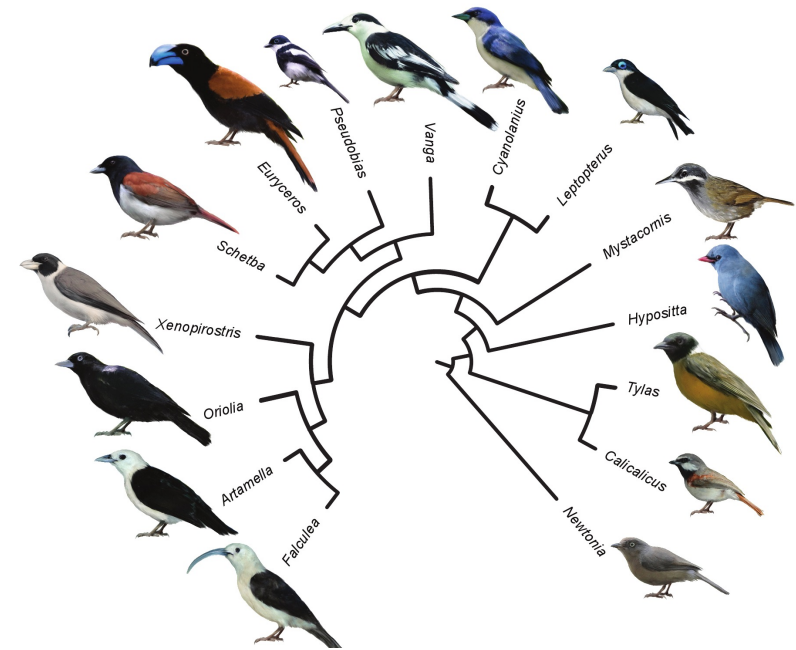
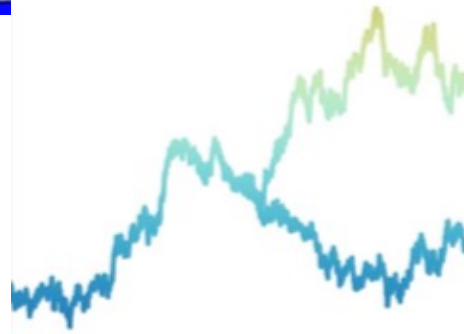
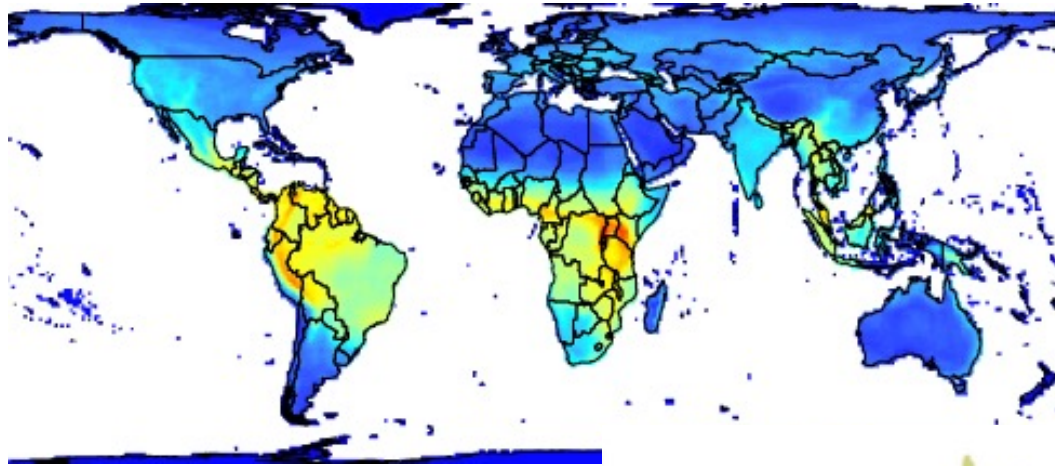


Application de modèles stochastiques à l'étude phylogénétique de l'évolution de la biodiversité

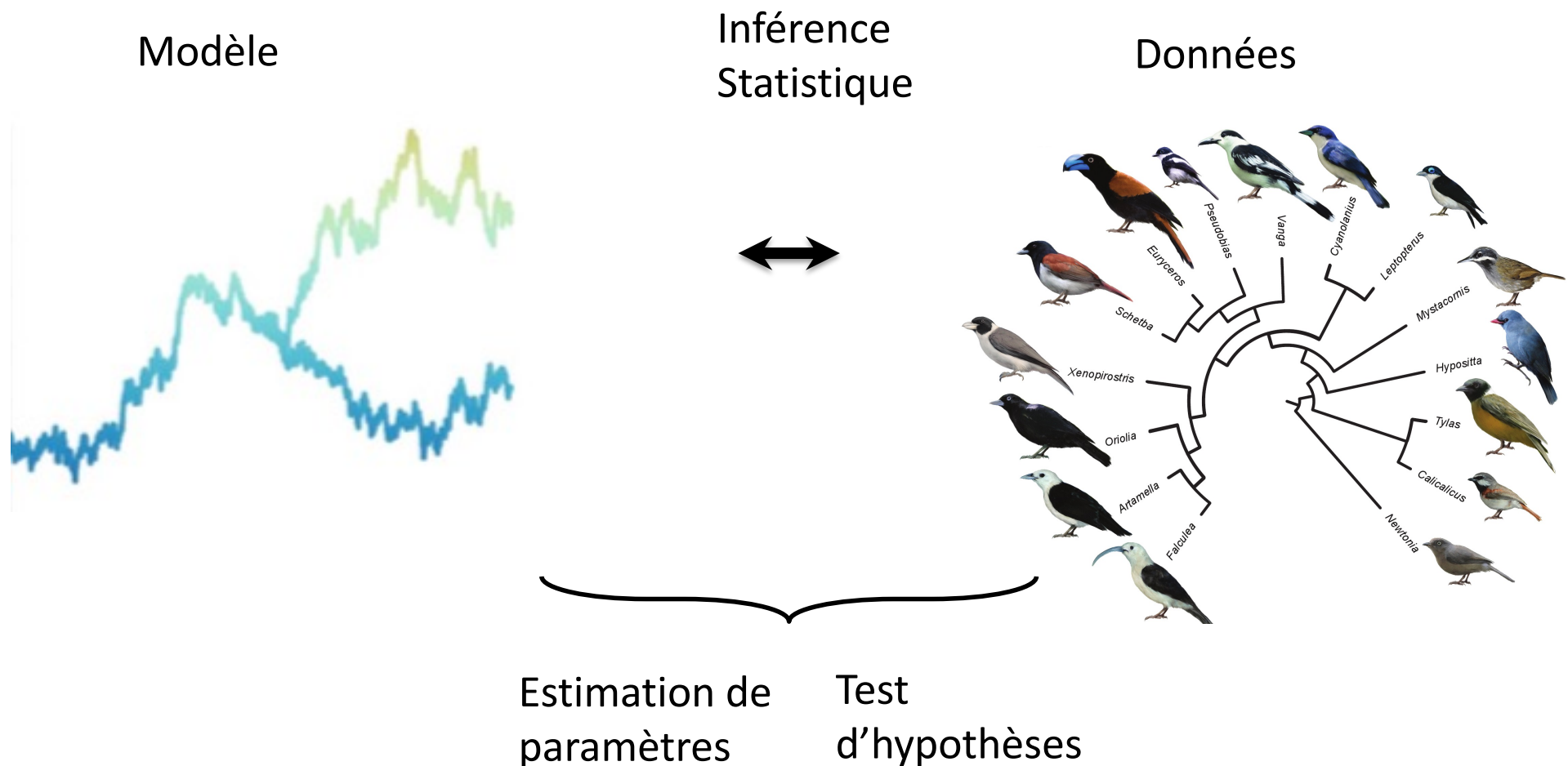


Hélène Morlon

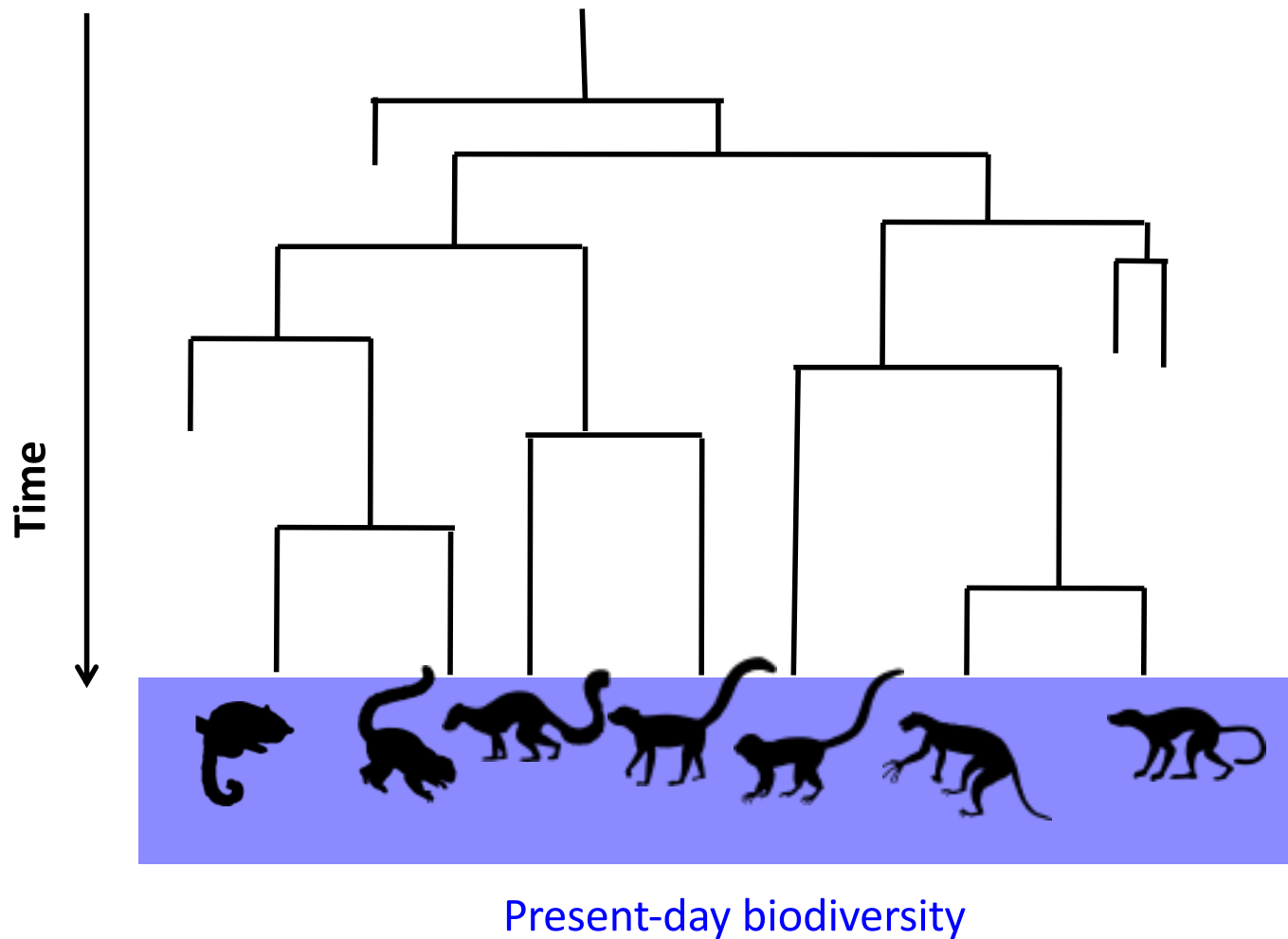
Ecole Normale Supérieure

BUT

Développer des **modèles** ET des approches d'**inférence statistique** qui permettent d'ajuster les modèles à des **données empiriques** de façon à **estimer des paramètres biologiques** et **tester différentes hypothèses biologiques**

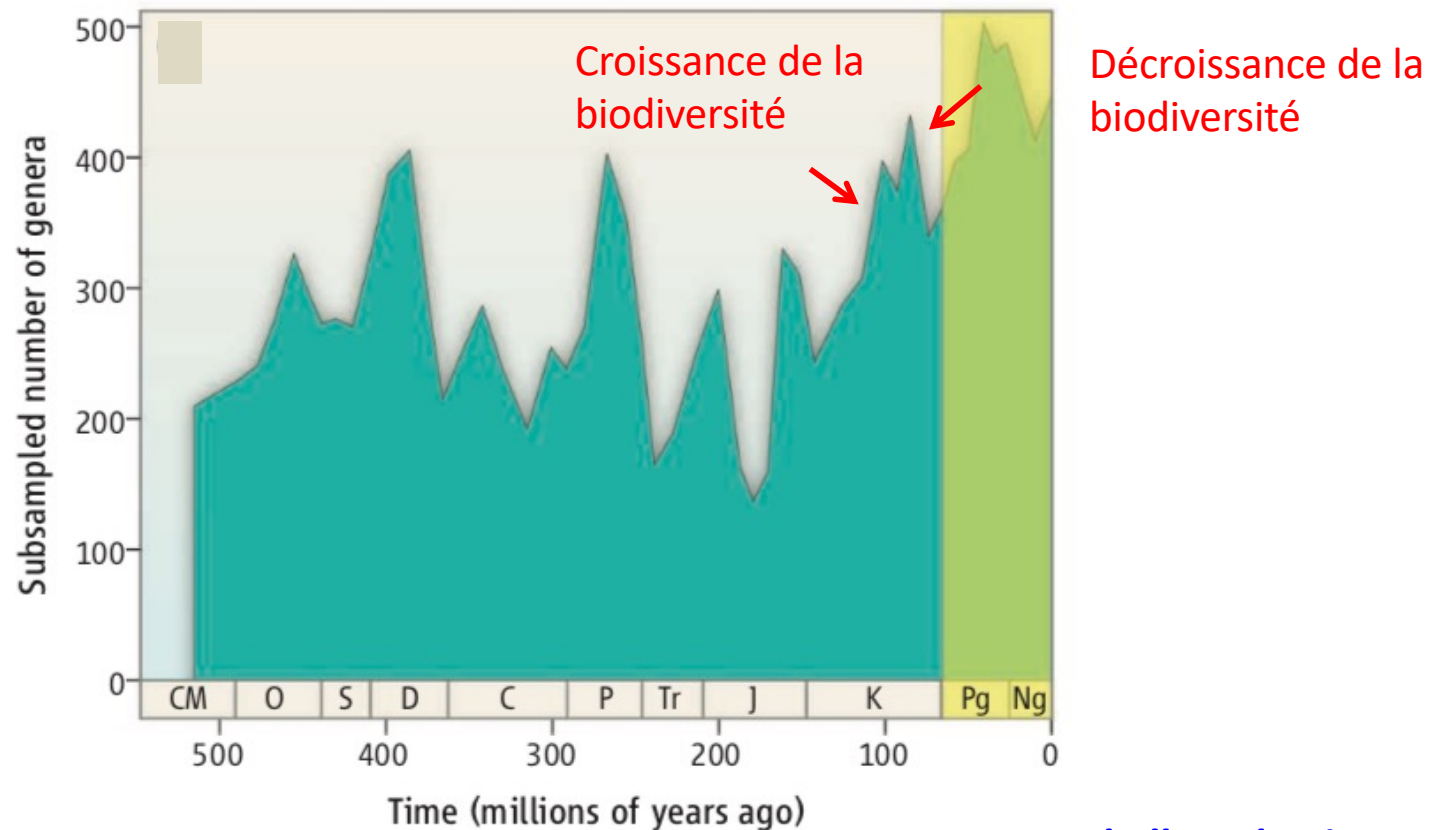


Application des modèles de naissance-mort à l'étude phylogénétique de la diversification



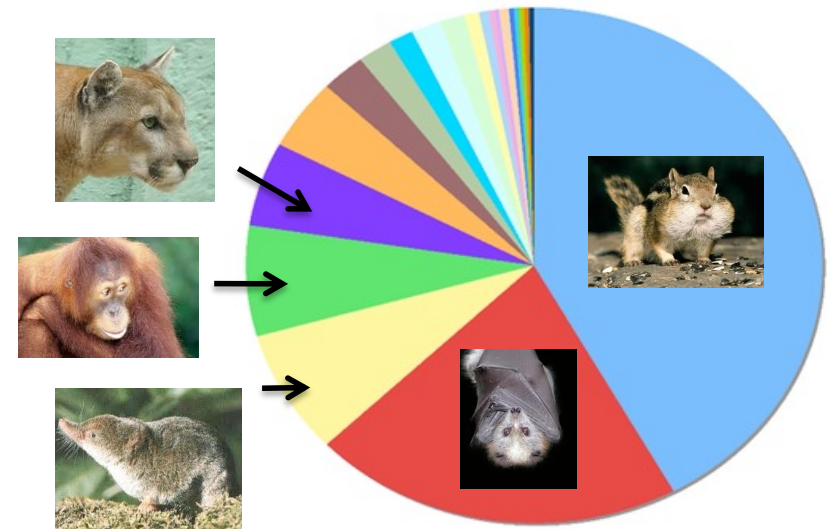
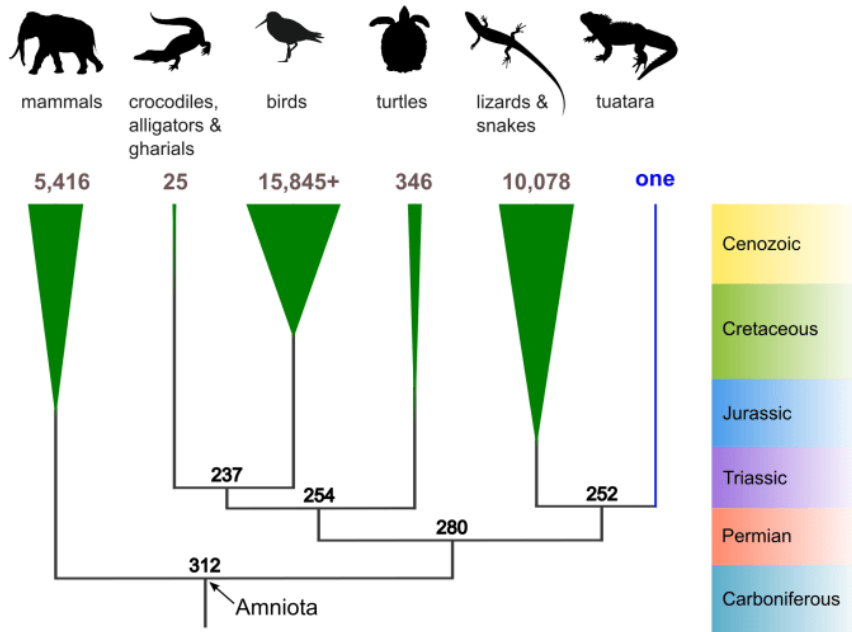
La diversification des espèces

Les espèces s'éteignent (extinction) et donnent naissance à de nouvelles espèces (spéciation). La balance entre événements de spéciation et d'extinction explique les variations en nombre d'espèces au cours du temps.



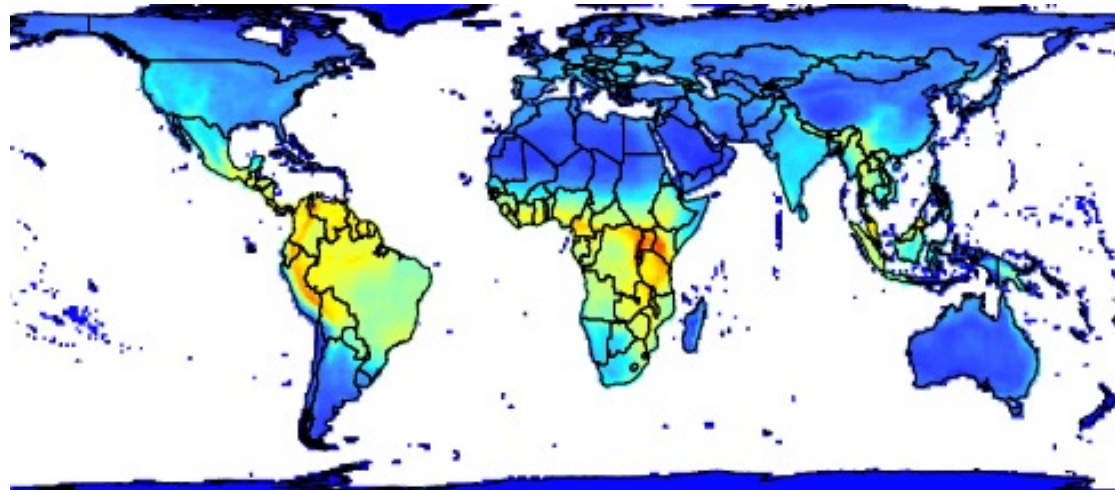
La diversification des espèces

La balance entre évènements de spéciation et d'extinction explique pourquoi certains groupes d'espèces sont beaucoup plus riches que d'autres



La diversification des espèces

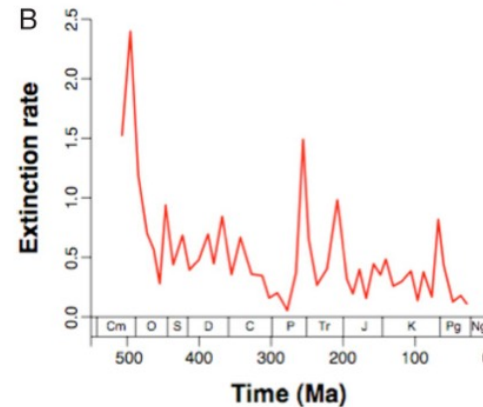
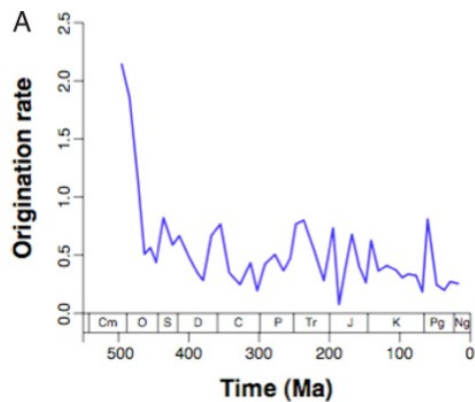
La balance entre évènements de spéciation, d'extinction et de dispersion explique pourquoi certaines régions de la planète sont beaucoup plus riches que d'autres



Comment estimer, à partir de données empiriques, les taux de spéciation et d'extinction des espèces?

Comment estimer, à partir de données empiriques, les taux de spéciation et d'extinction des espèces?

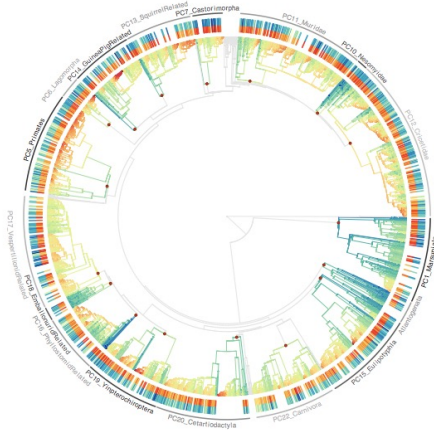
A partir des données fossiles



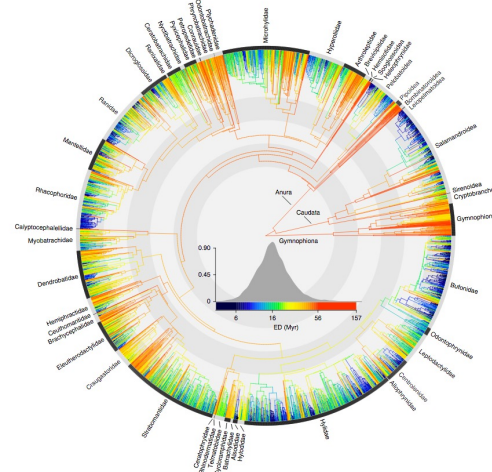
Alroy PNAS 2008

A partir des données phylogénétiques

Etude de la diversification des espèces à partir des arbres phylogénétiques des espèces actuelles



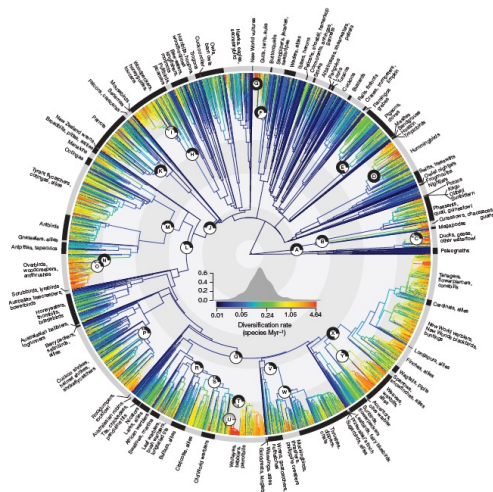
6 000 mammal species
Upham *et al.* PloSB 2019



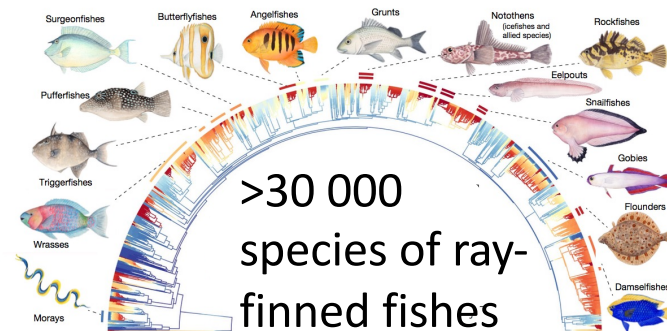
> 7 000 amphibian species
Jetz & Pyron *Nat Ecol Evol* 2018



2.2 Million species
 on a single tree

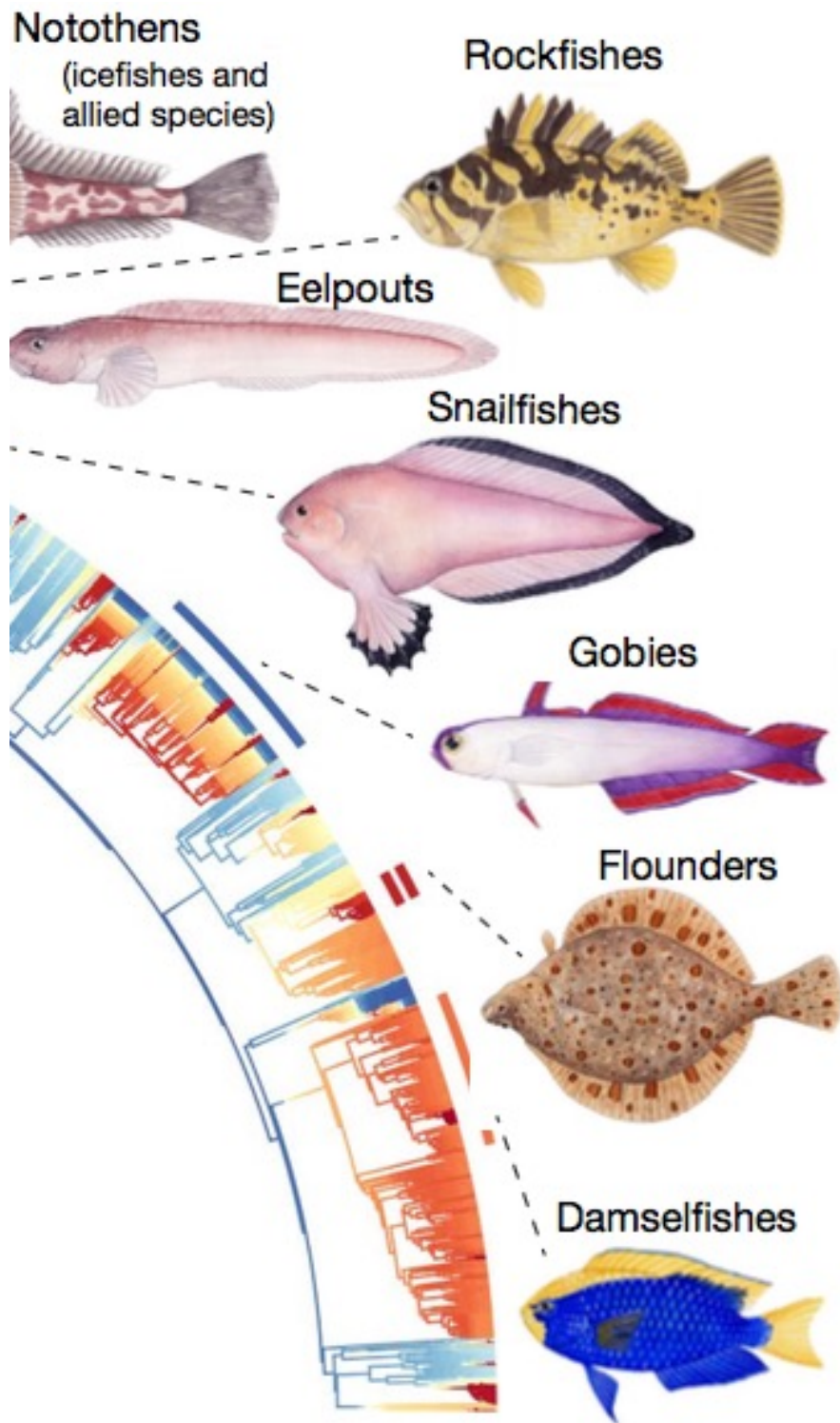


10 000 bird species
Jetz *et al.* Nature 2012



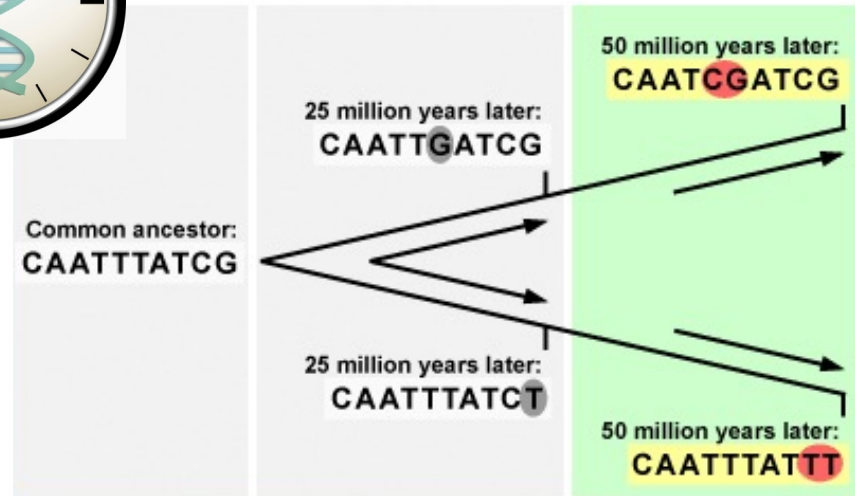
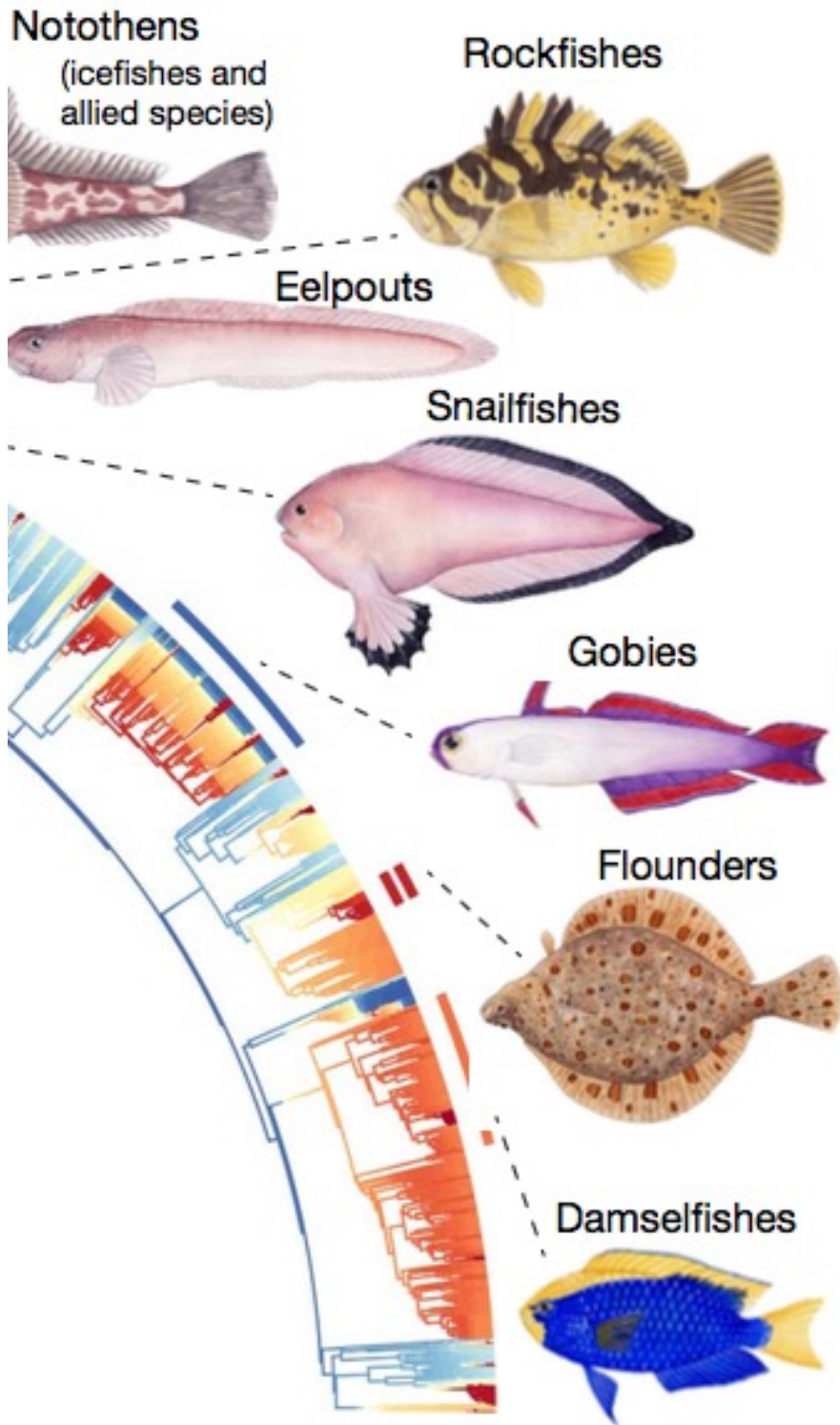
>30 000 species of ray-finned fishes
Rabosky *et al.* Nature 2018

