

# Package ‘RPANDA’

January 27, 2015

**Version** 1.0

**Date** 2014-02-04

**Type** Package

**Title** Phylogenetic ANalyses of DiversificAtion

**Depends** R (>= 2.14.2), ape, picante, pspline, deSolve

**Suggests** testthat

**Encoding** UTF-8

**Author** H el ene Morlon [aut, cre, cph]

**Maintainer** H el ene Morlon <morlon@biologie.ens.fr>

**Description** Implements fits of diversification models to phylogenetic data. See Morlon et al. PLoSB (2010), Morlon et al. PNAS (2011), Condamine et al. Eco Lett (2013), Morlon et al. Eco Lett (2014) and Manceau et al. (in prep).

**License** GPL-2

**URL** <https://github.com/hmorlon/PANDA>

**NeedsCompilation** no

**Repository** CRAN

**Date/Publication** 2014-12-02 14:36:08

## R topics documented:

RPANDA-package . . . . .	2
Balaenopteridae . . . . .	3
Calomys . . . . .	3
Cetacea . . . . .	4
fit_bd . . . . .	5
fit_coal_cst . . . . .	8
fit_coal_var . . . . .	10
fit_env . . . . .	12
fit_sgd . . . . .	15
InfTemp . . . . .	16
likelihood_bd . . . . .	17

likelihood_coal_cst . . . . .	18
likelihood_coal_var . . . . .	20
likelihood_sgd . . . . .	21
Phocoenidae . . . . .	22
plot_dtt . . . . .	23
plot_fit_bd . . . . .	24
plot_fit_env . . . . .	25
sim_sgd . . . . .	26

<b>Index</b>	<b>28</b>
--------------	-----------

---

RPANDA-package	<i>RPANDA</i>
----------------	---------------

---

## Description

Implements fits of various diversification models to phylogenetic data

## Details

Package: RPANDA  
 Type: Package  
 Version: 1.0  
 Date: 2013-12-27  
 License: GPL (>= 2)

## Author(s)

H el ene Morlon <morlon@biologie.ens.fr>  
 Fabien Condamine <fabien.condamine@gmail.com>  
 Marc Manceau <manceau@biologie.ens.fr>

## References

- Morlon, H., Potts, M.D., Plotkin, J.B. (2010) Inferring the dynamics of diversification: a coalescent approach, PLoS B 8(9): e1000493
- Morlon, H., Kemp, B., Plotkin, J.B., Brisson, D. (2012) Explosive radiation of a bacterial species group, Evolution, 66: 2577-2586
- Morlon, H., Parsons, T.L. and Plotkin, J.B. (2011) Reconciling molecular phylogenies with the fossil record, Proc Nat Acad Sci 108: 16327-16332
- Condamine, F.L., Rolland, J., and Morlon, H. (2013) Macroevolutionary perspectives to environmental change, Eco Lett 16: 72-85

Morlon, H. (2014) Phylogenetic approaches for studying diversification, *Eco Lett* 7: 508-525  
Manceau, M., Lambert, A., Morlon, H. (submitted)

---

Balaenopteridae      *Balaenopteridae phylogeny*

---

### **Description**

Ultrametric phylogenetic tree of the 9 extant Balaenopteridae species

### **Usage**

```
data(Balaenopteridae)
```

### **Details**

This phylogeny was extracted from Steeman et al. *Syst Bio* 2009 cetacean phylogeny

### **References**

Steeman ME et al.(2009) Radiation of extant cetaceans driven by restructuring of the oceans *Syst Biol* 58:573-585  
Morlon, H., Parsons, T.L., Plotkin, J.B. (2011) Reconciling molecular phylogenies with the fossil record *Proc Nat Acad Sci* 108: 16327-16332

### **Examples**

```
data(Balaenopteridae)  
print(Balaenopteridae)  
plot(Balaenopteridae)
```

---

Calomys      *Calomys phylogeny*

---

### **Description**

Ultrametric phylogenetic tree of 11 of the 13 extant Calomys species

### **Usage**

```
data(Calomys)
```

### **Details**

This phylogeny is from Pigot et al. *PloS Biol* 2012

**References**

- Pigot et al.(2012) Speciation and extinction drive the appearance of directional range size evolution in phylogenies and the fossil record *PloS Biol* 10:1-9
- Manceau, M., Lambert, A., Morlon, H. (submitted)

**Examples**

```
data(Calomys)
print(Calomys)
plot(Calomys)
```

---

Cetacea

*Cetacean phylogeny*

---

**Description**

Ultrametric phylogenetic tree for 87 of the 89 extant cetacean species

**Usage**

```
data(Cetacea)
```

**Details**

This phylogeny was constructed by Bayesian phylogenetic inference from six mitochondrial and nine nuclear genes. It was calibrated using seven paleontological age constraints and a relaxed molecular clock approach. See Steeman et al. (2009) for details.

**Source**

Steeman ME et al.(2009) Radiation of extant cetaceans driven by restructuring of the oceans, *Syst Biol* 58:573-585

**References**

- Steeman ME et al.(2009) Radiation of extant cetaceans driven by restructuring of the oceans *Syst Biol* 58:573-585
- Morlon, H., Parsons, T.L., Plotkin, J.B. (2011) Reconciling molecular phylogenies with the fossil record *Proc Nat Acad Sci* 108: 16327-16332
- Condamine, F.L., Rolland, J., Morlon, H. (2013) Macroevolutionary perspectives to environmental change *Eco Lett* 16: 72-85

**Examples**

```
data(Cetacea)
print(Cetacea)
plot(Cetacea)
```

fit\_bd

*Maximum likelihood fit of the general birth-death model***Description**

Fits the birth-death model with potentially time-varying rates and potentially missing extant species to a phylogeny, by maximum likelihood. Notations follow Morlon et al. PNAS 2011.

