

Package ‘TreeSim’

February 14, 2012

Type Package

Title Simulating trees under the birth-death model

Version 1.5

Date 2011-06-14

Author Tanja Stadler

Maintainer Tanja Stadler <tanja.stadler@env.ethz.ch>

Depends ape, geiger

Description The package simulates phylogenetic trees under a constant-rate birth-death process, conditioned on having a fixed number of final tips (`sim.bd.taxa`), or a fixed age (`sim.bd.age`), or a fixed age and number of tips (`sim.bd.taxa.age`). When conditioning on the number of final tips, the method allows for shifts in rates and mass extinction events during the birth-death process (`sim.rateshift.taxa`). When fixing on the age, the method further allows the speciation rate to change in a density-dependent way (`sim.bd.age`), and one can plot the average LTT plot (`LTT.average.root`). TreeSim further samples appropriately trees with n final tips from a set of trees generated by the common sampling algorithm stopping when a fixed number $m \gg n$ of leaves is first reached (`bd.gsa.taxa`). This latter method is appropriate for m -tip trees generated under a big class of models (details in the `bd.gsa.taxa` man page). For incomplete phylogeny, the missing speciation events can be added through simulations (`corsim`).

License GPL-2

LazyLoad yes

Repository CRAN

Date/Publication 2011-11-02 17:17:25

R topics documented:

TreeSim-package	2
corsim	3
LTT.average.root	4
sim.bd.age	5
sim.bd.taxa	7
sim.bd.taxa.age	8
sim.gsa.taxa	10
sim.rateshift.taxa	11

Index	14
--------------	-----------

TreeSim-package	<i>TreeSim: Simulating phylogenetic trees</i>
-----------------	---

Description

The package simulates phylogenetic trees under a constant-rate birth-death process, conditioned on having a fixed number of final tips (`sim.bd.taxa`), or a fixed age (`sim.bd.age`), or a fixed age and number of tips (`sim.bd.taxa.age`). When conditioning on the number of final tips, the method allows for shifts in rates and mass extinction events during the birth-death process (`sim.rateshift.taxa`). When fixing on the age, the method further allows the speciation rate to change in a density-dependent way (`sim.bd.age`), and one can plot the average LTT plot (`LTT.average.root`). TreeSim further samples appropriately trees with n final tips from a set of trees generated by the common sampling algorithm stopping when a fixed number $m \gg n$ of leaves is first reached (`bd.gsa.taxa`). This latter method is appropriate for m -tip trees generated under a big class of models (details in the `bd.gsa.taxa` man page). For incomplete phylogeny, the missing speciation events can be added through simulations (`corsim`).

Details

Package:	TreeSim
Type:	Package
Version:	1.5
Date:	2011-06-14
License:	GPL-2
LazyLoad:	yes

Author(s)

Tanja Stadler

<<http://www.tb.ethz.ch/people/tstadler>>

References

- Simulating trees on a fixed number of extant species. *Syst. Biol.* (2011).
- K. Hartmann, D. Wong, T. Stadler: Sampling trees from evolutionary models. *Syst. Biol.* (2010) 59: 465-476.
- T. Stadler: On incomplete sampling under birth-death models and connections to the sampling-based coalescent. *J. Theo. Biol.* (2009) 261: 58-66.
- N. Cusimano, T. Stadler, S. Renner: A new method for handling missing species in diversification analysis applicable to randomly or non-randomly sampled phylogenies. Submitted.

See Also

ape geiger

corsim	<i>corsim: Simulating the missing speciation events in an incomplete phylogenies</i>
--------	--

Description

corsim simulates the missing speciation event in an incomplete phylogeny assuming a constant speciation and extinction rate. These rates can be estimated with the functions `bd.shifts.optim` (if random speciation events are missing) and `bd.groups.optim` (if only young speciation events are missing) provided in `TreePar`. `corsim` allows to specify an upper and lower bound for the times of the missing speciation events.

Usage

```
corsim(x,lambda,mu,missing,told=0,tyoung=0)
```

Arguments

x	Vector of the speciation times in the incomplete phylogeny (where time is measured such that 0 is the present and increasing going into the past).
lambda	Speciation rate
mu	Extinction rate
missing	Number of missing species (i.e. speciation events).
told	Upper bound for the time of missing speciation events. Default is 0 which means no upper bound.
tyoung	Lower bound for the time of missing speciation events. Default is 0 which means no lower bound. <code>tyoung < told</code> unless <code>tyoung=told=0</code> is required.

Value

x	Vector of input and simulated speciation times.
---	---

Author(s)

Tanja Stadler

References

N. Cusimano, T. Stadler, S. Renner: A new method for handling missing species in diversification analysis applicable to randomly or non-randomly sampled phylogenies. Submitted.