Phylogenies represent the order and timing of divergences between ancestral species that led to present-day species



Phylogenies are constructed from the genetic (and sometimes phenotypic) similarity between present-day species

The datation is performed with fossil calibrations and the molecular clock



Notothens

(icefishes and allied species)



Rockfishes



Eelpouts



Snailfishes



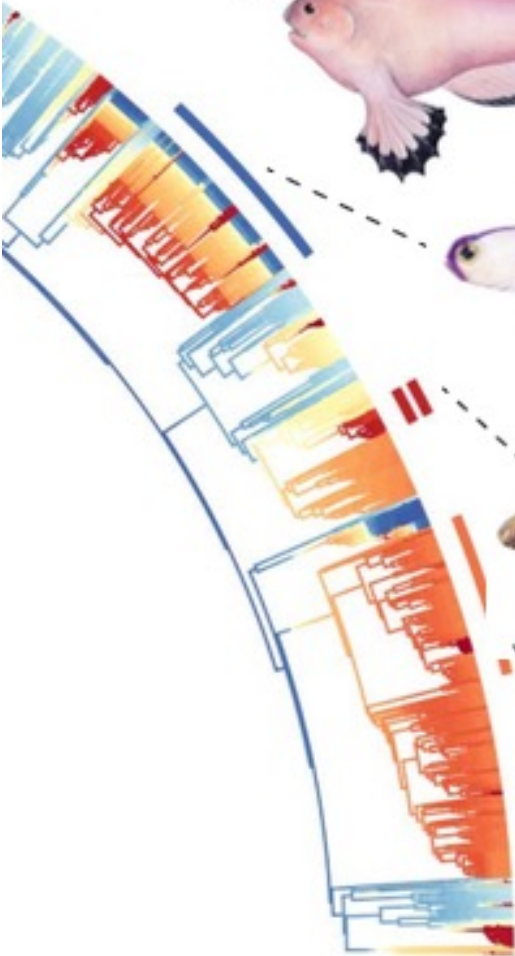
Gobies



Flounders



Damselfishes

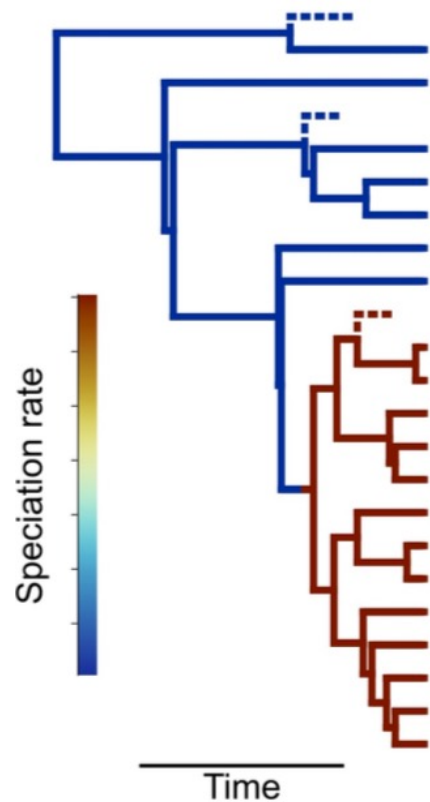


Dans ce qui suit, nous considérons l'arbre phylogénétique (daté) comme une **donnée**

BUT

Développer des **modèles de diversification** ET des approches d'**inférence statistique** qui permettent d'ajuster les modèles à des **arbres phylogénétiques** de façon à **estimer les taux de spéciation et extinction**

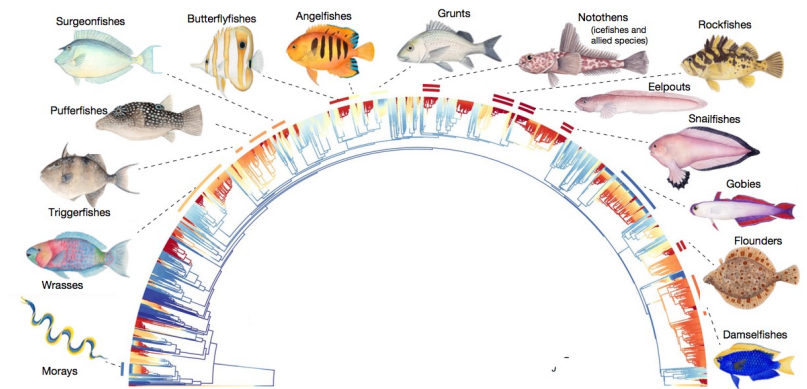
Modèle de diversification



Inférence
Statistique



Arbre phylogénétique

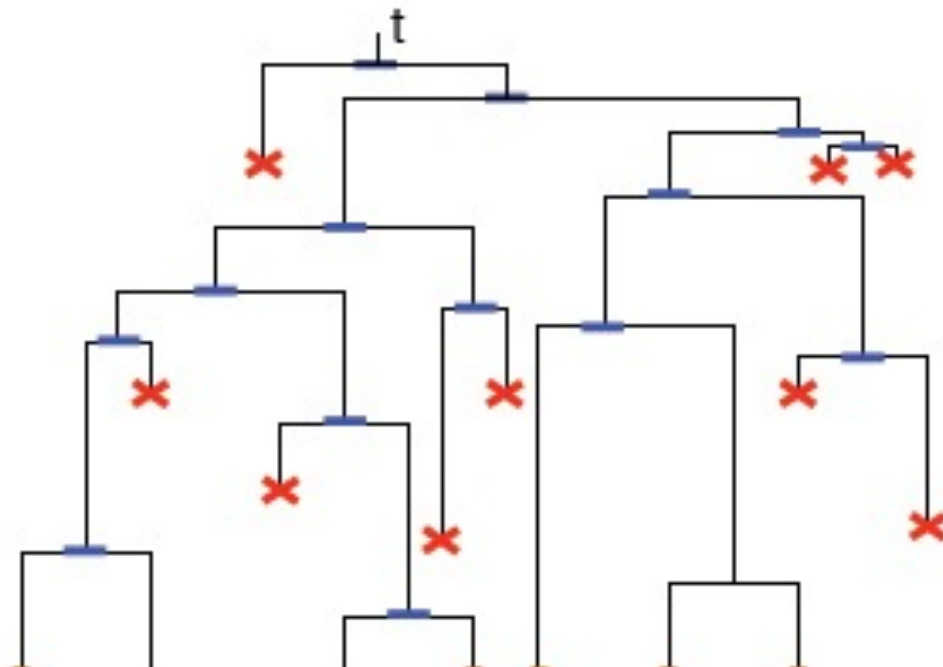


Estimation des taux de
spéciation et extinction

Birth-death model for the analysis of diversification

We model the evolution of a given clade (group of species that contains a common ancestor and all its descendants) with stochastic birth-death models:

- the clade starts with a single ancestor at time T
- species give rise to new species with the **per lineage per unit time rate of speciation λ** , and they go extinct with the **per lineage per unit time rate of extinction μ**



Ici une lignée est une espèce



Estimating diversification from the age and size of a clade

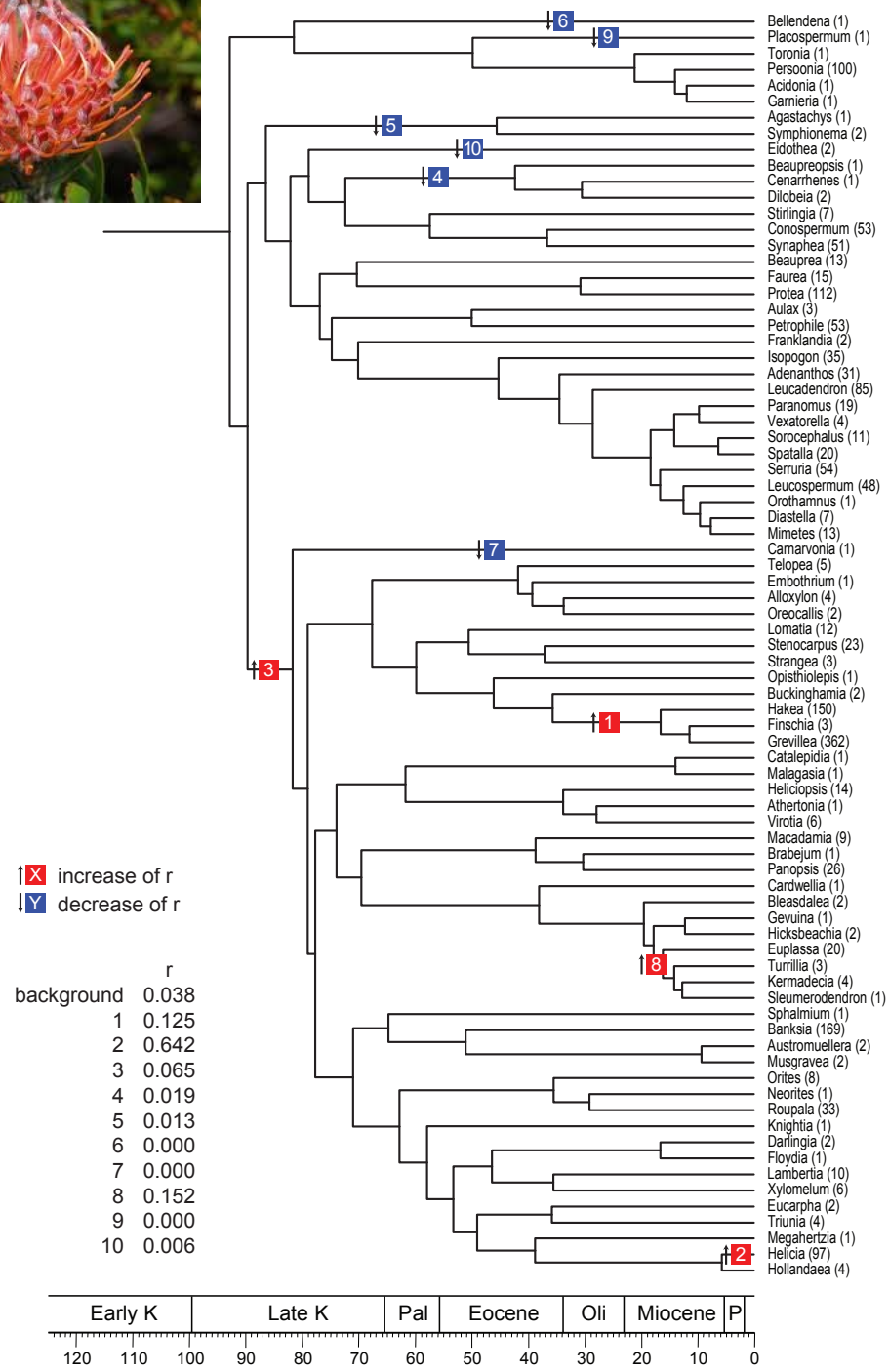
In the absence of extinction

$$\hat{\lambda} = \frac{\log(n)}{T}$$

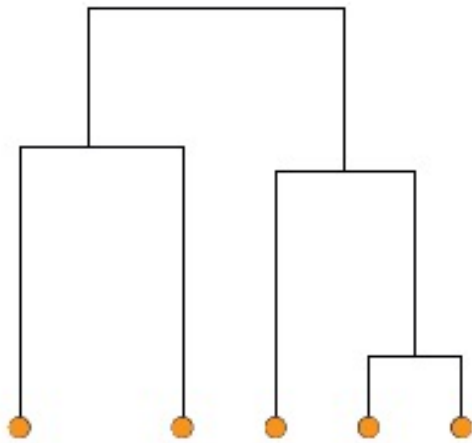
With extinction

$$\hat{r} = \frac{1}{T} \log(n(1 - \varepsilon) + \varepsilon) \quad \varepsilon = \frac{\mu}{\lambda}$$

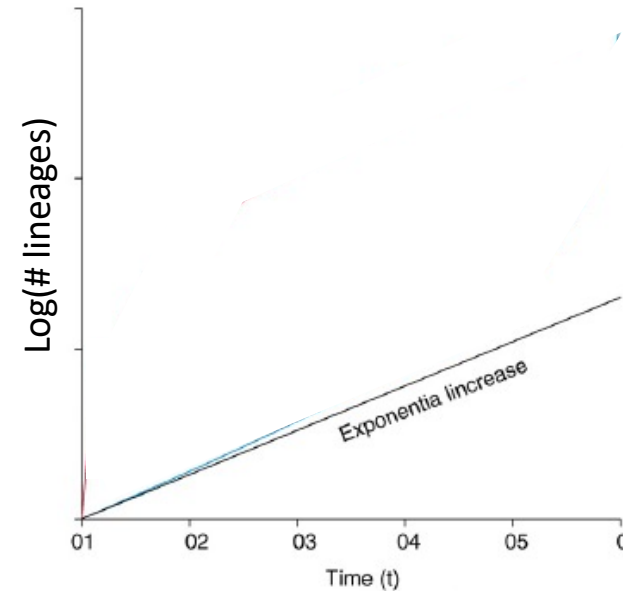
\nearrow
 $r = \lambda - \mu$



Estimating diversification using Lineage Through Time plots (LTT)

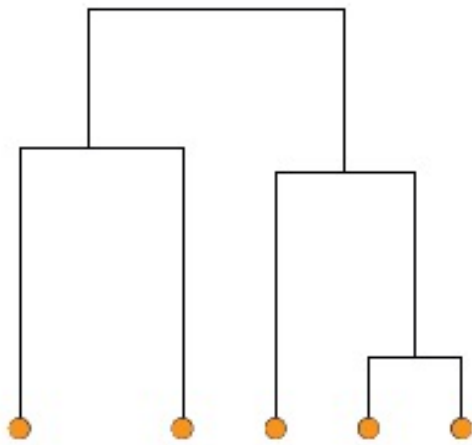


arbre phylogénétique des espèces
actuelles

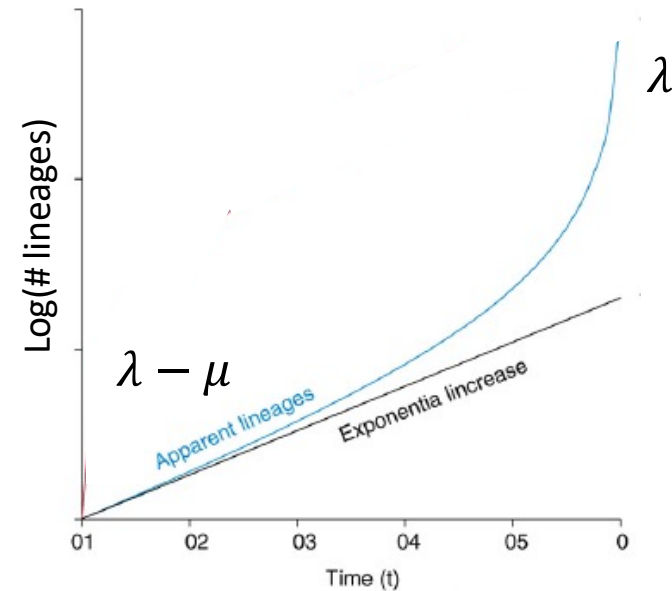


In the absence of extinction, an estimate of the speciation rate is given by the slope of the Lineage Through Time plot

Estimating diversification using Lineage Through Time plots (LTT)



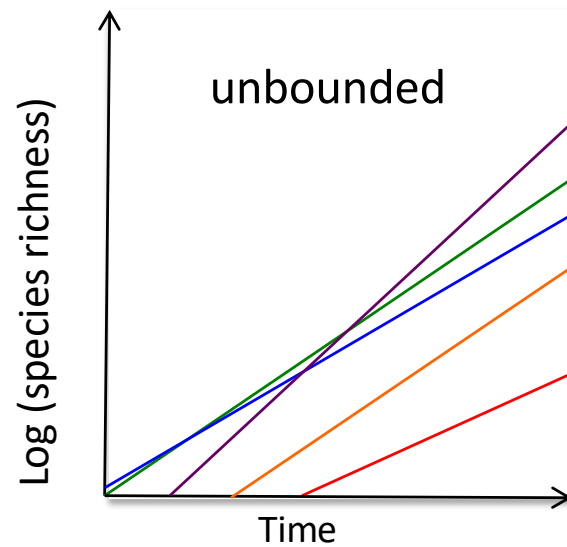
arbre phylogénétique des espèces actuelles



With extinction, the Lineage Through Time plot is characterized by an acceleration towards the present known as « the pull of the present »

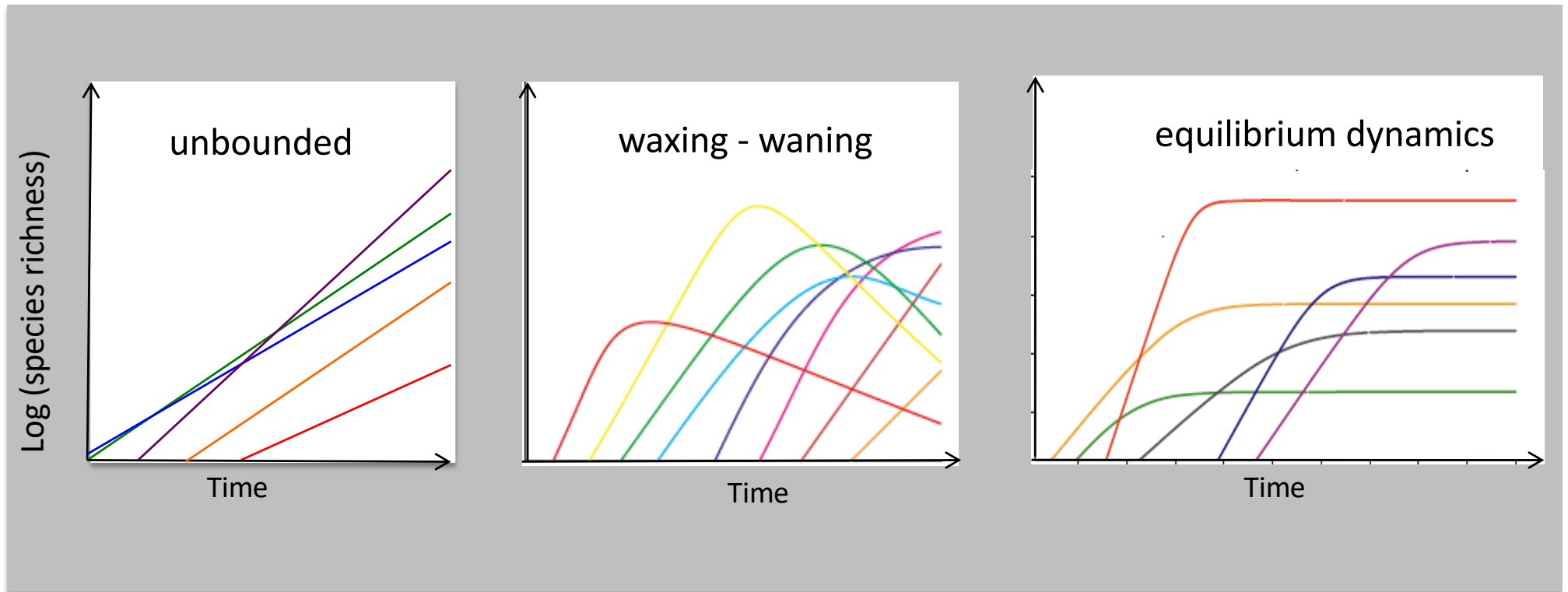
The slope towards the present provides an estimate of the speciation rate, the slope in the past provides an estimate of the net diversification rate.

All the above methods assume speciation and extinction rates are constant through time and homogeneous across lineages

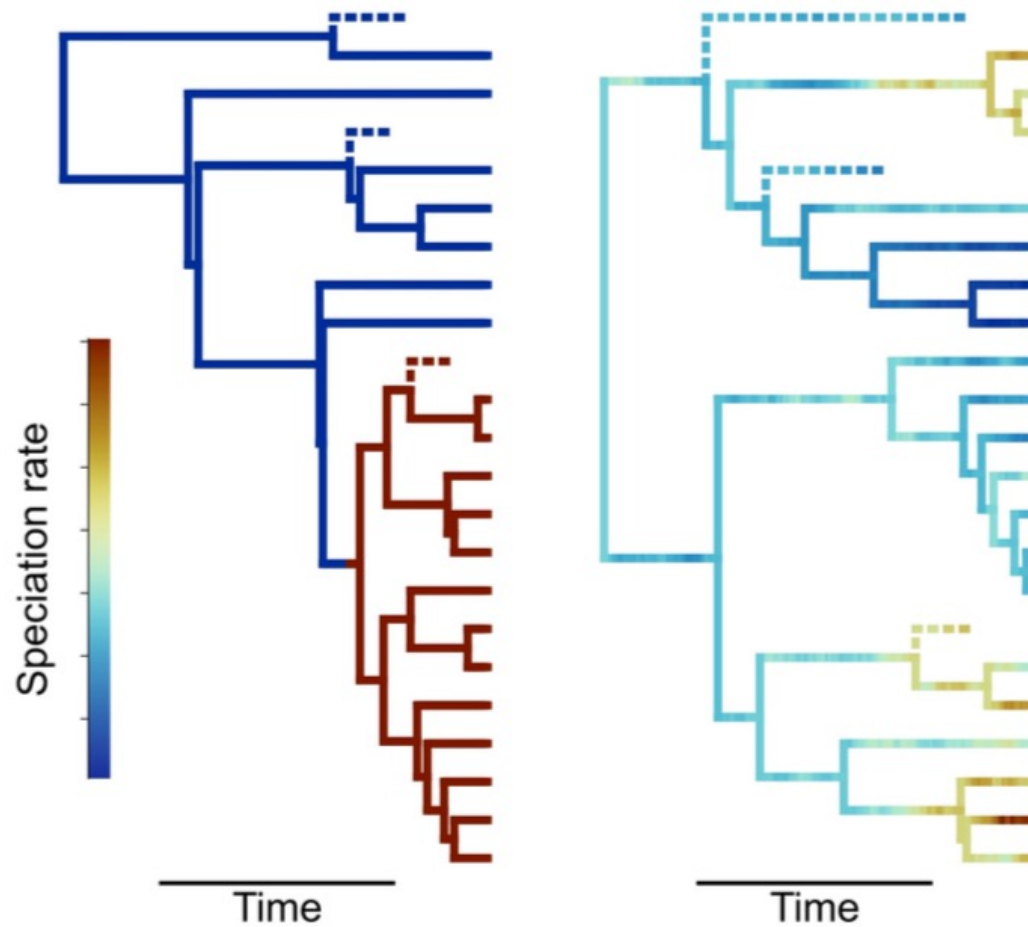


What if biodiversity dynamics are not unbounded?

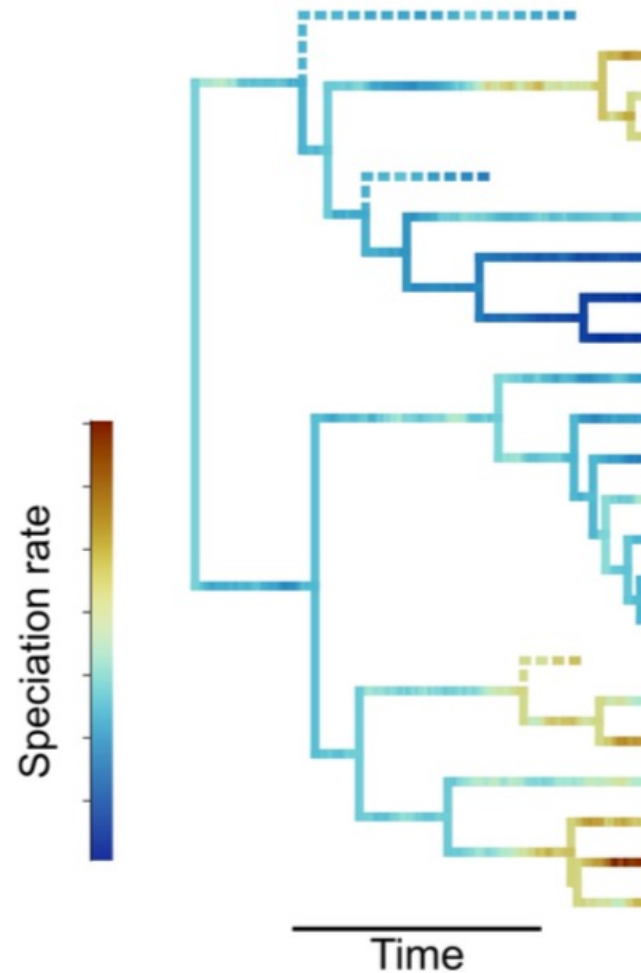
What if we want to test the support of different visions of how the evolution of diversity proceeds?



What if we rates are not homogeneous across lineages and we want to test the support of different visions of how rate variation occurs?



Intuitivement, la distribution des temps de branchement informe sur la variation des taux dans le temps, la topologie informe sur la variation des taux entre lignées

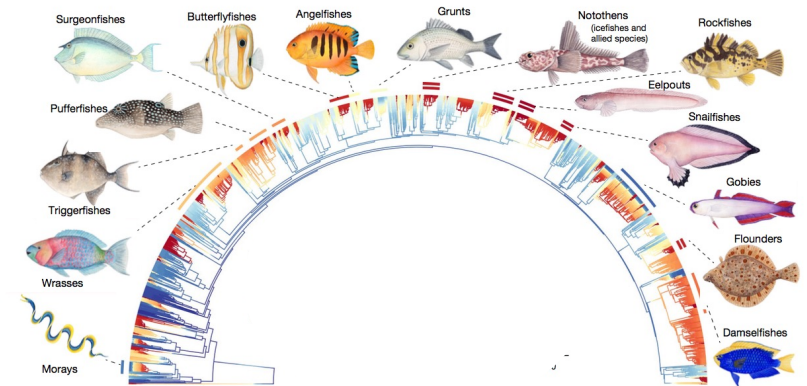
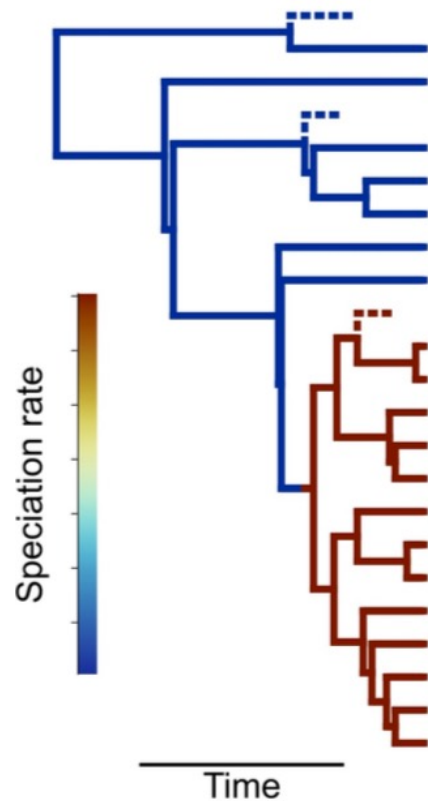


Inférence Statistique

Modèle de diversification



Arbre phylogénétique



Estimation des taux de
spéciation et extinction,
Test d'hypothèses

Inférence par vraisemblance

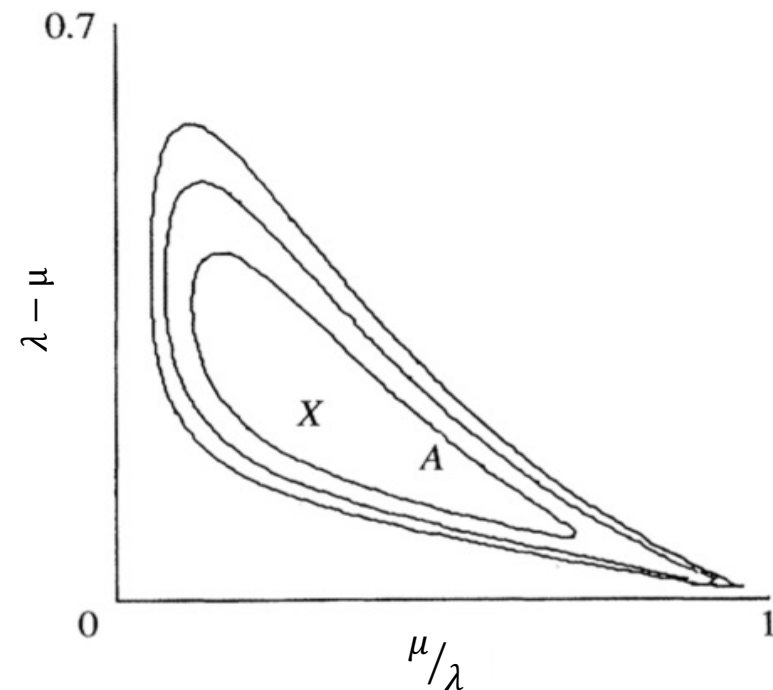
The likelihood is defined as

$$\mathcal{L}_x(\theta) = f_\theta(x)$$

where $f_\theta(x)$ is the probability of observing x under the model for parameters θ

The ML estimate is the parameter θ that maximizes $\mathcal{L}_x(\theta)$

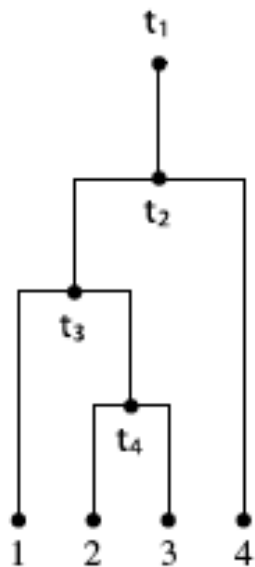
Contour de vraisemblance d'un modèle de naissance-mort sur une phylogénie simulée



Likelihood of a phylogeny of extant species under the birth-death process

stem age t_1
 speciation rate λ
 extinction rate μ

+ λ and μ can vary over time
 sampling probability f



sampling probability

probability that a lineage survives from t_1 to t_2 and leaves no descendant in the sample

$$\mathcal{L}(t_1, \dots, t_n) = \frac{f^n \Psi(t_2, t_1) \prod_{i=2}^n \lambda(t_i) \Psi(s_{i,1}, t_i) \Psi(s_{i,2}, t_i)}{1 - \Phi(t_1)},$$

speciation rate at time t_i

probability that a lineage alive at time t_1 has no descendant in the sample

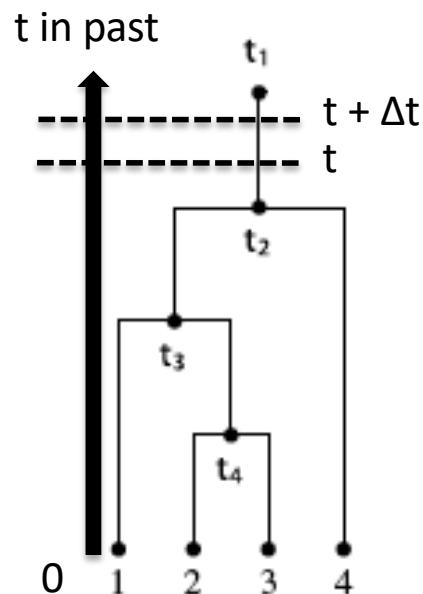
Likelihood of a phylogeny of extant species under the birth-death process

birth-death model

stem age t_1
 speciation rate λ
 extinction rate μ

+

λ and μ can vary over time
 sampling fraction f



$\Phi(t) = \text{P}\{\text{a lineage is not in the sample} \mid \text{it was alive at the time } t\}$

$$\begin{aligned} \Phi(t + \Delta t) &= \mu(t)\Delta t \\ &+ (1 - \mu(t)\Delta t)\lambda(t)\Phi^2(t) \\ &+ (1 - \mu(t)\Delta t)(1 - \lambda(t)\Delta t)\Phi(t) \\ &+ o(\Delta t). \end{aligned}$$