**Usage**

```
fit_bd(phylo, tot_time, f.lamb, f.mu, lamb_par, mu_par, f = 1,
      meth = "Nelder-Mead", cst.lamb = FALSE, cst.mu = FALSE,
      expo.lamb = FALSE, expo.mu = FALSE, fix.mu = FALSE,
      dt=0, cond = "crown")
```

**Arguments**

phylo	an object of type 'phylo' (see ape documentation)
tot_time	the age of the phylogeny (crown age, or stem age if known). If working with crown ages, tot_time is given by max(node.age(phylo)\$ages).
f.lamb	a function specifying the hypothesized functional form (e.g. constant, linear, exponential, etc.) of the variation of the speciation rate $\lambda$ with time. Any functional form may be used. This function has two arguments: the first argument is time; the second argument is a numeric vector of the parameters of the time-variation (to be estimated).
f.mu	a function specifying the hypothesized functional form (e.g. constant, linear, exponential, etc.) of the variation of the extinction rate $\mu$ with time. Any functional form may be used. This function has two arguments: the first argument is time; the second argument is a numeric vector of the parameters of the time-variation (to be estimated).
lamb_par	a numeric vector of initial values for the parameters of f.lamb to be estimated (these values are used by the optimization algorithm). The length of this vector is used to compute the total number of parameters in the model, so to fit a model with constant speciation rate (for example), lamb_par should be a vector of length 1. Otherwise aic values will be wrong.
mu_par	a numeric vector of initial values for the parameters of f.mu to be estimated (these values are used by the optimization algorithm). The length of this vector is used to compute the total number of parameters in the model, so to fit a model without extinction (for example), mu_par should be empty (vector of length 0). Otherwise aic values will be wrong.
f	the fraction of extant species included in the phylogeny
meth	optimization to use to maximize the likelihood function, see <i>optim</i> for more details.

cst.lamb	logical: should be set to TRUE only if f.lamb is constant (i.e. does not depend on time) to use analytical instead of numerical computation in order to reduce computation time.
cst.mu	logical: should be set to TRUE only if f.mu is constant (i.e. does not depend on time) to use analytical instead of numerical computation in order to reduce computation time.
expo.lamb	logical: should be set to TRUE only if f.lamb is exponential to use analytical instead of numerical computation in order to reduce computation time.
expo.mu	logical: should be set to TRUE only if f.mu is exponential to use analytical instead of numerical computation in order to reduce computation time.
fix.mu	logical: if set to TRUE, the extinction rate $\mu$ is fixed and will not be optimized.
dt	the default value is 0. In this case, integrals in the likelihood are computed using R "integrate" function, which can be quite slow. If a positive dt is given as argument, integrals are computed using a piece-wise constant approximation, and dt represents the length of the intervals on which functions are assumed to be constant. For an exponential dependency of the speciation rate with time, we found that dt=1e-3 gives a good trade-off between precision and computation time.
cond	conditioning to use to fit the model: <ul style="list-style-type: none"> <li>• FALSE: no conditioning (not recommended);</li> <li>• "stem": conditioning on the survival of the stem lineage (use when the stem age is known, in this case tot_time should be the stem age);</li> <li>• "crown" (default): conditioning on a speciation event at the crown age and survival of the 2 daughter lineages (use when the stem age is not known, in this case tot_time should be the crown age).</li> </ul>

## Details

The lengths of lamb\_par and mu\_par are used to compute the total number of parameters in the model, so to fit a model with constant speciation rate (for example), lamb\_par should be a vector of length 1. Otherwise aic values will be wrong. In the f.lamb and f.mu functions, the first argument (time) runs from the present to the past. Hence, if the parameter controlling the variation of  $\lambda$  with time is estimated to be positive (for example), this means that the speciation rate decreases from past to present.

## Value

a list with the following components

model	the name of the fitted model
LH	the maximum log-likelihood value
aicc	the second order Akaike's Information Criterion
lamb_par	a numeric vector of estimated f.lamb parameters, in the same order as defined in f.lamb
mu_par	a numeric vector of estimated f.mu parameters, in the same order as defined in f.mu (if fix.mu is FALSE)

**Author(s)**

H Morlon

**References**

Morlon, H., Parsons, T.L. and Plotkin, J.B. (2011) Reconciling molecular phylogenies with the fossil record *Proc Nat Acad Sci* 108: 16327-16332

Morlon, H. (2014) Phylogenetic approaches for studying diversification, *Eco Lett* 17:508-525

**See Also**

[plot\\_fit\\_bd](#), [plot\\_dtt](#), [likelihood\\_bd](#)

