

Package ‘DDD’

July 11, 2023

Type Package

Title Diversity-Dependent Diversification

Version 5.2.2

Date 2023-07-10

Depends R (>= 3.5.0)

Imports deSolve, ape, phytools, subplex, DEoptim, Matrix, expm,
SparseM, Rcpp (>= 1.0.10)

LinkingTo Rcpp (>= 1.0.10), RcppEigen, BH (>= 1.81.0-1)

Suggests testthat (>= 3.0.0), testit, igraph

License GPL-3

Encoding UTF-8

Description Implements maximum likelihood and bootstrap methods based on the diversity-dependent birth-death process to test whether speciation or extinction are diversity-dependent, under various models including various types of key innovations. See Etienne et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, <[DOI:10.1098/rspb.2011.1439](https://doi.org/10.1098/rspb.2011.1439)>, Etienne & Haegeman 2012, Am. Nat. 180: E75-E89, <[DOI:10.1086/667574](https://doi.org/10.1086/667574)>, Etienne et al. 2016. Meth. Ecol. Evol. 7: 1092-1099, <[DOI:10.1111/2041-210X.12565](https://doi.org/10.1111/2041-210X.12565)> and Laudanno et al. 2021. Syst. Biol. 70: 389-407, <[DOI:10.1093/sysbio/syaa048](https://doi.org/10.1093/sysbio/syaa048)>. Also contains functions to simulate the diversity-dependent process.

RoxygenNote 7.2.3

Config/testthat/edition 3

NeedsCompilation yes

Author Rampal S. Etienne [aut, cre] (<<https://orcid.org/0000-0003-2142-7612>>),
Bart Haegeman [aut] (<<https://orcid.org/0000-0003-2325-4727>>),
Hanno Hildenbrandt [ctb] (<<https://orcid.org/0000-0002-6784-1037>>),
Giovanni Laudanno [ctb] (<<https://orcid.org/0000-0002-2952-3345>>)

Maintainer Rampal S. Etienne <r.s.etienne@rug.nl>

Repository CRAN

Date/Publication 2023-07-10 23:40:47 UTC

R topics documented:

bd_loglik	2
bd_ML	4
brts2phylo	7
conv	8
dd_KI_loglik	8
dd_KI_ML	11
dd_KI_sim	14
dd_loglik	16
dd_LR	18
dd_ML	21
dd_MS_loglik	25
dd_MS_ML	27
dd_MS_sim	30
dd_multiple_KI_loglik	32
dd_sim	34
dd_SR_loglik	35
dd_SR_ML	37
dd_SR_sim	41
L2brts	42
L2phylo	43
optimizer	44
phylo2L	45
rng_respecting_sample	46
roundn	48
sample2	49
simplex	49
td_sim	50
transform_pars	51
untransform_pars	52
Index	53

bd_loglik

Loglikelihood for diversity-independent diversification model

Description

This function computes loglikelihood of a diversity-independent diversification model for a given set of branching times and parameter values.

Usage

```

bd_loglik(
  pars1,
  pars2,
  brts,
  missspec,
  methode = "odeint::runge_kutta_cash_karp54"
)

```

Arguments

pars1	<p>Vector of parameters:</p> <p>pars1[1] corresponds to lambda0 (speciation rate) pars1[2] corresponds to mu0 (extinction rate) pars1[3] corresponds to lambda1 (decline parameter in speciation rate) or K in diversity-dependence-like models pars1[4] corresponds to mu1 (decline parameter in extinction rate)</p>
pars2	<p>Vector of model settings:</p> <p>pars2[1] sets the model of time-dependence: - pars2[1] == 0 no time dependence - pars2[1] == 1 speciation and/or extinction rate is exponentially declining with time - pars2[1] == 2 stepwise decline in speciation rate as in diversity-dependence without extinction - pars2[1] == 3 decline in speciation rate following deterministic logistic equation for dmodel = 1 - pars2[1] == 4 decline in speciation rate such that the expected number of species matches with that of dmodel = 1 with the same mu</p> <p>pars2[2] sets the conditioning: - pars[2] == 0 conditioning on stem or crown age - pars[2] == 1 conditioning on stem or crown age and non-extinction of the phylogeny - pars[2] == 2 conditioning on stem or crown age and on the total number of extant taxa (including missing species) - pars[2] == 3 conditioning on the total number of extant taxa (including missing species)</p> <p>pars2[3] sets whether the likelihood is for the branching times (0) or the phylogeny (1)</p> <p>pars2[4] sets whether the parameters and likelihood should be shown on screen (1) or not (0)</p> <p>pars2[5] sets whether the first data point is stem age (1) or crown age (2)</p>
brts	<p>A set of branching times of a phylogeny, all positive</p>

missnumspec The number of species that are in the clade but missing in the phylogeny
 methode The method used to solve the master equation, default is 'odeint::runge_kutta_cash_karp54'.

Value

The loglikelihood

Author(s)

Rampal S. Etienne, Bart Haegeman & Cesar Martinez

References

- Etienne, R.S. et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. Am. Nat. 180: E75-E89, doi: 10.1086/667574

See Also

[bd_ML](#)

Examples

```
bd_loglik(pars1 = c(0.5,0.1), pars2 = c(0,1,1,0,2), brts = 1:10,
missnumspec = 0)
```

bd_ML	<i>Maximization of the loglikelihood under the diversity-independent, possibly time-dependent diversification model</i>
-------	---

Description

This function computes the maximum likelihood estimates of the parameters of a diversity-independent diversification model for a given set of phylogenetic branching times. It also outputs the corresponding loglikelihood that can be used in model comparisons.

Usage

```
bd_ML(
  brts,
  initparsopt = c(0.1, 0.05 * (tdmodel <= 1) + 10 * (length(brts) + missnumspec) *
    (tdmodel > 1)),
  idparsopt = c(1, 2 + (tdmodel > 1)),
  idparsfix = (1:4)[-idparsopt],
  parsfix = rep(0, 4)[idparsfix],
  missnumspec = 0,
  tdmodel = 0,
  cond = 1,
  btorph = 1,
```

```

soc = 2,
tol = c(0.001, 1e-04, 1e-06),
maxiter = 1000 * round((1.25)^length(idparsopt)),
changeloglikifnoconv = FALSE,
optimmethod = "subplex",
num_cycles = 1,
methode = "odeint::runge_kutta_cash_karp54",
verbose = FALSE
)

```

Arguments

brts	A set of branching times of a phylogeny, all positive
initparsopt	The initial values of the parameters that must be optimized
idparsopt	The ids of the parameters that must be optimized, e.g. 1:3 for intrinsic speciation rate, extinction rate and carrying capacity. The ids are defined as follows: id == 1 corresponds to lambda0 (speciation rate) id == 2 corresponds to mu0 (extinction rate) id == 3 corresponds to lamda1 (parameter controlling decline in speciation rate with time) id == 4 corresponds to mu1 (parameter controlling decline in extinction rate with time)
idparsfix	The ids of the parameters that should not be optimized, e.g. c(1,3) if lambda0 and lambda1 should not be optimized, but only mu0 and mu1. In that case idparsopt must be c(2,4). The default is to fix all parameters not specified in idparsopt.
parsfix	The values of the parameters that should not be optimized
missnumspec	The number of species that are in the clade but missing in the phylogeny
tdmodel	Sets the model of time-dependence: tdmodel == 0 : constant speciation and extinction rates tdmodel == 1 : speciation and/or extinction rate is exponentially declining with time tdmodel == 2 : stepwise decline in speciation rate as in diversity-dependence without extinction tdmodel == 3 : decline in speciation rate following deterministic logistic equation for dmodel = 1 tdmodel == 4 : decline in speciation rate such that the expected number of species matches with that of dmodel = 1 with the same mu
cond	Conditioning: cond == 0 : conditioning on stem or crown age cond == 1 : conditioning on stem or crown age and non-extinction of the phylogeny cond == 2 : conditioning on stem or crown age and on the total number of extant taxa (including missing species) cond == 3 : conditioning on the total number of extant taxa (including missing species)

btorph	Sets whether the likelihood is for the branching times (0) or the phylogeny (1)
soc	Sets whether stem or crown age should be used (1 or 2)
tol	Sets the tolerances in the optimization. Consists of: reltolx = relative tolerance of parameter values in optimization reltolf = relative tolerance of function value in optimization abstolx = absolute tolerance of parameter values in optimization
maxiter	Sets the maximum number of iterations in the optimization
changeloglikifnoconv	if TRUE the loglik will be set to -Inf if ML does not converge
optimmethod	Method used in optimization of the likelihood. Current default is 'subplex'. Alternative is 'simplex' (default of previous versions)
num_cycles	the number of cycles of optimization. If set at Inf, it will do as many cycles as needed to meet the tolerance set for the target function.
methode	The method used to solve the master equation under <code>tdmodel = 4</code> , default is 'odeint::runge_kutta_cash_karp54'.
verbose	Show the parameters and loglikelihood for every call to the loglik function

Details

The output is a dataframe containing estimated parameters and maximum loglikelihood. The computed loglikelihood contains the factor $q! m! / (q + m)!$ where q is the number of species in the phylogeny and m is the number of missing species, as explained in the supplementary material to Etienne et al. 2012.

Value

A dataframe with the following elements:

lambda0	gives the maximum likelihood estimate of lambda0
mu0	gives the maximum likelihood estimate of mu0
lambda1	gives the maximum likelihood estimate of lambda1
mu1	gives the maximum likelihood estimate of mu1
loglik	gives the maximum loglikelihood
df	gives the number of estimated parameters, i.e. degrees of freedom
conv	gives a message on convergence of optimization; conv = 0 means convergence

Author(s)

Rampal S. Etienne & Bart Haegeman

References

- Etienne, R.S. et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. Am. Nat. 180: E75-E89, doi: 10.1086/667574

See Also[bd_loglik](#)**Examples**

```
cat("Estimating parameters for a set of branching times brts with the default settings:")
brts = 1:20
bd_ML(brts = brts, cond = 1)
```

brts2phylo	<i>Function to convert a set of branching times into a phylogeny with random topology This code is taken from the package TESS by Sebastian Hoehna, where the function is called tess.create.phylo</i>
------------	--

Description

Converting a set of branching times to a phylogeny

Usage

```
brts2phylo(times, root = FALSE, tip.label = NULL)
```

Arguments

times	Set of branching times
root	When root is FALSE, the largest branching time will be assumed to be the crown age. When root is TRUE, it will be the stem age.
tip.label	Tip labels. If set to NULL, the labels will be t1, t2, etc.

Value

phy	A phylogeny of the phylo type
-----	-------------------------------

Author(s)

Rampal S. Etienne

References

- Etienne, R.S. et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. Am. Nat. 180: E75-E89, doi: 10.1086/667574

conv

Function to do convolution of two vectors

Description

Convolution of two vectors

Usage

```
conv(x, y)
```

Arguments

x	first vector
y	second vector

Value

vector that is the convolution of x and y

Author(s)

Rampal S. Etienne

References

- Etienne, R.S. et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. Am. Nat. 180: E75-E89, doi: 10.1086/667574

Examples

```
conv(1:10, 1:10)
```

dd_KI_loglik*Loglikelihood for diversity-dependent diversification models with decoupling of a subclade from a main clade at time $t = t_d$*

Description

This function computes loglikelihood of a diversity-dependent diversification model for a given set of branching times and parameter values where the diversity-dependent dynamics of a subclade decouple from the dynamics of the main clade at time t_d , potentially accompanied by a shift in parameters.