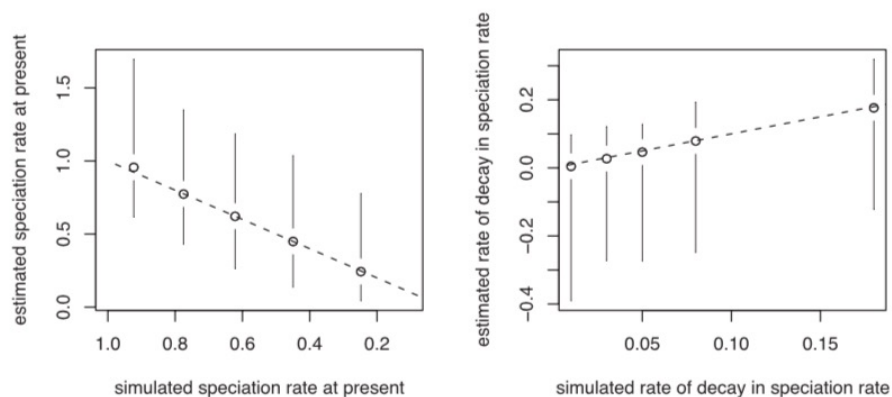
$\Delta t \rightarrow 0$

$$\frac{d\Phi}{dt} = \mu(t) - (\lambda(t) + \mu(t))\Phi(t) + \lambda(t)\Phi^2(t)$$

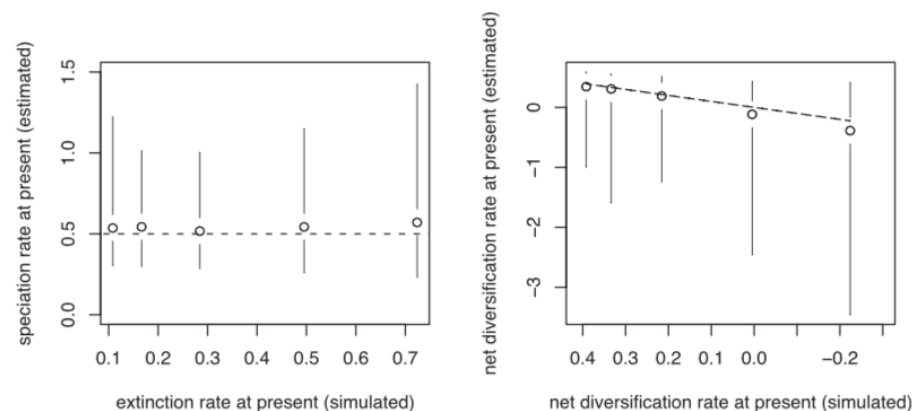
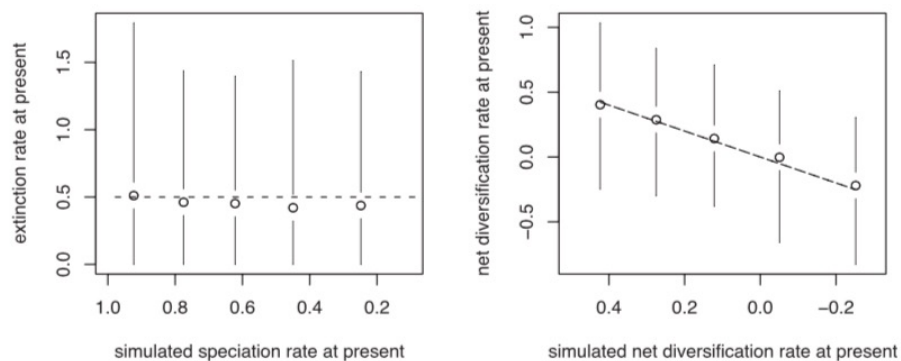
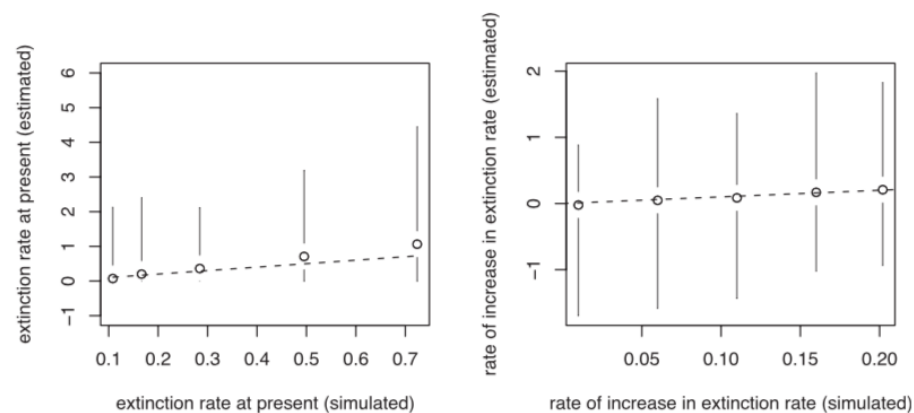
$$\Phi(t) = 1 - \frac{e^{\int_0^t \lambda(u) - \mu(u) du}}{\frac{1}{f} + \int_0^t e^{\int_0^s \lambda(u) - \mu(u) du} \lambda(s) ds}$$

Unbiased parameter estimates

Exponential decay of speciation rate,
constant extinction



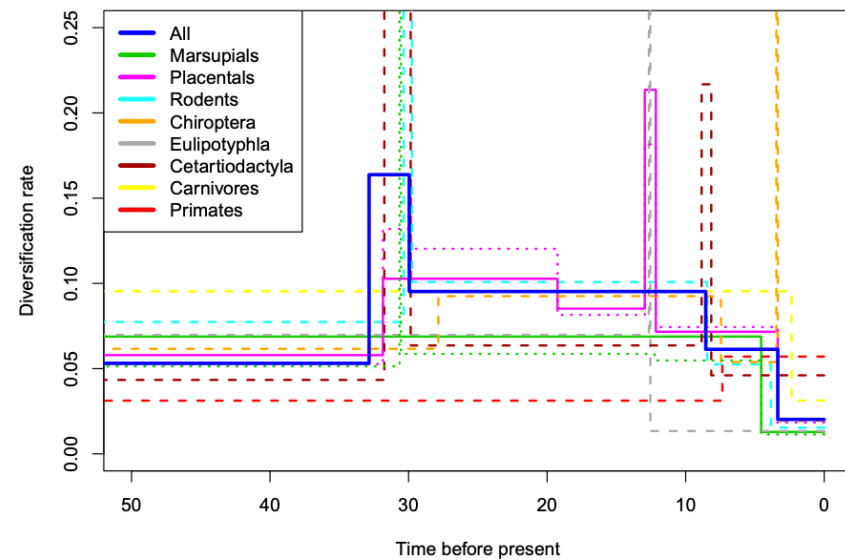
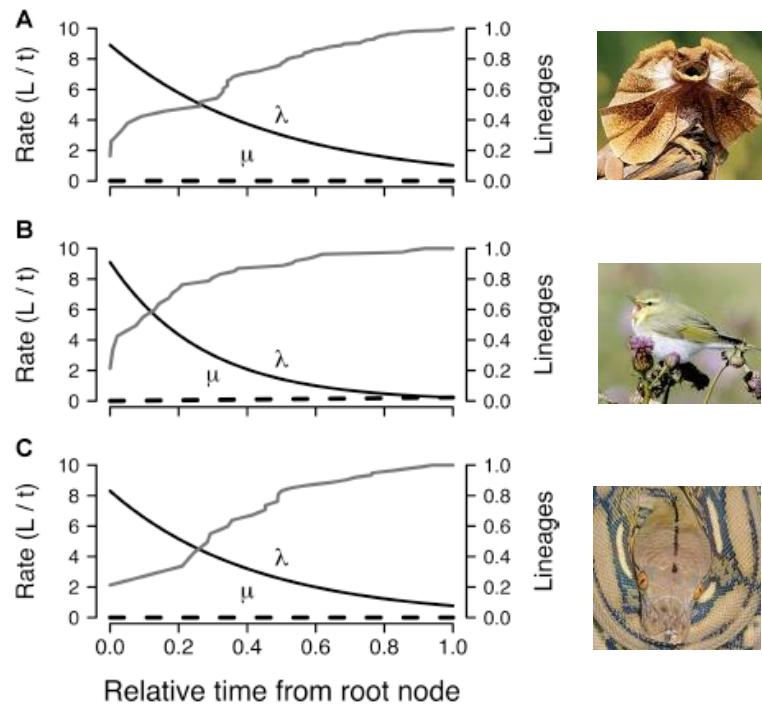
Exponential increase of extinction rate,
constant speciation



**Do rates of species diversification vary through time?
How?**

Speciation rates often decline through time

Estimates obtained by maximum-likelihood fit of birth-death models to phylogenies of present-day species



Stadler *et al.*
PNAS 2011

Rabosky & Lovette Evolution 2008

Why do rates of species diversification vary through time?



George Gaylord Simpson

Most diversity of life on Earth arises from adaptive radiations, the fast diversification of ecological traits in a rapidly speciating group of organisms

Under this hypothesis, diversification is fast in the presence of ecological opportunities, but slows down as these opportunities are exploited by an increasing number of species, generating so-called « Early burst » patterns.



George Gaylord Simpson

Diversity slowdowns linked to limited ecological opportunities and the diversity-dependent diversification model

PROCEEDINGS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Diversity-dependence brings molecular phylogenies closer to agreement with the fossil record

Rampal S. Etienne, Bart Haegeman, Tanja Stadler, Tracy Aze, Paul N. Pearson, Andy Purvis and Albert B. Phillimore

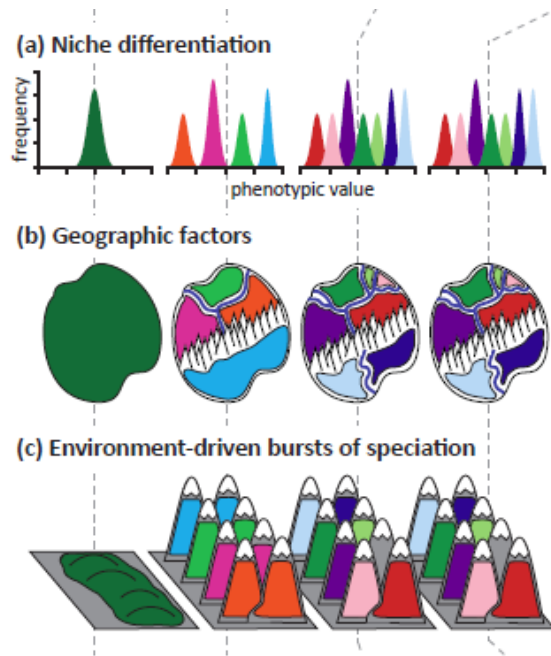
$$\lambda_n = \max\left(0, \lambda_0 - (\lambda_0 - \mu) \frac{n}{K}\right) \quad \text{and} \quad \mu_n = \mu$$



Maximum number of species

Why does diversification slow down?

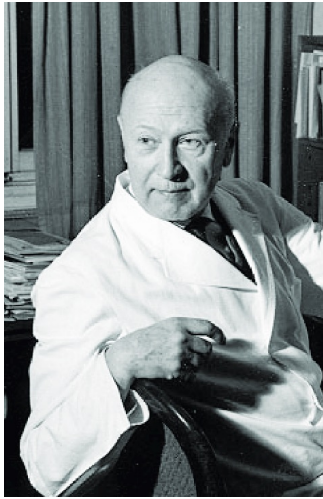
Daniel Moen and H el ene Morlon



Trends Ecol Evol 2014

Why are diversification slowdowns observed even when species do not coexist (and therefore do not compete for ecological opportunities)?

Do interactions with closely-related species necessarily impede diversification?



The evolutionary speed hypothesis and other climate-related hypotheses

Bernhard Rensch

Warmer climates select for small-bodied species, which have fast generation times, high mutation rates, and therefore potentially higher speciation rates

Metabolic activity is faster under warm climates, potentially increasing the action of free radicals and increasing mutation rates

Warmer climates are often associated with increased climatic stability and productivity, which could promote speciation

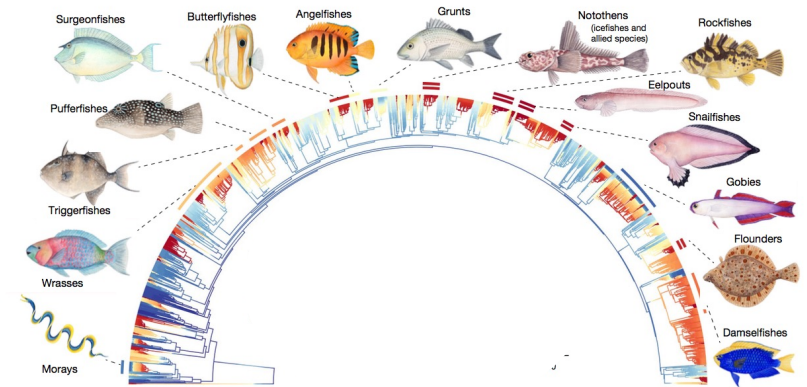
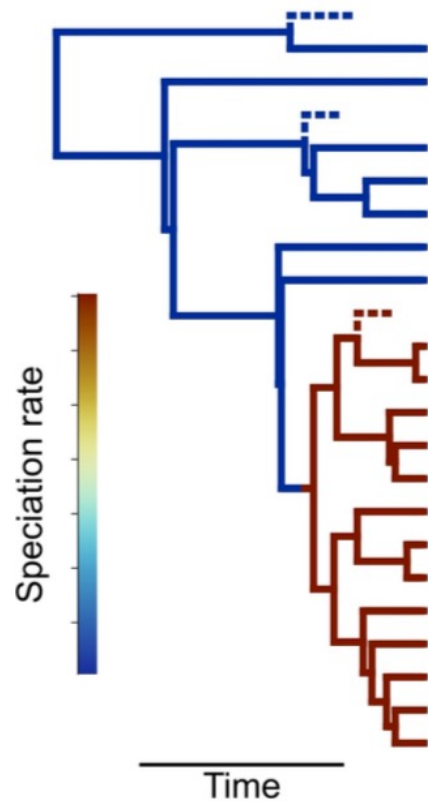
Could diversification slowdowns be related to the cooling of the Earth during the Cenozoic?

Inférence Statistique

Modèle de diversification

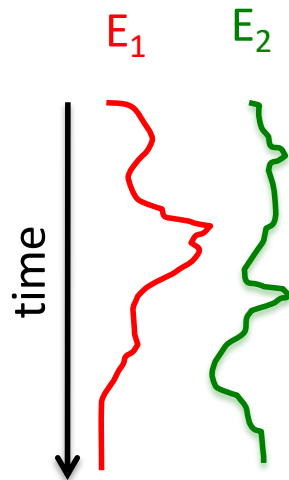


Arbre phylogénétique



**Test d'hypothèses par
comparaison de
vraisemblances**

Models of diversification with rates that depend on measured (a)biotic variables



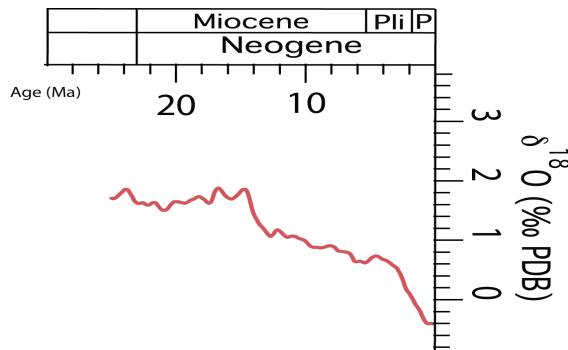
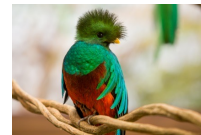
$$\tilde{\lambda}(t) = \lambda(t, E_1(t), E_2(t), \dots, E_k(t))$$

$$\tilde{\mu}(t) = \mu(t, E_1(t), E_2(t), \dots, E_k(t))$$

$$\mathcal{L}(t_1, \dots, t_n) = \frac{f^n \Psi(t_2, t_1) \prod_{i=2}^n \lambda(t_i) \Psi(s_{i,1}, t_i) \Psi(s_{i,2}, t_i)}{1 - \Phi(t_1)},$$

Did past climatic changes affect diversification rates? How?

218 phylogenies of tetrapod families

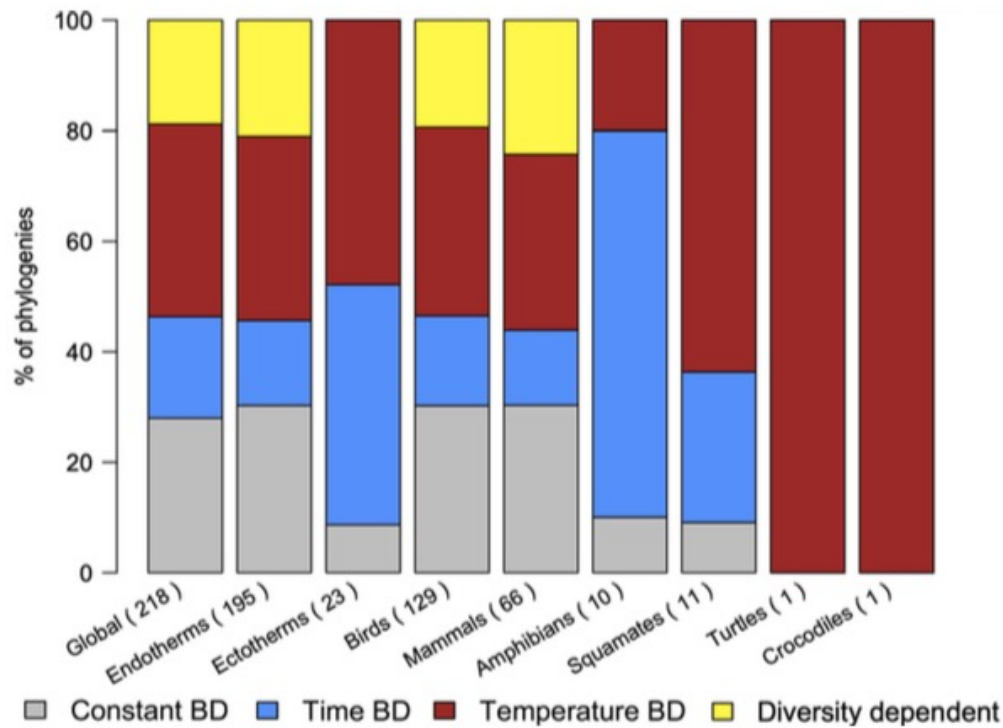


Zachos *et al.* 2008

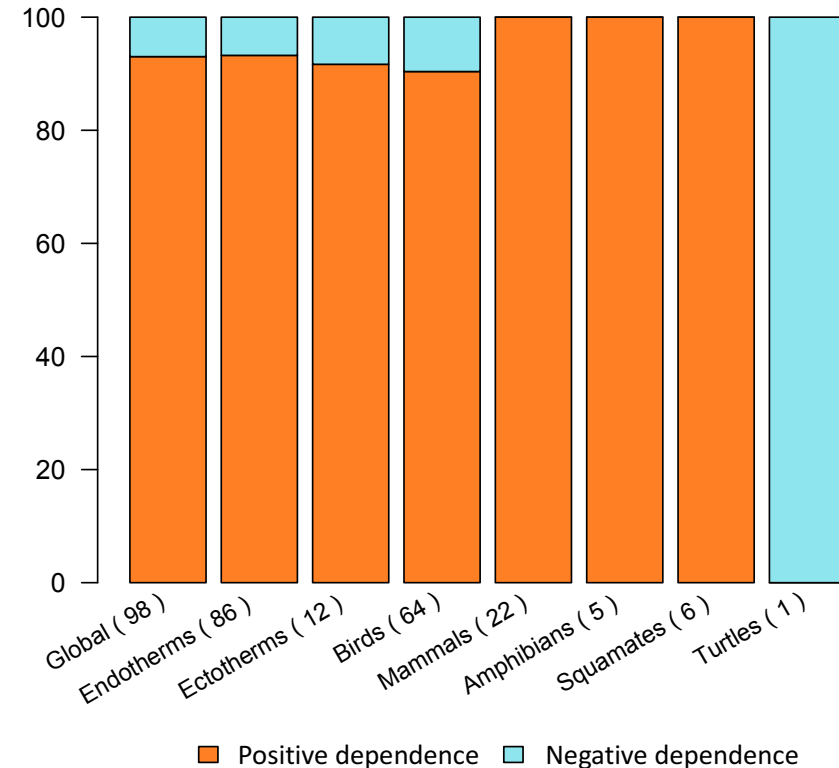
Comparison of 22 models including constant rate diversification models, models with time-varying rates, diversity-dependent rates, and temperature-dependent models

Condamine *et al.* Eco Lett 2019

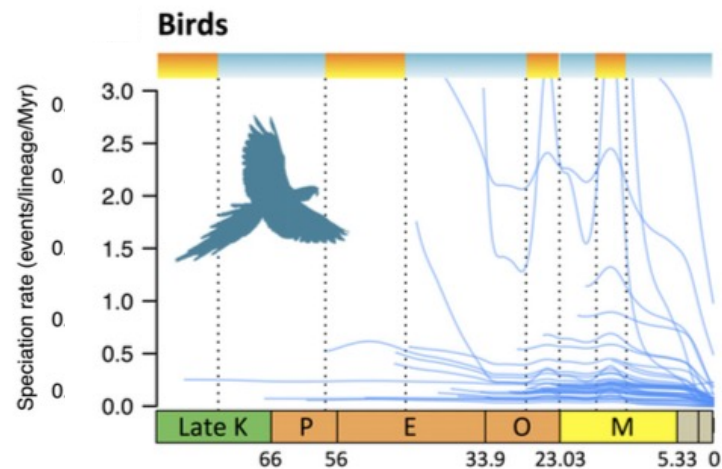
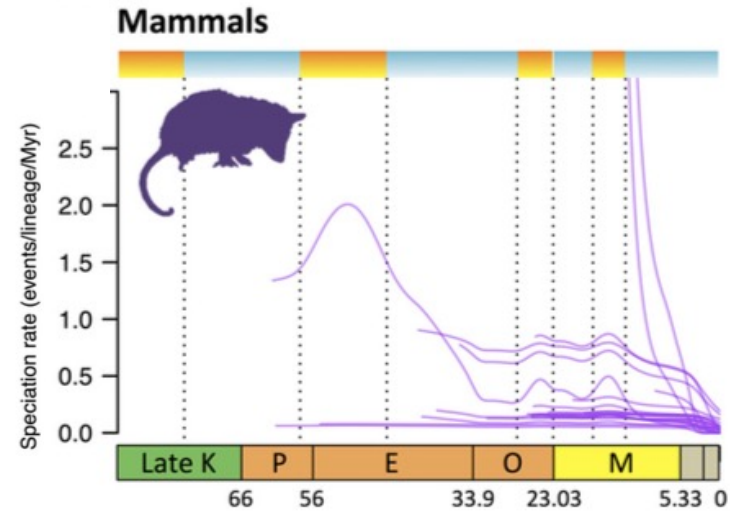
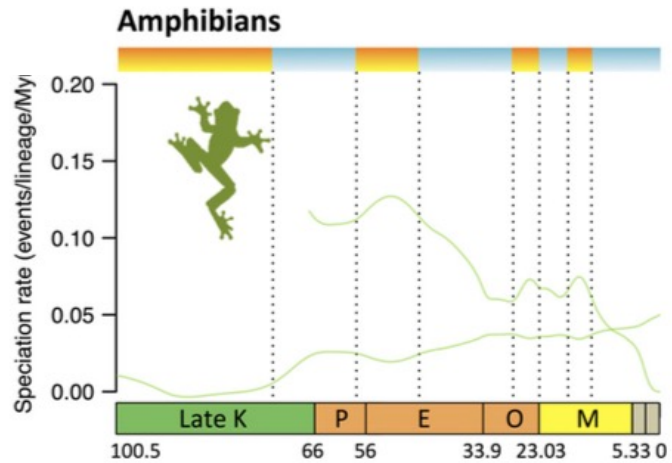
Speciation rates often vary with temperature



Speciation rates are higher during warm geological periods



Climate cooling during the Cenozoic results in a slowdown in diversification



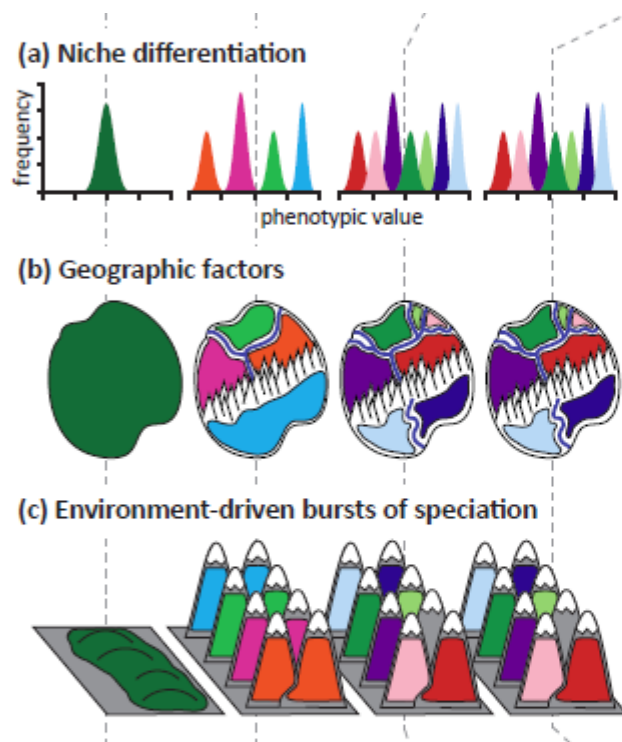
Climate cooling provides a potential alternative explanation for speciation rates slowdowns to the often invoked filling of niche space

Opinion

CellPress

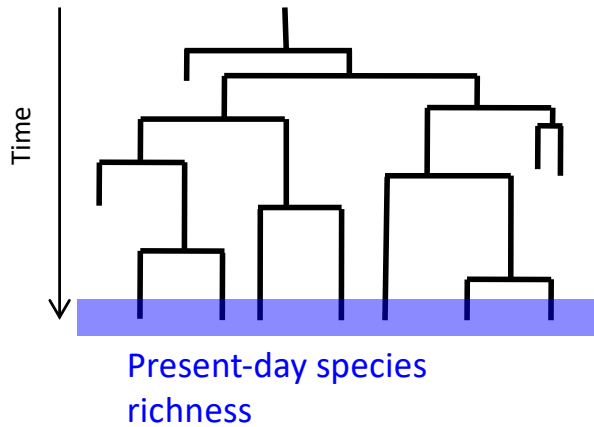
Why does diversification slow down?

Daniel Moen and H el ene Morlon



Trends Ecol Evol 2014

From estimates of speciation and extinction rates to estimates of past diversity



$$N(t) = N_0 e^{\int_0^t (-\lambda(s) + \bar{\mu}(s)) ds}$$

speciation rate extinction rate
↑
number of species today

Morlon *et al.* PNAS 2011

probability there were m species at time t
 given there were x species at time s
 and there are n species today

↓ ↓ ↓

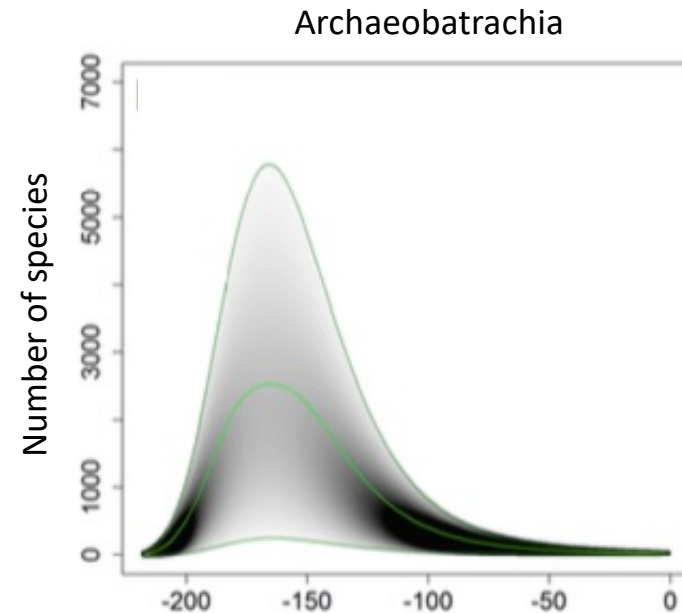
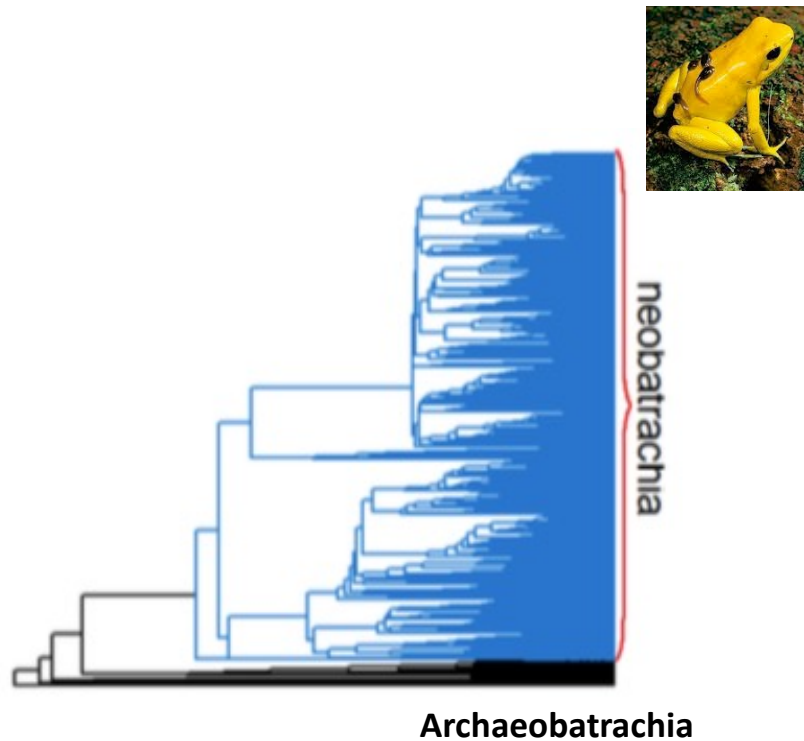
$$\mathbb{P}(N(t) = m \mid N(s) = x, N(T_{pres}) = n)$$



Olivier Billaud

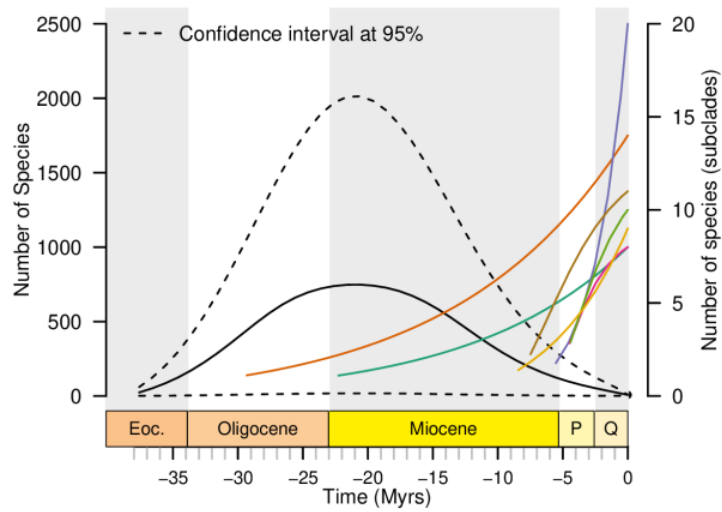
Billaud *et al.* Syst Bio 2019

Are old and species-poor groups groups that have always be poor, or are they the remnants of a diverse past?