See Also

sim.bd.age, sim.rateshift.taxa, sim.gsa.taxa, birthdeath.tree

Examples

```
# Speciation times of a five species tree:
x<-c(1,1.5,3,5)
# We simulate using the following parameters:
lambda<-2
mu<-1
tyoung<-0.5
told<-4.5
# We simulate 5 additional speciation times (i.e. five additional species):
missing<-5

# xcompleted is x plus 5 additional speciation events between 0.5 and 4.5 timesteps in the past. xcompleted corresponds to
xcompleted<-corsim(x,lambda,mu,missing,told,tyoung)
```

LTT.average.root	<i>LTT.average.root: plots the average LTT plot of a set of phylogenetic trees</i>
------------------	--

Description

LTT.average.root plots the average LTT plot of a set of phylogenetic trees of the same age (ie. time since root). The trees may be simulated using sim.bd.age in TreeSim.

Usage

```
LTT.average.root(trees)
```

Arguments

trees	list of phylogenetic trees
-------	----------------------------

Value

l tt	Vector containing average LTT plot. Visualize with plot(ltt,type='l',col='black',log="y",xlab="time",ylab="of species")
------	---

Author(s)

Tanja Stadler

References

T. Stadler: Simulating trees on a fixed number of extant species. Syst. Biol. (2011).

See Also

sim.bd.age, sim.rateshift.taxa, sim.gsa.taxa, birthdeath.tree

Examples

```

numbsim<-10
age<-10
lambda<-0.3
mu<-0.1
K<-40

tree<- sim.bd.age(age,numbsim,lambda,mu,mrca=TRUE,complete=FALSE,K=K)
littrees<-LTT.average.root(tree)
plot(littrees,type='l',col='black',log="y",xlab="time",ylab="number of species")

```

sim.bd.age

*sim.bd.age: Simulating birth-death trees of a fixed age***Description**

sim.bd.age simulates trees conditioned on (i) the time since origin or (ii) the time since the most recent common ancestor of the extant tips under the constant rate birth-death process. The method allows for incomplete sampling: only a fixed fraction of all tips is included in the sampled tree. The method assumes constant birth and death rates, or allows for a density-dependent birth rate.

Usage

```

sim.bd.age(age, numbsim, lambda, mu, frac = 1, mrca = FALSE,
  complete = TRUE, K = 0)

```

Arguments

age	Time since origin / most recent common ancestor
numbsim	Number of trees to simulate
lambda	Speciation rate
K	If K=0, then lambda is constant. If K>0, density-dependent speciation is assumed, with speciation rate $\lambda(1-m/K)$ when there are m extant species.
mu	Extinction rate

frac	Sampling fraction: The actual number of tips is n/frac, but only n tips are included (incomplete sampling).
mrca	If mrca=FALSE: age is time since origin. If mrca=TRUE: age is time since most recent common ancestor of the extant tips.
complete	If complete = TRUE, the tree with the extinct and non-sampled lineages is returned. If complete = FALSE, the extinct and non-sampled lineages are suppressed.

Value

treearray	Array of numbsim trees with the time since origin / most recent common ancestor being age. If tree goes extinct (only possible when mrca = FALSE), returns 0. If only one extant and no extinct tips, returns 1.
-----------	--

Author(s)

Tanja Stadler

References

Simulating trees on a fixed number of extant species. Syst. Biol. (2011).

See Also

sim.bd.taxa, sim.rateshift.taxa, sim.gsa.taxa, birthdeath.tree

Examples

```
age<-2
lambda <- 2.0
mu <- 0.5
frac <-0.6
numbsim<-3

##
# Simulating trees with time age since the origin:

sim.bd.age(age,numbsim,lambda,mu,mrca=FALSE,complete=TRUE)
sim.bd.age(age,numbsim,lambda,mu,frac,mrca=FALSE,complete=FALSE)

# Simulating trees with time age since the time of the most recent common
# ancestor of the extant and sampled species:

sim.bd.age(age,numbsim,lambda,mu,mrca=TRUE,complete=TRUE)
sim.bd.age(age,numbsim,lambda,mu,frac,mrca=TRUE,complete=FALSE)
```

sim.bd.taxa	<i>sim.bd.taxa: Simulating birth-death trees on a fixed number of extant taxa</i>
-------------	---

Description

sim.bd.taxa simulates trees on n species under the constant rate birth-death process. The method allows for incomplete sampling, i.e. (i) only a fixed fraction of all extant tips is included in the sampled tree or (ii) each tip from a big tree is included with a fixed probability. In both cases, the tree is conditioned to have n tips after sampling.