Usage

```
dd_KI_loglik(
  pars1,
  pars2,
  brtsM,
  brtsS,
  misnumspec,
  methode = "odeint::runge_kutta_cash_karp54"
)
```

Arguments

pars1 Vector of parameters:

pars1[1] corresponds to λ_M (speciation rate) of the main clade
 pars1[2] corresponds to μ_M (extinction rate) of the main clade
 pars1[3] corresponds to K_M (clade-level carrying capacity) of the main clade
 pars1[4] corresponds to λ_S (speciation rate) of the subclade
 pars1[5] corresponds to μ_S (extinction rate) of the subclade
 pars1[6] corresponds to K_S (clade-level carrying capacity) of the subclade
 pars1[7] corresponds to t_d (the time of decoupling)

pars2 Vector of model settings:

pars2[1] sets the maximum number of species for which a probability must be computed. This must be larger than $1 + \text{misnumspec} + \text{length}(\text{brts})$.

pars2[2] sets the model of diversity-dependence:

- pars2[2] == 1 linear dependence in speciation rate with parameter K (= diversity where speciation = extinction)
- pars2[2] == 1.3 linear dependence in speciation rate with parameter K' (= diversity where speciation = 0)
- pars2[2] == 2 exponential dependence in speciation rate with parameter K (= diversity where speciation = extinction)
- pars2[2] == 2.1 variant of exponential dependence in speciation rate with offset at infinity
- pars2[2] == 2.2 $1/n$ dependence in speciation rate
- pars2[2] == 2.3 exponential dependence in speciation rate with parameter x (= exponent)
- pars2[2] == 3 linear dependence in extinction rate
- pars2[2] == 4 exponential dependence in extinction rate
- pars2[2] == 4.1 variant of exponential dependence in extinction rate with offset at infinity
- pars2[2] == 4.2 $1/n$ dependence in extinction rate

pars2[3] sets the conditioning:

- pars2[3] == 0 no conditioning (or just crown age)
- pars2[3] == 1 conditioning on non-extinction of the phylogeny
- pars2[3] == 2 conditioning on number of species and crown age; not yet im-

	plemented
	- pars2[3] == 3 conditioning on number of species only; not yet implemented
	- pars2[3] == 4 conditioning on survival of the subclade
	- pars2[3] == 5 conditioning on survival of all subclades and of both crown lineages in the main clade. This assumes that subclades that have already shifted do not undergo another shift, i.e. shifts only occur in the main clade.
	pars2[4] Obsolete.
	pars2[5] sets whether the parameters and likelihood should be shown on screen (1) or not (0)
	pars2[6] sets whether the first data point is stem age (1) or crown age (2)
	pars2[7] sets whether the old (incorrect) likelihood should be used (0), or whether the new corrected likelihood should be used (1).
brtsM	A set of branching times of the main clade in the phylogeny, all positive
brtsS	A set of branching times of the subclade in the phylogeny, all positive
missnumspec	The number of species that are in the clade but missing in the phylogeny. One can specify the sum of the missing species in main clade and subclade or a vector <code>c(missnumspec_M,missnumspec_S)</code> with missing species in main clade and subclade respectively.
methode	The method used to solve the master equation, default is 'analytical' which uses matrix exponentiation; alternatively numerical ODE solvers can be used, such as 'odeint::runge_kutta_cash_karp54'. These were used in the package before version 3.1.

Value

The loglikelihood

Author(s)

Rampal S. Etienne & Bart Haegeman

References

- Etienne, R.S. et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. Am. Nat. 180: E75-E89, doi: 10.1086/667574

See Also

[dd_KI_ML](#), [dd_loglik](#) [dd_SR_loglik](#)

Examples

```
pars1 = c(0.25,0.12,25.51,1.0,0.16,8.61,9.8)
pars2 = c(200,1,0,18.8,1,2)
```

```

missnumspec = 0
brtsM = c(25.2,24.6,24.0,22.5,21.7,20.4,19.9,19.7,18.8,17.1,15.8,11.8,9.7,8.9,5.7,5.2)
brtsS = c(9.6,8.6,7.4,4.9,2.5)
dd_KI_loglik(pars1,pars2,brtsM,brtsS,missnumspec)

```

dd_KI_ML	<i>Maximization of the loglikelihood under a diversity-dependent diversification model with decoupling of a subclade's diversification dynamics from the main clade's dynamics</i>
----------	--

Description

This function computes the maximum likelihood estimates of the parameters of a diversity-dependent diversification model with decoupling of the diversification dynamics of a subclade from the dynamics of the main clade for a given set of phylogenetic branching times of main clade and subclade and the time of splitting of the lineage that will form the subclade. It also outputs the corresponding loglikelihood that can be used in model comparisons.

Usage

```

dd_KI_ML(
  brtsM,
  brtsS,
  tsplit,
  initparsopt = c(0.5, 0.1, 2 * (1 + length(brtsM) + missnumspec[1]), 2 * (1 +
    length(brtsS) + missnumspec[length(missnumspec)]), (tsplit + max(brtsS))/2),
  parsfix = NULL,
  idparsopt = c(1:3, 6:7),
  idparsfix = NULL,
  idparsnoshift = (1:7)[c(-idparsopt, (-1)^(length(idparsfix) != 0) * idparsfix)],
  res = 10 * (1 + length(c(brtsM, brtsS)) + sum(missnumspec)),
  ddmodel = 1,
  missnumspec = 0,
  cond = 1,
  soc = 2,
  tol = c(0.001, 1e-04, 1e-06),
  maxiter = 1000 * round((1.25)^length(idparsopt)),
  changeloglikifnoconv = FALSE,
  optimmethod = "subplex",
  num_cycles = 1,
  methode = "analytical",
  correction = TRUE,
  verbose = FALSE
)

```

Arguments

<code>brtsM</code>	A set of branching times of the main clade in a phylogeny, all positive
<code>brtsS</code>	A set of branching times of the subclade in a phylogeny, all positive
<code>tsplit</code>	The branching time at which the lineage forming the subclade branches off, positive
<code>initparsopt</code>	The initial values of the parameters that must be optimized
<code>parsfix</code>	The values of the parameters that should not be optimized
<code>idparsopt</code>	The ids of the parameters that must be optimized, e.g. 1:7 for all parameters. The ids are defined as follows: id == 1 corresponds to λ_M (speciation rate) of the main clade id == 2 corresponds to μ_M (extinction rate) of the main clade id == 3 corresponds to K_M (clade-level carrying capacity) of the main clade id == 4 corresponds to λ_S (speciation rate) of the subclade id == 5 corresponds to μ_S (extinction rate) of the subclade id == 6 corresponds to K_S (clade-level carrying capacity) of the subclade id == 7 corresponds to t_d (the time of decoupling)
<code>idparsfix</code>	The ids of the parameters that should not be optimized, e.g. c(1,3,4,6) if λ and K should not be optimized, but only μ . In that case <code>idparsopt</code> must be c(2,5,7). The default is to fix all parameters not specified in <code>idparsopt</code> .
<code>idparsnoshift</code>	The ids of the parameters that should not shift; This can only apply to ids 4, 5 and 6, e.g. <code>idparsnoshift = c(4,5)</code> means that λ and μ have the same values before and after <code>tshift</code>
<code>res</code>	sets the maximum number of species for which a probability must be computed, must be larger than $1 + \max(\text{length}(\text{brtsM}), \text{length}(\text{brtsS}))$
<code>ddmodel</code>	sets the model of diversity-dependence: ddmodel == 1 : linear dependence in speciation rate with parameter K (= diversity where speciation = extinction) ddmodel == 1.3 : linear dependence in speciation rate with parameter K' (= diversity where speciation = 0) ddmodel == 2 : exponential dependence in speciation rate with parameter K (= diversity where speciation = extinction) ddmodel == 2.1 : variant of exponential dependence in speciation rate with offset at infinity ddmodel == 2.2 : 1/n dependence in speciation rate ddmodel == 2.3 : exponential dependence in speciation rate with parameter x (= exponent) ddmodel == 3 : linear dependence in extinction rate ddmodel == 4 : exponential dependence in extinction rate ddmodel == 4.1 : variant of exponential dependence in extinction rate with offset at infinity ddmodel == 4.2 : 1/n dependence in extinction rate with offset at infinity
<code>missnumspec</code>	The number of species that are in the clade but missing in the phylogeny. One can specify the sum of the missing species in main clade and subclade or a vector c(missnumspec_M,missnumspec_S) with missing species in main clade and subclade respectively.

cond	Conditioning: cond == 0 : no conditioning cond == 1 : conditioning on non-extinction of the phylogeny
soc	Sets whether stem or crown age should be used (1 or 2); stem age only works when cond = 0
tol	Sets the tolerances in the optimization. Consists of: reltolx = relative tolerance of parameter values in optimization reltolf = relative tolerance of function value in optimization abstolx = absolute tolerance of parameter values in optimization
maxiter	Sets the maximum number of iterations in the optimization
changeloglikifnoconv	if TRUE the loglik will be set to -Inf if ML does not converge
optimmethod	Method used in optimization of the likelihood. Current default is 'subplex'. Alternative is 'simplex' (default of previous versions)
num_cycles	the number of cycles of optimization. If set at Inf, it will do as many cycles as needed to meet the tolerance set for the target function.
methode	The method used to solve the master equation, default is 'analytical' which uses matrix exponentiation; alternatively numerical ODE solvers can be used, such as 'odeint::runge_kutta_cash_karp54'. These were used in the package before version 3.1.
correction	Sets whether the correction should be applied (TRUE) or not (FALSE)
verbose	Show the parameters and loglikelihood for every call to the loglik function

Details

The output is a dataframe containing estimated parameters and maximum loglikelihood. The computed loglikelihood contains the factor $q! / m!(q + m)!$ where q is the number of species in the phylogeny and m is the number of missing species, as explained in the supplementary material to Etienne et al. 2012.

Value

lambda_M	gives the maximum likelihood estimate of lambda of the main clade
mu_M	gives the maximum likelihood estimate of mu of the main clade
K_M	gives the maximum likelihood estimate of K of the main clade
lambda_2	gives the maximum likelihood estimate of lambda of the subclade
mu_S	gives the maximum likelihood estimate of mu of the subclade
K_S	gives the maximum likelihood estimate of K of the subclade
t_d	gives the time of the decoupling event
loglik	gives the maximum loglikelihood
df	gives the number of estimated parameters, i.e. degrees of freedom
conv	gives a message on convergence of optimization; conv = 0 means convergence

Note

The optimization may get trapped in local optima. Try different starting values to search for the global optimum.