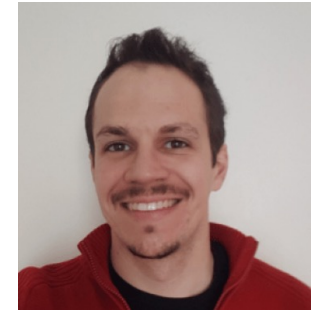


Old and poor frog families are the remnant of a diverse past

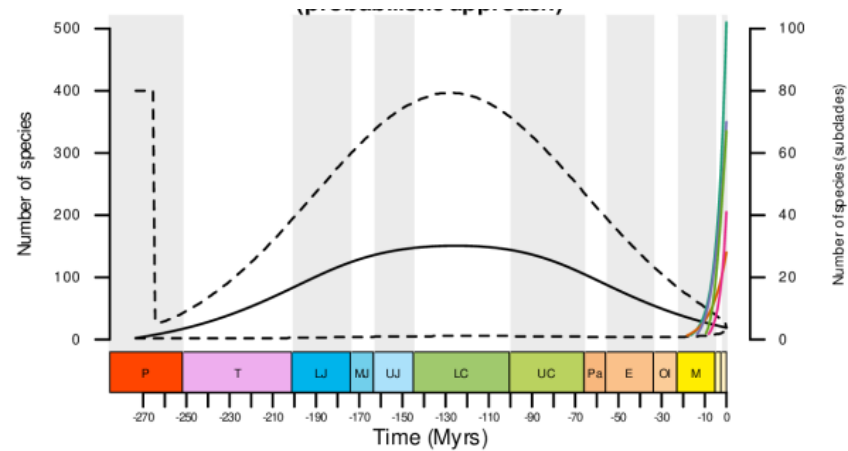
Other examples of old and poor groups that are the remnant of a diverse past



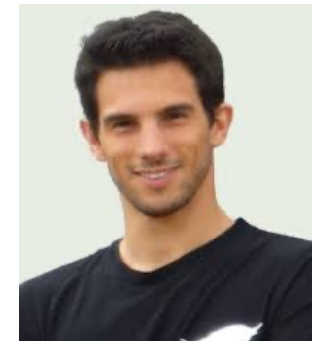
Nathan
Mazet



Pierre-Henri
Fabre



Fabien
Condamine



**Do rates of species diversification vary across lineages?
How & Why?**

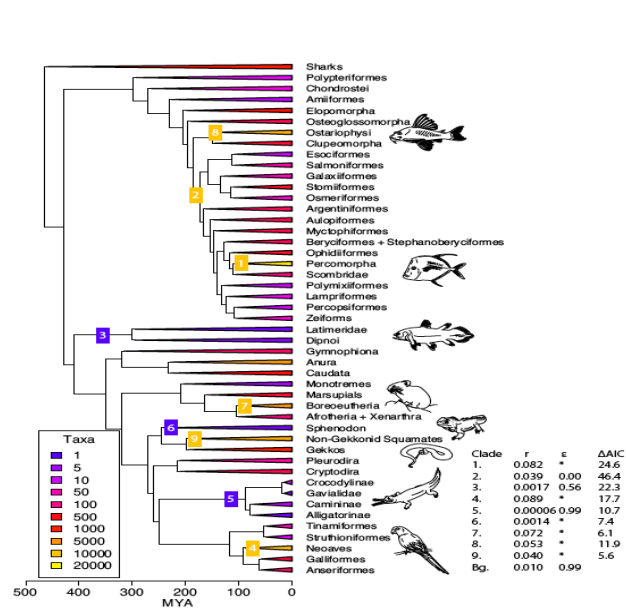


George Gaylord Simpson

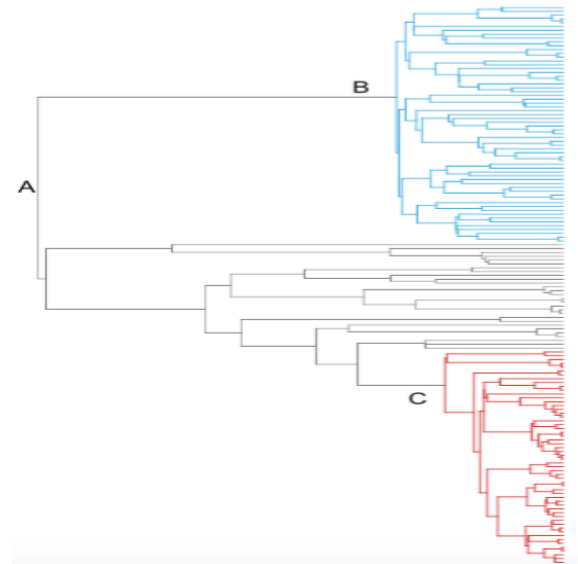
Most diversity of life on Earth arises from adaptive radiations, the fast diversification of ecological traits in a rapidly speciating group of organisms

Under this hypothesis, bursts of diversification are clade-wide, linked to the rapid filling of a niche space that has been freed from other occupants (e.g. by major environmental changes) or opened by a major key innovation.

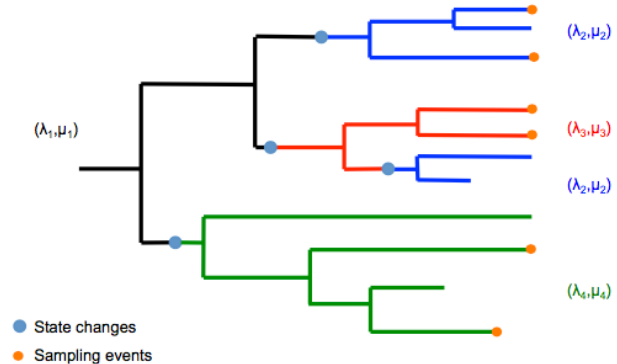
Heterogeneous birth-death models with clade-wide rate variation: few rate shifts with large effects



MEDUSA
Alfaro et al. PNAS 2009



BAMM
Rabosky PloS One 2014

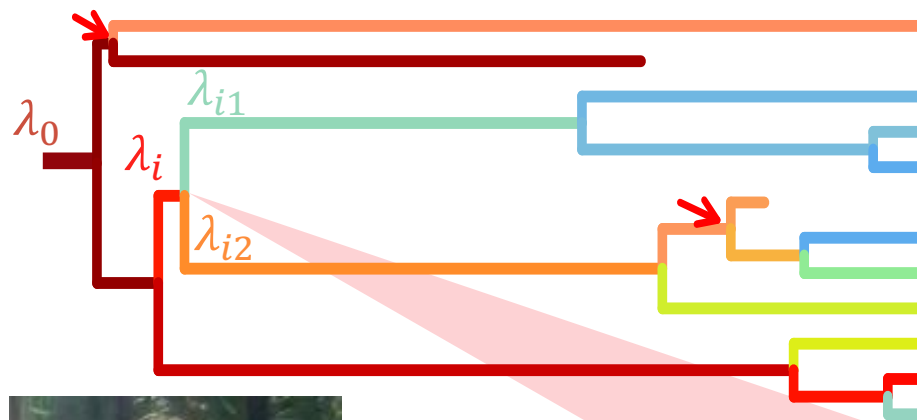


MSBD
Barido-Sottani et al. Syst Bio 2020

Another view of diversification

The complex interplay between species evolving ecologies and their specific spatial and environmental context results in highly dynamic diversification rates

The Cladogenetic Diversification rate Shift model (ClaDS):
a new model with lineage-specific rate variation

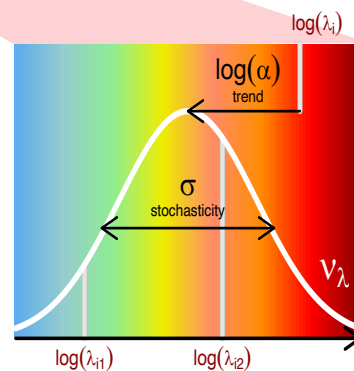


λ_0 initial speciation rate

α deterministic trend

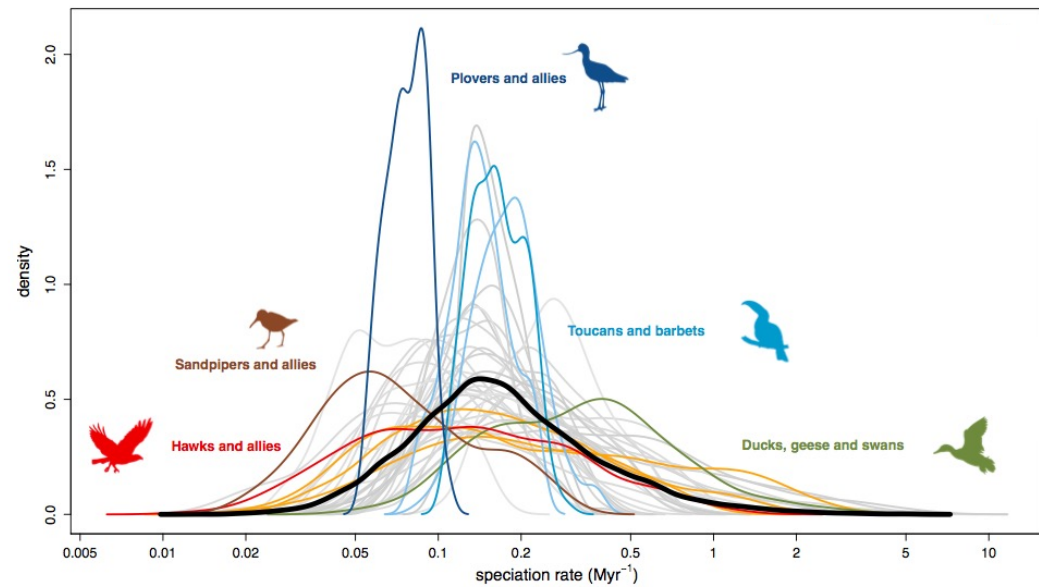
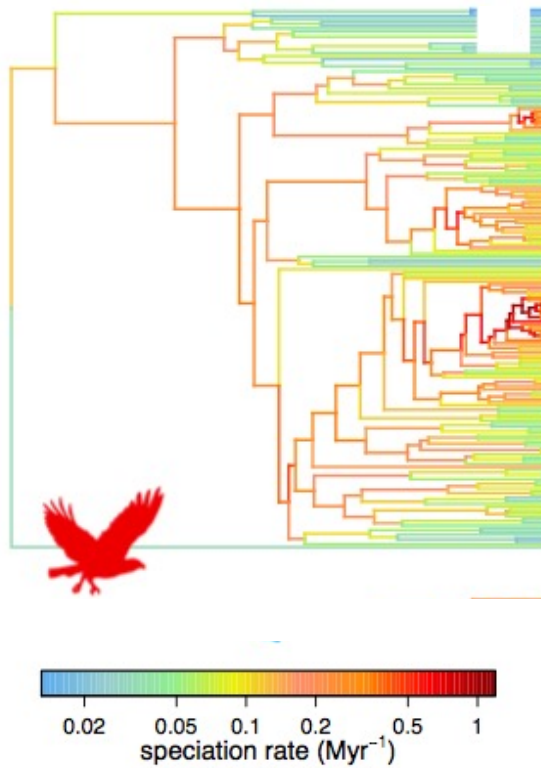
σ^2 stochastic variation

$\varepsilon = \mu_i / \lambda_i$ relative extinction rate

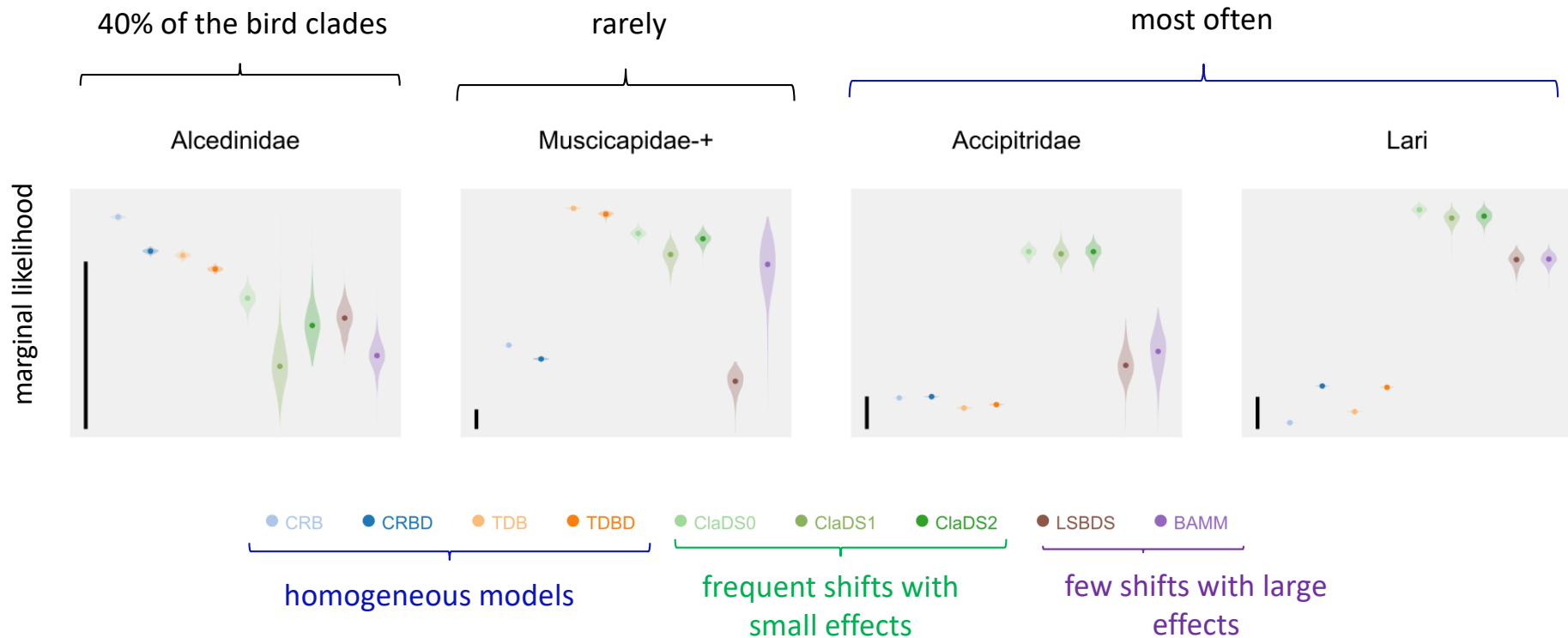


Maliet et al.
Nature Ecology & Evolution 2019

Speciation rates vary widely across lineages



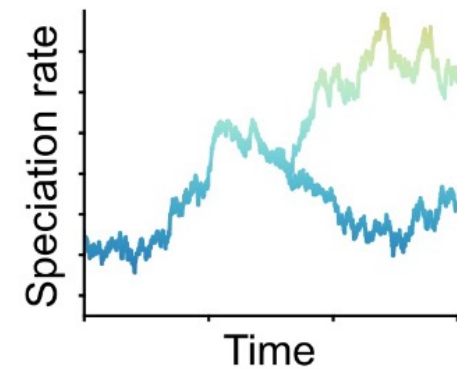
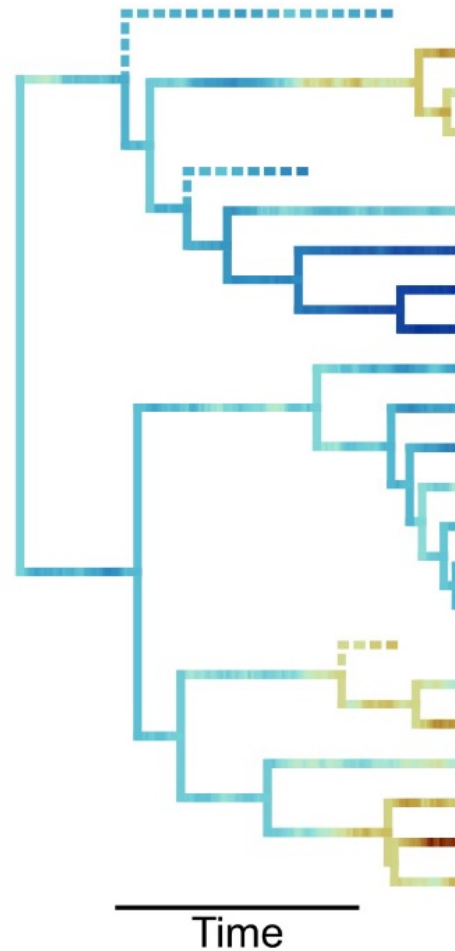
Models with many small shifts are better supported than models with few large shifts



The Birth-Death Diffusion model (BDD): continuous diffusion of diversification rates

Speciation

$$d\ln(\lambda_i(t)) = \alpha dt + \sigma_\lambda dW(t),$$



The Birth-Death Diffusion model (BDD): continuous diffusion of diversification rates

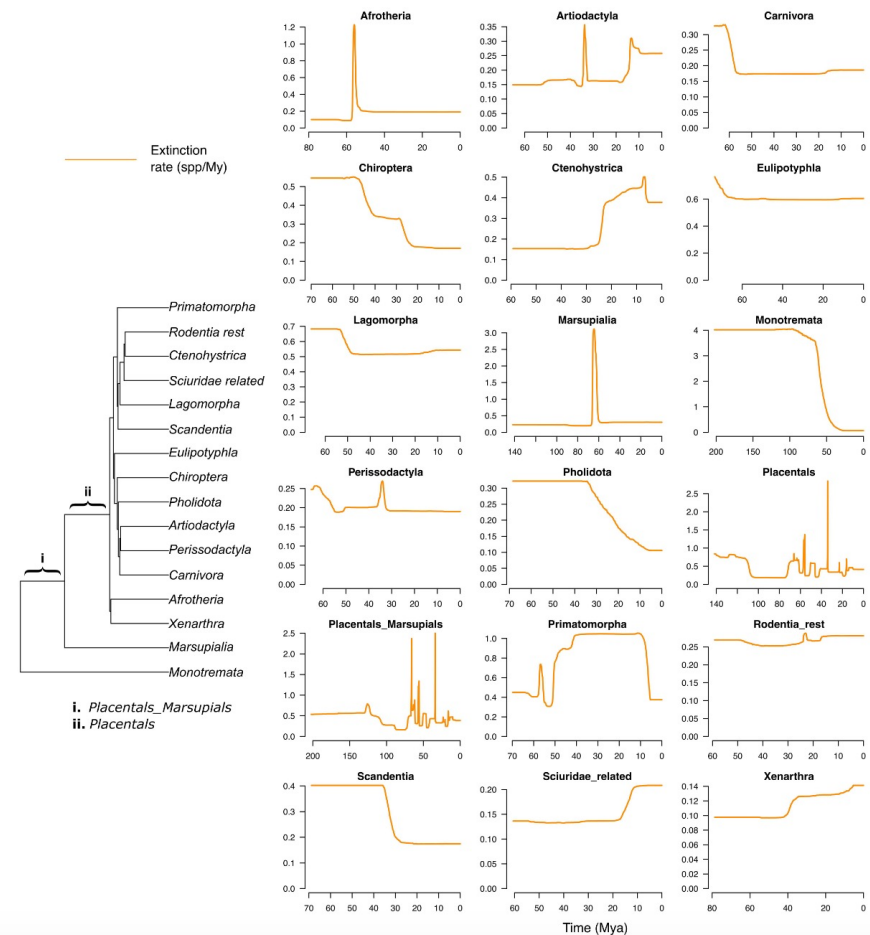
Extinction not constrained by fossil information

- No extinction
- Constant extinction
- Constant turnover
- Follows a Geometric Brownian Motion (GBM)

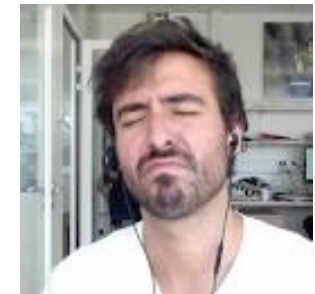
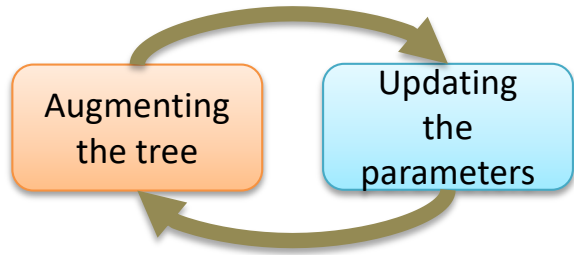
$$d\ln(\mu_i(t)) = \sigma_\mu dW(t),$$

where $W(t)$ denotes the Wiener process.

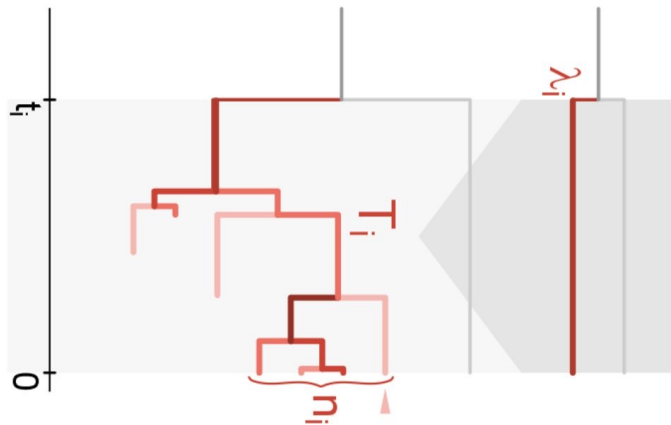
Extinction constrained by fossil information



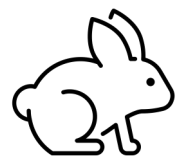
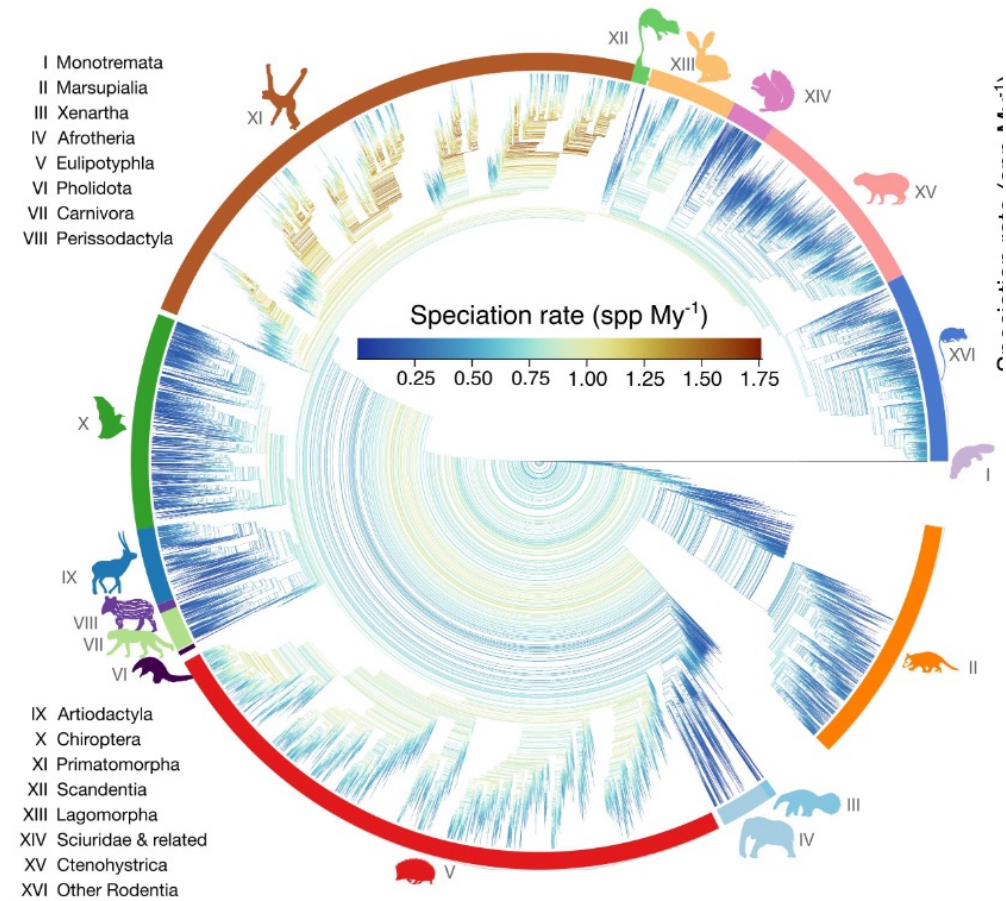
Inference with data augmentation



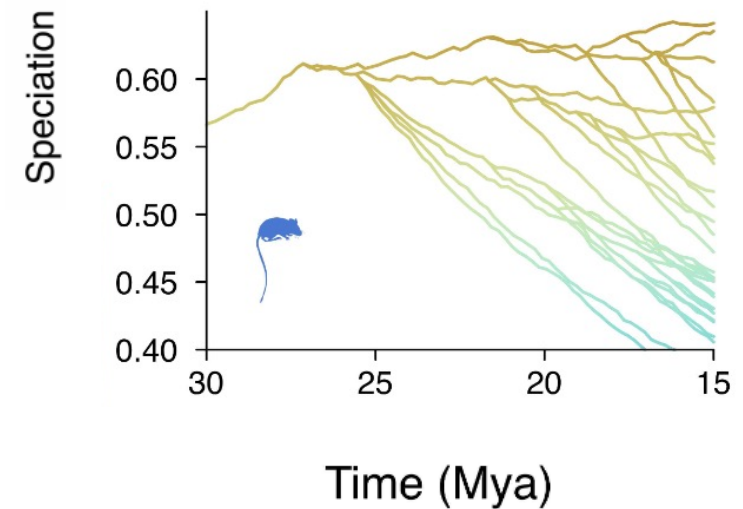
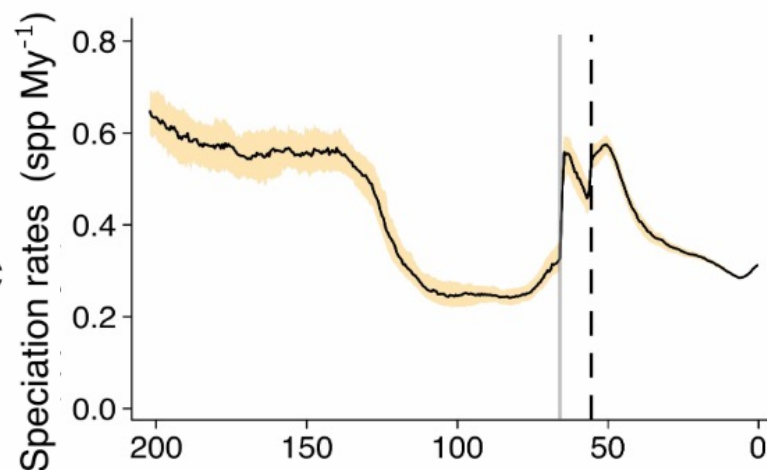
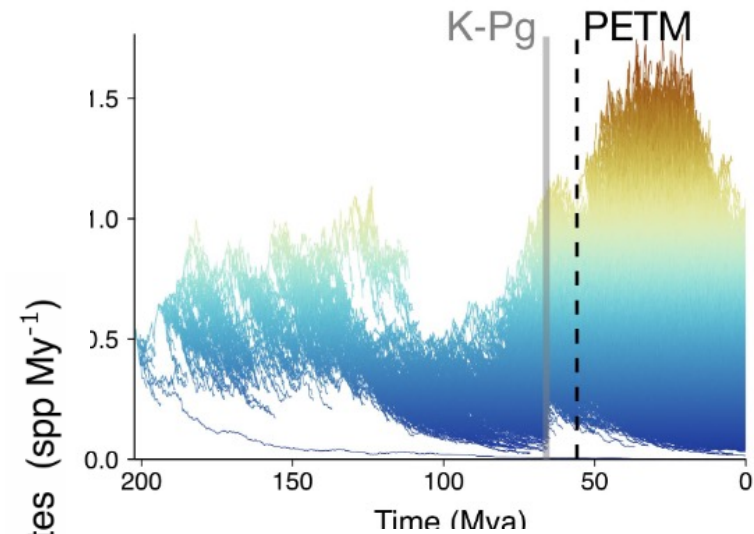
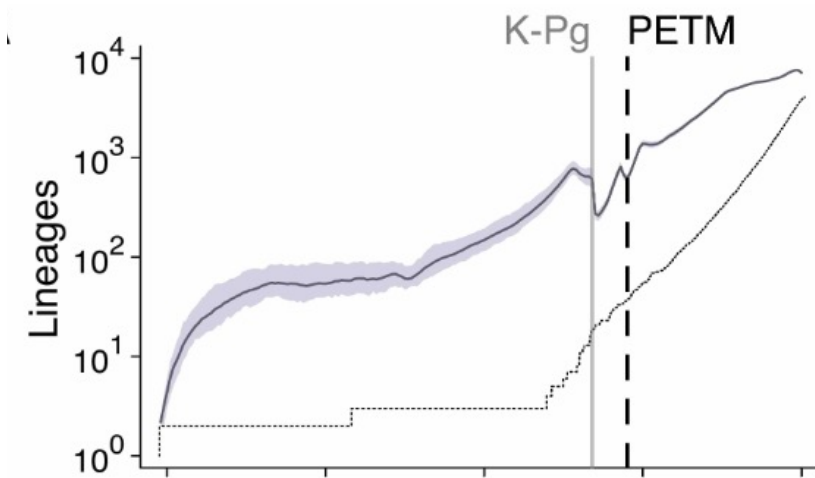
ClaDS

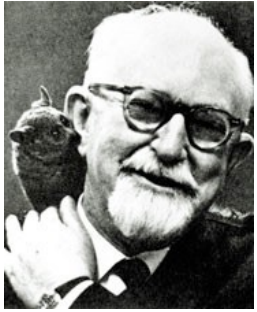


BDD

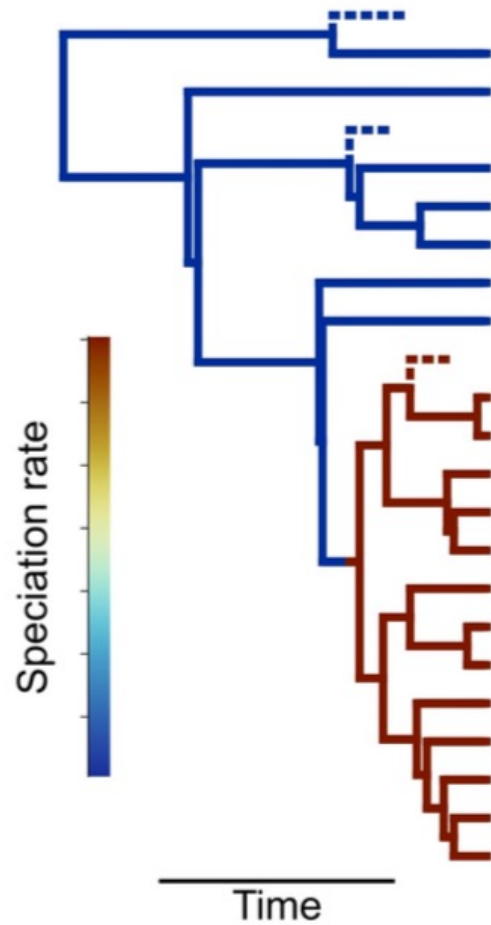


Reconstruction of Mammals past diversification and diversity dynamics under the Birth-Death Diffusion model using data augmentation

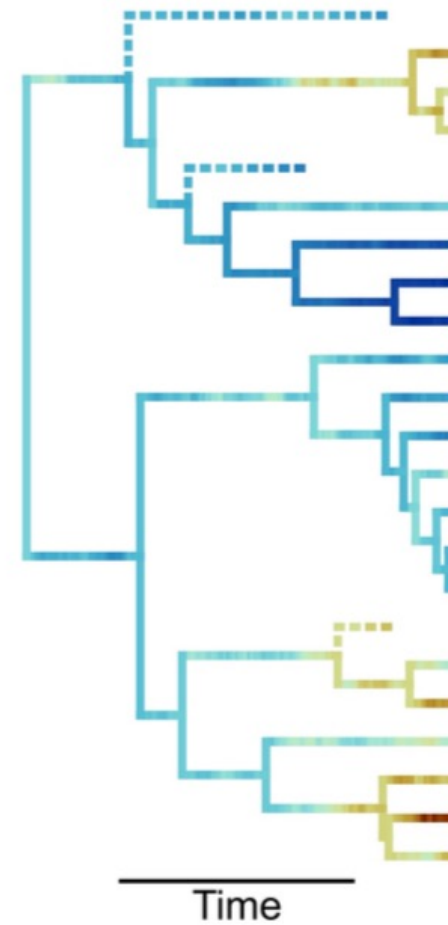




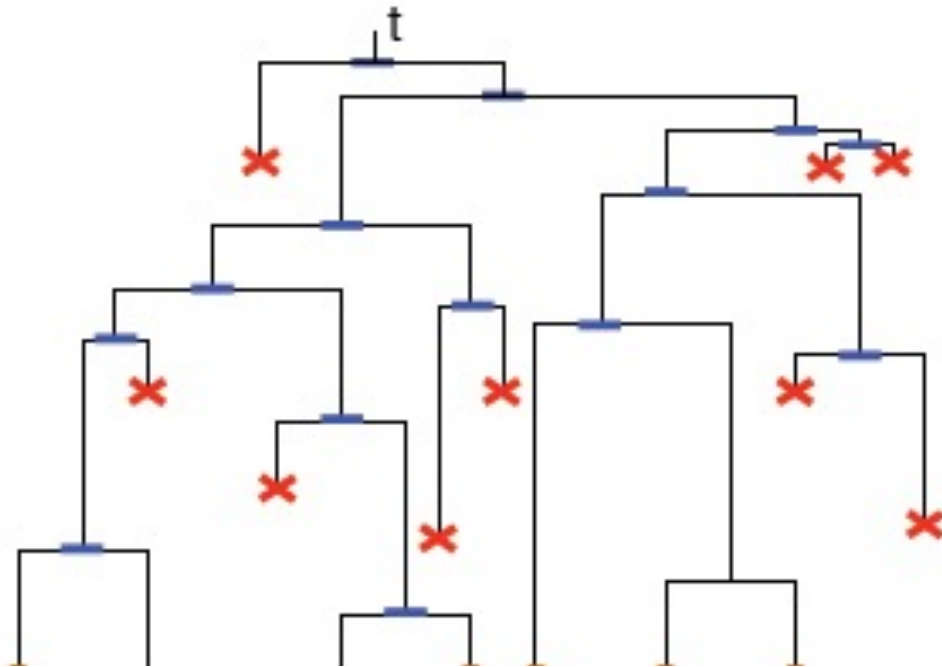
Major key
innovations
Adaptive radiations



Complex interplay between species
evolving ecologies and their specific
spatial and environmental context



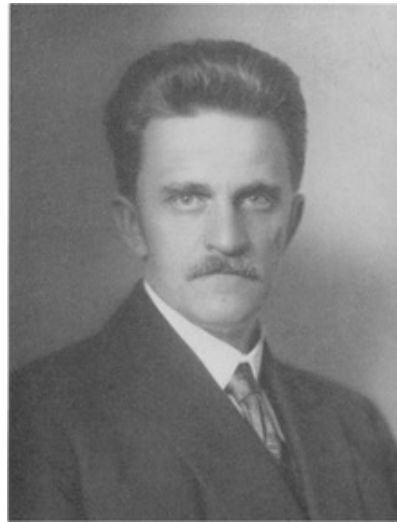
Birth-death model for the analysis of diversification



Ici une lignée est
une espèce

Macroevolution: evolution above the species level

Macroevolution distinguished from microevolution, which refers to evolutionary change within a species or population.