Usage

```
sim.bd.taxa(n, numbsim, lambda, mu, frac = 1, complete = TRUE,
  stochsampling = FALSE)
```

Arguments

n	Number of extant sampled leaves
numbsim	Number of trees to simulate
lambda	Speciation rate
mu	Extinction rate
frac	When complete = FALSE: Sampling fraction: If stochsampling=FALSE: The actual number of tips is n/frac , but only n tips are included (incomplete sampling). If stochsampling=TRUE: Each tip is included into the final tree with probability frac . When complete = TRUE: all extinct and non-sampled lineages are included, i.e. the tree has n/frac extant leaves.
complete	If TRUE, the tree with the extinct and non-sampled lineages is returned. If FALSE, the extinct lineages are suppressed.
stochsampling	See frac.

Value

out[[1]]	List of numbsim simulated trees with n extant sampled leaves
out[[2]]	Vector of age of trees (time to root plus length of edge above root)

Note

For stochsampling = TRUE: The algorithm is fast for the critical process, $\lambda=\mu$.

Author(s)

Tanja Stadler

References

- T. Stadler: Simulating trees on a fixed number of extant species. *Syst. Biol.* (2011).
 T. Stadler: On incomplete sampling under birth-death models and connections to the sampling-based coalescent. *J. Theo. Biol.* (2009) 261: 58-66.

See Also

sim.bd.age, sim.rateshift.taxa, sim.gsa.taxa, birthdeath.tree

Examples

```
n<-10
lambda <- 2.0
mu <- 0.5
frac <-0.6
numbsim<-2

##
# Simulating numbsim trees with n species under a birth-death process with
# speciation rate lambda an extinction rate mu:

sim.bd.taxa(n,numbsim,lambda,mu)

# Each extant species is included in final tree with probability frac
# (the tree has n species AFTER sampling):

sim.bd.taxa(n,numbsim,lambda,mu,frac,complete=FALSE,stochsampling=TRUE)

# A fraction frac of the extant species is included into the final tree
# (the tree has n species AFTER sampling):

sim.bd.taxa(n,numbsim,lambda,mu,frac,complete=FALSE,stochsampling=FALSE)
```

sim.bd.taxa.age	<i>sim.bd.taxa.age: Simulating birth-death trees with a given age on a fixed number of extant taxa</i>
-----------------	--

Description

sim.bd.taxa.age simulates trees on n species with a (i) fixed time since origin or (ii) fixed time since the most recent common ancestor of the sampled species under the constant rate birth-death process. The method allows for incomplete sampling, i.e. each tip from a big tree is included with a fixed probability. The tree is conditioned to have n tips after sampling and a fixed time since origin or since the most recent common ancestor of the sampled species.

Usage

```
sim.bd.taxa.age(n, numbsim, lambda, mu, frac = 1, age, mrca = FALSE)
```

Arguments

n	Number of extant sampled leaves
numbsim	Number of trees to simulate
lambda	Speciation rate
mu	Extinction rate
frac	Each tip is included into the final tree with probability frac.
age	The time since origin / most recent common ancestor
mrca	If mrca = FALSE: The time since the origin of the process. If mrca = TRUE: The time since the most recent common ancestor of the sampled species.

Value

treearray	Array of numbsim trees with n>1 leaves with a given age. The extinct lineages are not included.
-----------	---

Note

The algorithm is fast for the critical process, lambda=mu.

Author(s)

Tanja Stadler

References

T. Stadler: On incomplete sampling under birth-death models and connections to the sampling-based coalescent. *J. Theo. Biol.* (2009) 261: 58-66.

See Also

sim.bd.age, sim.rateshift.taxa, sim.gsa.taxa, birthdeath.tree

Examples

```
n<-10
lambda <- 2.0
mu <- 0.5
frac <-0.6
numbsim<-2
age<-2

##
# Simulating numbsim reconstructed trees with n sampled species under a
# birth-death process with speciation rate lambda, extinction rate mu,
# sampling probability frac, and time age since origin:

sim.bd.taxa.age(n, numbsim, lambda, mu, frac, age, mrca = FALSE)

# Simulating numbsim reconstructed trees with n sampled species under a
```

```
# birth-death process with speciation rate lambda, extinction rate mu,
# sampling probability frac, and time age since the most recent
# common ancestor of the extant sampled species:

sim.bd.taxa.age(n, numbsim, lambda, mu, frac, age, mrca = TRUE)
```

 sim.gsa.taxa

sim.gsa.taxa: Sampling trees on n tips from bigger trees

Description

sim.gsa.taxa samples trees on n tips appropriately (using the GSA approach, see references) from bigger trees simulated under the simple sampling approach (i.e. simulating until first $m \gg n$ tips are reached or the tree is extinct). This methods works for m-tip trees generated under a model where: (i) the number of tips eventually tends to zero or stays bigger than n and (ii) rate changes do not depend on the time between the change and the present - e.g. one cannot model a mass extinction event 1 million years BEFORE the present. But one can model a mass extinction event 1 million years AFTER the origin of the tree.