**Examples**

```
# Some examples may take a little bit of time. Be patient!

data(Cetacea)
tot_time<-max(node.age(Cetacea)$ages)

# Fit the pure birth model (no extinction) with a constant speciation rate
f.lamb <-function(t,y){y[1]}
f.mu<-function(t,y){0}
lamb_par<-c(0.09)
mu_par<-c()
#result_cst <- fit_bd(Cetacea,tot_time,f.lamb,f.mu,lamb_par,mu_par,
#                    f=87/89,cst.lamb=TRUE,fix.mu=TRUE,dt=1e-3)
#result_cst$model <- "pure birth with constant speciation rate"

# Fit the pure birth model (no extinction) with exponential variation
# of the speciation rate with time
f.lamb <-function(t,y){y[1] * exp(y[2] * t)}
f.mu<-function(t,y){0}
lamb_par<-c(0.05, 0.01)
mu_par<-c()
#result_exp <- fit_bd(Cetacea,tot_time,f.lamb,f.mu,lamb_par,mu_par,
#                    f=87/89,expo.lamb=TRUE,fix.mu=TRUE,dt=1e-3)
#result_exp$model <- "pure birth with exponential variation in speciation rate"

# Fit the a pure birth model (no extinction) with linear variation of
# the speciation rate with time
f.lamb <-function(t,y){y[1] + y[2] * t}
f.mu<-function(t,y){0}
lamb_par<-c(0.09, 0.001)
mu_par<-c()
#result_lin <- fit_bd(Cetacea,tot_time,f.lamb,f.mu,lamb_par,mu_par,f=87/89,fix.mu=TRUE,dt=1e-3)
#result_lin$model <- "pure birth with linear variation in speciation rate"

# Fit a birth-death model with exponential variation of the speciation
# rate with time and constant extinction
f.lamb<-function(t,y){y[1] * exp(y[2] * t)}
```

```
f.mu <-function(t,y){y[1]}
lamb_par <- c(0.05, 0.01)
mu_par <-c(0.005)
#result_bexp_dcst <- fit_bd(Cetacea,tot_time,f.lamb,f.mu,lamb_par,mu_par,
#                           f=87/89,expo.lamb=TRUE,cst.mu=TRUE,dt=1e-3)
#result_bexp_dcst$model <- "birth-death with exponential variation in speciation rate
#                           and constant extinction"

# Find the best model
#index <- which.min(c(result_cst$aicc, result_exp$aicc, result_lin$aicc,result_bexp_dcst$aicc))
#rbind(result_exp, result_lin, result_cst, result_bexp_dcst)[index,]
```

---

fit\_coal\_cst

*Maximum likelihood fit of the equilibrium model*


---

## Description

Fits the equilibrium diversity model with potentially time-varying turnover rate and potentially missing extant species to a phylogeny, by maximum likelihood. The implementation allows only exponential time variation of the turnover rate, although this could be modified using expressions in Morlon et al. PLoSB 2010. Notations follow Morlon et al. PLoSB 2010.

## Usage

```
fit_coal_cst(phylo, tau0 = 1e-2, gamma = 1, cst.rate = FALSE,
             meth = "Nelder-Mead", N0 = 0)
```

## Arguments

phylo	an object of type 'phylo' (see ape documentation)
tau0	initial value of the turnover rate at present (used by the optimization algorithm)
gamma	initial value of the parameter controlling the exponential variation in turnover rate (used by the optimization algorithm)
cst.rate	logical: should be set to TRUE to fit an equilibrium diversity model with time-constant turnover rate (known as the Hey model, model 1 in Morlon et al. PLoSB 2010). By default, a model with exponential time-varying rate exponential is fitted (model 2 in Morlon et al. PLoSB 2010).
meth	optimization to use to maximize the likelihood function, see <i>optim</i> for more details.
N0	Number of extant species. With default value(0), N0 is set to the number of tips in the phylogeny. That is, the phylogeny is assumed to be 100% complete.



## Details

This function fits models 1 (when `cst.rate=TRUE`) and 2 (when `cst.rate=FALSE`) from the PloSB 2010 paper. Likelihoods arising from these models are directly comparable to likelihoods from the `fit_coal_var` function, thus allowing to test support for equilibrium versus expanding diversity scenarios. Time runs from the present to the past. Hence, if `gamma` is estimated to be positive (for example), this means that the speciation rate decreases from past to present.

## Value

a list with the following components

<code>model</code>	the name of the fitted model
<code>LH</code>	the maximum log-likelihood value
<code>aicc</code>	the second order Akaike's Information Criterion
<code>tau0</code>	the estimated turnover rate at present
<code>gamma</code>	the estimated parameter controlling the exponential variation in turnover rate (if <code>cst.rate</code> is <code>FALSE</code> )

## Author(s)

H Morlon

## References

- Hey, J. (1992) Using phylogenetic trees to study speciation and extinction, *Evolution*, 46: 627-640
- Morlon, H., Potts, M.D., Plotkin, J.B. (2010) Inferring the dynamics of diversification: a coalescent approach, *PLoS B*, 8(9): e1000493
- Morlon, H., Kemp, B., Plotkin, J.B., Brisson, D. (2012) Explosive radiation of a bacterial species group, *Evolution*, 66: 2577-2586
- Morlon, H. (2014) Phylogenetic approaches for studying diversification, *Eco Lett*, 17:508-525

## See Also

[likelihood\\_coal\\_cst](#), [fit\\_coal\\_var](#)

## Examples

```
data(Cetacea)
result <- fit_coal_cst(Cetacea, tau0=1.e-3, gamma=-1, cst.rate=FALSE, N0=89)
print(result)
```

fit\_coal\_var

*Fit birth-death model using a coalescent approach***Description**

Fits the expanding diversity model with potentially time-varying rates and potentially missing extant species to a phylogeny, by maximum likelihood. The implementation allows only exponential time variation of the speciation and extinction rates, although this could be modified using expressions in Morlon et al. PLoSB 2010. Notations follow Morlon et al. PLoSB 2010.