Yuri Filipchenko (1927)

Large evolutionary transitions cannot be explained by Darwinian evolution

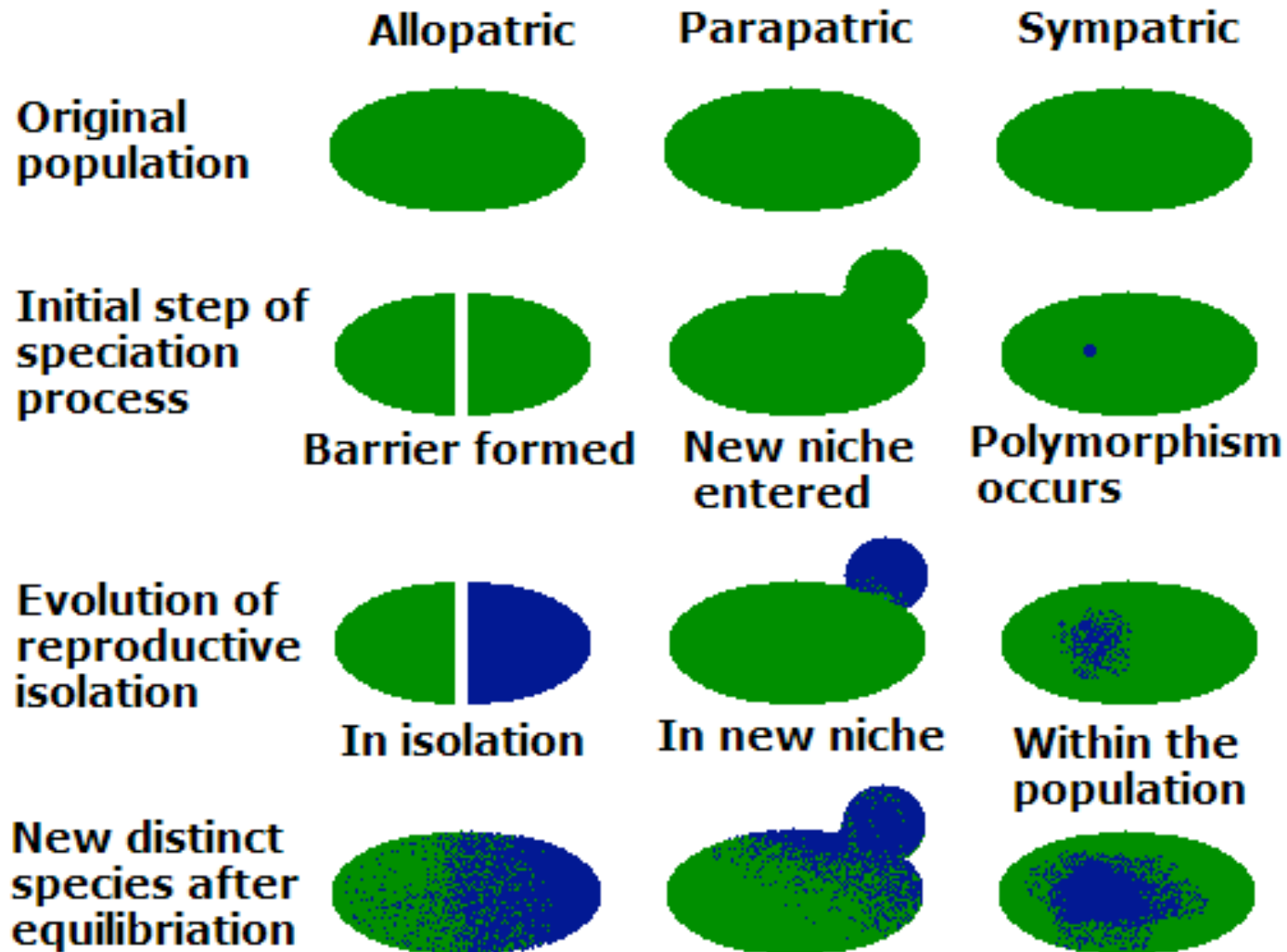
Macroevolution as the sum of microevolutionary changes over geological times?



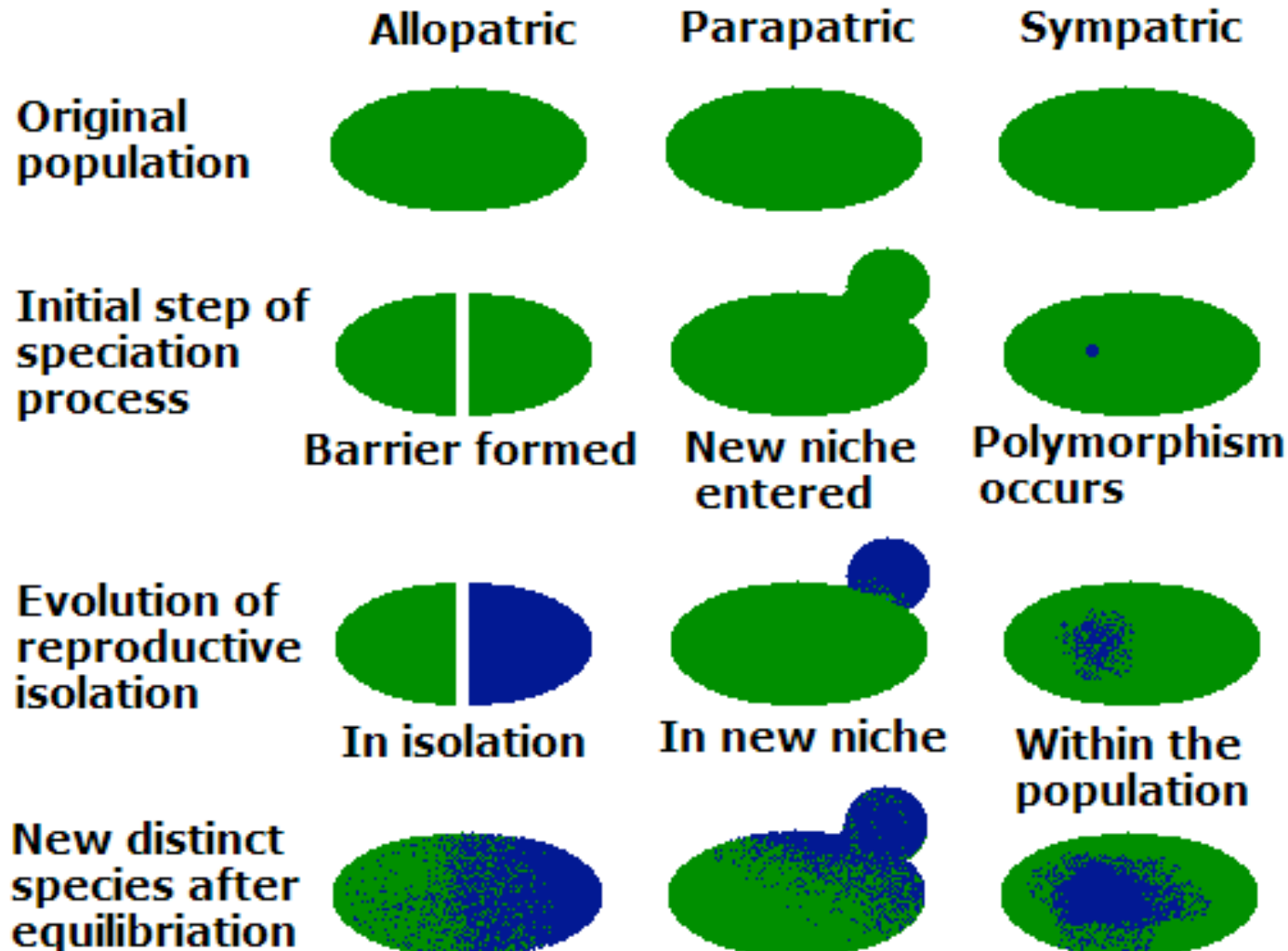
Theodosius Dobzhansky (1937)

What are the microevolutionary processes that modulate speciation and extinction rates?

The speciation process



Do we see a correlation between speciation rates (estimated from phylogenies) and intraspecific measures of differentiation?



Mixed support for an association between speciation rates and population geographic structure, genetic differentiation, or evolution of reproductive isolation



Macroevolutionary speciation rates are decoupled from the evolution of intrinsic reproductive isolation in *Drosophila* and birds

Daniel L. Rabosky^{a,1} and Daniel R. Matute^b



Positive association between population genetic differentiation and speciation rates in New World birds

Michael G. Harvey^{a,b,c,d,1}, Glenn F. Seeholzer^{a,b}, Brian Tilston Smith^{a,b,e}, Daniel L. Rabosky^{c,d}, Andrés M. Cuervo^{a,b,f}, and Robb T. Brumfield^{a,b}

updates

No link between population isolation and speciation rate in squamate reptiles

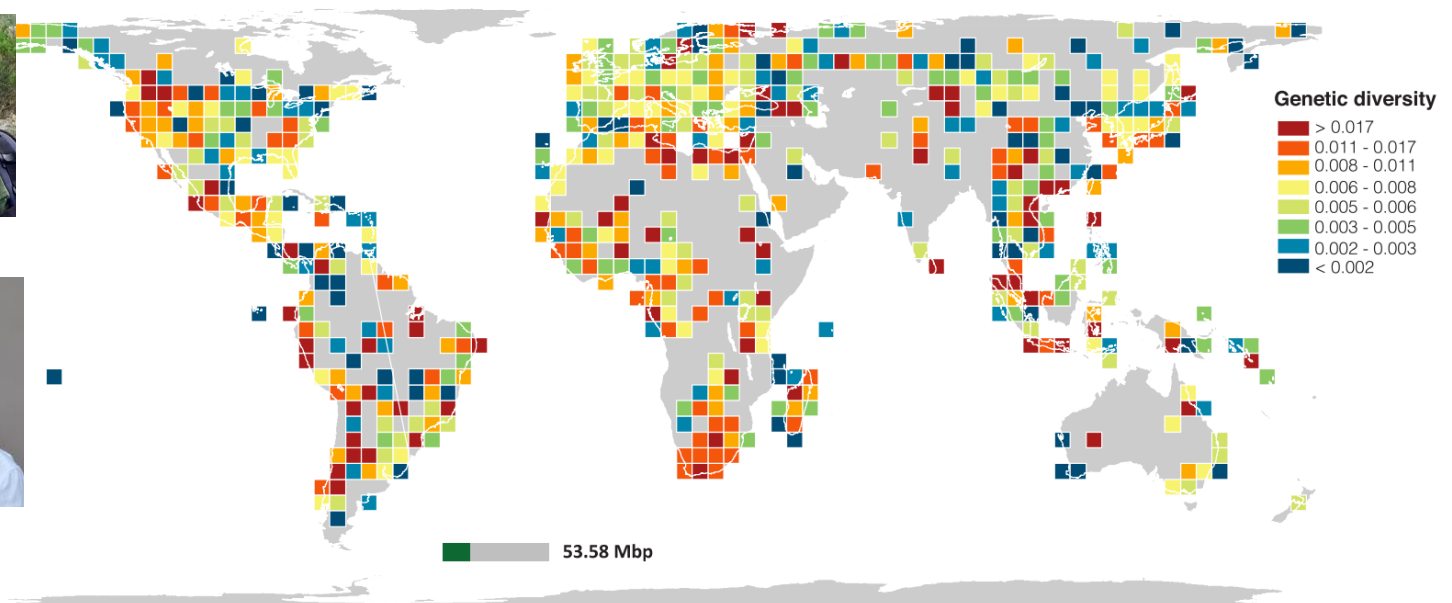
Sonal Singhal^{a,1}, Guarino R. Colli^b, Maggie R. Grundler^{c,d}, Gabriel C. Costa^e, Ivan Prates^{f,g}, and Daniel L. Rabosky^{f,g,1}

Illustration: test of the link between genetic diversity and speciation rates in mammals



Genetic diversity: level of polymorphism of DNA sequences among individuals within a species

Genetic diversity has major implications for the evolution and conservation of species



Genetic diversity: level of polymorphism of DNA sequences among individuals within a species

Watterson's theta

nombre de sites polymorphes

$$\hat{\theta}_w = \frac{K}{a_n}, \quad a_n = \sum_{i=1}^{n-1} \frac{1}{i}$$

Tajima's theta

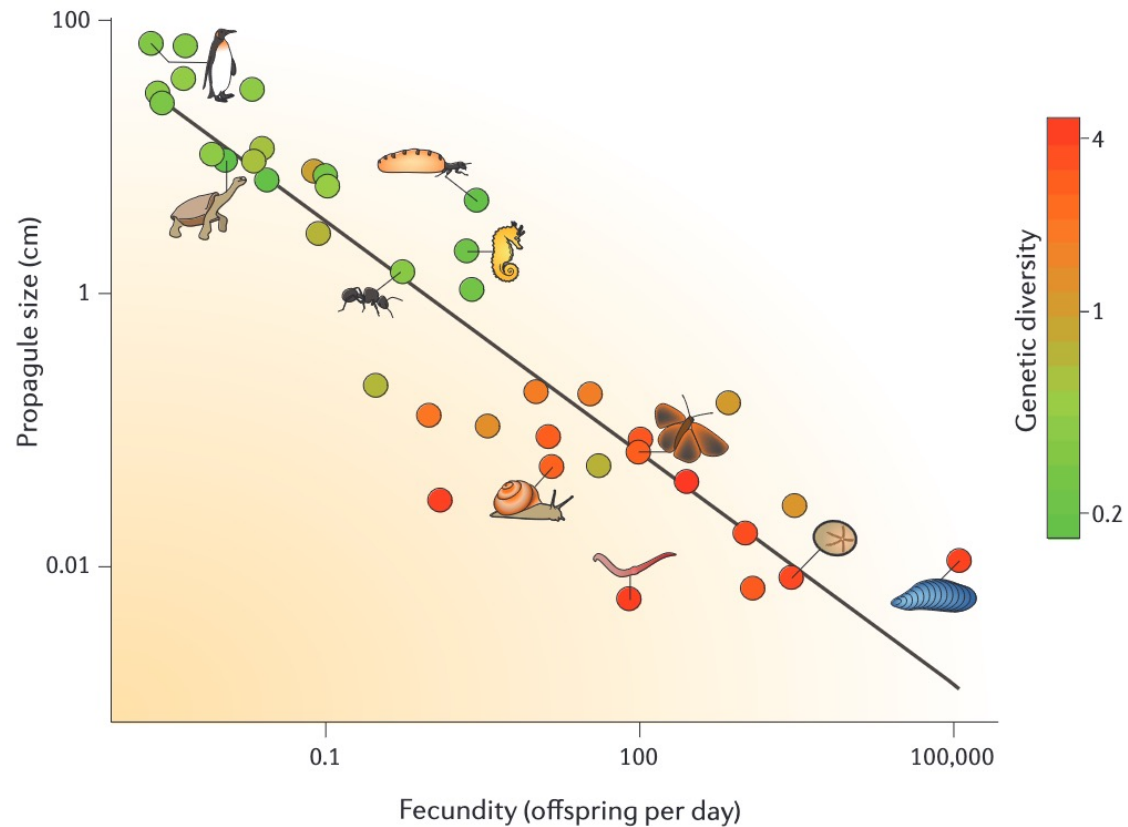
nombre de polymorphismes entre les séquences i et j

$$\hat{k} = \frac{\sum \sum_{i < j} k_{ij}}{\binom{n}{2}}$$

Under the Wright-Fisher model, these metrics are estimators of the expected number of polymorphisms between two DNA sequences

$$\vartheta = N_e \mu$$

Studies on the correlates of genetic diversity have focused on life history traits



Which correlation could we expect between genetic diversity and speciation rate?

Genetic diversity

$$\vartheta = N_e \mu$$

Speciation rate

Larger in species with:

Large population sizes

more likely hit by
geographic barriers



High mutation rates



Geographic structure (depends on
the details of migration process)



Favors:

Divergent selection if limited by the
availability of polymorphic alleles



Favors:

Competition-mediated divergent
selection

Which correlation could we expect between genetic diversity and speciation rate?

Genetic diversity

$$\vartheta = N_e \mu$$

Speciation rate



Generates:

Bottlenecks

Smaller in species with:

Small population sizes

faster accumulation of slightly deleterious mutations



Divergent selection



Geographic structure (depends on the details of migration process)



Estimating intraspecific genetic diversity across mammals

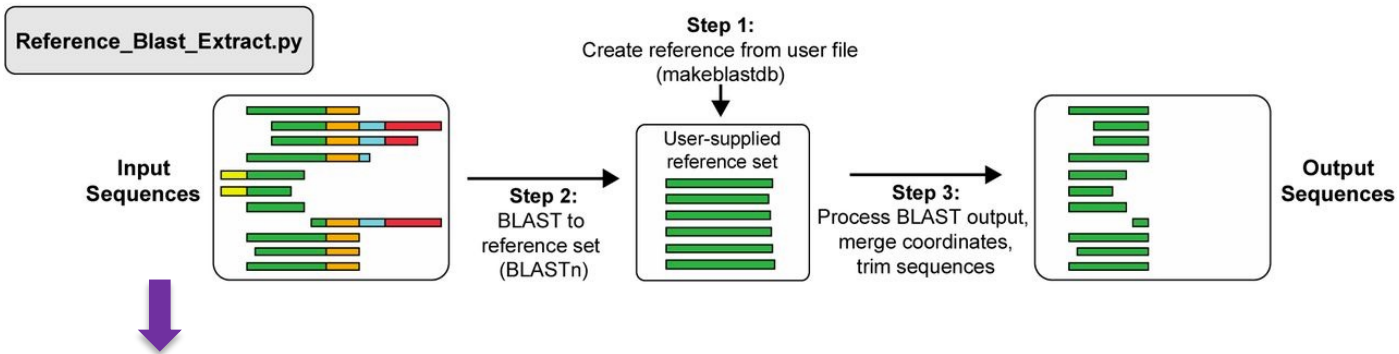
GenBank "Mammalia"[Organism] AND CYTB NOT "Homo sapiens"[Organism]



124,289 sequences of mammals Cytochrome b

Split into 138 families for better alignments

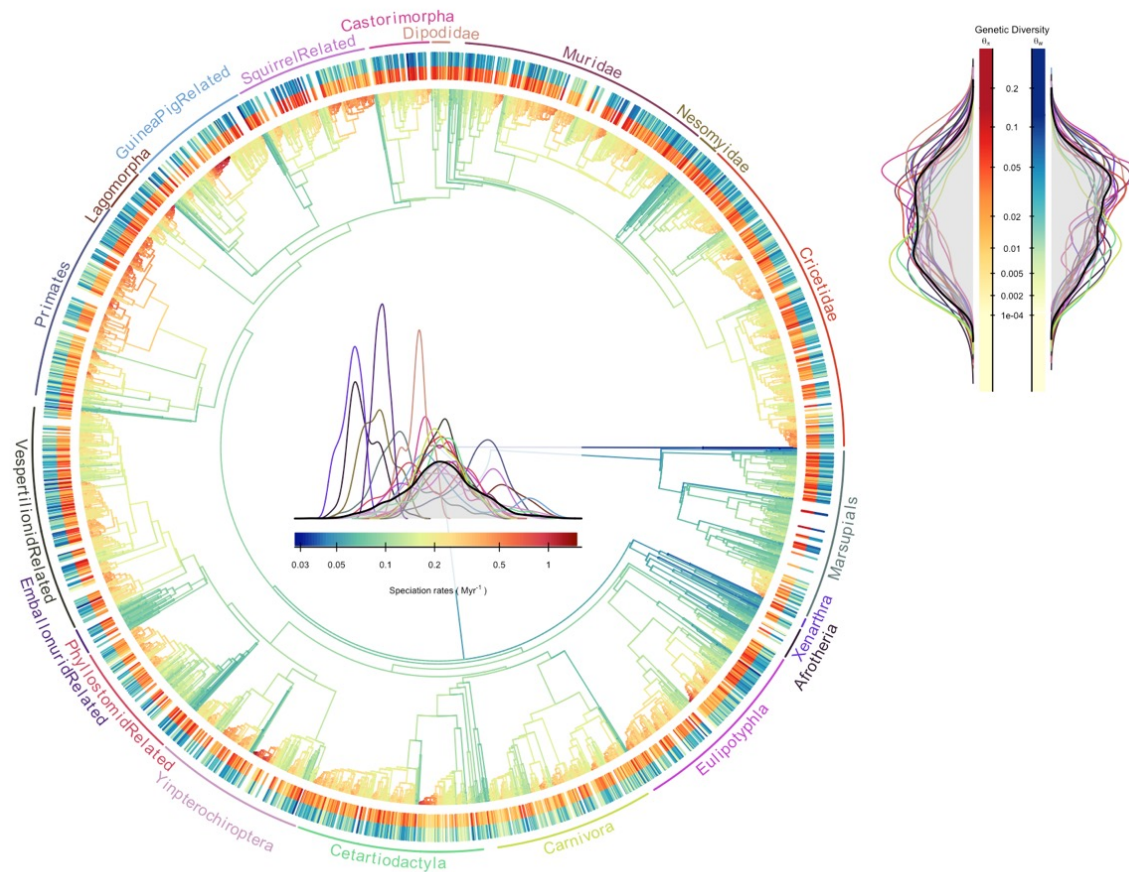
SuperCRUNCH 
For Phylogenetic Data
Portik and Wiens
(2019) MEE 2020



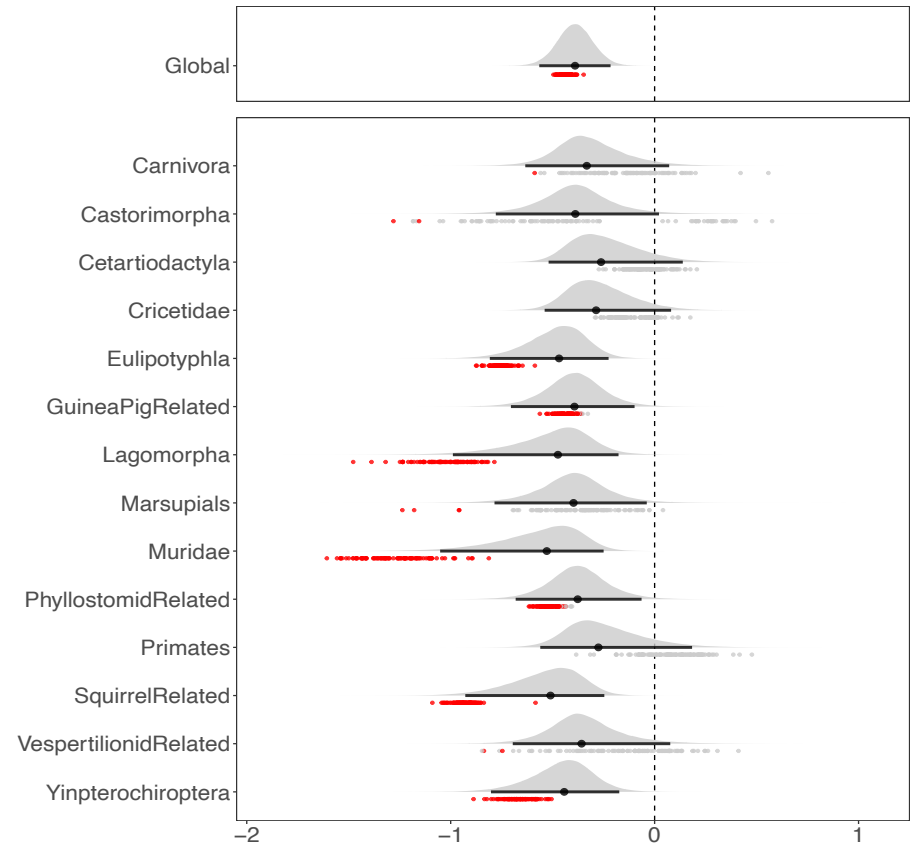
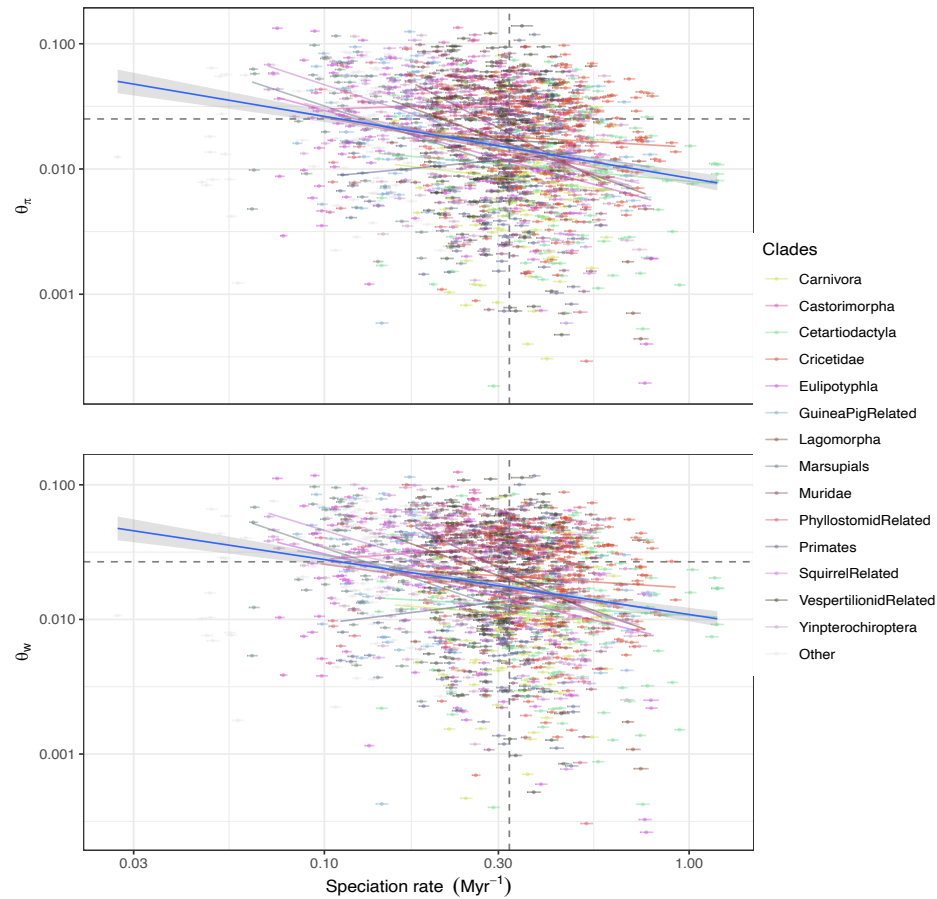
3899 species → 2004 species with at least 5 individuals → genetic diversity

Estimating branch-specific speciation rates

ClaDS estimates on Upham's PLoS Biology 2019 mammals phylogeny



Negative correlation between genetic diversity and speciation rates



The negative correlation between genetic diversity and speciation rates is not linked to an indirect effect of covariates

term	$\theta_T \sim \text{Traits}$				$\lambda \sim \text{Traits}$				$\theta_T \sim \lambda + \text{Traits}$			
	PGLS		BMLM		PGLS		BMLM		PGLS		BMLM	
	Estimate	SE	Estimate	95% CI	Estimate	SE	Estimate	95% CI	Estimate	SE	Estimate	95% CI
λ									-0.285	0.078	-0.273	[-0.448; -0.078]
Mean latitude	-0.113	0.029	-0.100	[-0.174; -0.023]	-0.002	0.006	0.005	[-0.013; 0.023]	-0.104	0.029	-0.096	[-0.173; -0.018]
Mean temperature	0.226	0.097	0.218	[0.005; 0.419]	-0.022	0.018	-0.020	[-0.055; 0.017]	0.212	0.096	0.196	[0.011; 0.392]
Body Mass	-0.129	0.026	-0.123	[-0.189; -0.058]	0.004	0.009	0.011	[-0.01; 0.032]	-0.124	0.025	-0.118	[-0.182; -0.055]
Generation length	-0.052	0.106	-0.105	[-0.353; 0.149]	-0.009	0.028	0.010	[-0.05; 0.074]	-0.068	0.103	-0.104	[-0.34; 0.136]
Litter size	-0.254	0.093	-0.370	[-0.594; -0.15]	0.055	0.027	0.049	[-0.015; 0.111]	-0.238	0.091	-0.359	[-0.592; -0.125]

Which correlation could we expect between genetic diversity and speciation rate?

Genetic diversity

$$\vartheta = N_e \mu$$

Speciation rate



Generates:

Bottlenecks

Smaller in species with:

Small population sizes

faster accumulation of
slightly deleterious
mutations



Divergent selection

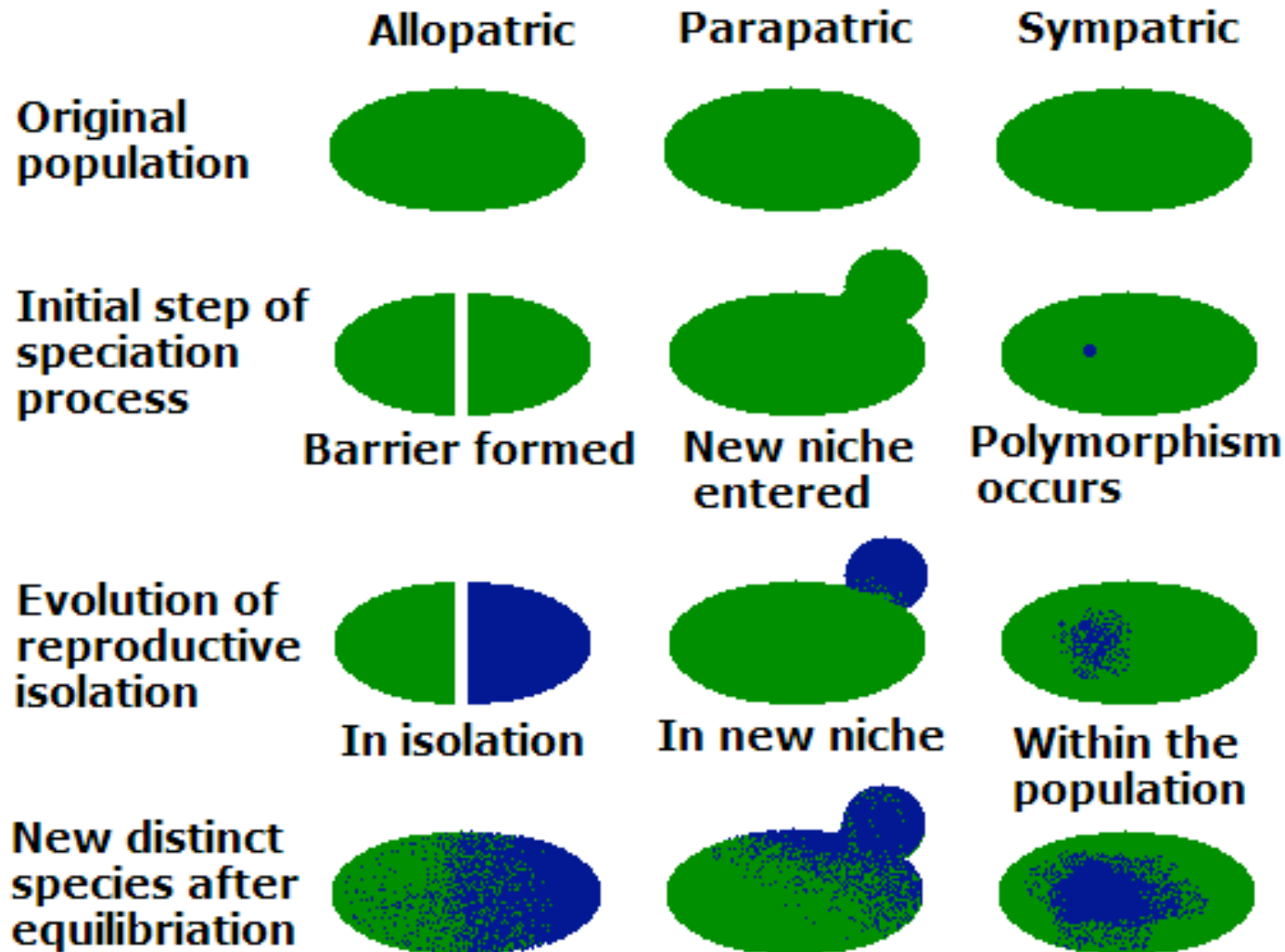


Geographic structure (depends on
the details of migration process)



What are the microevolutionary processes that modulate speciation rates?

The speciation process



Which correlation between intraspecific measures of differentiation and speciation rates?