Note that the number of trees in the output is significantly smaller than the number of trees in the input (in order to have correct sampling).

Usage

```
sim.gsa.taxa(treearray, n, frac = 1, sampling = 40, complete = TRUE)
```

Arguments

treearray	Array of trees with a fixed number of leaves
n	Number of tips in sampled trees
frac	Relevant when complete = FALSE: The actual number of tips is n/frac, but only n tips are included (incomplete sampling). When complete = TRUE: We set frac = 1.
sampling	Default is 40, the higher the closer the returned trees in treearray are to the "true" distribution, but the more trees are required in treearray to reach a big sample of output trees.
complete	If TRUE, the tree with the extinct lineages is returned. If FALSE, the extinct lineages are suppressed.

Value

treearray	Array of sampled trees with n extant sampled leaves
-----------	---

Author(s)

Tanja Stadler

References

- T. Stadler: Simulating trees on a fixed number of extant species. Submitted. 2010.
- K. Hartmann, D. Wong, T. Stadler: Sampling trees from evolutionary models. *Syst. Biol.* (2010) 59: 465-476.

See Also

sim.bd.age, sim.bd.taxa, sim.rateshift.taxa, birthdeath.tree

Examples

```
##
# First 100 trees on 9 leaves under a birth-death process are generated.
# Then trees on 5 species are sampled from these 100 trees using the GSA
# (see references).
# You can easily simulate trees on m species with the simple sampling
# approach (see references) under a variety of models. Then use the
# provided GSA algorithm to get a correct sample of trees on n<<m species:

m<-9
n<-5
numbsim<-100
lambda <- 2.0
mu <- 0.5

t<-sim.bd.taxa(m,numbsim,lambda,mu)
t2<-sim.gsa.taxa(t,n)
```

sim.rateshift.taxa	<i>sim.rateshift.taxa: Simulating trees incorporating mass extinction events and rate shifts</i>
--------------------	--

Description

sim.rateshift.taxa simulates trees on n species under the constant rate birth-death process. At user-specified points in the past, the rates can shift. Further, mass extinction events can be incorporated. The method further allows for incomplete sampling, i.e. only a fixed fraction of all tips is included in the sampled tree. The tree is conditioned to have n tips after sampling.

Usage

```
sim.rateshift.taxa(n, numbsim, lambda, mu, frac, times, complete = TRUE)
```

Arguments

n	Number of extant sampled leaves
numbsim	Number of trees to simulate
lambda	Vector of speciation rates, the rate in entry i is the speciation rate prior to time times[i]
mu	Vector of extinction rates, the rate in entry i is the extinction rate prior to time times[i].
frac	Vector of proportion of species surviving mass extinction event. Entry i corresponds to the mass extinction at time times[i]. If frac[i]=1, only rate shift but no mass extinction at time times[i].
times	Vector of mass extinction and rate shift times. Time is 0 today and increasing going backwards in time. Specify the vector as times[i]<times[i+1]. times[1]=0 (today).
complete	If TRUE, the tree with the extinct lineages and non-sampled lineages is returned (so the tree has round(n/frac[1]) extant tips). If FALSE, the extinct lineages and non-sampled lineages are suppressed.

Value

out[[1]]	List of numbsim simulated trees with n extant sampled leaves
out[[2]]	Vector of age of trees (time to root plus length of edge above root)

Author(s)

Tanja Stadler

References

T. Stadler: Simulating trees on a fixed number of extant species. Syst. Biol. (2011).

See Also

sim.bd.age, sim.bd.taxa, sim.gsa.taxa, birthdeath.tree

Examples

```
n<-10
numbsim<-1

##
# Simulating trees with a fixed number of species having shifts in rate
# and mass extinction events.
# Between today and time 0.3 in the past, we have speciation rate 2,
# extinction rate 0. At time 0.3, we have a mass extinction event which
# 10% of the species survive. Prior to 0.3, we have a speciation rate
# of 1 and an extinction rate of 0.3:

sim.rateshift.taxa(n,numbsim,c(2,1),c(0,0.3),
```

```
c(1,0.1),c(0,0.3),complete=TRUE)

# The fraction 0.6 of the extant species is included into the final tree
# (the tree has n species AFTER sampling, extinct and
# non-sampled lineages are not included):

sim.rateshift.taxa(n,numbsim,c(2,1),c(0,0.3),
c(0.6,0.1),c(0,0.3),complete=FALSE)
```

Index

`corsim`, [3](#)

`LTT.average.root`, [4](#)

`sim.bd.age`, [5](#)

`sim.bd.taxa`, [7](#)

`sim.bd.taxa.age`, [8](#)

`sim.gsa.taxa`, [10](#)

`sim.rateshift.taxa`, [11](#)

`TreeSim (TreeSim-package)`, [2](#)

`TreeSim-package`, [2](#)