasible tree, this will find one

Usage

```
GetEnclosingTaxonomy(phy_feasible)
```

Arguments

tips	A vector of taxon names
------	-------------------------

Value

A phylogeny of the enclosing taxonomy

GetSimilarity	<i>Get similarity Uses ape's cophenetic.phylo to get patristic distances between pairs of taxa, then makes the diagonals Inf. If the tree has no branch lengths, makes every edge length 1.</i>
---------------	---

Description

Get similarity Uses ape's cophenetic.phylo to get patristic distances between pairs of taxa, then makes the diagonals Inf. If the tree has no branch lengths, makes every edge length 1.

Usage

```
GetSimilarity(phy_full, taxa_feasible, truncate_full_to_mrca = FALSE)
```

Arguments

phy_full	A phylo object with all feasible taxa to sammple from
taxa_feasible	A vector of taxon names that are feasible to study (often on another tree)
truncate_full_to_mrca	If TRUE, prune the full tree to the node that is the MRCA of the

Value

A matrix with tips on the rows and columns and patristic distances in the cells. Rows are tips on the phy_full tree, columns are taxa in the taxa_feasible vector

LabelNodesWithChosenDescendants	<i>Include descendant taxon ids in node labels</i>
---------------------------------	--

Description

Include descendant taxon ids in node labels

Usage

```
LabelNodesWithChosenDescendants(node, phy, clean = FALSE, sep = ", ")
```

Arguments

node	node number, typically a tip
phy	phylo object
clean	wipe existing node labels
sep	Separator

Value

A phylogeny with terminal node numbers as node.labels. Unfortunately, each will start with NA.

LabelNodesWithFeasibleDescendants

Include descendant taxon ids in node labels

Description

Include descendant taxon ids in node labels

Usage

```
LabelNodesWithFeasibleDescendants(taxa_feasible, phy)
```

Arguments

taxa_feasible vector of taxon names that are feasible for later suse
 phy phylo object

Value

A phylogeny with terminal node numbers as node.labels.

SubsampleTree

Return a sampled tree

Description

Return a sampled tree

Usage

```
SubsampleTree(  
  phy_feasible,  
  n,  
  phy_full = NULL,  
  replace_full = TRUE,  
  replace_feasible = FALSE,  
  truncate_full_to_mrca = FALSE,  
  less_memory = FALSE,  
  verbose = TRUE  
)
```

Arguments

phy_feasible	The tree to subsample
n	The number of taxa to include
phy_full	The larger tree giving relationships
replace_full	If TRUE, will allow selecting the same taxon from the full tree more than once. If FALSE, forbids this. Both cases return n taxa
replace_feasible	If TRUE, will allow selecting the same taxon from the set of feasible taxa more than once (returning a smaller than desired tree). If FALSE, forbids this.
truncate_full_to_mrca	If TRUE, prune the full tree to the node that is the MRCA of the feasible tree
less_memory	If TRUE, uses a much slower approach that will not create giant matrices
verbose	If TRUE, all the output will print to the screen

Value

A phylo object where taxa are sampled based on representing flat sampling from taxonomy

Index

GetClosest, [2](#)

GetClosestSamples, [2](#)

GetEnclosingTaxonomy, [3](#)

GetSimilarity, [4](#)

LabelNodesWithChosenDescendants, [4](#)

LabelNodesWithFeasibleDescendants, [5](#)

SubsampleTree, [5](#)