AS

Macroevolutionary speciation rates are decoupled from the evolution of intrinsic reproductive isolation in *Drosophila* and birds

Daniel L. Rabosky^{a,1} and Daniel R. Matute^b

AS

Positive association between population genetic differentiation and speciation rates in New World birds

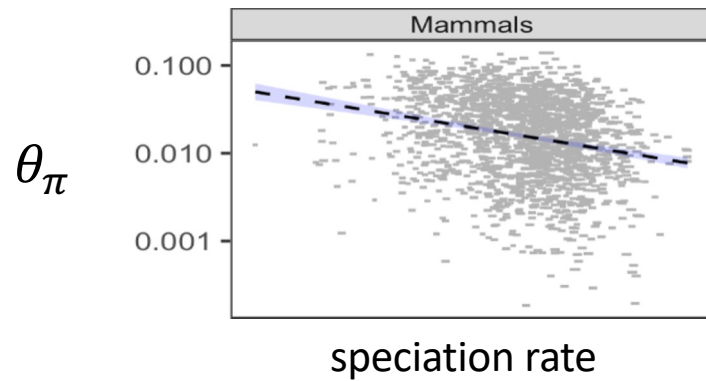
Michael G. Harvey^{a,b,c,d,1}, Glenn F. Seeholzer^{a,b}, Brian Tilston Smith^{a,b,e}, Daniel L. Rabosky^{c,d}, Andrés M. Cuervo^{a,b,f}, and Robb T. Brumfield^{a,b}

updates

No link between population isolation and speciation rate in squamate reptiles

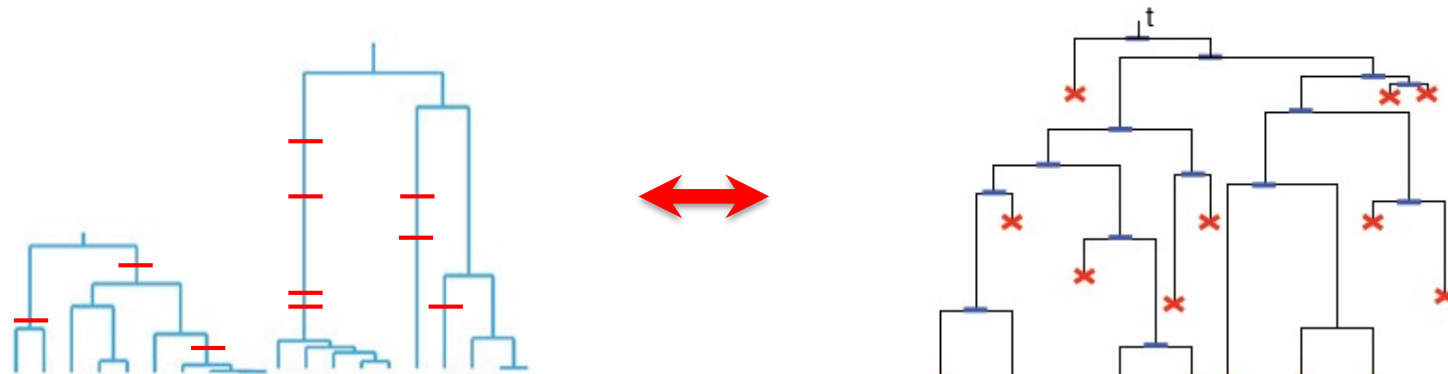
Sonal Singhal^{a,1}, Guarino R. Colli^b, Maggie R. Grundler^{c,d}, Gabriel C. Costa^e, Ivan Prates^{f,g}, and Daniel L. Rabosky^{f,g,1}

Which correlation between intraspecific measures of differentiation and speciation rates?



Silva et al. (under revision for Nature Comm)

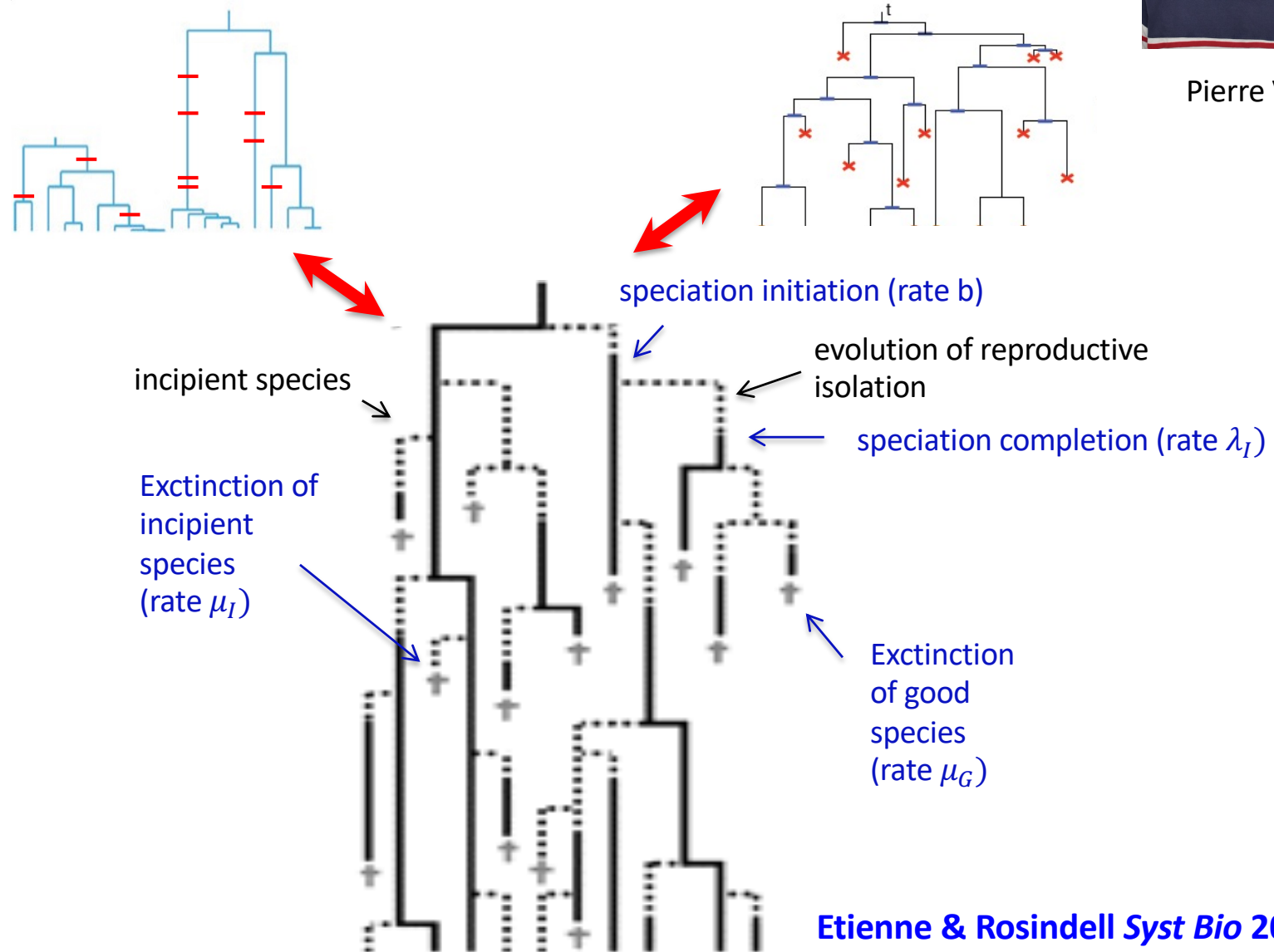
Need to go from verbal to quantitative predictions



The protracted birth-death model as a way to bridge micro and macroevolutionary speciation research



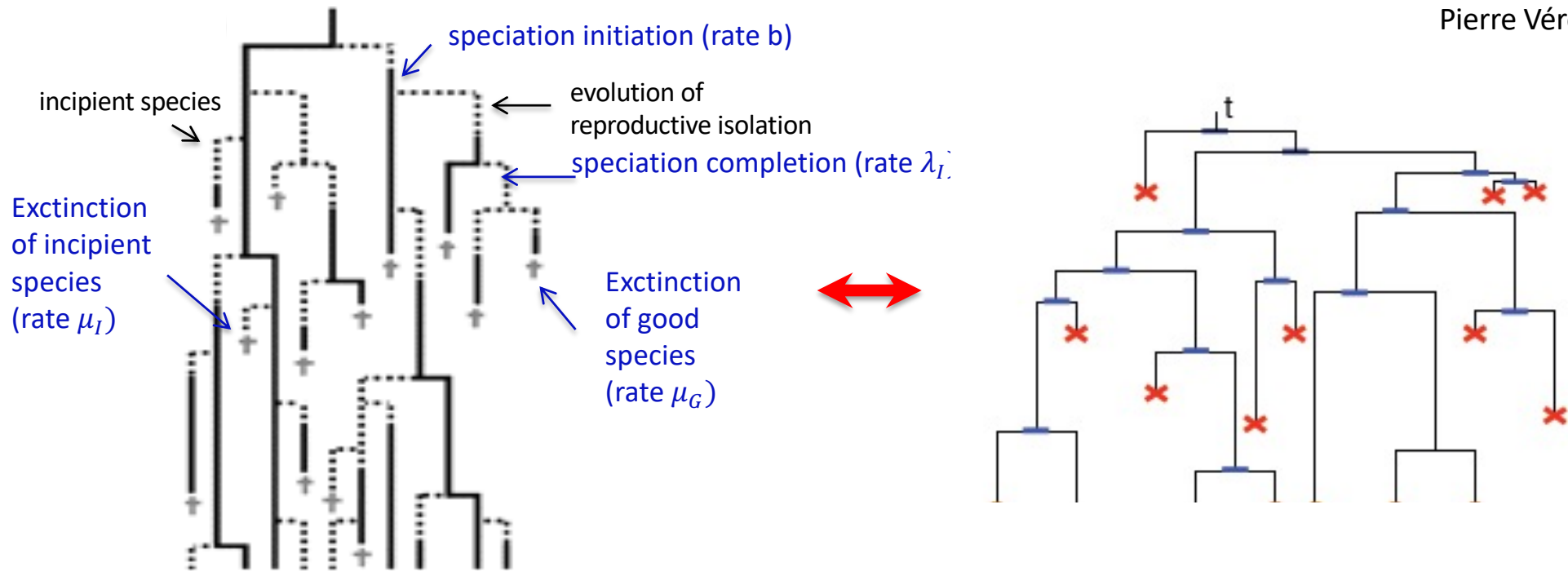
Pierre Véron



Linking the protracted birth-death model and the birth-death model



Pierre Véron



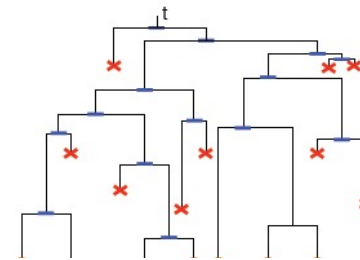
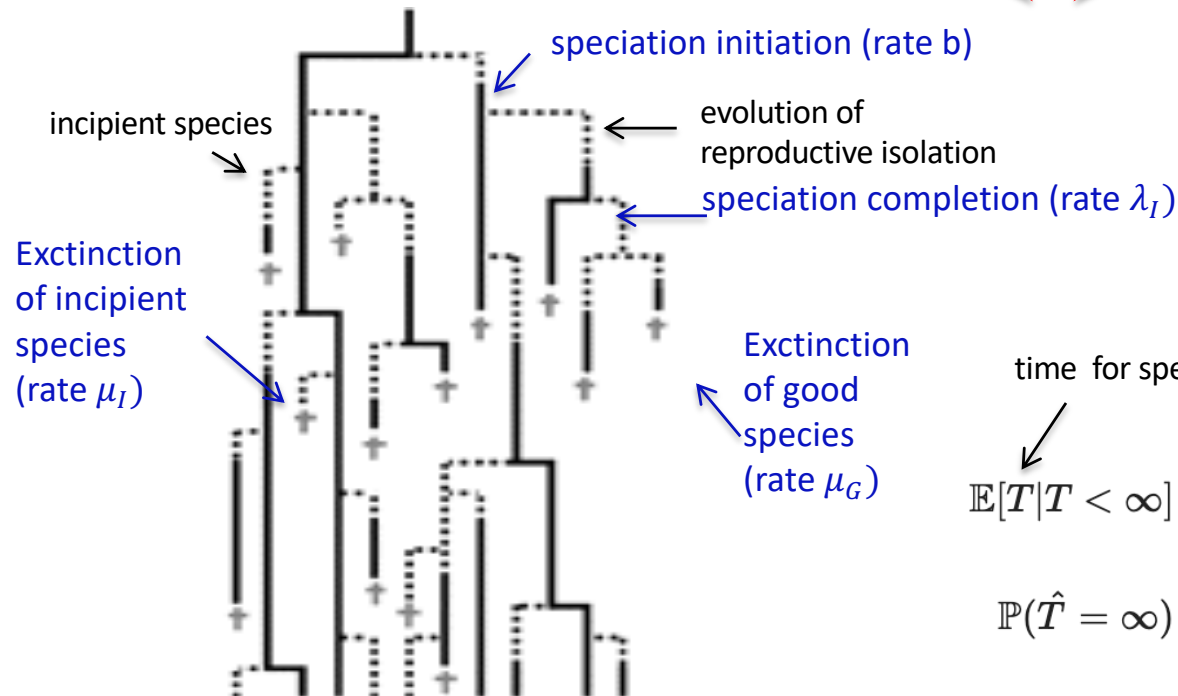
Etienne & Rosindell *Syst Bio* 2012

What are the 'speciation' ($\hat{\lambda}$) and 'extinction' ($\hat{\mu}$) rates that have the same expected time to speciation than under the PBD model with speciation initiation rate b , speciation completion rate (λ_I) and extinction rates of the ancestral (μ_G) and incipient species (μ_I)?

'Speciation' and 'extinction' rates under the protracted birth-death model



Pierre Véron



time for speciation

$$\mathbb{E}[T|T < \infty] = \frac{1}{\hat{\lambda} + \hat{\mu}} = \frac{bp + \lambda_I + \mu_I + 2\mu_G}{(bp + \mu_G)(\lambda_I + \mu_I + \mu_G)} \quad \text{and}$$

$$\mathbb{P}(\hat{T} = \infty) = \frac{\hat{\mu}}{\hat{\lambda} + \hat{\mu}} = \frac{\mu_G}{\mu_G + bp} \left(1 + \frac{pb}{\lambda_I + \mu_I + \mu_G} \right).$$

$$\hat{\mu} = \mu_G \left(1 - \frac{\mu_G}{bp + \lambda_I + \mu_I + 2\mu_G} \right)$$

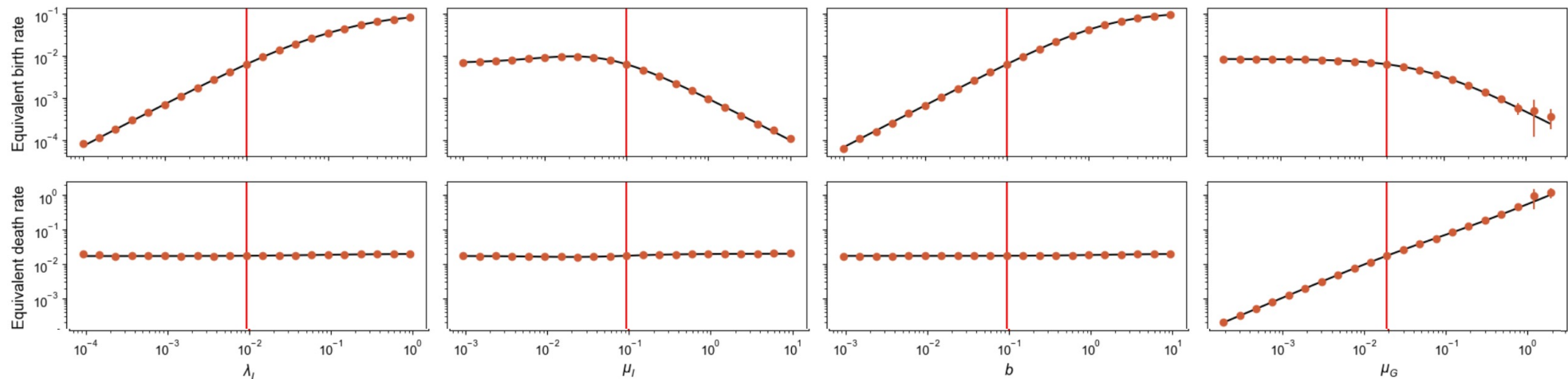
$$\hat{\lambda} = bp \frac{\lambda_I + \mu_I}{bp + \lambda_I + \mu_I + 2\mu_G}.$$

The **‘speciation’ rate** is directly related to the rates of speciation initiation and completion; it decreases with increasing rates of extinction (for incipient and good species), but only when they are high compared with the rates of speciation initiation and completion



Pierre Véron

The **‘extinction’ rate** is directly related to the rate of extinction of good species)



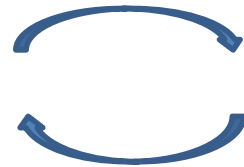
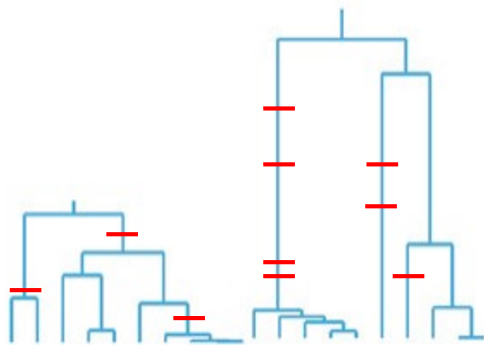
Linking demographic and genetic processes to the protracted birth-death model



Pierre Véron

Demographic process

Population genetics



Conclusions

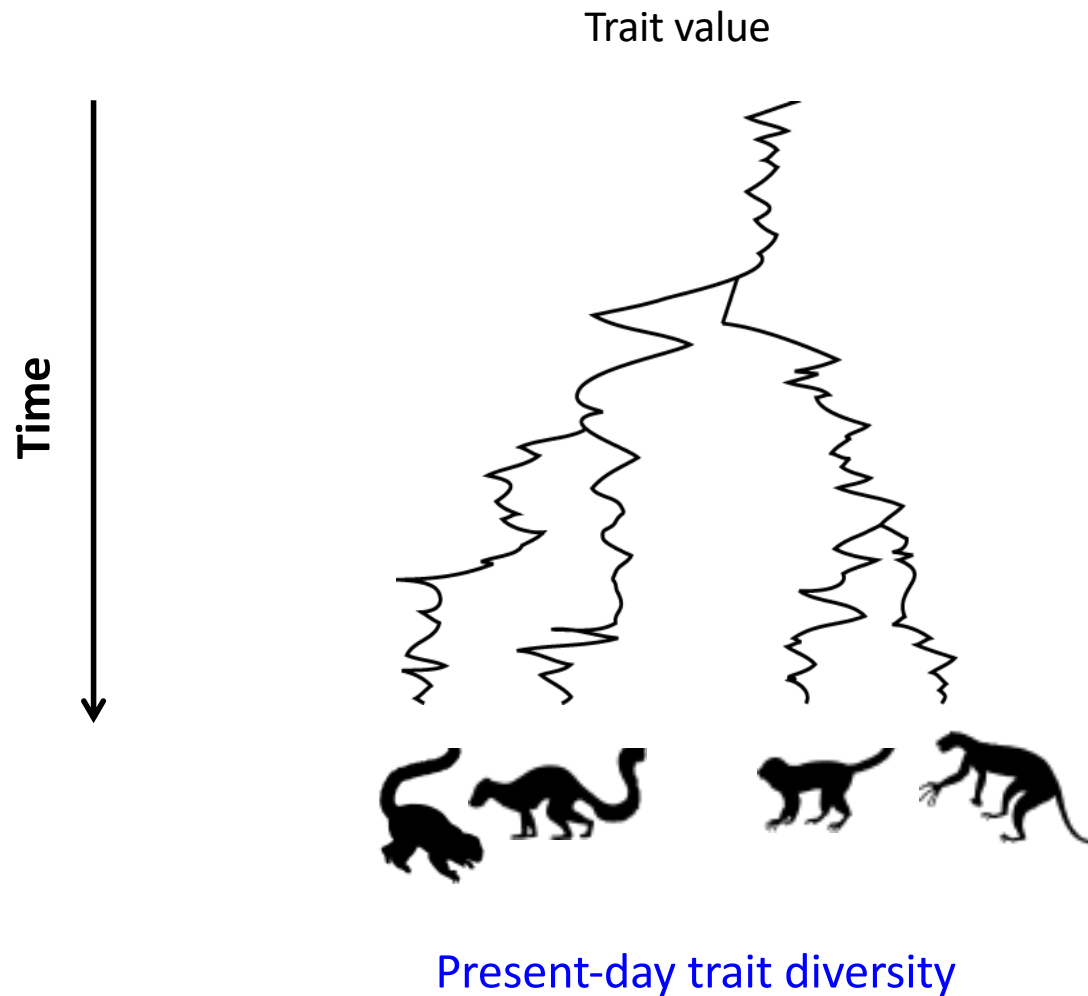
Speciation and extinction rates vary widely across lineages, explaining why some species groups are much more species rich than others

Differences in speciation and extinction rates can be linked to species specific traits as well as abiotic and biotic factors

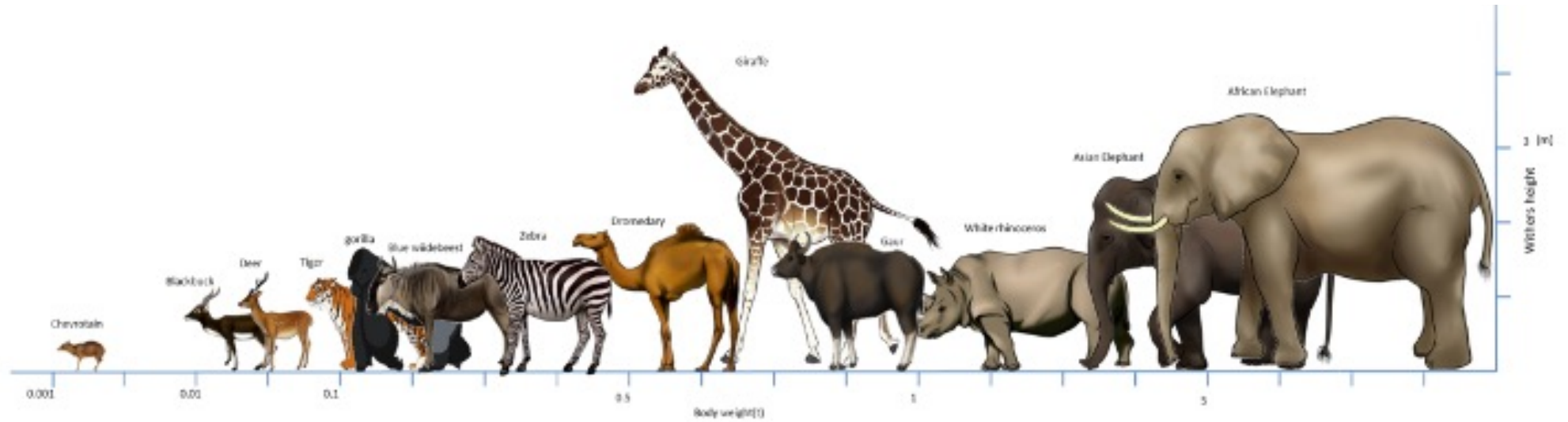
We have well-developed models to assess the effect of species-specific traits and abiotic factors on speciation and extinction rates

Understanding which microevolutionary processes act a rate-limiting step in speciation (and therefore drive present day species richness patterns) remains a major research frontier

Application des modèles de diffusion à l'étude phylogénétique de l'évolution des phénotypes



La diversité phénotypique

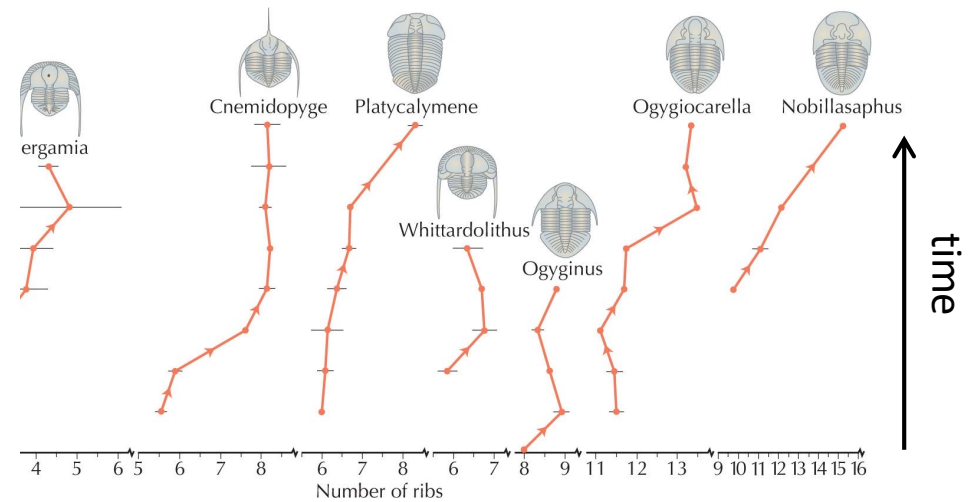
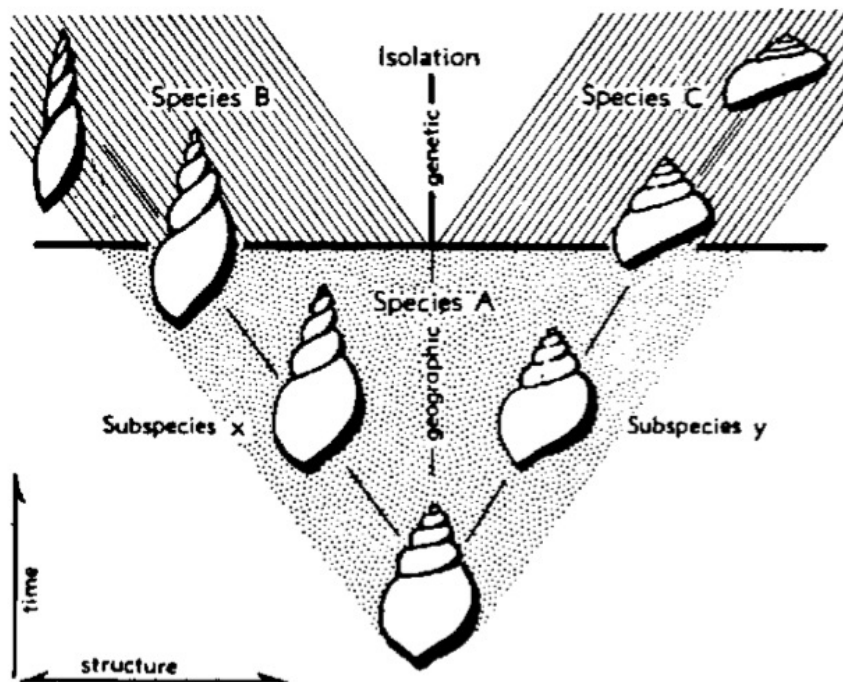


La diversité phénotypique



The mode and tempo of phenotypic evolution over geological time scales through the lens of fossil data

Phyletic gradualism



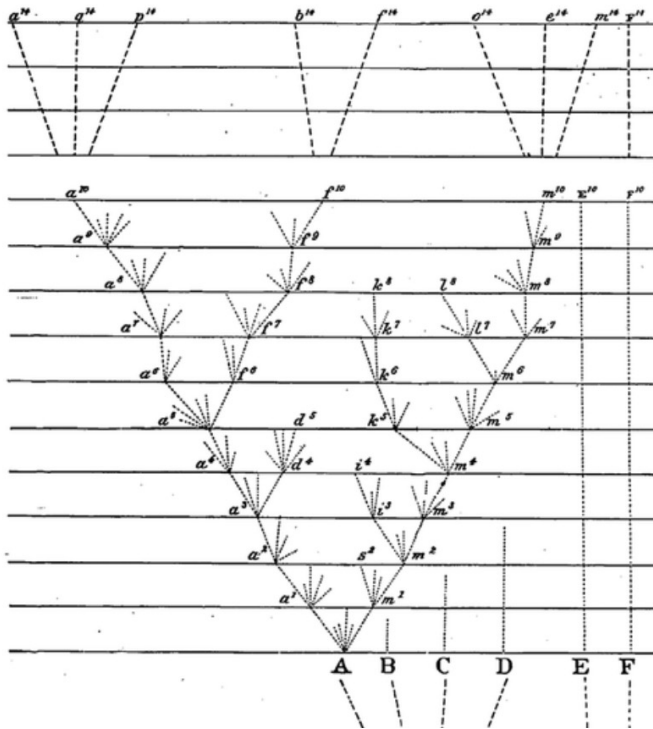
Phyletic gradualism and “gaps” in the fossil record

« Why do we not find the infinitely numerous transitional links that would illustrate the slow and steady operation of natural selection? Why is not every geological formation and every stratum full of such intermediate links? Geology assuredly does not reveal any such finely graduated organic chain; and this, perhaps, is the gravest objection which can be urged against my theory »

« The geological record is extremely imperfect and this fact will to a large extent explain why we do not find interminable varieties, connecting together all the extinct and existing forms of life by the finest graduated steps »

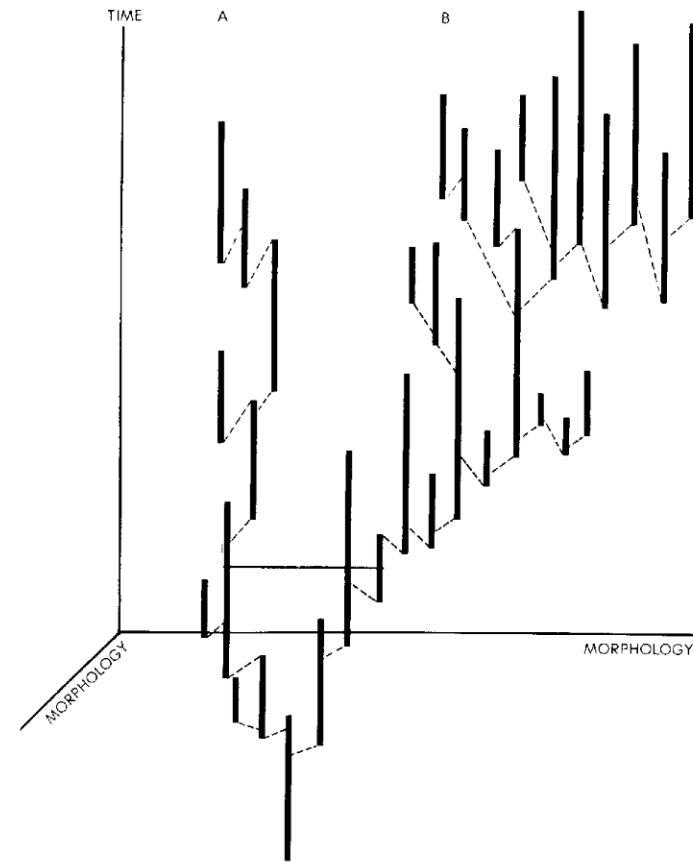
Punctuated equilibria: an alternative to phyletic gradualism