Author(s)

Rampal S. Etienne & Bart Haegeman

References

- Etienne, R.S. et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. Am. Nat. 180: E75-E89, doi: 10.1086/667574

See Also

[dd_KI_loglik](#), [dd_ML](#), [dd_SR_ML](#),

Examples

```
cat("This will estimate parameters for two sets of branching times brtsM, brtsS\n")
cat("without conditioning.\n")
cat("The tolerance of the optimization is set high so runtime is fast in this example.\n")
cat("In real applications, use the default or more stringent settings for tol.\n")
brtsM = 4:10
brtsS = seq(0.1,3.5,0.7)
tsplit = 5
dd_KI_ML(brtsM = brtsM, brtsS = brtsS, tsplit = tsplit, idparsopt = c(1:3,6,7),
  initparsopt = c(0.885, 2e-14, 6.999, 6.848, 4.001), idparsfix = NULL,
  parsfix = NULL, idparsnoshift = c(4,5), cond = 0, tol = c(3E-1,3E-1,3E-1),
  optimmethod = 'simplex')
```

dd_KI_sim

Function to simulate a key innovation in macro-evolution with the innovative clade decoupling from the diversity-dependent diversification dynamics of the main clade

Description

Simulating a diversity-dependent diversification process where at a given time a new clade emerges with different inherent speciation rate and extinction rate and clade-level carrying capacity and with decoupled dynamics

Usage

```
dd_KI_sim(pars, age, ddmodel = 1)
```

Arguments

pars	Vector of parameters: pars[1] corresponds to lambda_M (speciation rate of the main clade) pars[2] corresponds to mu_M (extinction rate of the main clade) pars[3] corresponds to K_M (clade-level carrying capacity of the main clade) pars[4] corresponds to lambda_S (speciation rate of the subclade) pars[5] corresponds to mu_S (extinction rate of the subclade) pars[6] corresponds to K_S (clade-level carrying capacity of the subclade) pars[7] tinn, the time the shift in rates occurs in the lineage leading to the subclade
age	Sets the crown age for the simulation
ddmodel	Sets the model of diversity-dependence: ddmodel == 1 : linear dependence in speciation rate with parameter K (= diversity where speciation = extinction) ddmodel == 1.3 : linear dependence in speciation rate with parameter K' (= diversity where speciation = 0) ddmodel == 2 : exponential dependence in speciation rate with parameter K (= diversity where speciation = extinction) ddmodel == 2.1 : variant of exponential dependence in speciation rate with offset at infinity ddmodel == 2.2 : 1/n dependence in speciation rate ddmodel == 2.3 : exponential dependence in speciation rate with parameter x (= exponent) ddmodel == 3 : linear dependence in extinction rate ddmodel == 4 : exponential dependence in extinction rate ddmodel == 4.1 : variant of exponential dependence in extinction rate with offset at infinity ddmodel == 4.2 : 1/n dependence in extinction rate with offset at infinity

Value

out	A list with the following elements: The first element is the tree of extant species in phylo format The second element is the tree of all species, including extinct species, in phylo format The third element is a matrix of all species where - the first column is the time at which a species is born - the second column is the label of the parent of the species; positive and negative values only indicate whether the species belongs to the left or right crown lineage - the third column is the label of the daughter species itself; positive and negative values only indicate whether the species belongs to the left or right crown lineage - the fourth column is the time of extinction of the species If the fourth element equals -1, then the species is still extant. - the fifth column indicates whether the species belong to the main clade (0) or the subclade (1)
-----	---

The fourth element is the subclade tree of extant species (without stem)
 The fifth element is the subclade tree of all species (without stem)
 The sixth element is the same as the first, except that it has attributed 0 for the main clade and 1 for the subclade
 The seventh element is the same as the Second, except that it has attributed 0 for the main clade and 1 for the subclade
 The sixth and seventh element will be NULL if the subclade does not exist (because it went extinct).

Author(s)

Rampal S. Etienne

References

- Etienne, R.S. et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. Am. Nat. 180: E75-E89, doi: 10.1086/667574

Examples

```
dd_KI_sim(c(0.2,0.1,20,0.1,0.05,30,4),10)
```

dd_loglik

Loglikelihood for diversity-dependent diversification models

Description

This function computes loglikelihood of a diversity-dependent diversification model for a given set of branching times and parameter values.

Usage

```
dd_loglik(pars1, pars2, brts, missnumspec, methode = "analytical")
```

Arguments

pars1 Vector of parameters:

pars1[1] corresponds to lambda (speciation rate)
 pars1[2] corresponds to mu (extinction rate)
 pars1[3] corresponds to K (clade-level carrying capacity)

pars2 Vector of model settings:

pars2[1] sets the maximum number of species for which a probability must be computed. This must be larger than 1 + missnumspec + length(brts).

pars2[2] sets the model of diversity-dependence:

- pars2[2] == 1 linear dependence in speciation rate with parameter K (= diversity where speciation = extinction)
- pars2[2] == 1.3 linear dependence in speciation rate with parameter K' (= diversity where speciation = 0)
- pars2[2] == 1.4 : positive diversity-dependence in speciation rate with parameter K' (= diversity where speciation rate reaches half its maximum); $\lambda = \lambda_0 * S/(S + K')$ where S is species richness
- pars2[2] == 1.5 : positive and negative diversity-dependence in speciation rate with parameter K' (= diversity where speciation = 0); $\lambda = \lambda_0 * S/K' * (1 - S/K')$ where S is species richness
- pars2[2] == 2 exponential dependence in speciation rate with parameter K (= diversity where speciation = extinction)
- pars2[2] == 2.1 variant of exponential dependence in speciation rate with offset at infinity
- pars2[2] == 2.2 1/n dependence in speciation rate
- pars2[2] == 2.3 exponential dependence in speciation rate with parameter x (= exponent)
- pars2[2] == 3 linear dependence in extinction rate
- pars2[2] == 4 exponential dependence in extinction rate
- pars2[2] == 4.1 variant of exponential dependence in extinction rate with offset at infinity
- pars2[2] == 4.2 1/n dependence in extinction rate
- pars2[2] == 5 linear dependence in speciation and extinction rate

pars2[3] sets the conditioning:

- pars2[3] == 0 conditioning on stem or crown age
- pars2[3] == 1 conditioning on stem or crown age and non-extinction of the phylogeny
- pars2[3] == 2 conditioning on stem or crown age and on the total number of extant taxa (including missing species)
- pars2[3] == 3 conditioning on the total number of extant taxa (including missing species)

pars2[4] sets whether the likelihood is for the branching times (0) or the phylogeny (1)

pars2[5] sets whether the parameters and likelihood should be shown on screen (1) or not (0)

pars2[6] sets whether the first data point is stem age (1) or crown age (2)

brts	A set of branching times of a phylogeny, all positive
misnumspec	The number of species that are in the clade but missing in the phylogeny
methode	The method used to solve the master equation, default is 'analytical' which uses matrix exponentiation; alternatively numerical ODE solvers can be used. Before version 3.1 these were solvers from the deSolve package such as 'lsoda' and 'ode45'. Currently solvers from odeint are used, such as 'odeint::runge_kutta_cash_karp54', 'odeint::runge_kutta_fehlberg78', 'odeint::runge_kutta_dopri5', or 'odeint::bulirsch_stoer'. The first two are recommended in most cases.

Value

The loglikelihood

Author(s)

Rampal S. Etienne & Bart Haegeman

References

- Etienne, R.S. et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. Am. Nat. 180: E75-E89, doi: 10.1086/667574

See Also

[dd_ML](#), [dd_SR_loglik](#), [dd_KI_loglik](#)

Examples

```
dd_loglik(pars1 = c(0.5,0.1,100), pars2 = c(100,1,1,1,0,2), brts = 1:10, missnumspec = 0)
```

dd_LR

Bootstrap likelihood ratio test of diversity-dependent diversification model

Description

This function computes the maximum likelihood and the associated estimates of the parameters of a diversity-dependent diversification model for a given set of phylogenetic branching times. It then performs a bootstrap likelihood ratio test of the diversity-dependent (DD) model against the constant-rates (CR) birth-death model. Finally, it computes the power of this test.

Usage

```
dd_LR(
  brts,
  initparsoptDD,
  initparsoptCR,
  missnumspec,
  outputfilename = NULL,
  seed = 42,
  endmc = 1000,
  alpha = 0.05,
  plotit = TRUE,
  res = 10 * (1 + length(brts) + missnumspec),
  ddmodel = 1,
  cond = 1,
  btorph = 1,
```

```

soc = 2,
tol = c(0.001, 1e-04, 1e-06),
maxiter = 2000,
changeloglikifnoconv = FALSE,
optimmethod = "subplex",
methode = "analytical"
)

```

Arguments

<code>brts</code>	A set of branching times of a phylogeny, all positive
<code>initparsoptDD</code>	The initial values of the parameters that must be optimized for the diversity-dependent (DD) model: <code>lambda_0</code> , <code>mu</code> and <code>K</code>
<code>initparsoptCR</code>	The initial values of the parameters that must be optimized for the constant-rates (CR) model: <code>lambda</code> and <code>mu</code>
<code>missnumspec</code>	The number of species that are in the clade but missing in the phylogeny
<code>outputfilename</code>	The name (and location) of the file where the output will be saved. Default is no save.
<code>seed</code>	The seed for the pseudo random number generator for simulating the bootstrap data
<code>endmc</code>	The number of bootstraps
<code>alpha</code>	The significance level of the test
<code>plotit</code>	Boolean to plot results or not
<code>res</code>	Sets the maximum number of species for which a probability must be computed, must be larger than $1 + \text{length}(\text{brts})$
<code>ddmodel</code>	Sets the model of diversity-dependence: <code>ddmodel == 1</code> : linear dependence in speciation rate with parameter <code>K</code> (= diversity where speciation = extinction) <code>ddmodel == 1.3</code> : linear dependence in speciation rate with parameter <code>K'</code> (= diversity where speciation = 0) <code>ddmodel == 2</code> : exponential dependence in speciation rate with parameter <code>K</code> (= diversity where speciation = extinction) <code>ddmodel == 2.1</code> : variant of exponential dependence in speciation rate with offset at infinity <code>ddmodel == 2.2</code> : 1/n dependence in speciation rate <code>ddmodel == 2.3</code> : exponential dependence in speciation rate with parameter <code>x</code> (= exponent) <code>ddmodel == 3</code> : linear dependence in extinction rate <code>ddmodel == 4</code> : exponential dependence in extinction rate <code>ddmodel == 4.1</code> : variant of exponential dependence in extinction rate with offset at infinity <code>ddmodel == 4.2</code> : 1/n dependence in extinction rate with offset at infinity <code>ddmodel == 5</code> : linear dependence in speciation and extinction rate

cond	<p>Conditioning:</p> <p>cond == 0 : conditioning on stem or crown age</p> <p>cond == 1 : conditioning on stem or crown age and non-extinction of the phylogeny</p> <p>cond == 2 : conditioning on stem or crown age and on the total number of extant taxa (including missing species)</p> <p>cond == 3 : conditioning on the total number of extant taxa (including missing species)</p> <p>Note: cond == 3 assumes a uniform prior on stem age, as is the standard in constant-rate birth-death models, see e.g. D. Aldous & L. Popovic 2004. Adv. Appl. Prob. 37: 1094-1115 and T. Stadler 2009. J. Theor. Biol. 261: 58-66.</p>
btorph	Sets whether the likelihood is for the branching times (0) or the phylogeny (1)
soc	Sets whether stem or crown age should be used (1 or 2)
tol	<p>Sets the tolerances in the optimization. Consists of:</p> <p>reltolx = relative tolerance of parameter values in optimization</p> <p>reltolf = relative tolerance of function value in optimization</p> <p>abstolx = absolute tolerance of parameter values in optimization</p>
maxiter	Sets the maximum number of iterations in the optimization
changeloglikifnoconv	if TRUE the loglik will be set to -Inf if ML does not converge
optimmethod	Method used in optimization of the likelihood. Current default is 'subplex'. Alternative is 'simplex' (default of previous versions)
methode	The method used to solve the master equation, default is 'analytical' which uses matrix exponentiation; alternatively numerical ODE solvers can be used, such as 'odeint::runge_kutta_cash_karp54'. These were used in the package before version 3.1.