**Usage**

```
fit_coal_var(phylo, lamb0 = 0.1, alpha = 1, mu0 = 0.01, beta = 0,
             meth = "Nelder-Mead", N0 = 0, cst.lamb = FALSE, cst.mu = FALSE,
             fix.eps = FALSE, mu.0 = FALSE, pos = TRUE)
```

**Arguments**

phylo	an object of type 'phylo' (see ape documentation)
lamb0	initial value of the speciation rate at present (used by the optimization algorithm)
alpha	initial value of the parameter controlling the exponential variation in speciation rate (used by the optimization algorithm)
mu0	initial value of the extinction rate at present (used by the optimization algorithm)
beta	initial value of the parameter controlling the exponential variation in extinction rate.
meth	optimization to use to maximize the likelihood function, see <i>optim</i> for more details.
N0	Number of extant species. With default value(0), N0 is set to the number of tips in the phylogeny. That is, the phylogeny is assumed to be 100% complete.
cst.lamb	logical: should be set to TRUE only if f.lamb is constant (i.e. does not depend on time, models 3, 4b & 5 in Morlon et al. PLoSB 2010) to use analytical instead of numerical computation in order to reduce computation time.
cst.mu	logical: should be set to TRUE only if f.mu is constant (i.e. does not depend on time, models 3 & 4a in Morlon et al. PLoSB 2010) to use analytical instead of numerical computation in order to reduce computation time.
fix.eps	logical: should be set to TRUE only if the extinction fraction is constant (i.e. does not depend on time, model 4c in Morlon et al. PLoSB 2010)
mu.0	logical: should be set to TRUE to force the extinction rate to 0 (models 5 & 6 in Morlon et al. PLoSB 2010)
pos	logical: should be set to FALSE only to not enforce positive speciation and extinction rates

## Details

The function fits models 3 to 6 from the PloSB 2010 paper. Likelihoods arising from these models are computed using the coalescent approximation and are directly comparable to likelihoods from the fit\_coal\_cst function, thus allowing to test support for equilibrium versus expanding diversity scenarios.

These models can be fitted using the options specified below:

- model 3:  
with `cst.lamb=TRUE` & `cst.mu=TRUE`
- model 4a:  
with `cst.lamb=FALSE` & `cst.mu=TRUE`
- model 4b:  
with `cst.lamb=TRUE` & `cst.mu=FALSE`
- model 4c:  
with `cst.lamb=FALSE`, `cst.mu=FALSE` & `fix.eps=TRUE`
- model 4d:  
with `cst.lamb=FALSE`, `cst.mu=FALSE` & `fix.eps=FALSE`
- model 5:  
with `cst.lamb=TRUE` & `mu.0=TRUE`
- model 6:  
with `cst.lamb=FALSE` & `mu.0=TRUE`

Time runs from the present to the past. Hence, if alpha is estimated to be positive (for example), this means that the speciation rate decreases from past to present.

## Value

a list with the following components

<code>model</code>	the name of the fitted model
<code>LH</code>	the maximum log-likelihood value
<code>aicc</code>	the second order Akaike's Information Criterion
<code>model.parameters</code>	the estimated parameters

## Author(s)

H Morlon

## References

- Morlon, H., Potts, M.D., Plotkin, J.B. (2010) Inferring the dynamics of diversification: a coalescent approach, PLoS B 8(9): e1000493
- Morlon, H., Kemps, B., Plotkin, J.B., Brisson, D. (2012) Explosive radiation of a bacterial species group, *Evolution*, 66: 2577-2586
- Morlon, H. (2014) Phylogenetic approaches for studying diversification, *Eco Lett*, 17:508-525

**See Also**

[likelihood\\_coal\\_var](#), [fit\\_coal\\_cst](#)

**Examples**

```
data(Cetacea)
result <- fit_coal_var(Cetacea, lamb0=0.01, alpha=-0.001, mu0=0.0, beta=0, N0=89)
print(result)
```

---

fit\_env

*Maximum likelihood fit of the environmental birth-death model*


---

**Description**

Fits the environmental birth-death model with potentially missing extant species to a phylogeny, by maximum likelihood. Notations follow Morlon et al. PNAS 2011 and Condamine et al. ELE 2013.

**Usage**

```
fit_env(phylo, env_data, tot_time, f.lamb, f.mu, lamb_par, mu_par, df= NULL, f = 1,
meth = "Nelder-Mead", cst.lamb = FALSE, cst.mu = FALSE,
expo.lamb = FALSE, expo.mu = FALSE, fix.mu = FALSE,
dt=0, cond = "crown")
```

**Arguments**

phylo	an object of type 'phylo' (see ape documentation)
env_data	environmental data, given as a data frame with two columns. The first column is time, the second column is the environmental data (temperature for instance).
tot_time	the age of the phylogeny (crown age, or stem age if known). If working with crown ages, tot_time is given by <code>max(node.age(phylo)\$ages)</code> .
f.lamb	a function specifying the hypothesized functional form of the variation of the speciation rate $\lambda$ with time and the environmental variable. Any functional form may be used. This function has three arguments: the first argument is time; the second argument is the environmental variable; the third argument is a numeric vector of the parameters controlling the time and environmental variation (to be estimated).
f.mu	a function specifying the hypothesized functional form of the variation of the extinction rate $\mu$ with time and the environmental variable. Any functional form may be used. This function has three arguments: the first argument is time; the second argument is the environmental variable; the second argument is a numeric vector of the parameters controlling the time and environmental variation (to be estimated).