Gradual changes

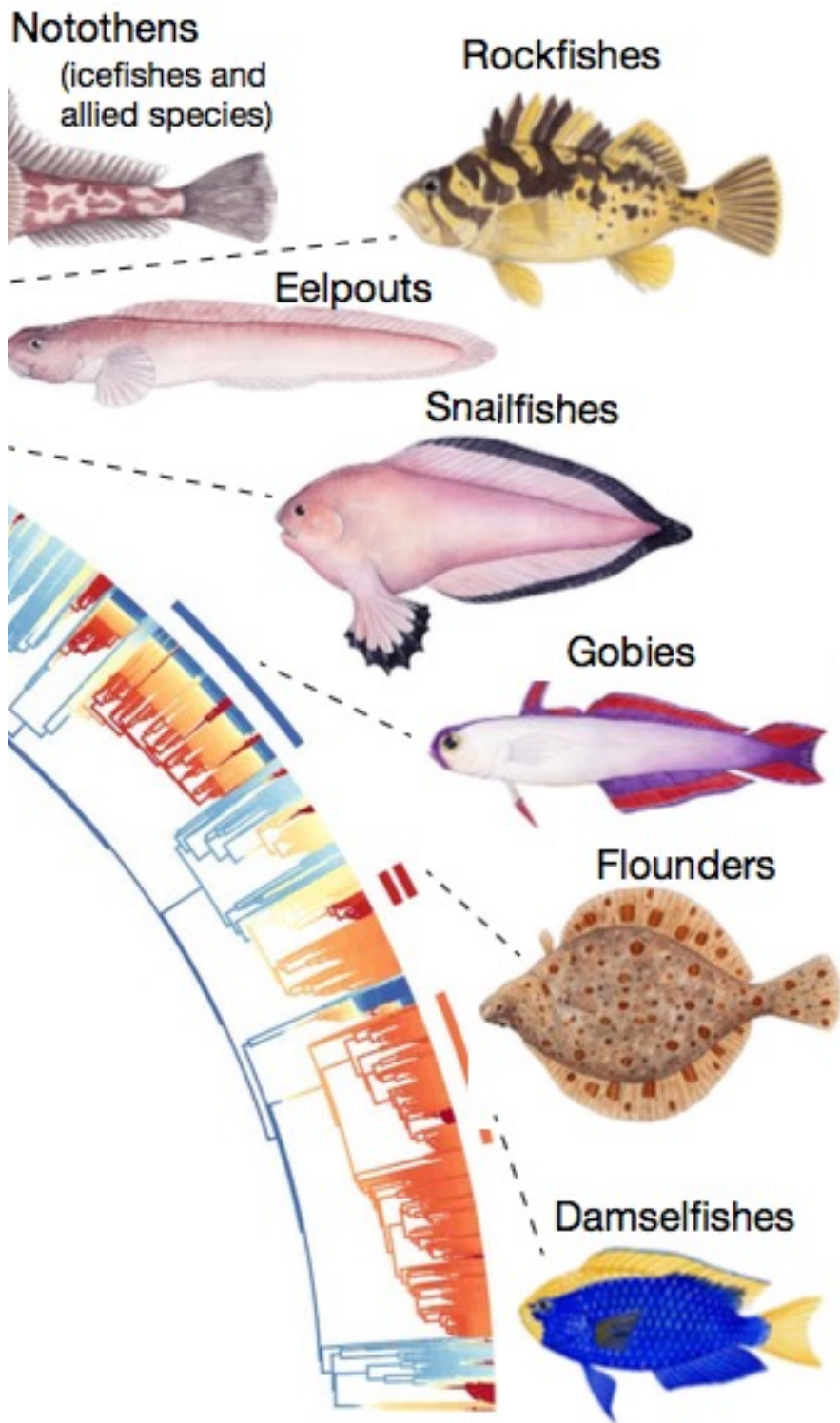


Darwin 1859

Periods of stases separated by periods
of fast evolutionary change



Eldredge & Gould 1972



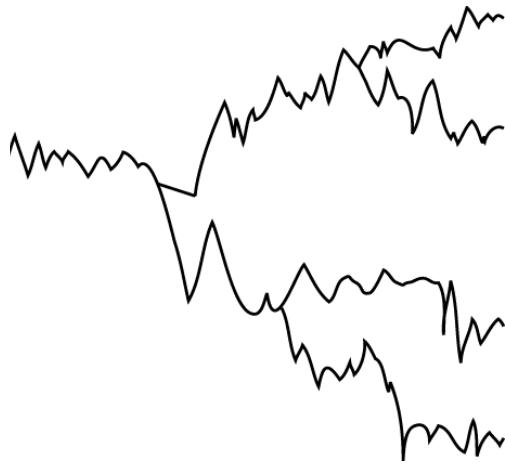
L'évolution phénotypique peut être étudiée à partir des traits des espèces actuelles et d'une phylogénie datée de ces espèces

Ici l'arbre phylogénétique est fixé, il n'est pas modélisé

BUT

Développer des **modèles d'évolution phénotypique** ET des approches d'**inférence statistique** qui permettent d'ajuster les modèles à des **données phénotypiques actuelles** de façon à **étudier les modes et vitesse d'évolution des phénotypes**

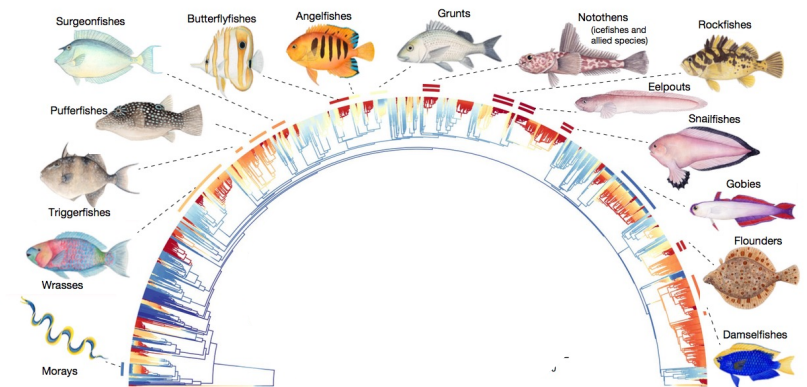
Modèle d'évolution
phénotypique



Inférence
Statistique



Données phénotypiques
actuelles



Mode et vitesse d'évolution
des phénotypes

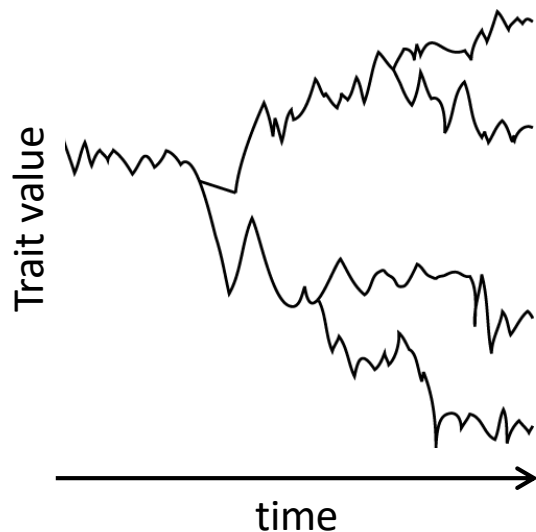
Le mouvement Brownien pour étudier l'évolution phénotypique

Evolution phénotypique pendant la vie d'une espèce



$$dX(t) = \sigma dB(t)$$

Evolution phénotypique pendant la diversification d'un groupe d'espèces

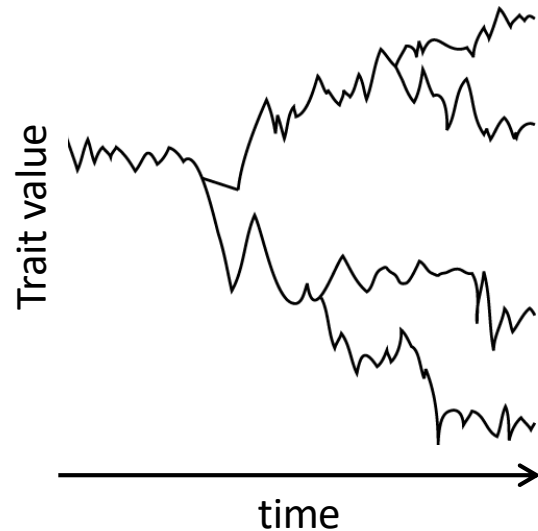


- Mouvement Brownien le long d'une lignée
- Evolution indépendante après la spéciation

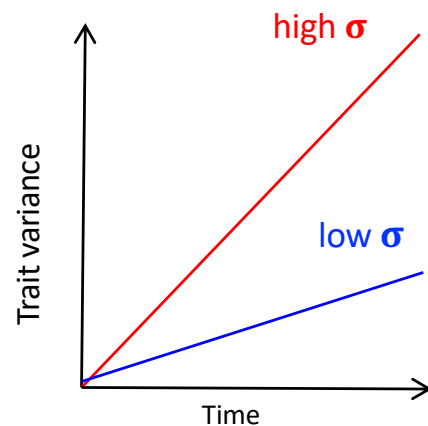


Ici une lignée est une espèce

Le mouvement Brownien pour étudier l'évolution phénotypique



- Mouvement Brownien le long d'une lignée
- Evolution indépendante après la spéciation



La variance phénotypique entre espèces augmente linéairement avec le temps depuis leur divergence

$$V = \sigma^2 t$$

σ comme mesure de la vitesse d'évolution phénotypique

Likelihood associated to the Brownian process

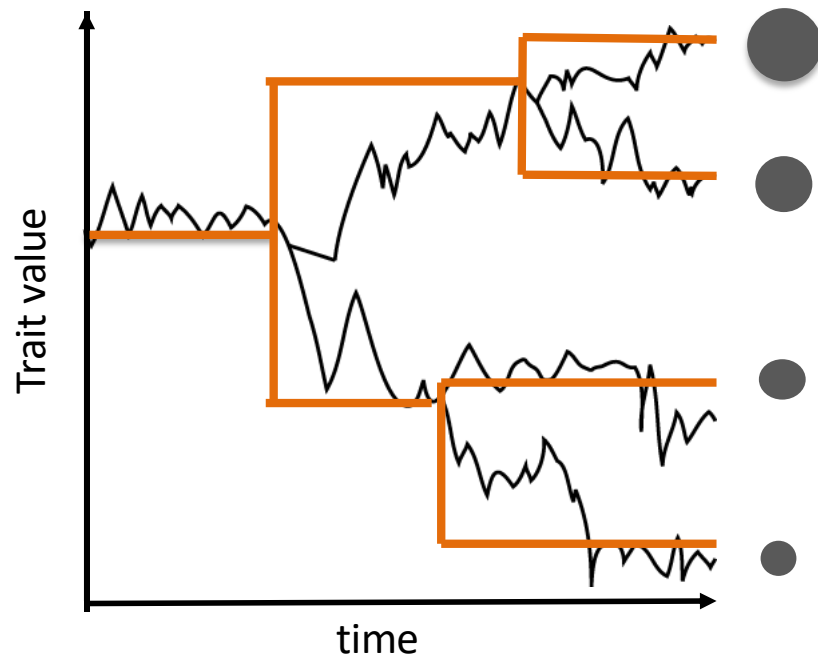
Let X be a quantitative trait assumed to evolve as a Brownian process with rate σ on a fixed, ultrametric, bifurcating species tree with branch lengths in unit of time, assumed to be known.

$$dX(t) = \sigma dB(t)$$

shared evolutionary history between species i and j

expected variance-covariance matrix

$$V_{ij} = \sigma^2 S_{ij}$$

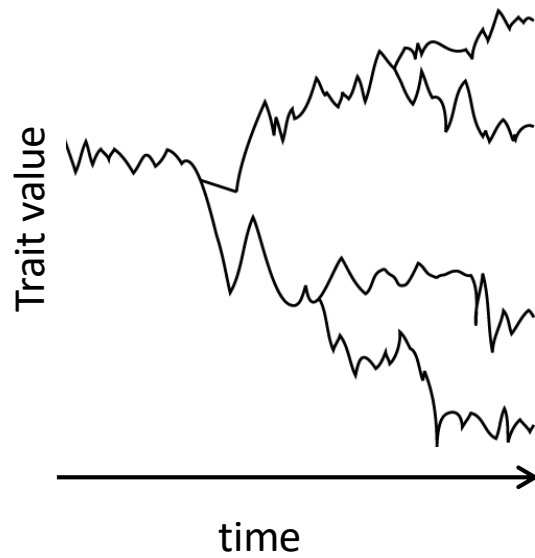


$$L = \frac{\exp \left\{ -\frac{1}{2} [\mathbf{X} - \mathbf{E}(\mathbf{X})]' (\mathbf{V}^{-1}) [\mathbf{X} - \mathbf{E}(\mathbf{X})] \right\}}{\sqrt{(2\pi)^{N*} \det(\mathbf{V})}}$$

Felsenstein 1973

→ σ estimated by maximum likelihood or Bayesian inference

Les traits évoluent-ils comme un mouvement Brownien?



Le processus d'Orstein-Uhlenbeck

$$dX(t) = \alpha[\theta - X(t)]dt + \sigma dB(t)$$

where

- α = strength of "selection"
- θ = adaptive optimum

Selection component (pointed to by a red arrow from the first term)

Drift (BM) component (pointed to by a blue arrow from the second term)

$$L = \frac{\exp\left\{-\frac{1}{2}[\mathbf{X} - \mathbf{E}(\mathbf{X})]'(\mathbf{V}^{-1})[\mathbf{X} - \mathbf{E}(\mathbf{X})]\right\}}{\sqrt{(2\pi)^N * \det(\mathbf{V})}}$$

expected covariance between species i and j

total height of the tree

shared evolutionary history between species i and j

$$V_{ij} = \frac{\sigma^2}{2\alpha} e^{-2\alpha(T-s_{ij})} (1 - e^{-2\alpha s_{ij}})$$

Models of phenotypic evolution with time-dependent rates

$dX(t)$: change in $X(t)$ in a small amount of time

The « Early Burst » model: trait evolution is fast in the initial phase of an evolutionary radiation, and slows down later on as an effect of limited ecological opportunity

$$dX(t) = \sigma(t) dB(t) \quad \text{with } \sigma^2(t) = \sigma_0^2 e^{rt} \quad \text{where } r \text{ is the rate of decline}$$

Blomberg *et al.* Am Nat 2003

The « Diversity-dependent » model: trait evolution slows down as species pile up in a clade, therefore reducing ecological opportunity

$$dX(t) = \sigma(t) dB(t) \quad \text{with } \sigma^2(t) = \sigma_0^2 e^{bn(t)} \quad \text{where } n(t) \text{ is the number of lineages in the clade at time } t$$

Weir & Mursleen *Evolution* 2013

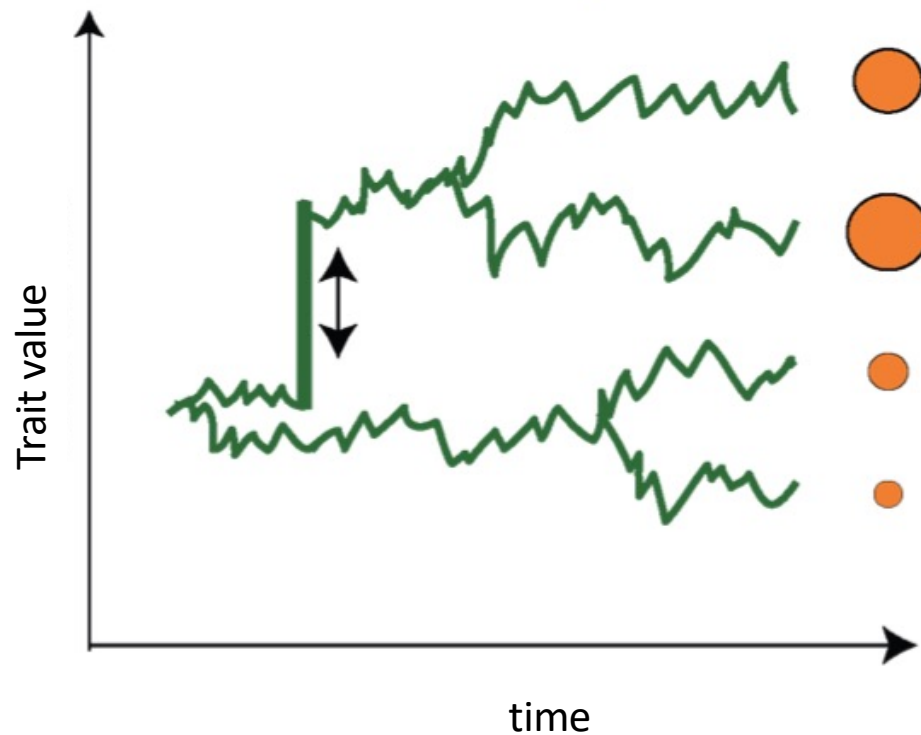
The « Environment-dependent » model: the rate of trait evolution depends on past environmental variables

$$dX(t) = \sigma(t) dB(t) \quad \text{with } \sigma^2(t) = f(t, E_1(t), E_2(t), \dots, E_k(t))$$

Clavel & Morlon *PNAS* 2017

Models of phenotypic evolution with jumps (punctuated equilibrium)

Levy process

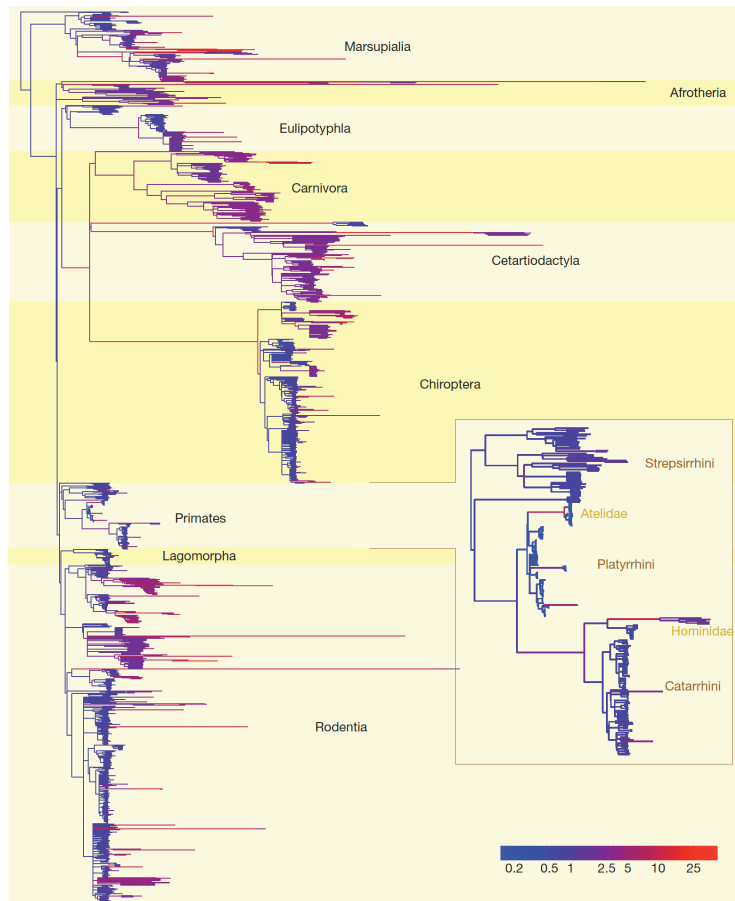


Bokma 2008, 2009

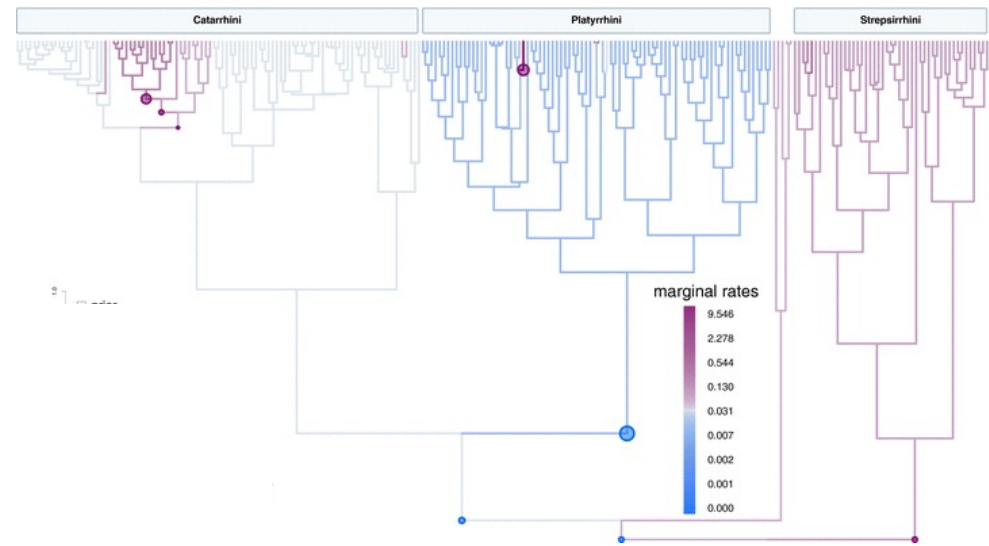
Landis et al. 2013

Duchen et al. 2017

Models of phenotypic evolution with rates that vary across lineages



Venditti *et al.* Nature 2011



Eastman *et al.* Evolution 2011

Models of phenotypic evolution with rates that vary according to interspecific interactions

Syst. Biol. 65(4):700–710, 2016
© The Author(s) 2016. Published by Oxford University Press, on behalf of the Society of Systematic Biologists. All rights reserved.
For Permissions, please email: journals.permissions@oup.com
DOI:10.1093/sysbio/syw020
Advance Access publication March 9, 2016



Jonathan Drury

Estimating the Effect of Competition on Trait Evolution Using Maximum Likelihood Inference

JONATHAN DRURY*, JULIEN CLAVEL, MARC MANCEAU, AND HÉLÈNE MORLON

average trait value in the clade at time t

$$dX(t) = S(\bar{X}_t - X_t)dt + \sigma dB(t)$$

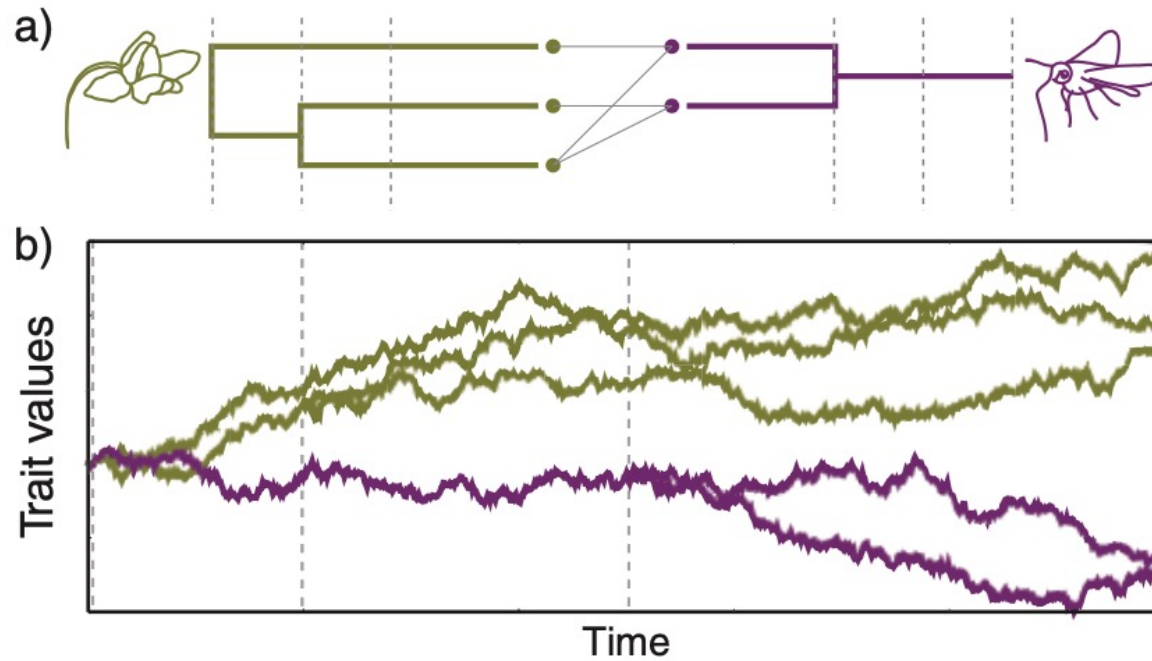
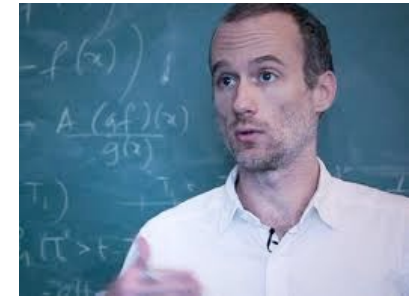


strength of interaction

$S < 0$: matching competition (beak shape in competition for food)

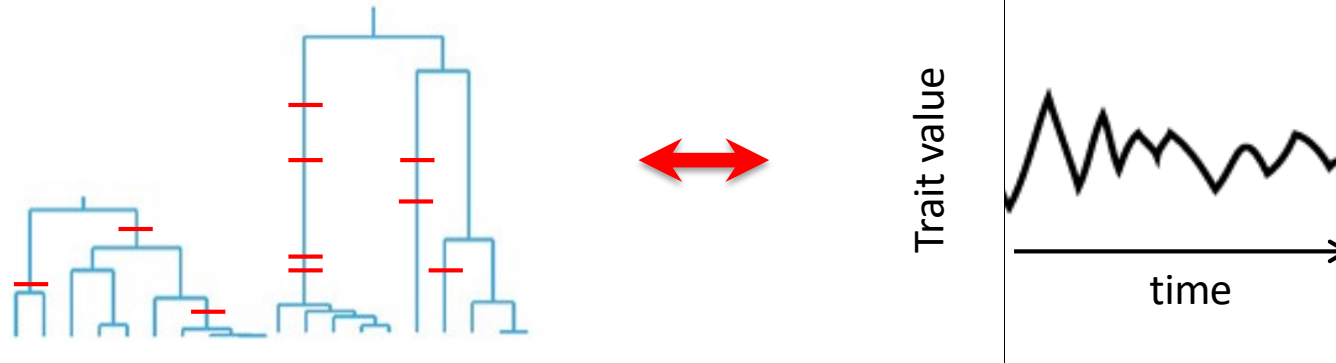
$S > 0$: matching mutualism (Mullerian mimicry)

Models of phenotypic evolution with rates that vary according to interspecific interactions



What are the microevolutionary processes that modulate phenotypic evolution on macroevolutionary timescales?

A microscopic model of phenotypic evolution with macroevolutionary predictions



Lande (*Evolution* 1976)

**Hypothesis 1: trait distribution is Gaussian
with constant variance**

$$p(z, t) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{[z - \bar{z}(t)]^2}{2\sigma^2}\right\}$$

Hypothesis 2: fitness function is Gaussian

$$W(z) = \exp\left\{-\frac{(z - \vartheta)^2}{2w^2}\right\}$$

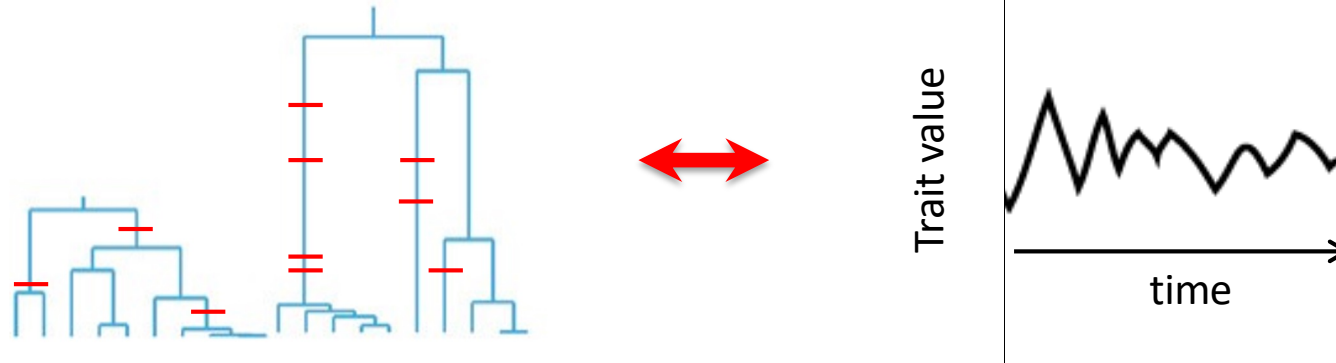
Hypothesis 3: constant population size N

$$d\bar{z}(t) = A[\theta - \bar{z}(t)]dt + \Sigma dB(t)$$

$$A = \frac{\sigma^2}{w^2 + \sigma^2}$$

$$\Sigma = \frac{\sigma^2}{N}$$

A microscopic model of phenotypic evolution with macroevolutionary predictions



In the framework of adaptive dynamics?

The Annals of Applied Probability
2007, Vol. 17, No. 1, 102–155
DOI: 10.1214/105051606000000628
© Institute of Mathematical Statistics, 2007

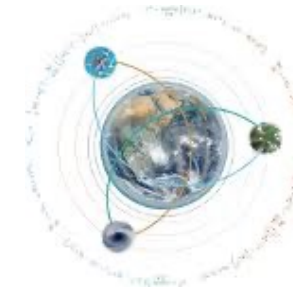
EVOLUTION OF DISCRETE POPULATIONS AND THE CANONICAL DIFFUSION OF ADAPTIVE DYNAMICS¹

BY NICOLAS CHAMPAGNAT AND AMAURY LAMBERT

Birth-death model + rare mutations + artificially avoid extinction

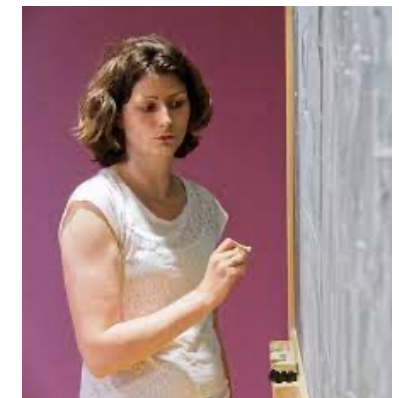
→ Diffusion equation for the trait

Phenotypic evolution in a Moran model



Thuy Vo

- N individuals following a Moran process
- individuals characterized by quantitative trait x
- at rate $c(x,y)$, an individual of trait x dies and is replaced by an individual of trait y



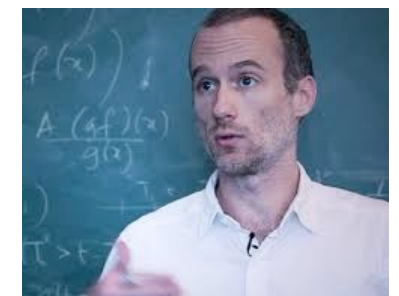
Hélène Leman

- quantitative trait z results from genetic architecture

$$z^i = \frac{1}{L} \sum_{l=1}^L w_l \eta_l^i$$

number of loci \rightarrow
 \uparrow genetic effect
allelic state \downarrow

- mutations occur at rate $\gamma\mu$ with γ between 0 and 1 ; when they occur, each nucleotide mutates with proba α



Amaury Lambert

Phenotypic evolution in a Moran model

First result Accelerate time $t \mapsto t/\gamma$
Take the limit of rare mutations $\gamma \rightarrow 0$

**Trait evolves as jump process
with a rate that depends on N and c**

Second result Accelerate time
Take the limit of small mutational effect $L \rightarrow \infty$

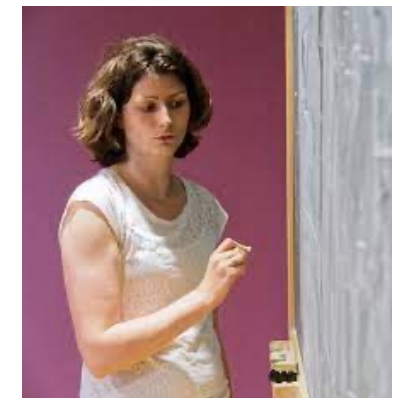
Canonical diffusion equation

$$dZ_t = \frac{1}{2} N^2 \sigma^2 \mu \partial_2 f_N(Z_t, Z_t) dt + \sigma(Z_t) \sqrt{N \mu c_N(Z_t, Z_t)} dB_t$$

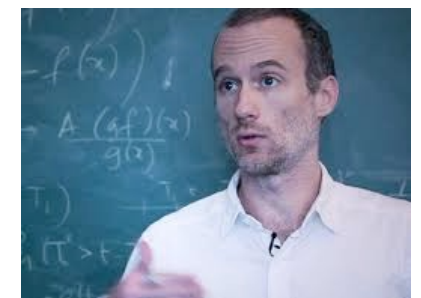
with $f_N(x, y) := c_N(y, x) - c_N(x, y)$. the fitness



Thuy Vo



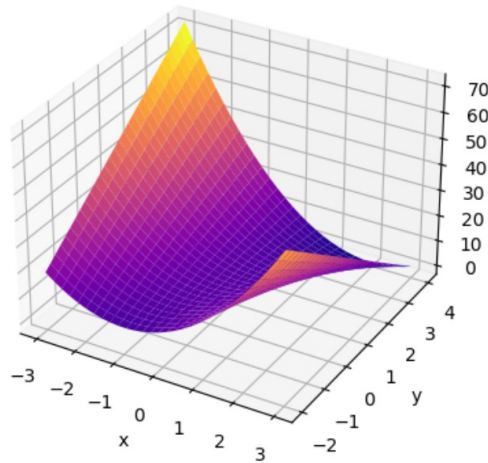
Hélène Leman



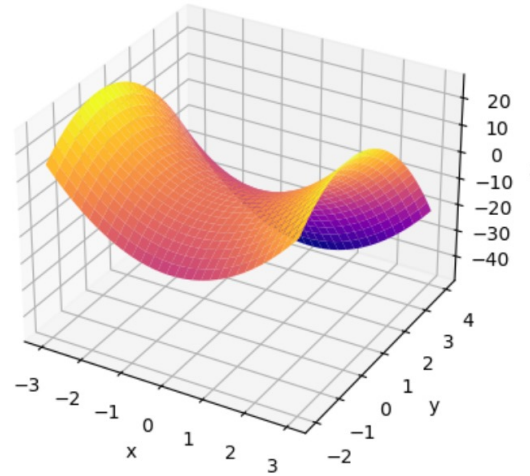
Amaury Lambert

Phenotypic evolution in a Moran model

Case that leads to the Ornstein-Uhlenbeck



$$c_N(x, y) = \frac{1}{N} (b(x - a)(x - y) + k_1) + k_2$$



$$f_N(x, y)$$

$$d\tilde{Z}_u = \underbrace{\sigma^2 \mu b(a - \tilde{Z}_u)}_{\text{strength of selection}} du + \underbrace{\sigma \sqrt{\mu \left(\frac{k_1}{N} + k_2 \right)}}_{\text{drift}} dB_u.$$

strength of selection

drift

→ Can be estimated by fit of the OU model to phylogenetic comparative data



Thuy Vo

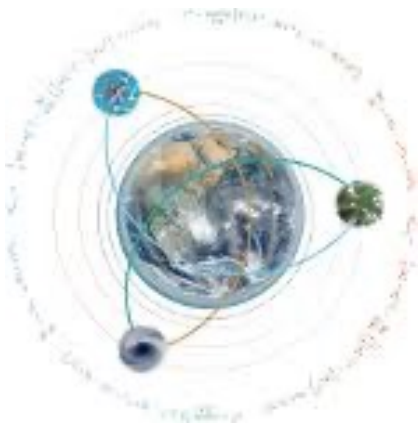


Hélène Leman



Amaury Lambert

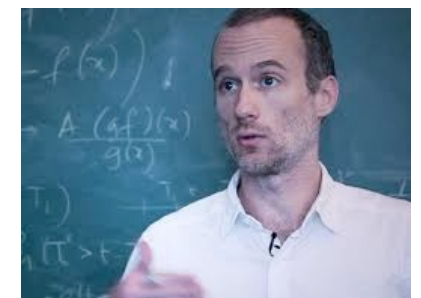
Post-doctoral position open!