Details

The output is a list with 3 elements:

Value

treeCR	a list of trees generated under the constant-rates model using the ML parameters under the CR model
treeDD	a list of trees generated under the diversity-dependent model using the ML parameters under the diversity-dependent model
out	<p>a dataframe with the parameter estimates and maximum likelihoods for diversity-dependent and constant-rates models \$model - the model used to generate the data. 0 = unknown (for real data), 1 = CR, 2 = DD</p> <p>\$mc - the simulation number for each model</p> <p>\$lambda_CR - speciation rate estimated under CR</p> <p>\$mu_CR - extinction rate estimated under CR</p> <p>\$LL_CR - maximum likelihood estimated under CR</p> <p>\$conv_CR - convergence code for likelihood optimization; conv = 0 means convergence</p>

λ_{DD1} - initial speciation rate estimated under DD for first set of initial values

μ_{DD1} - extinction rate estimated under DD for first set of initial values

K_{DD1} - clade-wide carrying-capacity estimated under DD for first set of initial values

LL_{DD1} - maximum likelihood estimated under DD for first set of initial values

$conv_{DD1}$ - convergence code for likelihood optimization for first set of initial values; $conv = 0$ means convergence

λ_{DD2} - initial speciation rate estimated under DD for second set of initial values

μ_{DD2} - extinction rate estimated under DD for second set of initial values

K_{DD2} - clade-wide carrying-capacity estimated under DD for second set of initial values

LL_{DD2} - maximum likelihood estimated under DD for second set of initial values

$conv_{DD2}$ - convergence code for likelihood optimization for second set of initial values; $conv = 0$ means convergence

LR - likelihood ratio between DD and CR

pvalue	p-value of the test
LRalpha	Likelihood ratio at the significance level alpha
powerof test	power of the test for significance level alpha

Author(s)

Rampal S. Etienne & Bart Haegeman

References

- Etienne, R.S. et al. 2016. *Meth. Ecol. Evol.* 7: 1092-1099, doi: 10.1111/2041-210X.12565
- Etienne, R.S. et al. 2012, *Proc. Roy. Soc. B* 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. *Am. Nat.* 180: E75-E89, doi: 10.1086/667574

See Also

[dd_loglik](#), [dd_ML](#)

dd_ML

Maximization of the loglikelihood under a diversity-dependent diversification model

Description

This function computes the maximum likelihood estimates of the parameters of a diversity-dependent diversification model for a given set of phylogenetic branching times. It also outputs the corresponding loglikelihood that can be used in model comparisons.

Usage

```
dd_ML(
  brts,
  initparsopt = initparsoptdefault(ddmodel, brts, missnumspec),
  idparsopt = 1:length(initparsopt),
  idparsfix = (1:(3 + (ddmodel == 5)))[-idparsopt],
  parsfix = parsfixdefault(ddmodel, brts, missnumspec, idparsopt),
  res = 10 * (1 + length(brts) + missnumspec),
  ddmodel = 1,
  missnumspec = 0,
  cond = 1,
  btorph = 1,
  soc = 2,
  tol = c(0.001, 1e-04, 1e-06),
  maxiter = 1000 * round((1.25)^length(idparsopt)),
  changeloglikifnoconv = FALSE,
  optimmethod = "subplex",
  num_cycles = 1,
  methode = "analytical",
  verbose = FALSE
)
```

Arguments

<code>brts</code>	A set of branching times of a phylogeny, all positive
<code>initparsopt</code>	The initial values of the parameters that must be optimized
<code>idparsopt</code>	The ids of the parameters that must be optimized, e.g. 1:3 for intrinsic speciation rate, extinction rate and carrying capacity. The ids are defined as follows: <code>id == 1</code> corresponds to λ (speciation rate) <code>id == 2</code> corresponds to μ (extinction rate) <code>id == 3</code> corresponds to K (clade-level carrying capacity) <code>id == 4</code> corresponds to r ($r = b/a$ where $\mu = \mu_0 + b * N$ and $\lambda = \lambda_0 - a * N$) (This is only available when <code>ddmodel = 5</code>)
<code>idparsfix</code>	The ids of the parameters that should not be optimized, e.g. <code>c(1,3)</code> if λ and K should not be optimized, but only μ . In that case <code>idparsopt</code> must be 2. The default is to fix all parameters not specified in <code>idparsopt</code> .
<code>parsfix</code>	The values of the parameters that should not be optimized
<code>res</code>	Sets the maximum number of species for which a probability must be computed, must be larger than <code>1 + length(brts)</code>
<code>ddmodel</code>	Sets the model of diversity-dependence: <code>ddmodel == 1</code> : linear dependence in speciation rate with parameter K (= diversity where speciation = extinction) <code>ddmodel == 1.3</code> : linear dependence in speciation rate with parameter K' (= diversity where speciation = 0) <code>ddmodel == 1.4</code> : positive diversity-dependence in speciation rate with parameter K' (= diversity where speciation rate reaches half its maximum); $\lambda =$

$\lambda_0 * S / (S + K')$ where S is species richness
 ddmodel == 1.5 : positive and negative dependence in speciation rate with parameter K' (= diversity where speciation = 0); $\lambda = \lambda_0 * S / K' * (1 - S / K')$ where S is species richness
 ddmodel == 2 : exponential dependence in speciation rate with parameter K (= diversity where speciation = extinction)
 ddmodel == 2.1 : variant of exponential dependence in speciation rate with offset at infinity
 ddmodel == 2.2 : 1/n dependence in speciation rate
 ddmodel == 2.3 : exponential dependence in speciation rate with parameter x (= exponent)
 ddmodel == 3 : linear dependence in extinction rate
 ddmodel == 4 : exponential dependence in extinction rate
 ddmodel == 4.1 : variant of exponential dependence in extinction rate with offset at infinity
 ddmodel == 4.2 : 1/n dependence in extinction rate with offset at infinity
 ddmodel == 5 : linear dependence in speciation and extinction rate

missnumspec	The number of species that are in the clade but missing in the phylogeny
cond	Conditioning: cond == 0 : conditioning on stem or crown age cond == 1 : conditioning on stem or crown age and non-extinction of the phylogeny cond == 2 : conditioning on stem or crown age and on the total number of extant taxa (including missing species) cond == 3 : conditioning on the total number of extant taxa (including missing species) Note: cond == 3 assumes a uniform prior on stem age, as is the standard in constant-rate birth-death models, see e.g. D. Aldous & L. Popovic 2004. Adv. Appl. Prob. 37: 1094-1115 and T. Stadler 2009. J. Theor. Biol. 261: 58-66.
btorph	Sets whether the likelihood is for the branching times (0) or the phylogeny (1)
soc	Sets whether stem or crown age should be used (1 or 2)
tol	Sets the tolerances in the optimization. Consists of: reltolx = relative tolerance of parameter values in optimization reltolf = relative tolerance of function value in optimization abstolx = absolute tolerance of parameter values in optimization
maxiter	Sets the maximum number of iterations in the optimization
changeloglikifnoconv	if TRUE the loglik will be set to -Inf if ML does not converge
optimmethod	Method used in optimization of the likelihood. Current default is 'subplex'. Alternative is 'simplex' (default of previous versions)
num_cycles	the number of cycles of optimization. If set at Inf, it will do as many cycles as needed to meet the tolerance set for the target function.
methode	The method used to solve the master equation, default is 'analytical' which uses matrix exponentiation; alternatively numerical ODE solvers can be used, such

as 'odeint::runge_kutta_cash_karp54'. These were used in the package before version 3.1.

verbose Show the parameters and loglikelihood for every call to the loglik function

Details

The output is a dataframe containing estimated parameters and maximum loglikelihood. The computed loglikelihood contains the factor $q! m! / (q + m)!$ where q is the number of species in the phylogeny and m is the number of missing species, as explained in the supplementary material to Etienne et al. 2012.

Value

lambda	gives the maximum likelihood estimate of lambda
mu	gives the maximum likelihood estimate of mu
K	gives the maximum likelihood estimate of K
r	(only if ddmodel == 5) gives the ratio of linear dependencies in speciation and extinction rates
loglik	gives the maximum loglikelihood
df	gives the number of estimated parameters, i.e. degrees of freedom
conv	gives a message on convergence of optimization; conv = 0 means convergence

Author(s)

Rampal S. Etienne & Bart Haegeman

References

- Etienne, R.S. et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. Am. Nat. 180: E75-E89, doi: 10.1086/667574

See Also

[dd_loglik](#), [dd_SR_ML](#), [dd_KI_ML](#),

Examples

```
cat("Estimating the intrinsic speciation rate lambda and the carrying capacity K")
cat("for a fixed extinction rate of 0.1, conditioning on clade survival and two missing species:")
brts = 1:5
dd_ML(brts = brts, initparsopt = c(1.3078, 7.4188), idparsopt = c(1, 3), parsfix = 0.1,
      cond = 1, missnumspec = 2, tol = c(1E-3, 1E-3, 1E-4), optimmethod = 'simplex')
```

dd_MS_loglik	<i>Loglikelihood for macro-evolutionary succession under diversity-dependent diversification with the key innovation at time $t = t_d$</i>
--------------	---

Description

This function computes the loglikelihood of a diversity-dependent diversification model for a given set of branching times and parameter values where the diversity-dependent dynamics of an innovative subclade have different parameters from the dynamics of the main clade from time t_d , but both are governed by the same carrying capacity and experience each other's diversity.