lamb_par	a numeric vector of initial values for the parameters of f.lamb to be estimated (these values are used by the optimization algorithm). The length of this vector is used to compute the total number of parameters in the model, so to fit a model with constant speciation rate (for example), lamb_par should be a vector of length 1. Otherwise aic values will be wrong.
mu_par	a numeric vector of initial values for the parameters of f.mu to be estimated (these values are used by the optimization algorithm). The length of this vector is used to compute the total number of parameters in the model, so to fit a model without extinction (for example), mu_par should be empty (vector of length 0). Otherwise aic values will be wrong.
df	the degree of freedom to use to define the spline. As a default, smooth.spline(env_data[,1], env_data[,2])\$df is used. See <i>sm.spline</i> for details.
f	the fraction of extant species included in the phylogeny
meth	optimization to use to maximize the likelihood function, see <i>optim</i> for more details.
cst.lamb	logical: should be set to TRUE only if f.lamb is constant (i.e. does not depend on time or the environmental variable) to use analytical instead of numerical computation in order to reduce computation time.
cst.mu	logical: should be set to TRUE only if f.mu is constant (i.e. does not depend on time or the environmental variable) to use analytical instead of numerical computation in order to reduce computation time.
expo.lamb	logical: should be set to TRUE only if f.lamb is an exponential function of time (and does not depend on the environmental variable) to use analytical instead of numerical computation in order to reduce computation time.
expo.mu	logical: should be set to TRUE only if f.mu is an exponential function of time (and does not depend on the environmental variable) to use analytical instead of numerical computation in order to reduce computation time.
fix.mu	logical: if set to TRUE, the extinction rate $\mu$ is fixed and will not be optimized.
dt	the default value is 0. In this case, integrals in the likelihood are computed using R "integrate" function, which can be quite slow. If a positive dt is given as argument, integrals are computed using a piece-wise constant approximation, and dt represents the length of the intervals on which functions are assumed to be constant. We found that 1e-3 generally provides a good trade-off between precision and computation time.
cond	conditioning to use to fit the model: <ul style="list-style-type: none"> <li>• FALSE: no conditioning (not recommended);</li> <li>• "stem": conditioning on the survival of the stem lineage (use when the stem age is known, in this case tot_time should be the stem age);</li> <li>• "crown" (default): conditioning on a speciation event at the crown age and survival of the 2 daughter lineages (use when the stem age is not known, in this case tot_time should be the crown age).</li> </ul>

## Details

The lengths of lamb\_par and mu\_par are used to compute the total number of parameters in the model, so to fit a model with constant speciation rate (for example), lamb\_par should be a vector of

length 1. Otherwise aic values will be wrong. In the f.lamb and f.mu functions, time runs from the present to the past.

### Value

a list with the following components

model	the name of the fitted model
LH	the maximum log-likelihood value
aicc	the second order Akaike's Information Criterion
lamb_par	a numeric vector of estimated f.lamb parameters, in the same order as defined in f.lamb
mu_par	a numeric vector of estimated f.mu parameters, in the same order as defined in f.mu (if fix.mu is FALSE)

### Note

The speed of convergence of the fit might depend on the degree of freedom chosen to define the spline.

### Author(s)

H Morlon and F Condamine

### References

Morlon, H., Parsons, T.L. and Plotkin, J.B. (2011) Reconciling molecular phylogenies with the fossil record *Proc Nat Acad Sci* 108: 16327-16332

Condamine, F.L., Rolland, J., and Morlon, H. (2013) Macroevolutionary perspectives to environmental change, *Eco Lett* 16: 72-85

Morlon, H. (2014) Phylogenetic approaches for studying diversification, *Eco Lett*, 17:508-525

### See Also

[plot\\_fit\\_env](#), [fit\\_bd](#), [likelihood\\_bd](#)

### Examples

```
data(Cetacea)
tot_time<-max(node.age(Cetacea)$ages)
data(InfTemp)
dof<-smooth.spline(InfTemp[,1], InfTemp[,2])$df

# Fits a model with lambda varying as an exponential function of temperature
# and mu fixed to 0 (no extinction). Here t stands for time and x for temperature.
f.lamb <-function(t,x,y){y[1] * exp(y[2] * x)}
f.mu<-function(t,x,y){0}
lamb_par<-c(0.10, 0.01)
mu_par<-c()
```

```
#result_exp <- fit_env(Cetacea,InfTemp,tot_time,f.lamb,f.mu,lamb_par,mu_par,
#                      f=87/89,fix.mu=TRUE,df=dof,dt=1e-3)
```

---

fit\_sgd *Maximum likelihood fit of the SGD model*

---

### Description

Fits the SGD model with exponential growth of the metacommunity, by maximum likelihood. Notations follow Manceau et al. (submitted)

### Usage

```
fit_sgd(phylo, tot_time, par, f=1, meth = "Nelder-Mead")
```

### Arguments

phylo	an object of type 'phylo' (see ape documentation)
tot_time	the age of the phylogeny (crown age, or stem age if known). If working with crown ages, tot_time is given by <code>max(node.age(phylo)\$ages)</code>
par	a numeric vector of initial values for the parameters (b,d,nu) to be estimated (these values are used by the optimization algorithm)
f	the fraction of extant species included in the phylogeny
meth	optimization to use to maximize the likelihood function, see <i>optim</i> for more details.

### Value

a list with the following components

model	the name of the fitted model
LH	the maximum log-likelihood value
aicc	the second order Akaike's Information Criterion
par	a numeric vector of estimated values of b (birth), b-d (growth) and nu (mutation)

### Author(s)

M Manceau

### References

Manceau, M., Lambert, A., Morlon, H. (submitted)

### See Also

[likelihood\\_sgd](#)

## Examples

```
# Some examples may take a little bit of time. Be patient!
data(Calomys)
tot_time <- max(node.age(Calomys)$ages)
par_init <- c(10, 1, 1)
#fit_sgd(Calomys, tot_time, par_init, f=11/13)
```

---

InfTemp

*Paleotemperature data across the Cenozoic*

---

## Description

Paleotemperature data across the Cenozoic inferred from delta O18 measurements

## Usage

```
data(InfTemp)
```

## Details

Paleotemperature data inferred from delta O18 measurements using the equation of Epstein et al. (1953). The format is a dataframe with the two following variables:

Age a numeric vector corresponding to the geological age, in Myrs before the present

Temperature a numeric vector corresponding to the inferred temperature at that age

## References

Epstein, S., Buchsbaum, R., Lowenstam, H.A., Urey, H.C. (1953) Revised carbonate-water isotopic temperature scale *Geol. Soc. Am. Bull.* 64: 1315-1326

Zachos, J.C., Dickens, G.R., Zeebe, R.E. (2008) An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics *Nature* 451: 279-283

Condamine, F.L., Rolland, J., Morlon, H. (2013) Macroevolutionary perspectives to environmental change *Eco Lett* 16: 72-85

## Examples

```
data(InfTemp)
plot(InfTemp)
```



likelihood\_bd

*Likelihood of a phylogeny under the general birth-death model***Description**

Computes the likelihood of a phylogeny under a birth-death model with potentially time-varying rates and potentially missing extant species. Notations follow Morlon et al. PNAS 2011.

**Usage**

```
likelihood_bd(phylo, tot_time, f.lamb, f.mu, f, cst.lamb = FALSE, cst.mu = FALSE,
              expo.lamb = FALSE, expo.mu = FALSE, dt=0, cond = "crown")
```

**Arguments**

phylo	an object of type 'phylo' (see ape documentation)
tot_time	the age of the phylogeny (crown age, or stem age if known). If working with crown ages, tot_time is given by <code>max(node.age(phylo)\$ages)</code> .
f.lamb	a function specifying the time-variation of the speciation rate $\lambda$ . This function as a single argument (time). Any function may be used.
f.mu	a function specifying the time-variation of the speciation rate $\mu$ . This function as a single argument (time). Any function may be used.
f	the fraction of extant species included in the phylogeny
cst.lamb	logical: should be set to TRUE only if f.lamb is constant (i.e. does not depend on time) to use analytical instead of numerical computation in order to reduce computation time.
cst.mu	logical: should be set to TRUE only if f.mu is constant (i.e. does not depend on time) to use analytical instead of numerical computation in order to reduce computation time.
expo.lamb	logical: should be set to TRUE only if f.lamb is exponential to use analytical instead of numerical computation in order to reduce computation time.
expo.mu	logical: should be set to TRUE only if f.mu is exponential to use analytical instead of numerical computation in order to reduce computation time.
dt	the default value is 0. In this case, integrals in the likelihood are computed using R "integrate" function, which can be quite slow. If a positive dt is given as argument, integrals are computed using a piece-wise constant approximation, and dt represents the length of the intervals on which functions are assumed to be constant. For an exponential dependency of the speciation rate with time, we found that <code>dt=1e-3</code> gives a good trade-off between precision and computation time.
cond	conditioning to use to fit the model: <ul style="list-style-type: none"> <li>• FALSE: no conditioning (not recommended);</li> <li>• "stem": conditioning on the survival of the stem lineage (use when the stem age is known, in this case tot_time should be the stem age);</li> </ul>

- "crown" (default): conditioning on a speciation event at the crown age and survival of the 2 daughter lineages (use when the stem age is not known, in this case tot\_time should be the crown age).

### Details

When specifying f.lamb and f.mu, time runs from the present to the past (hence if the speciation rate decreases with time, f.lamb must be a positive function of time).

### Value

the loglikelihood value of the phylogeny, given f.lamb and f.mu

### Author(s)

H Morlon

### References

Morlon, H., Parsons, T.L. and Plotkin, J.B. (2011) Reconciling molecular phylogenies with the fossil record *Proc Nat Acad Sci* 108: 16327-16332

### Examples

```
data(Cetacea)
tot_time <- max(node.age(Cetacea)$ages)
# Compute the likelihood for a pure birth model (no extinction) with
# an exponential variation of speciation rate with time
lamb_par <- c(0.1, 0.01)
f.lamb <- function(t){lamb_par[1] * exp(lamb_par[2] * t)}
f.mu <- function(t){0}
f <- 87/89
lh <- likelihood_bd(Cetacea,tot_time,f.lamb,f.mu,f,cst.mu=TRUE,expo.lamb=TRUE, dt=1e-3)
```

---

likelihood\_coal\_cst     *Likelihood of a phylogeny under the equilibrium diversity model*

---

### Description

Computes the likelihood of a phylogeny under the equilibrium diversity model with potentially time-varying rates and potentially missing extant species. Notations follow Morlon et al. PloSB 2010.

### Usage

```
likelihood_coal_cst(Vtimes, ntips, tau0, gamma, N0)
```

**Arguments**

Vtimes	a vector of branching times (sorted from present to past)
ntips	the number of tips in the phylogeny
tau0	the turnover rate at present
gamma	the parameter controlling the exponential variation in turnover rate. With gamma=0, the turnover rate is constant over time.
N0	the number of extant species

**Details**

Time runs from the present to the past. Hence, a positive gamma (for example) means that the turnover rate declines from past to present.