Usage

```
dd_MS_loglik(
  pars1,
  pars2,
  brtsM,
  brtsS,
  misnumspec,
  methode = "odeint::runge_kutta_cash_karp54"
)
```

Arguments

pars1	<p>Vector of parameters:</p> <p>pars1[1] corresponds to λ_M (speciation rate) of the main clade pars1[2] corresponds to μ_M (extinction rate) of the main clade pars1[3] corresponds to K_M (clade-level carrying capacity) of the main clade pars1[4] corresponds to λ_S (speciation rate) of the subclade pars1[5] corresponds to μ_S (extinction rate) of the subclade pars1[6] corresponds to t_d (the time of the key innovation)</p>
pars2	<p>Vector of model settings:</p> <p>pars2[1] sets the maximum number of species for which a probability must be computed. This must be larger than $1 + \text{misnumspec} + \text{length}(\text{brts})$.</p> <p>pars2[2] sets the model of diversity-dependence:</p> <ul style="list-style-type: none"> - pars2[2] == 1 linear dependence in speciation rate with parameter K (= diversity where speciation = extinction) - pars2[2] == 1.3 linear dependence in speciation rate with parameter K' (= diversity where speciation = 0) - pars2[2] == 2 exponential dependence in speciation rate with parameter K (= diversity where speciation = extinction) - pars2[2] == 2.1 variant of exponential dependence in speciation rate with offset at infinity - pars2[2] == 2.2 $1/n$ dependence in speciation rate

- pars2[2] == 2.3 exponential dependence in speciation rate with parameter x (= exponent)
- pars2[2] == 3 linear dependence in extinction rate
- pars2[2] == 4 exponential dependence in extinction rate
- pars2[2] == 4.1 variant of exponential dependence in extinction rate with offset at infinity
- pars2[2] == 4.2 1/n dependence in extinction rate

pars2[3] sets the conditioning:

- pars2[3] == 0 no conditioning
- pars2[3] == 1 conditioning on non-extinction of the phylogeny

pars2[4] sets the time of splitting of the branch that will undergo the key innovation leading to different parameters

pars2[5] sets whether the parameters and likelihood should be shown on screen (1) or not (0)

pars2[6] sets whether the first data point is stem age (1) or crown age (2)

pars2[7] sets whether the old (incorrect) likelihood should be used (0) or whether new corrected version should be used (1)

brtsM	A set of branching times of the main clade in the phylogeny, all positive
brtsS	A set of branching times of the subclade in the phylogeny, all positive
misnumspec	The number of species that are in the clade but missing in the phylogeny. One can specify the sum of the missing species in main clade and subclade or a vector c(misnumspec_M,misnumspec_S) with missing species in main clade and subclade respectively.
methode	The method used to solve the master equation, default is 'analytical' which uses matrix exponentiation; alternatively numerical ODE solvers can be used, such as 'odeint::runge_kutta_cash_karp54'. These were used in the package before version 3.1.

Value

The loglikelihood

Author(s)

Rampal S. Etienne & Bart Haegeman

References

- Etienne, R.S. et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. Am. Nat. 180: E75-E89, doi: 10.1086/667574

See Also

[dd_MS_ML](#), [dd_loglik](#), [dd_KI_loglik](#), [dd_SR_loglik](#)

Examples

```

pars1 = c(0.2,0.1,40,1.0,0.1,9.8)
pars2 = c(200,1,0,18.8,1,2)
misnumspec = 0
brtsM = c(25.2,24.6,24.0,22.5,21.7,20.4,19.9,19.7,18.8,17.1,15.8,11.8,9.7,8.9,5.7,5.2)
brtsS = c(9.6,8.6,7.4,4.9,2.5)
dd_MS_loglik(pars1,pars2,brtsM,brtsS,misnumspec)

```

dd_MS_ML

Maximization of the loglikelihood under a diversity-dependent diversification model with decoupling of a subclade's diversification dynamics from the main clade's dynamics

Description

This function computes the maximum likelihood estimates of the parameters of a diversity-dependent diversification model where the diversity-dependent dynamics of an innovative subclade have different parameters from the dynamics of the main clade from time t_d , but both are governed by the same carrying capacity and experience each other's diversity. Required is a given set of phylogenetic branching times of main clade and subclade and the time of splitting of the lineage that will form the subclade. The function also outputs the corresponding loglikelihood that can be used in model comparisons.

Usage

```

dd_MS_ML(
  brtsM,
  brtsS,
  tsplit,
  initparsopt = c(0.5, 0.1, 2 * (1 + length(brtsM) + length(brtsS) + sum(misnumspec)),
    (tsplit + max(brtsS))/2),
  parsfix = NULL,
  idparsopt = c(1:3, 6),
  idparsfix = NULL,
  idparsnoshift = (1:6)[c(-idparsopt, (-1)^(length(idparsfix) != 0) * idparsfix)],
  res = 10 * (1 + length(c(brtsM, brtsS)) + sum(misnumspec)),
  ddmodel = 1.3,
  misnumspec = 0,
  cond = 0,
  soc = 2,
  tol = c(0.001, 1e-04, 1e-06),
  maxiter = 1000 * round((1.25)^length(idparsopt)),
  changeloglikifnoconv = FALSE,
  optimmethod = "subplex",
  num_cycles = 1,

```

```

methode = "ode45",
correction = FALSE,
verbose = FALSE
)

```

Arguments

<code>brtsM</code>	A set of branching times of the main clade in a phylogeny, all positive
<code>brtsS</code>	A set of branching times of the subclade in a phylogeny, all positive
<code>tsplit</code>	The branching time at which the lineage forming the subclade branches off, positive
<code>initparsopt</code>	The initial values of the parameters that must be optimized
<code>parsfix</code>	The values of the parameters that should not be optimized
<code>idparsopt</code>	The ids of the parameters that must be optimized, e.g. 1:7 for all parameters. The ids are defined as follows: id == 1 corresponds to λ_M (speciation rate) of the main clade id == 2 corresponds to μ_M (extinction rate) of the main clade id == 3 corresponds to K_M (clade-level carrying capacity) of the main clade id == 4 corresponds to λ_S (speciation rate) of the subclade id == 5 corresponds to μ_S (extinction rate) of the subclade id == 6 corresponds to t_d (the time of the key innovation)
<code>idparsfix</code>	The ids of the parameters that should not be optimized, e.g. c(1,3,4,6) if λ and K should not be optimized, but only μ . In that case <code>idparsopt</code> must be c(2,5,7). The default is to fix all parameters not specified in <code>idparsopt</code> .
<code>idparsnoshift</code>	The ids of the parameters that should not shift; This can only apply to ids 4, 5 and 6, e.g. <code>idparsnoshift = c(4,5)</code> means that λ and μ have the same values before and after <code>tshift</code>
<code>res</code>	sets the maximum number of species for which a probability must be computed, must be larger than $1 + \max(\text{length}(\text{brtsM}), \text{length}(\text{brtsS}))$
<code>ddmodel</code>	sets the model of diversity-dependence: ddmodel == 1 : linear dependence in speciation rate with parameter K (= diversity where speciation = extinction) ddmodel == 1.3 : linear dependence in speciation rate with parameter K' (= diversity where speciation = 0) ddmodel == 2 : exponential dependence in speciation rate with parameter K (= diversity where speciation = extinction) ddmodel == 2.1 : variant of exponential dependence in speciation rate with offset at infinity ddmodel == 2.2 : 1/n dependence in speciation rate ddmodel == 2.3 : exponential dependence in speciation rate with parameter x (= exponent) ddmodel == 3 : linear dependence in extinction rate ddmodel == 4 : exponential dependence in extinction rate ddmodel == 4.1 : variant of exponential dependence in extinction rate with offset at infinity ddmodel == 4.2 : 1/n dependence in extinction rate with offset at infinity

missnumspec	The number of species that are in the clade but missing in the phylogeny. One can specify the sum of the missing species in main clade and subclade or a vector <code>c(missnumspec_M,missnumspec_S)</code> with missing species in main clade and subclade respectively.
cond	Conditioning: <code>cond == 0</code> : no conditioning <code>cond == 1</code> : conditioning on non-extinction of the phylogeny
soc	Sets whether stem or crown age should be used (1 or 2); stem age only works when <code>cond = 0</code>
tol	Sets the tolerances in the optimization. Consists of: <code>reltolx</code> = relative tolerance of parameter values in optimization <code>reltolf</code> = relative tolerance of function value in optimization <code>abstolx</code> = absolute tolerance of parameter values in optimization
maxiter	Sets the maximum number of iterations in the optimization
changeloglikifnoconv	if TRUE the loglik will be set to -Inf if ML does not converge
optimmethod	Method used in optimization of the likelihood. Current default is 'subplex'. Alternative is 'simplex' (default of previous versions)
num_cycles	the number of cycles of optimization. If set at Inf, it will do as many cycles as needed to meet the tolerance set for the target function.
methode	The method used in the ode solver. This can be either 'analytical' for explicit matrix exponentiation or any of the solvers in the deSolve package.
correction	Sets whether the correction should be applied (TRUE) or not (FALSE)
verbose	Show the parameters and loglikelihood for every call to the loglik function

Details

The output is a dataframe containing estimated parameters and maximum loglikelihood. The computed loglikelihood contains the factor $q! m! / (q + m)!$ where q is the number of species in the phylogeny and m is the number of missing species, as explained in the supplementary material to Etienne et al. 2012.

Value

<code>lambda_M</code>	gives the maximum likelihood estimate of lambda of the main clade
<code>mu_M</code>	gives the maximum likelihood estimate of mu of the main clade
<code>K_M</code>	gives the maximum likelihood estimate of K of the main clade
<code>lambda_2</code>	gives the maximum likelihood estimate of lambda of the subclade
<code>mu_S</code>	gives the maximum likelihood estimate of mu of the subclade
<code>t_d</code>	gives the time of the key innovation event
<code>loglik</code>	gives the maximum loglikelihood
<code>df</code>	gives the number of estimated parameters, i.e. degrees of freedom
<code>conv</code>	gives a message on convergence of optimization; <code>conv = 0</code> means convergence

Note

The optimization may get trapped in local optima. Try different starting values to search for the global optimum.