**Value**

a list containing the following components:

res	the loglikelihood value of the phylogeny, given tau0 and gamma
all	vector of all the individual loglikelihood values corresponding to each branching event

**Author(s)**

H Morlon

**References**

Morlon, H., Potts, M.D., Plotkin, J.B. (2010) Inferring the dynamics of diversification: a coalescent approach, PLoS B 8(9): e1000493

**Examples**

```
data(Cetacea)
Vtimes <- sort(branching.times(Cetacea))
tau0 <- 0.1
gamma <- 0.001
ntips <- Ntip(Cetacea)
N0 <- 89
likelihood <- likelihood_coal_cst(Vtimes,ntips,tau0,gamma,N0)
```

---

likelihood\_coal\_var    *Likelihood of a birth-death model using a coalescent approach*

---

### Description

Computes the likelihood of a phylogeny under the expanding diversity model with potentially time-varying rates and potentially missing extant species to a phylogeny. Notations follow Morlon et al. PLoSB 2010.

### Usage

```
likelihood_coal_var(Vtimes, ntips, lamb0, alpha, mu0, beta, N0, pos = TRUE)
```

### Arguments

Vtimes	a vector of branching times (sorted from present to past)
ntips	number of species in the phylogeny
lamb0	the speciation rate at present
alpha	the parameter controlling the exponential variation in speciation rate.
mu0	the extinction rate at present
beta	the parameter controlling the exponential variation in extinction rate.
N0	the number of extant species
pos	logical: should be set to FALSE only to not enforce positive speciation and extinction rates

### Details

Time runs from the present to the past. Hence, a positive alpha (for example) means that the speciation rate declines from past to present.

### Value

a list containing the following components:

res	the loglikelihood value of the phylogeny, given the parameters
all	vector of all the individual loglikelihood values corresponding to each branching event

### Author(s)

H Morlon

### References

Morlon, H., Potts, M.D., Plotkin, J.B. (2010) Inferring the dynamics of diversification: a coalescent approach, PLoS B 8(9): e1000493

**Examples**

```

data(Cetacea)
Vtimes <- sort(branching.times(Cetacea))
lamb0 <- 0.1
alpha <- 0.001
mu0<-0
beta<-0
ntips <- Ntip(Cetacea)
N0 <- 89
likelihood <- likelihood_coal_var(Vtimes, ntips, lamb0, alpha, mu0, beta, N0)

```

---

likelihood\_sgd

*Likelihood of a phylogeny under the SGD model*


---

**Description**

Computes the likelihood of a phylogeny under the SGD model with exponential increasing of the metacommunity, and potentially missing extant species. Notations follow Manceau et al. (in press).

**Usage**

```
likelihood_sgd(phylo, tot_time, b, d, nu, f)
```

**Arguments**

phylo	an object of type 'phylo' (see ape documentation)
tot_time	the age of the phylogeny (crown age, or stem age if known). If working with crown ages, tot_time is given by max(node.age(phylo)\$ages).
b	the (constant) birth rate of individuals in the model.
d	the (constant) death rate of individuals in the model.
nu	the (constant) mutation rate of individuals in the model.
f	the fraction of extant species included in the phylogeny

**Value**

the likelihood value of the phylogeny, given the model and the parameter values b, d, nu.

**Author(s)**

M Manceau

**References**

Manceau M., Lambert A., Morlon H. , submitted.

**Examples**

```
data(Cetacea)
tot_time <- max(node.age(Cetacea)$ages)
b <- 1e6
d <- 1e6-0.5
nu <- 0.6
f <- 87/89
#lh <- likelihood_sgd(Cetacea, tot_time, b, d, nu, f)
```

---

Phocoenidae

*Phocoenidae phylogeny*

---

**Description**

Ultrametric phylogenetic tree of the 6 extant Phocoenidae (porpoise) species

**Usage**

```
data(Phocoenidae)
```

**Details**

This phylogeny was extracted from Steeman et al. Syst Bio 2009 cetacean phylogeny

**References**

Steeman ME et al.(2009) Radiation of extant cetaceans driven by restructuring of the oceans *Syst Biol* 58:573-585

Morlon, H., Parsons, T.L., Plotkin, J.B. (2011) Reconciling molecular phylogenies with the fossil record *Proc Nat Acad Sci* 108: 16327-16332

**Examples**

```
data(Phocoenidae)
print(Phocoenidae)
plot(Phocoenidae)
```

---

plot_dtt	<i>Plot diversity through time</i>
----------	------------------------------------

---

**Description**

Plot the estimated number of species through time

**Usage**

```
plot_dtt(fit.bd, tot_time, N0)
```

**Arguments**

fit.bd	an object of class 'fit.bd', output of the 'fit_bd' function
tot_time	the age of the underlying phylogeny (crown age, or stem age if known). If working with crown ages, tot_time is given by <code>max(node.age(phylo)\$ages)</code> .
N0	number of extant species. If all extant species are represented in the phylogeny, N0 is given by <code>length(phylo\$tip.label)</code>

**Value**

Plot representing how the estimated number of species vary through time

**Author(s)**

H Morlon

**References**

Morlon, H., Parsons, T.L. and Plotkin, J.B. (2011) Reconciling molecular phylogenies with the fossil record *Proc Nat Acad Sci* 108: 16327-16332

Morlon, H. (2014) Phylogenetic approaches for studying diversification, *Eco Lett* 17:508-525

**See Also**

[fit\\_bd](#)