Author(s)

Rampal S. Etienne & Bart Haegeman

References

- Etienne, R.S. et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. Am. Nat. 180: E75-E89, doi: 10.1086/667574

See Also

[dd_MS_loglik](#), [dd_ML](#), [dd_KI_ML](#), [dd_SR_ML](#),

Examples

```
cat("This will estimate parameters for two sets of branching times brtsM, brtsS\n")
cat("without conditioning.\n")
cat("The tolerance of the optimization is set high so runtime is fast in this example.\n")
cat("In real applications, use the default or more stringent settings for tol.\n")
brtsM = 4:10
brtsS = seq(0.1,3.5,0.7)
tsplit = 5
dd_MS_ML(brtsM = brtsM, brtsS = brtsS, tsplit = tsplit, idparsopt = c(1:3,6),
         initparsopt = c(0.885, 2e-14, 10, 4.001), idparsfix = NULL, parsfix = NULL,
         idparsnoshift = c(4,5), cond = 0, tol = c(3E-1,3E-1,3E-1))
```

dd_MS_sim

Function to simulate the macro-evolutionary succession process assuming diversity-dependent diversification

Description

Simulating a diversity-dependent diversification process where at a given time a new clade emerges with different inherent speciation rate and extinction rate

Usage

```
dd_MS_sim(pars, age, ddmodel = 1.3)
```

Arguments

pars	<p>Vector of parameters:</p> <p>pars[1] corresponds to lambda_M (speciation rate of the main clade) pars[2] corresponds to mu_M (extinction rate of the main clade) pars[3] corresponds to K' (maximum number of species or a proxy for it in case of exponential decline in speciation rate) pars[4] corresponds to lambda_S (speciation rate of the novel subclade) pars[5] corresponds to mu_S (extinction rate) pars[6] tinn, the time the shift in rates occurs in the lineage leading to the subclade</p>
age	Sets the crown age for the simulation
ddmodel	<p>Sets the model of diversity-dependence:</p> <p>ddmodel == 1.3 : linear dependence in speciation rate with parameter K' (= diversity where speciation = 0); ddmodel = 1 will be interpreted as this model ddmodel == 2.1 : variant of exponential dependence in speciation rate with offset at infinity; ddmodel = 2 will be interpreted as this model ddmodel == 2.2 : 1/n dependence in speciation rate ddmodel == 2.3 : exponential dependence in speciation rate with parameter x (= exponent)</p>

Value

out	<p>A list with the following elements: The first element is the tree of extant species in phylo format The second element is the tree of all species, including extinct species, in phylo format The third element is a matrix of all species where</p> <ul style="list-style-type: none"> - the first column is the time at which a species is born - the second column is the label of the parent of the species; positive and negative values only indicate whether the species belongs to the left or right crown lineage - the third column is the label of the daughter species itself; positive and negative values only indicate whether the species belongs to the left or right crown lineage - the fourth column is the time of extinction of the species <p>If the fourth element equals -1, then the species is still extant.</p> <ul style="list-style-type: none"> - the fifth column indicates whether the species belong to the main clade (0) or the subclade (1) <p>The fourth element is the subclade tree of extant species (without stem) The fifth element is the subclade tree of all species (without stem) The sixth element is the same as the first, except that it has attributed 0 for the main clade and 1 for the subclade The seventh element is the same as the Second, except that it has attributed 0 for the main clade and 1 for the subclade The sixth and seventh element will be NULL if the subclade does not exist (because it went extinct).</p>
-----	---

Author(s)

Rampal S. Etienne

References

- Etienne, R.S. et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. Am. Nat. 180: E75-E89, doi: 10.1086/667574

Examples

```
dd_MS_sim(c(0.2,0.1,20,0.1,0.05,4),10)
```

dd_multiple_KI_loglik *Loglikelihood for diversity-dependent diversification models with multiple decoupling (rate shift) events*

Description

This function computes loglikelihood of a diversity-dependent diversification model for a given set of branching times and parameter values where the diversity-dependent dynamics of subclades decouple from the dynamics of main clades, potentially accompanied by a shift in parameters.

Usage

```
dd_multiple_KI_loglik(
  pars1_list,
  pars2,
  brts_k_list,
  misnumspec_list,
  reltol = 1e-14,
  abstol = 1e-16,
  methode = "odeint::runge_kutta_cash_karp54"
)
```

Arguments

- | | |
|------------|---|
| pars1_list | list of parameter sets one for each rate regime (subclade). The parameters are: lambda (speciation rate), mu (extinction rate), and K (clade-level carrying capacity). |
| pars2 | Vector of model settings:

pars2[1] sets the maximum number of species for which a probability must be computed. This must be larger than 1 + misnumspec + length(brts).

pars2[2] sets the model of diversity-dependence:
- pars2[2] == 1 linear dependence in speciation rate with parameter K (= diversity where speciation = extinction) |

- pars2[2] == 1.3 linear dependence in speciation rate with parameter K' (= diversity where speciation = 0)
- pars2[2] == 2 exponential dependence in speciation rate with parameter K (= diversity where speciation = extinction)
- pars2[2] == 2.1 variant of exponential dependence in speciation rate with offset at infinity
- pars2[2] == 2.2 1/n dependence in speciation rate
- pars2[2] == 2.3 exponential dependence in speciation rate with parameter x (= exponent)
- pars2[2] == 3 linear dependence in extinction rate
- pars2[2] == 4 exponential dependence in extinction rate
- pars2[2] == 4.1 variant of exponential dependence in extinction rate with offset at infinity
- pars2[2] == 4.2 1/n dependence in extinction rate

pars2[3] sets the conditioning:

- pars2[3] == 0 no conditioning (or just crown age)
- pars2[3] == 1 conditioning on non-extinction of the phylogeny
- pars2[3] == 2 conditioning on number of species and crown age; not yet implemented
- pars2[3] == 3 conditioning on number of species only; not yet implemented
- pars2[3] == 4 conditioning on survival of the subclade
- pars2[3] == 5 conditioning on survival of all subclades and of both crown lineages in the main clade. This assumes that subclades that have already shifted do not undergo another shift, i.e. shifts only occur in the main clade.

pars2[4] Obsolete.

pars2[5] sets whether the parameters and likelihood should be shown on screen (1) or not (0)

pars2[6] sets whether the first data point is stem age (1) or crown age (2)

pars2[7] sets whether the old (incorrect) likelihood should be used (0), or whether the new corrected likelihood should be used (1).

`brts_k_list` list of matrices, one for each rate regime (subclade). Each matrix has in the first row the branching times including the shift/decoupling time and the present time (0) in negative time (i.e. 10 mya = -10). In the second row it has the number of lineages, i.e. starting at 2 for a phylogeny with a crown and increasing by one at each branching time and decreasing by one at each decoupling/shift time. The last element is the same as the second last.

`misnumspec_list` list containing the number of missing species for each clade. If only a single number m of missing species is known for the entire phylogeny, then each element of the list should be 0:m. One can also create this from m using the function `create_misnumspec_list`

`reltol` relative tolerance in integration of the ODE system, default at 1e-14

`abstol` tolerance tolerance in integration of the ODE system, default at 1e-16

methode The method used to solve the master equation, default is 'analytical' which uses matrix exponentiation; alternatively numerical ODE solvers can be used, such as 'odeint::runge_kutta_cash_karp54'. These were used in the package before version 3.1.

dd_sim *Function to simulate the diversity-dependent diversification process*

Description

Simulating the diversity-dependent diversification process

Usage

```
dd_sim(pars, age, ddmodel = 1)
```

Arguments

pars Vector of parameters:

pars[1] corresponds to lambda (speciation rate)
 pars[2] corresponds to mu (extinction rate)
 pars[3] corresponds to K (clade-level carrying capacity)

age Sets the crown age for the simulation

ddmodel Sets the model of diversity-dependence:
 ddmodel == 1 : linear dependence in speciation rate with parameter K (= diversity where speciation = extinction)
 ddmodel == 1.3 : linear dependence in speciation rate with parameter K' (= diversity where speciation = 0)
 ddmodel == 2 : exponential dependence in speciation rate with parameter K (= diversity where speciation = extinction)
 ddmodel == 2.1 : variant of exponential dependence in speciation rate with offset at infinity
 ddmodel == 2.2 : 1/n dependence in speciation rate
 ddmodel == 2.3 : exponential dependence in speciation rate with parameter x (= exponent)
 ddmodel == 3 : linear dependence in extinction rate
 ddmodel == 4 : exponential dependence in extinction rate
 ddmodel == 4.1 : variant of exponential dependence in extinction rate with offset at infinity
 ddmodel == 4.2 : 1/n dependence in extinction rate with offset at infinity
 ddmodel == 5 : linear dependence in speciation and extinction rate

Value

out

A list with the following four elements: The first element is the tree of extant species in phylo format
 The second element is the tree of all species, including extinct species, in phylo format
 The third element is a matrix of all species where
 - the first column is the time at which a species is born
 - the second column is the label of the parent of the species; positive and negative values only indicate whether the species belongs to the left or right crown lineage
 - the third column is the label of the daughter species itself; positive and negative values only indicate whether the species belongs to the left or right crown lineage
 - the fourth column is the time of extinction of the species. If this equals -1, then the species is still extant.
 The fourth element is the set of branching times of the tree of extant species.

Author(s)

Rampal S. Etienne

References

- Etienne, R.S. et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. Am. Nat. 180: E75-E89, doi: 10.1086/667574

Examples

```
dd_sim(c(0.2, 0.1, 20), 10)
```

dd_SR_loglik	<i>Loglikelihood for diversity-dependent diversification models with a shift in the parameters at time $t = tshift$</i>
--------------	--

Description

This function computes loglikelihood of a diversity-dependent diversification model for a given set of branching times and parameter values where the parameters are allowed to shift at time $t = tshift$

Usage

```
dd_SR_loglik(pars1, pars2, brts, misnumspec, methode = "analytical")
```

Arguments

- pars1** Vector of parameters:
- pars1[1] corresponds to lambda (speciation rate) before the shift
 pars1[2] corresponds to mu (extinction rate) before the shift
 pars1[3] corresponds to K (clade-level carrying capacity) before the shift
 pars1[4] corresponds to lambda (speciation rate) after the shift
 pars1[5] corresponds to mu (extinction rate) after the shift
 pars1[6] corresponds to K (clade-level carrying capacity) after the shift
 pars1[7] corresponds to tshift (the time of shift)
- pars2** Vector of model settings:
- pars2[1] sets the maximum number of species for which a probability must be computed. This must be larger than $1 + \text{missnumspec} + \text{length}(\text{brts})$.
- pars2[2] sets the model of diversity-dependence:
 - pars2[2] == 1 linear dependence in speciation rate with parameter K (= diversity where speciation = extinction)
 - pars2[2] == 1.3 linear dependence in speciation rate with parameter K' (= diversity where speciation = 0)
 - pars2[2] == 2 exponential dependence in speciation rate with parameter K (= diversity where speciation = extinction)
 - pars2[2] == 2.1 variant of exponential dependence in speciation rate with offset at infinity
 - pars2[2] == 2.2 1/n dependence in speciation rate
 - pars2[2] == 2.3 exponential dependence in speciation rate with parameter x (= exponent)
 - pars2[2] == 3 linear dependence in extinction rate
 - pars2[2] == 4 exponential dependence in extinction rate
 - pars2[2] == 4.1 variant of exponential dependence in extinction rate with offset at infinity
 - pars2[2] == 4.2 1/n dependence in extinction rate
- pars2[3] sets the conditioning:
 - pars2[3] == 0 no conditioning
 - pars2[3] == 1 conditioning on non-extinction of the phylogeny
 - pars2[3] == 2 conditioning on non-extinction of the phylogeny and on the total number of extant taxa (including missing species)
- pars2[4] sets whether the likelihood is for the branching times (0) or the phylogeny (1)
- pars2[5] sets whether the parameters and likelihood should be shown on screen (1) or not (0)
- pars2[6] sets whether the first data point is stem age (1) or crown age (2)
- brts** A set of branching times of a phylogeny, all positive

missnumspec The number of species that are in the clade but missing in the phylogeny

methode The method used to solve the master equation, default is 'analytical' which uses matrix exponentiation; alternatively numerical ODE solvers can be used, such as 'odeint::runge_kutta_cash_karp54'. These were used in the package before version 3.1.