**Examples**

```
data(Balaenopteridae)
tot_time<-max(node.age(Balaenopteridae)$ages)

# Fit the pure birth model (no extinction) with exponential variation of the speciation rate
# with time
f.lamb <-function(t,y){y[1] * exp(y[2] * t)}
f.mu<-function(t,y){0}
lamb_par<-c(0.08, 0.01)
```

```

mu_par<-c()
result <- fit_bd(Balaenopteridae,tot_time,f.lamb,f.mu,lamb_par,mu_par,f=1,
                expo.lamb = TRUE, fix.mu=TRUE)

# plot estimated number of species through time
# plot_dtt(result, tot_time, N0=9)

```

---

plot_fit_bd	<i>Plot speciation, extinction &amp; net diversification rate functions of a fitted model</i>
-------------	---

---

### Description

Plot estimated speciation, extinction & net diversification rates through time

### Usage

```
plot_fit_bd(fit.bd, tot_time)
```

### Arguments

fit.bd	an object of class 'fit.bd', output of the 'fit_bd' function
tot_time	the age of the phylogeny (crown age, or stem age if known). If working with crown ages, tot_time is given by max(node.age(phylo)\$ages).

### Value

Plots representing how the estimated speciation, extinction & net diversification rate functions vary through time

### Author(s)

H Morlon

### See Also

[fit\\_bd](#)

### Examples

```

data(Balaenopteridae)
tot_time<-max(node.age(Balaenopteridae)$ages)

# Fit the pure birth model (no extinction) with exponential variation of the speciation rate
# with time
f.lamb <-function(t,y){y[1] * exp(y[2] * t)}
f.mu<-function(t,y){0}
lamb_par<-c(0.08, 0.01)

```



```

mu_par<-c()
result <- fit_bd(Balaenopteridae,tot_time,f.lamb,f.mu,lamb_par,mu_par,
                expo.lamb = TRUE, fix.mu=TRUE)
# plot fitted rates
#plot_fit_bd(result, tot_time)

```

---

plot_fit_env	<i>Plot speciation, extinction &amp; net diversification rate functions of a fitted environmental model</i>
--------------	---

---

### Description

Plot estimated speciation, extinction & net diversification rates as a function of the environmental data and time

### Usage

```
plot_fit_env(fit.env, env_data, tot_time)
```

### Arguments

fit.env	an object of class 'fit.env', output of the 'fit_env' function
env_data	environmental data, given as a data frame with two columns. The first column is time, the second column is the environmental data (temperature for instance).
tot_time	the age of the phylogeny (crown age, or stem age if known). If working with crown ages, tot_time is given by max(node.age(phylo)\$ages).

### Value

Plots representing how the estimated speciation, extinction & net diversification rate functions vary as a function of the environmental data & time

### Author(s)

H Morlon and FL Condamine

### See Also

[fit\\_env](#)

### Examples

```

data(Balaenopteridae)
tot_time<-max(node.age(Balaenopteridae)$ages)
data(InfTemp)
dof<-smooth.spline(InfTemp[,1], InfTemp[,2])$df

```

```

# Fit the pure birth model (no extinction) with exponential variation of the speciation rate
# with time
f.lamb <-function(t,x,y){y[1] * exp(y[2] * x)}
f.mu<-function(t,x,y){0}
lamb_par<-c(0.10, 0.01)
mu_par<-c()
result <- fit_env(Balaenopteridae,InfTemp,tot_time,f.lamb,f.mu,
  lamb_par,mu_par,f=1, fix.mu=TRUE, df=dof, dt=1e-3)

# plot fitted rates
#plot_fit_env(result, InfTemp, tot_time)

```

---

sim\_sgd

*Algorithm for simulating a phylogenetic tree under the SGD model*


---

### Description

Simulates a phylogeny arising from the SGD model with exponentially increasing metapopulation size. Notations follow Manceau et al. (submitted).

### Usage

```
sim_sgd(tau, b, d, nu)
```

### Arguments

tau	the simulation time, which corresponds to the length of the phylogeny
b	the (constant) per-individual birth rate
d	the (constant) per-individual death rate
nu	the (constant) per-individual mutation rate

### Value

a phylogenetic tree of class "phylo" (see ape documentation)

### Author(s)

M Manceau

### References

Manceau M., Lambert A., Morlon H. (submitted)

**Examples**

```
tau <- 10
b <- 1e6
d <- b-0.5
nu <- 0.6
tree <- sim_sgd(tau,b,d,nu)
plot(tree)
```

# Index

## \*Topic **datasets**

Balaenopteridae, [3](#)

Calomys, [3](#)

Cetacea, [4](#)

InfTemp, [16](#)

Phocoenidae, [22](#)

Balaenopteridae, [3](#)

Calomys, [3](#)

Cetacea, [4](#)

fit\_bd, [5](#), [14](#), [23](#), [24](#)

fit\_coal\_cst, [8](#), [12](#)

fit\_coal\_var, [9](#), [10](#)

fit\_env, [12](#), [25](#)

fit\_sgd, [15](#)

InfTemp, [16](#)

likelihood\_bd, [7](#), [14](#), [17](#)

likelihood\_coal\_cst, [9](#), [18](#)

likelihood\_coal\_var, [12](#), [20](#)

likelihood\_sgd, [15](#), [21](#)

Phocoenidae, [22](#)

plot\_dtt, [7](#), [23](#)

plot\_fit\_bd, [7](#), [24](#)

plot\_fit\_env, [14](#), [25](#)

RPANDA (RPANDA-package), [2](#)

RPANDA-package, [2](#)

sim\_sgd, [26](#)