Value

The loglikelihood

Author(s)

Rampal S. Etienne & Bart Haegeman

References

- Etienne, R.S. et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. Am. Nat. 180: E75-E89, doi: 10.1086/667574

See Also

[dd_SR_ML](#), [dd_loglik](#), [dd_KI_loglik](#)

Examples

```
dd_SR_loglik(pars1 = c(0.2,0.1,50,0.2,0.1,70,5), pars2 = c(100,1,1,1,0,2),
  brts = 1:10, missnumspec = 0)
```

dd_SR_ML	<i>Maximization of the loglikelihood under a diversity-dependent diversification model with a shift in the parameters</i>
----------	---

Description

This function computes the maximum likelihood estimates of the parameters of a diversity-dependent diversification model with shifting parameters at time $t = tshift$ for a given set of phylogenetic branching times. It also outputs the corresponding loglikelihood that can be used in model comparisons.

Usage

```
dd_SR_ML(
  brts,
  initparsopt = c(0.5, 0.1, 2 * (1 + length(brts) + missnumspec), 2 * (1 + length(brts) +
    missnumspec), max(brts)/2),
  parsfix = NULL,
  idparsopt = c(1:3, 6:7),
  idparsfix = NULL,
```

```

idparsnoshift = (1:7)[c(-idparsopt, (-1)^(length(idparsfix) != 0) * idparsfix)],
res = 10 * (1 + length(brts) + missnumspec),
ddmodel = 1,
missnumspec = 0,
cond = 1,
btorph = 1,
soc = 2,
allbp = FALSE,
tol = c(0.001, 1e-04, 1e-06),
maxiter = 1000 * round((1.25)^length(idparsopt)),
changeloglikifnoconv = FALSE,
optimmethod = "subplex",
num_cycles = 1,
methode = "analytical",
verbose = FALSE
)

```

Arguments

<code>brts</code>	A set of branching times of a phylogeny, all positive
<code>initparsopt</code>	The initial values of the parameters that must be optimized
<code>parsfix</code>	The values of the parameters that should not be optimized
<code>idparsopt</code>	The ids of the parameters that must be optimized, e.g. 1:7 for all parameters. The ids are defined as follows: id == 1 corresponds to lambda (speciation rate) before the shift id == 2 corresponds to mu (extinction rate) before the shift id == 3 corresponds to K (clade-level carrying capacity) before the shift id == 4 corresponds to lambda (speciation rate) after the shift id == 5 corresponds to mu (extinction rate) after the shift id == 6 corresponds to K (clade-level carrying capacity) after the shift id == 7 corresponds to tshift (the time of shift)
<code>idparsfix</code>	The ids of the parameters that should not be optimized, e.g. c(1,3,4,6) if lambda and K should not be optimized, but only mu. In that case idparsopt must be c(2,5,7). The default is to fix all parameters not specified in idparsopt.
<code>idparsnoshift</code>	The ids of the parameters that should not shift; This can only apply to ids 4, 5 and 6, e.g. idparsnoshift = c(4,5) means that lambda and mu have the same values before and after tshift
<code>res</code>	sets the maximum number of species for which a probability must be computed, must be larger than 1 + length(brts)
<code>ddmodel</code>	sets the model of diversity-dependence: ddmodel == 1 : linear dependence in speciation rate ddmodel == 2 : exponential dependence in speciation rate ddmodel == 2.1 : variant of exponential dependence in speciation rate with offset at infinity ddmodel == 2.2 : 1/n dependence in speciation rate ddmodel == 3 : linear dependence in extinction rate ddmodel == 4 : exponential dependence in extinction rate

	ddmodel == 4.1 : variant of exponential dependence in extinction rate with offset at infinity
	ddmodel == 4.2 : 1/n dependence in extinction rate with offset at infinity
missnumspec	The number of species that are in the clade but missing in the phylogeny
cond	Conditioning: cond == 0 : no conditioning cond == 1 : conditioning on non-extinction of the phylogeny cond == 2 : conditioning on non-extinction of the phylogeny and on the total number of extant taxa (including missing species) cond == 3 : conditioning on the total number of extant taxa (including missing species)
	Note: cond == 3 assumes a uniform prior on stem age, as is the standard in constant-rate birth-death models, see e.g. D. Aldous & L. Popovic 2004. Adv. Appl. Prob. 37: 1094-1115 and T. Stadler 2009. J. Theor. Biol. 261: 58-66.
btorph	Sets whether the likelihood is for the branching times (0) or the phylogeny (1)
soc	Sets whether stem or crown age should be used (1 or 2)
allbp	Sets whether a search should be done with various initial conditions, with tshift at each of the branching points (TRUE/FALSE)
tol	Sets the tolerances in the optimization. Consists of: reltolx = relative tolerance of parameter values in optimization reltolf = relative tolerance of function value in optimization abstolx = absolute tolerance of parameter values in optimization
maxiter	Sets the maximum number of iterations in the optimization
changeloglikifnoconv	if TRUE the loglik will be set to -Inf if ML does not converge
optimmethod	Method used in optimization of the likelihood. Current default is 'subplex'. Alternative is 'simplex' (default of previous versions)
num_cycles	the number of cycles of optimization. If set at Inf, it will do as many cycles as needed to meet the tolerance set for the target function.
methode	The method used to solve the master equation, default is 'analytical' which uses matrix exponentiation; alternatively numerical ODE solvers can be used, such as 'odeint::runge_kutta_cash_karp54'. These were used in the package before version 3.1.
verbose	Show the parameters and loglikelihood for every call to the loglik function

Details

The output is a dataframe containing estimated parameters and maximum loglikelihood. The computed loglikelihood contains the factor $q! / m!(q + m)!$ where q is the number of species in the phylogeny and m is the number of missing species, as explained in the supplementary material to Etienne et al. 2012.

Value

lambda_1	gives the maximum likelihood estimate of lambda before the shift
mu_1	gives the maximum likelihood estimate of mu before the shift
K_1	gives the maximum likelihood estimate of K before the shift
lambda_2	gives the maximum likelihood estimate of lambda after the shift
mu_2	gives the maximum likelihood estimate of mu after the shift
K_2	gives the maximum likelihood estimate of K after the shift
t_shift	gives the time of the shift
loglik	gives the maximum loglikelihood
df	gives the number of estimated parameters, i.e. degrees of freedom
conv	gives a message on convergence of optimization; conv = 0 means convergence

Note

The optimization may get trapped in local optima. Try different starting values to search for the global optimum.

Author(s)

Rampal S. Etienne & Bart Haegeman

References

- Etienne, R.S. et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. Am. Nat. 180: E75-E89, doi: 10.1086/667574

See Also

[dd_SR_loglik](#), [dd_ML](#), [dd_KI_ML](#),

Examples

```
cat("This will estimate parameters for a sets of branching times brts without conditioning.\n")
cat("The tolerance of the optimization is set ridiculously high to make runtime fast.\n")
cat("In real applications, use the default or more stringent settings for tol.\n")
brts = 1:10
dd_SR_ML(brts = brts, initparsopt = c(0.4581, 1E-6, 17.69, 11.09, 8.9999), idparsopt = c(1:3,6,7),
         idparsfix = NULL, parsfix = NULL, idparsnoshift = c(4,5), cond = 0,
         tol = c(1E-1,1E-1,1E-1),optimmethod = 'simplex'
)
```

dd_SR_sim	<i>Function to simulate the diversity-dependent diversification process with a shift in one or more of the parameters</i>
-----------	---

Description

Simulating the diversity-dependent diversification process with a parameter shift at a certain time

Usage

```
dd_SR_sim(pars, age, ddmodel = 1)
```

Arguments

pars	<p>Vector of parameters:</p> <p>pars[1] corresponds to lambda1 (speciation rate before the rate shift) pars[2] corresponds to mu1 (extinction rate before the rate shift) pars[3] corresponds to K1 (clade-level carrying capacity before the rate shift) pars[4] corresponds to lambda2 (speciation rate after the rate shift) pars[5] corresponds to mu2 (extinction rate after the rate shift) pars[6] corresponds to K2 (clade-level carrying capacity after the rate shift) pars[7] corresponds to the time of shift</p>
age	Sets the crown age for the simulation
ddmodel	<p>Sets the model of diversity-dependence:</p> <p>ddmodel == 1 : linear dependence in speciation rate with parameter K (= diversity where speciation = extinction) ddmodel == 1.3 : linear dependence in speciation rate with parameter K' (= diversity where speciation = 0) ddmodel == 2 : exponential dependence in speciation rate with parameter K (= diversity where speciation = extinction) ddmodel == 2.1 : variant of exponential dependence in speciation rate with offset at infinity ddmodel == 2.2 : 1/n dependence in speciation rate ddmodel == 2.3 : exponential dependence in speciation rate with parameter x (= exponent) ddmodel == 3 : linear dependence in extinction rate ddmodel == 4 : exponential dependence in extinction rate ddmodel == 4.1 : variant of exponential dependence in extinction rate with offset at infinity ddmodel == 4.2 : 1/n dependence in extinction rate with offset at infinity ddmodel == 5 : linear dependence in speciation and extinction rate</p>

Value

out	A list with the following four elements: The first element is the tree of extant species in phylo format
-----	--

The second element is the tree of all species, including extinct species, in phylo format

The third element is a matrix of all species where

- the first column is the time at which a species is born
- the second column is the label of the parent of the species; positive and negative values only indicate whether the species belongs to the left or right crown lineage
- the third column is the label of the daughter species itself; positive and negative values only indicate whether the species belongs to the left or right crown lineage
- the fourth column is the time of extinction of the species. If this equals -1, then the species is still extant.

The fourth element is the set of branching times of the tree of extant species.

Author(s)

Rampal S. Etienne

References

- Etienne, R.S. et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. Am. Nat. 180: E75-E89, doi: 10.1086/667574

Examples

```
dd_SR_sim(c(0.2, 0.1, 20, 0.2, 0.1, 40, 5), 10)
```

L2brts

Function to convert a table with speciation and extinction events to a set of branching times

Description

Converting a table with speciation and extinction events to a set of branching times

Usage

```
L2brts(L, dropextinct = T)
```

Arguments

L Matrix of events as produced by dd_sim:

- the first column is the time at which a species is born in Mya
- the second column is the label of the parent of the species; positive and negative values indicate whether the species belongs to the left or right crown lineage
- the third column is the label of the daughter species itself; positive and negative

values indicate whether the species belongs to the left or right crown lineage
 - the fourth column is the time of extinction of the species; if the fourth element equals -1, then the species is still extant.

dropextinct Sets whether the phylogeny should drop species that are extinct at the present

Value

brts A set of branching times

Author(s)

Rampal S. Etienne

References

- Etienne, R.S. et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. Am. Nat. 180: E75-E89, doi: 10.1086/667574

Examples

```
sim = dd_sim(c(0.2,0.1,20),10)
phy = L2brts(sim$L)
plot(phy)
```

L2phylo	<i>Function to convert a table with speciation and extinction events to a phylogeny</i>
---------	---

Description

Converting a table with speciation and extinction events to a phylogeny

Usage

```
L2phylo(L, dropextinct = T)
```

Arguments

L Matrix of events as produced by dd_sim:

- the first column is the time at which a species is born in Mya
- the second column is the label of the parent of the species; positive and negative values indicate whether the species belongs to the left or right crown lineage
- the third column is the label of the daughter species itself; positive and negative values indicate whether the species belongs to the left or right crown lineage
- the fourth column is the time of extinction of the species; if the fourth element equals -1, then the species is still extant.

dropextinct Sets whether the phylogeny should drop species that are extinct at the present

Value

phy A phylogeny of the phylo type

Author(s)

Rampal S. Etienne

References

- Etienne, R.S. et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. Am. Nat. 180: E75-E89, doi: 10.1086/667574

Examples

```
sim = dd_sim(c(0.2,0.1,20),10)
phy = L2phylo(sim$L)
plot(phy)
```

optimizer

Carries out optimization (finding a minimum)

Description

A wrapper to use several optimization routines, currently only 'simplex' (a method adopted from Matlab, or 'subplex', from the R package subplex). The function is called from several packages by the same author.

Usage

```
optimizer(
  optimmethod = "simplex",
  optimpars = c(1e-04, 1e-04, 1e-06, 1000),
  num_cycles = 1,
  fun,
  trparsopt,
  jitter = 0,
  ...
)
```

Arguments

optimmethod The method to use for optimization, either 'simplex' or 'subplex'

optimpars Parameters of the optimization: relative tolerance in function arguments, relative tolerance in function value, absolute tolerance in function arguments, and maximum number of iterations

num_cycles	Number of cycles of the optimization. When set to Inf, the optimization will be repeated until the result is, within the tolerance, equal to the starting values, with a maximum of 10 cycles.
fun	Function to be optimized
trparsopt	Initial guess of the parameters to be optimized
jitter	Perturbation of an initial parameter value when precisely equal to 0.5; this is only relevant when subplex is chosen. The default value is 0, so no jitter is applied. A recommended value when using it is 1E-5.
...	Any other arguments of the function to be optimized, or settings of the optimization routine

Value

out	A list containing optimal function arguments (par, the optimal function value (fvalues) and whether the optimization converged (conv)
-----	---

.

Author(s)

Rampal S. Etienne

Examples

```
cat("No examples")
```

phylo2L	<i>Function to convert phylogeny to a table with speciation and extinction events</i>
---------	---

Description

Converting a phylogeny to a table with speciation and extinction events

Usage

```
phylo2L(phy)
```

Arguments

phy	A phylogeny of the phylo type
-----	-------------------------------

Value

L Matrix of events as produced by `dd_sim`:

- the first column is the time at which a species is born in Mya
- the second column is the label of the parent of the species; positive and negative values indicate whether the species belongs to the left or right crown lineage
- the third column is the label of the daughter species itself; positive and negative values indicate whether the species belongs to the left or right crown lineage
- the fourth column is the time of extinction of the species; if the fourth element equals -1, then the species is still extant.

Author(s)

Liang Xu

References

- Etienne, R.S. et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. Am. Nat. 180: E75-E89, doi: 10.1086/667574

Examples

```
sim = dd_sim(c(0.2,0.1,20),10)
phy = sim$tas
L = phylo2L(phy)
phy2 = L2phylo(L, dropextinct = FALSE)
graphics::par(mfrow = c(1,3))
graphics::plot(phy)
graphics::plot(phy2)
graphics::plot(L2phylo(sim$L, dropextinct = FALSE))
```

`rng_respecting_sample` *Sampling in which zero probabilities are removed*

Description

Sampling in which zero probabilities are removed

Usage

```
rng_respecting_sample(x, size, replace, prob)
```

Arguments

x	either a vector of one or more elements from which to choose, or a positive integer. See ‘Details.’
size	a non-negative integer giving the number of items to choose.
replace	should sampling be with replacement?
prob	a vector of probability weights for obtaining the elements of the vector being sampled.

Value

a vector of length size with elements drawn from either x or from the integers 1:x.

Note

thanks to Pedro Neves for finding this feature in base::sample

Author(s)

Richel J.C. Bilderbeek

See Also

See [sample](#) for more details

Examples

```
# Number of draws
n <- 1000

# Do normal sampling
set.seed(42)
draws_1 <- DDD:::rng_respecting_sample(
  1:3, size = n, replace = TRUE, prob = c(1.0, 1.0, 1.0)
)

# Do a sampling with one element of probability zero
set.seed(42)
draws_2 <- DDD:::rng_respecting_sample(
  1:4, size = n, replace = TRUE, prob = c(1.0, 1.0, 1.0, 0.0)
)
testit::assert(sum(draws_2 == 4) == 0)
testit::assert(draws_1 == draws_2)

# Use base sampling will give different results,
# as it results in different RNG values
set.seed(42)
draws_3 <- sample(
  1:4, size = n, replace = TRUE, prob = c(1.0, 1.0, 1.0, 0.0)
)
testit::assert(sum(draws_3 == 4) == 0)
```

```
testit::assert(!all(draws_1 == draws_3))
```

roundn	<i>Rounds up in the usual manner</i>
--------	--------------------------------------

Description

The standard round function in R rounds x.5 to the nearest even integer. This is odd behavior that is corrected in roundn

Usage

```
roundn(x, digits = 0)
```

Arguments

x	Number to be rounded
digits	Sets the number of decimals in rounding.

Value

n	A number
---	----------

Author(s)

Rampal S. Etienne

Examples

```
round(2.5)
roundn(2.5)
round(3.5)
roundn(3.5)
round(2.65,digits = 1)
roundn(2.65,digits = 1)
round(2.75,digits = 1)
roundn(2.75,digits = 1)
```

sample2	<i>Takes samples in the usual manner</i>
---------	--

Description

The standard sample function in R samples from n numbers when x = n. This is unwanted behavior when the size of the vector to sample from changes dynamically. This is corrected in sample2

Usage

```
sample2(x, size, replace = FALSE, prob = NULL)
```

Arguments

x	A vector of one or more elements
size	A non-negative integer giving the number of items to choose.
replace	Should sampling be with replacement?
prob	A vector of probability weights for obtaining the elements of the vector being sampled.

Value

sam	A vector of length size that is sampled from x.
-----	---

Author(s)

Rampal S. Etienne

Examples

```
sample(x = 10,size = 5,replace = TRUE)
sample2(x = 10,size = 5,replace = TRUE)
```

simplex	<i>Carries out optimization using a simplex algorithm (finding a minimum)</i>
---------	---

Description

Function to optimize target function using a simplex method adopted from Matlab

Usage

```
simplex(fun, trparsopt, optimpars, ...)
```

Arguments

fun	Function to be optimized
trparsopt	Initial guess of the parameters to be optimized
optimpars	Parameters of the optimization: relative tolerance in function arguments, relative tolerance in function value, absolute tolerance in function arguments, and maximum number of iterations
...	Any other arguments of the function to be optimized, or settings of the optimization routine

Value

out	A list containing optimal function arguments (par), the optimal function value (fvalues) and whether the optimization converged (conv)
-----	--

Author(s)

Rampal S. Etienne

Examples

```
cat("No examples")
```

 td_sim |

Simulation of a diversity-dependent-like time-dependent process

Description

Simulates a phylogenetic tree branching according to a time-dependent process calibrated on the expected number of species under a diversity-dependent process over time.

Usage

```
td_sim(pars, age, ddmodel = 1, methode = "ode45")
```

Arguments

pars	Vector of parameters: pars[1] corresponds to lambda0 (speciation rate) pars[2] corresponds to mu0 (extinction rate) pars[3] corresponds to lambda1 (decline parameter in speciation rate) or K in diversity-dependence-like models pars[4] corresponds to mu1 (decline parameter in extinction rate)
------	--

age	crown age of the tree to simulate, i.e. the simulation time.
ddmodel	the diversity-dependent model used as reference for the time-dependent model.
methode	The method used to solve the master equation. See <code>deSolve::ode()</code> documentation for possible inputs

Value

A list with the following four elements: The first element is the tree of extant species in phylo format

The second element is the tree of all species, including extinct species, in phylo format

The third element is a matrix of all species where - the first column is the time at which a species is born

- the second column is the label of the parent of the species; positive and negative values only indicate whether the species belongs to the left or right crown lineage

- the third column is the label of the daughter species itself; positive and negative values only indicate whether the species belongs to the left or right crown lineage

- the fourth column is the time of extinction of the species. If this equals -1, then the species is still extant.

Author(s)

César Martinez, Rampal S. Etienne

transform_pars	<i>Transforming parameters from -Inf to Inf into parameters from -1 to 1</i>
----------------	--

Description

Function to transform pars in a way that is more useful for optimization: `trpars <- sign(pars) * pars/(sign(pars) + pars)`;

Usage

```
transform_pars(pars)
```

Arguments

pars	Parameters to be transformed
------	------------------------------

Value

Transformed parameters

Author(s)

Rampal S. Etienne

untransform_pars	<i>Untransforming parameters from -1 to 1 into parameters from -Inf to Inf.</i>
------------------	---

Description

Function to untransform pars after optimization: $\text{pars} \leftarrow \text{sign}(\text{trpars}) * \text{trpars} / (\text{sign}(\text{trpars}) - \text{trpars});$

Usage

```
untransform_pars(trpars)
```

Arguments

trpars Parameters to be untransformed

Value

Untransformed parameters

Author(s)

Rampal S. Etienne

Index

* models

- bd_loglik, [2](#)
- bd_ML, [4](#)
- brts2phylo, [7](#)
- conv, [8](#)
- dd_KI_loglik, [8](#)
- dd_KI_ML, [11](#)
- dd_KI_sim, [14](#)
- dd_loglik, [16](#)
- dd_LR, [18](#)
- dd_ML, [21](#)
- dd_MS_loglik, [25](#)
- dd_MS_ML, [27](#)
- dd_MS_sim, [30](#)
- dd_sim, [34](#)
- dd_SR_loglik, [35](#)
- dd_SR_ML, [37](#)
- dd_SR_sim, [41](#)
- L2brts, [42](#)
- L2phylo, [43](#)
- optimizer, [44](#)
- phylo2L, [45](#)
- roundn, [48](#)
- sample2, [49](#)
- simplex, [49](#)

- dd_MS_sim, [30](#)
- dd_multiple_KI_loglik, [32](#)
- dd_sim, [34](#)
- dd_SR_loglik, [10, 18, 26, 35, 40](#)
- dd_SR_ML, [14, 24, 30, 37, 37](#)
- dd_SR_sim, [41](#)

- L2brts, [42](#)
- L2phylo, [43](#)

- optimizer, [44](#)

- phylo2L, [45](#)

- rng_respecting_sample, [46](#)
- roundn, [48](#)

- sample, [47](#)
- sample2, [49](#)
- simplex, [49](#)

- td_sim, [50](#)
- transform_pars, [51](#)

- untransform_pars, [52](#)

- bd_loglik, [2, 7](#)
- bd_ML, [4, 4](#)
- brts2phylo, [7](#)

- conv, [8](#)

- dd_KI_loglik, [8, 14, 18, 26, 37](#)
- dd_KI_ML, [10, 11, 24, 30, 40](#)
- dd_KI_sim, [14](#)
- dd_loglik, [10, 16, 21, 24, 26, 37](#)
- dd_LR, [18](#)
- dd_ML, [14, 18, 21, 21, 30, 40](#)
- dd_MS_loglik, [25, 30](#)
- dd_MS_ML, [26, 27](#)