Thuy Vo



Hélène Leman



Amaury Lambert



Fabien Condamine



Olivier Billaud



Joelle Barido-Sottani

THANKS!



Nathan Mazet



Ignacio Quintero



Jonathan Rolland



Dan Moen



Jérémy
Andréoletti



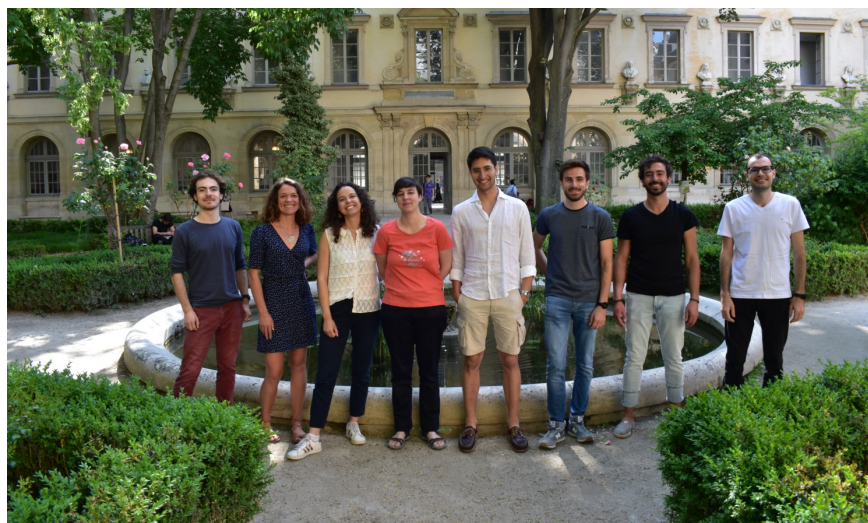
Sophia Lambert



Jakub Voznica



Odile Maliet



What are the factors that modulate rates of species diversification?

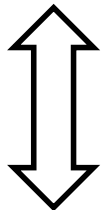


The Court Jester

Abiotic factors

climatic variation
geological context

EXTRINSIC



The Red Queen

Biotic factors

competition
mutualistic and antagonistic interactions

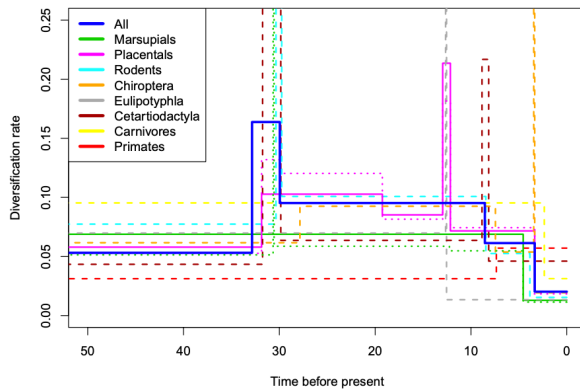
INTRINSIC

Species-specific traits

reproduction mode
life-history traits
dispersal capacity
developmental constraints

Phylogenetic birth-death models allow testing if and how diversification rates vary:

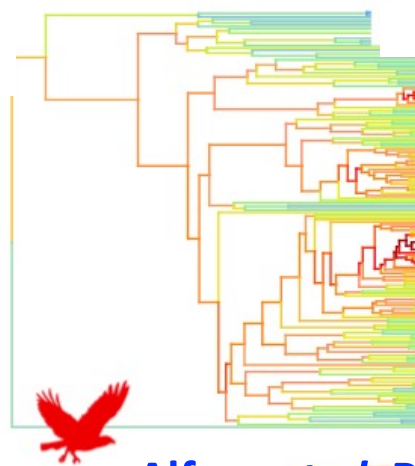
Through time



Stadler *et al.* PNAS 2011,
Morlon *et al.* PNAS 2011,
May *et al.* MEE 2016, etc...

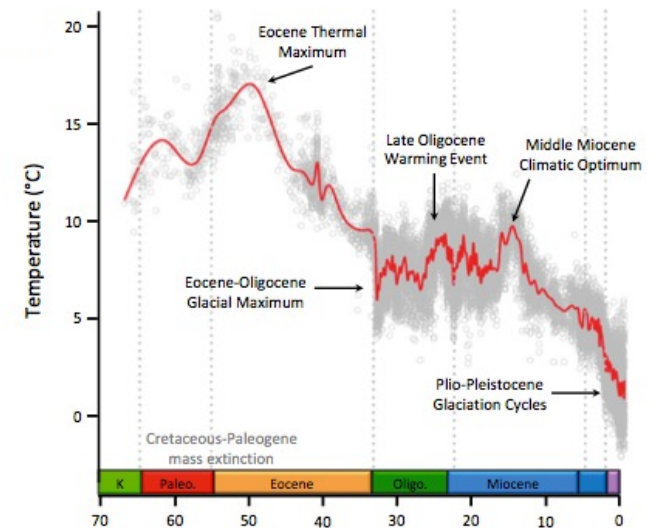
Across geographic regions

Across lineages

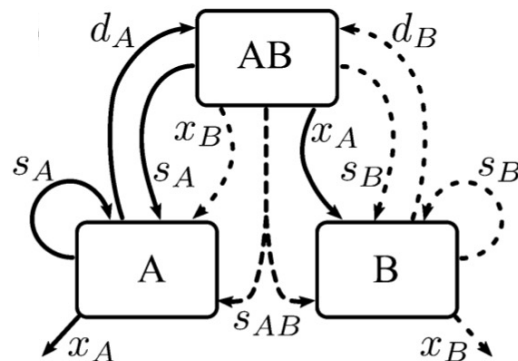


Alfaro *et al.* PNAS 2009,
Morlon *et al.* PNAS 2011
Rabosky *et al.* PloS One 2014,
Maliet *et al.* NEE 2019,
Barido-Sottani *et al.* Syst Bio 2020, etc...

With past environmental changes



Condamine *et al.* Eco Lett 2013,
Lewitus *et al.* Syst Bio 2017



Goldberg *et al.* Syst Bio 2011,
Quintero *et al.* in press,
etc...

**REVIEW AND
SYNTHESIS**

Phylogenetic approaches for studying diversification

Hélène Morlon*

Abstract

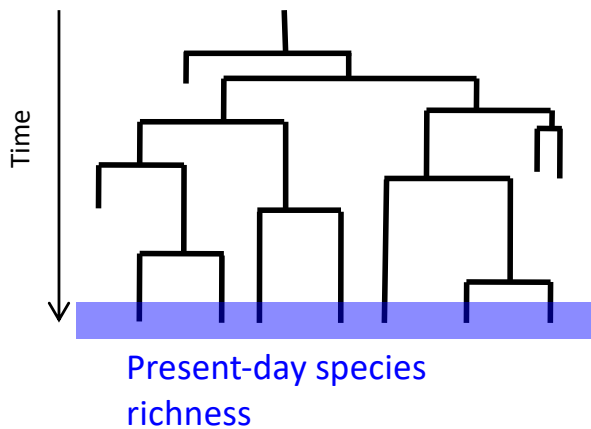
Estimating rates of speciation and extinction, and understanding how and why they vary over evolutionary time, geographical space and species groups, is a key to understanding how ecologi-

Conclusions

Stochastic models are powerful for understanding modes and rates of diversification and trait evolution, i.e. how biodiversity evolves

This is fundamental for understanding current patterns of species and phenotypic diversity

Current levels of species richness result from the balance of speciation and extinction events



Speciation rate
average number of
speciation events per
Myr per lineage

Extinction rate
average number of
extinction events per
Myr per lineage

Net diversification rate
speciation rate – extinction rate

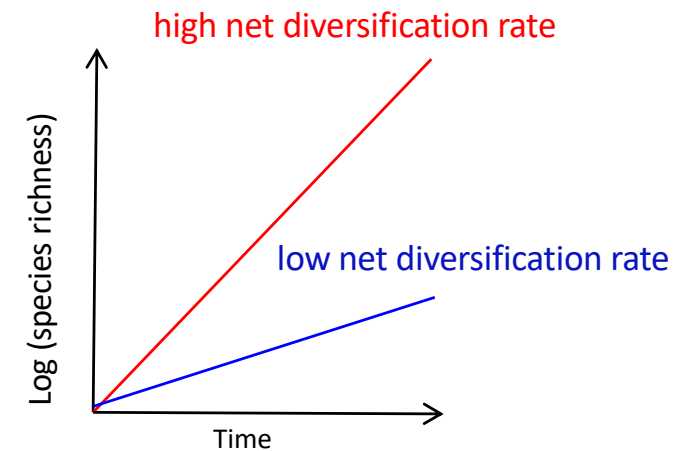


Illustration: the diversification of cetaceans

We can test the statistical support of models where diversification rates “shift” at the base of major families

Support for a 4-shift rate model in the cetacean phylogeny

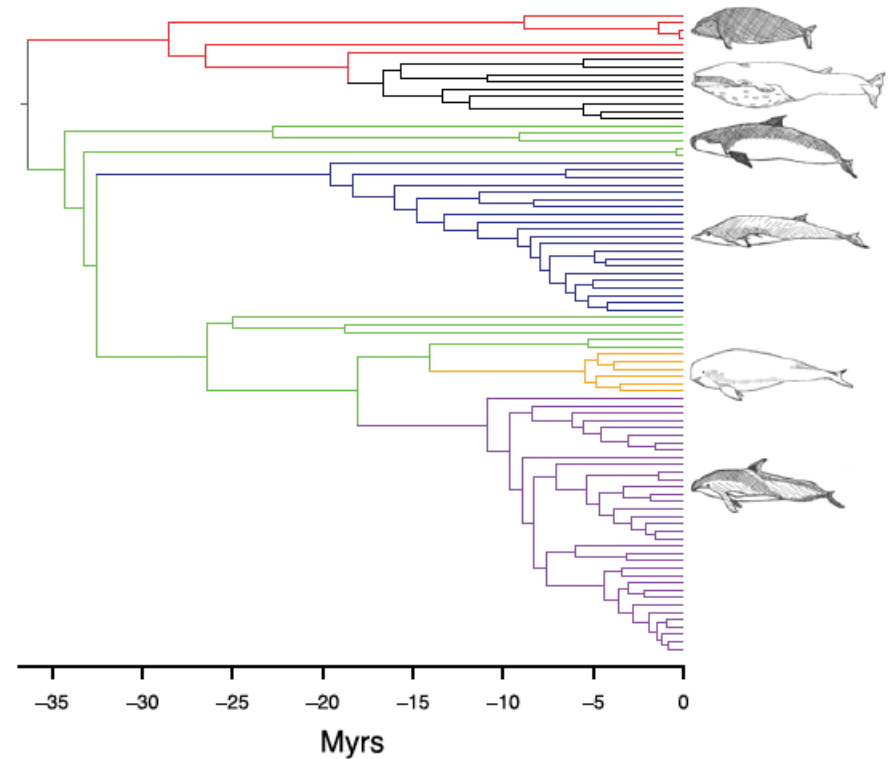
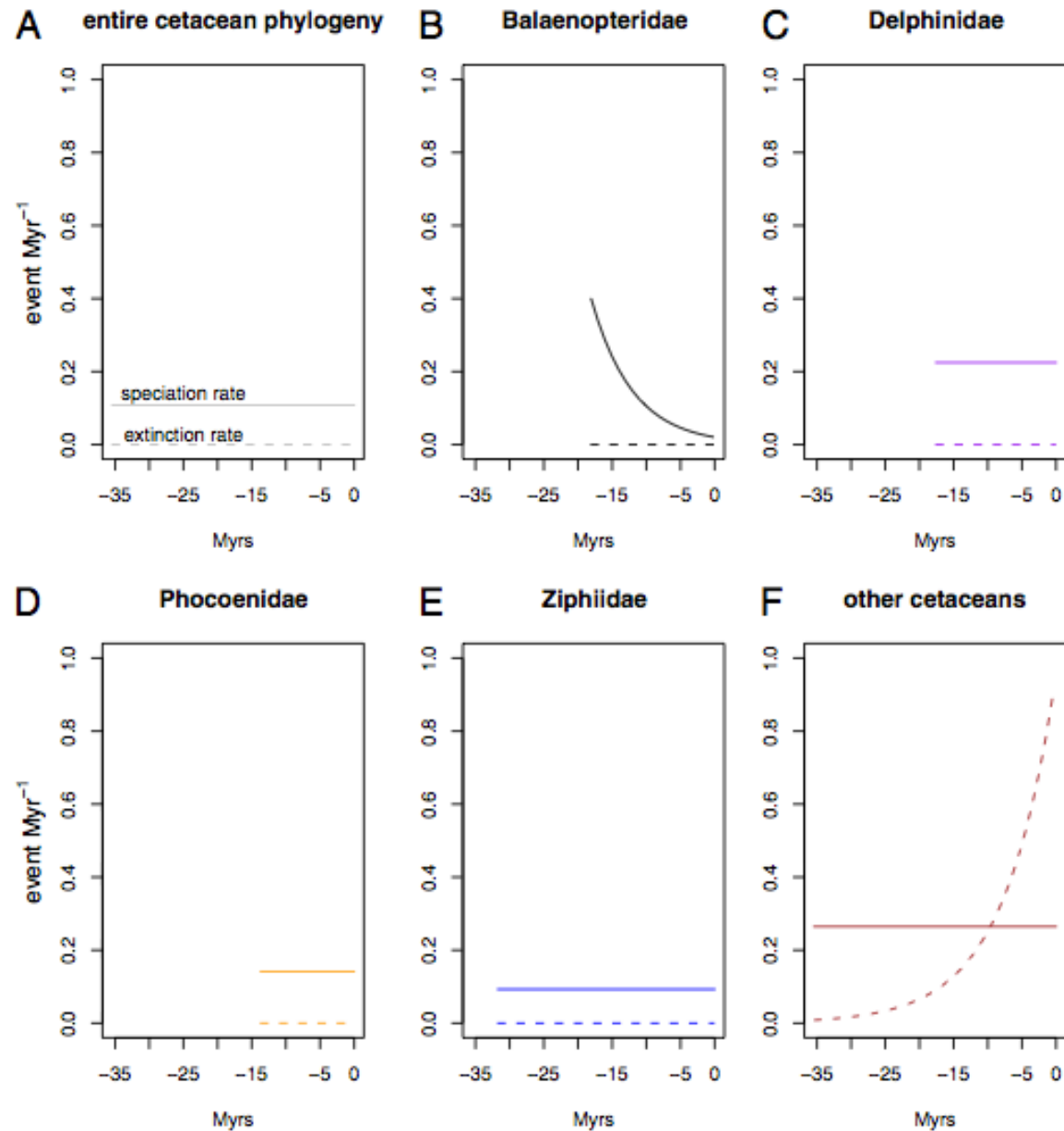


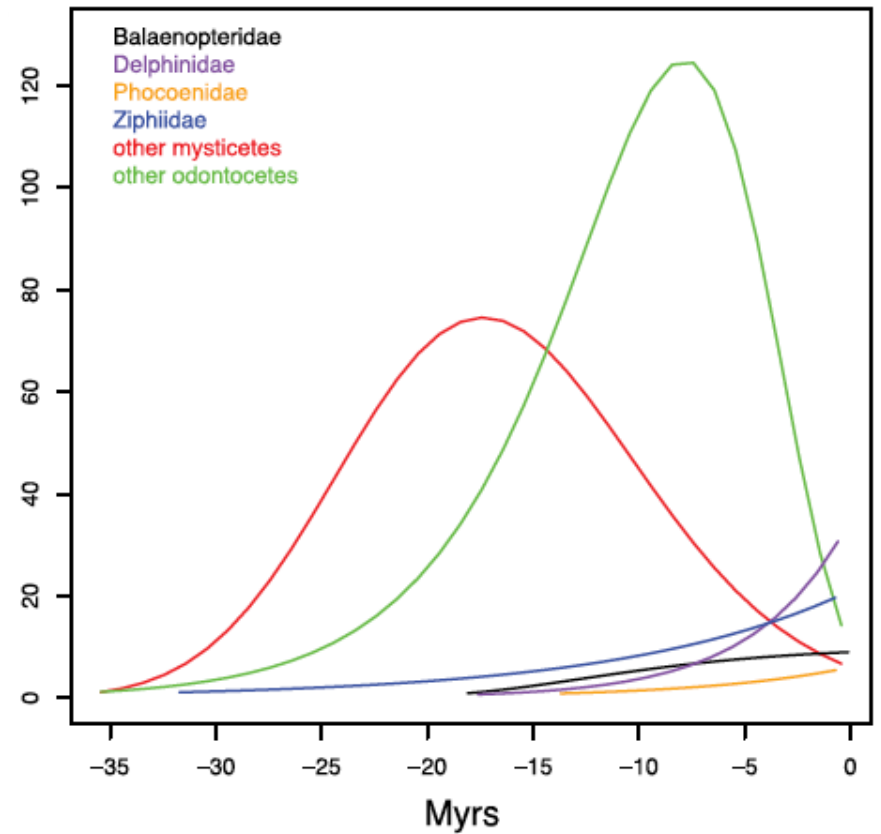
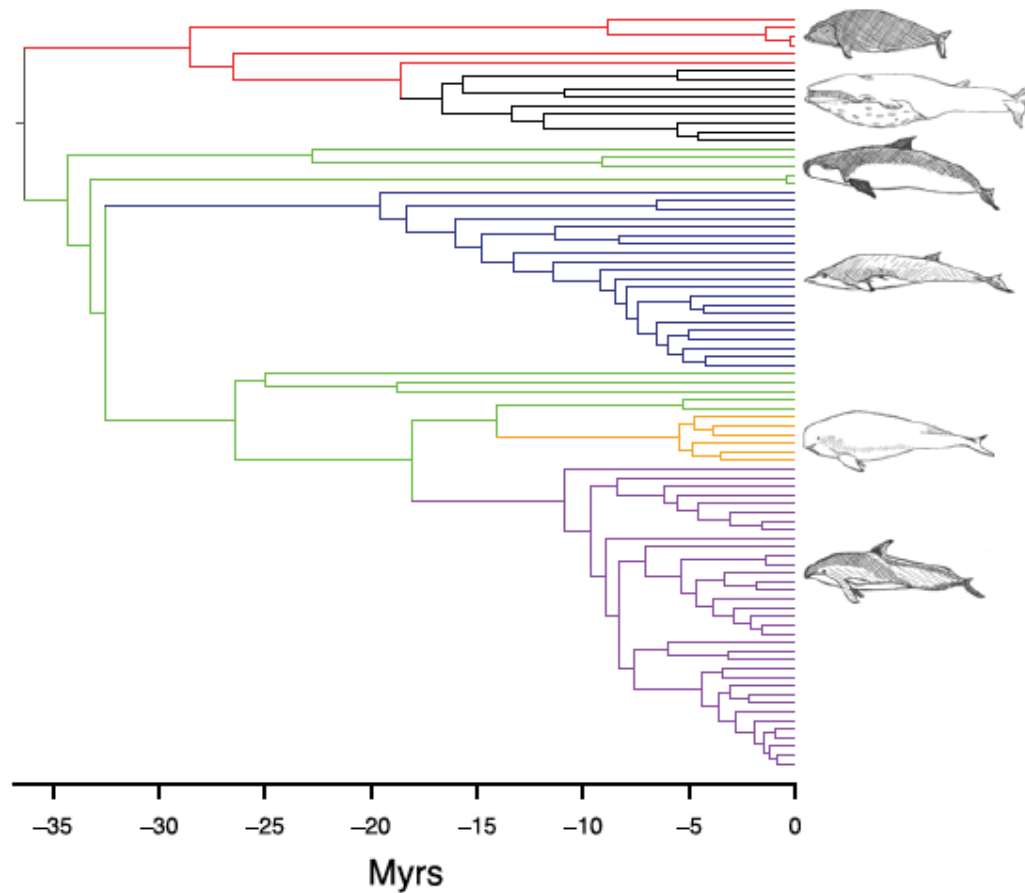
Table S2. Statistical support for rate shifts in the cetacean phylogeny

Model	nb	Description	LogL	AIC _c
No shift	1	Best fit model	-279.03	560.08
One shift	5	Best fit model: shift in the Delphinidae	-262.93*	536.22
Two shifts	6	Best fit model: shifts in the Delphinidae and Phocoenidae	-260.17 [†]	532.85
Three shifts	7	Best fit model: shifts in the Delphinidae, Phocoenidae and Ziphiidae	-256.13 [‡]	526.94
Four shifts	8	Best fit model: shifts in the Delphinidae, Phocoenidae, Ziphiidae, and Balaenopteridae	-250.13	517.14

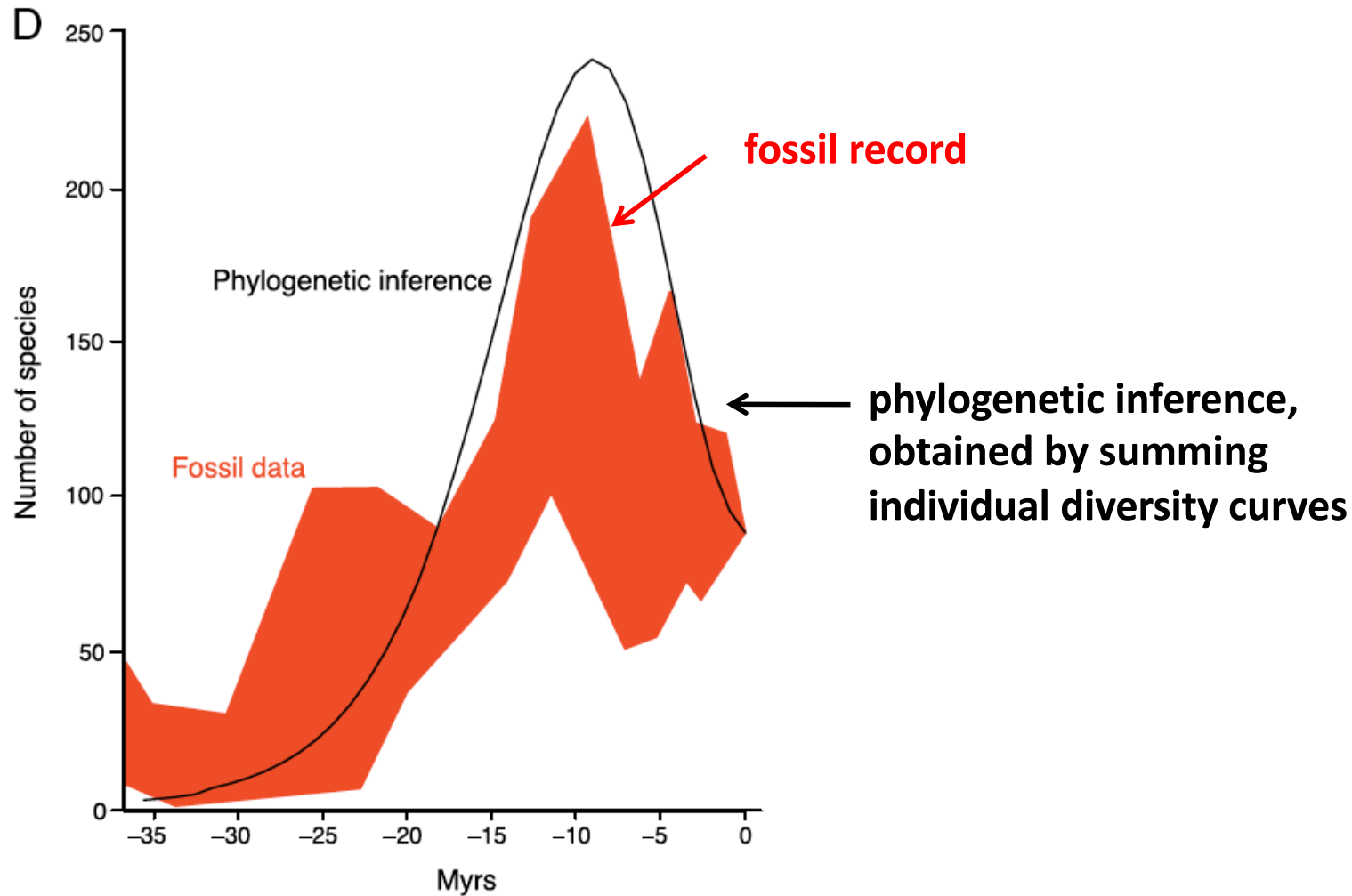
We can compute estimates of speciation and extinction rates through time



Resulting estimates of diversity through time : boom-then-bust diversity dynamics



The resulting diversity curve is consistent with the fossil record

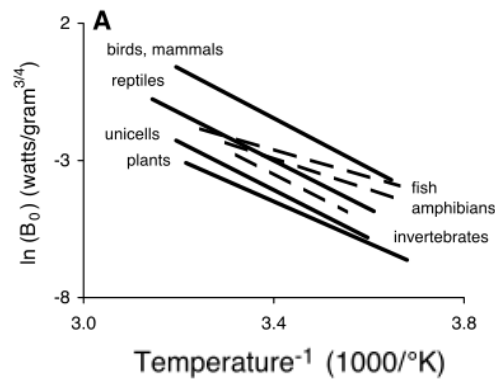


Temperature affects major aspects of biology through its effect on metabolic rates, body-size, and productivity

Kleiber's law

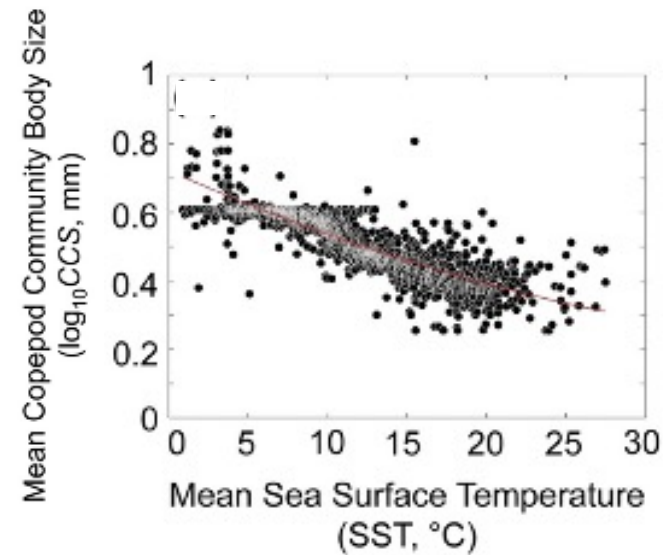
$$B \sim M^{3/4} e^{-E_a/kT}$$

metabolic rate \rightarrow B
 $M^{3/4}$ \rightarrow body-size
 $e^{-E_a/kT}$ \rightarrow activation energy Boltzmann's constant



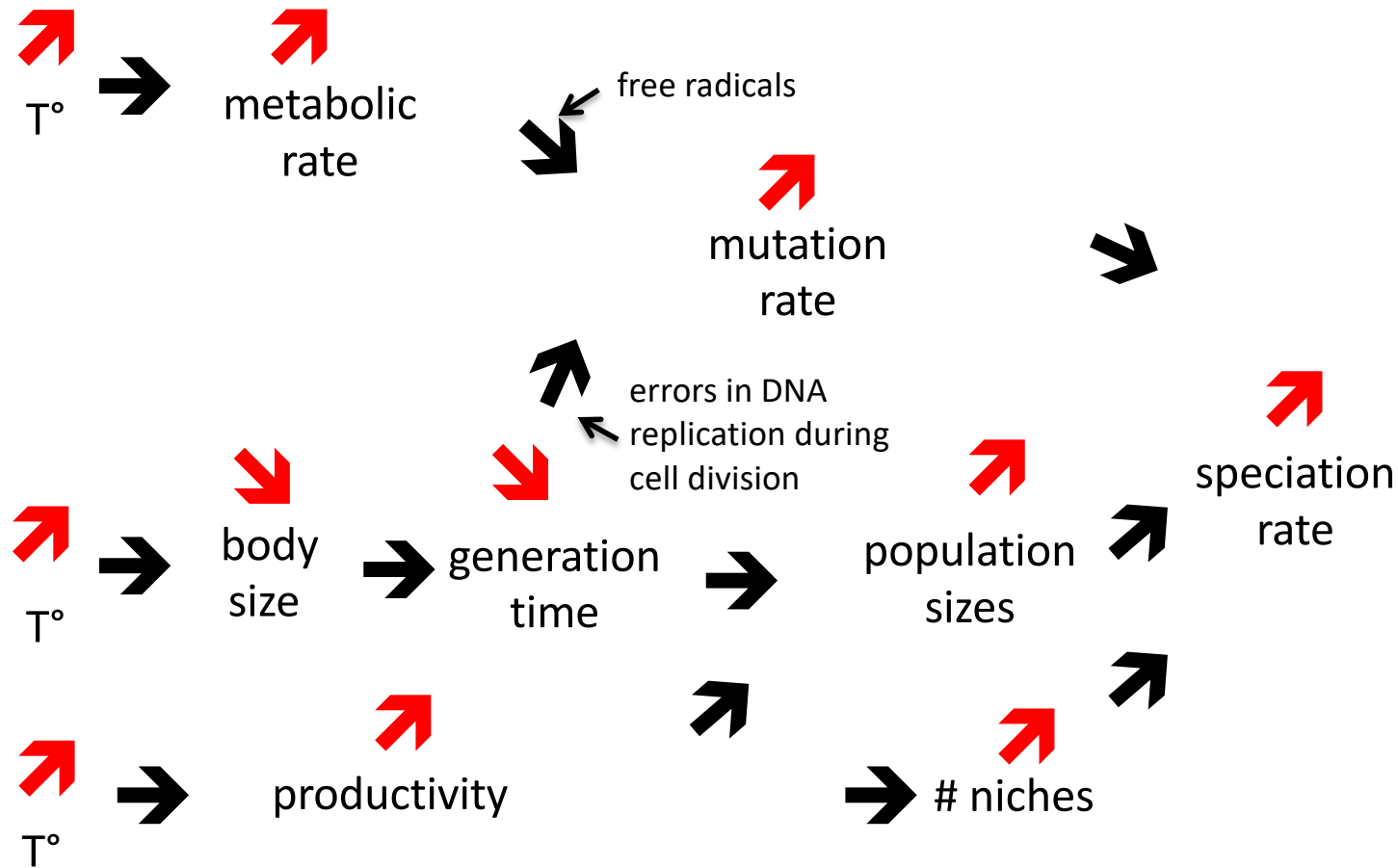
Gillooly et al. Science 2001

Bergmann's rule



Evans et al. Ecography 2019

Based on the metabolic, body-size and productivity hypotheses, temperature should (positively) affect speciation rates



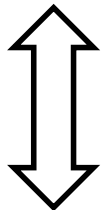
What are the factors that modulate speciation and extinction rates?



The Court Jester

Abiotic factors
climatic variation
geological context

EXTRINSIC



The Red Queen

Biotic factors
competition
mutualistic and antagonistic interactions

INTRINSIC

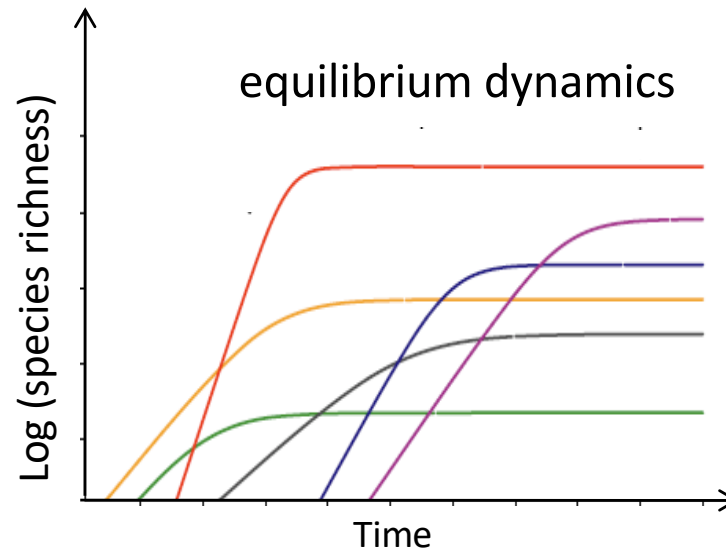


Species-specific traits
reproduction mode
life-history traits
dispersal capacity



Biotic factors
competition

Diversity-dependent birth-death models



PROCEEDINGS
OF
THE ROYAL
SOCIETY **B**

Proc. R. Soc. B (2012) 279, 1300–1309
doi:10.1098/rspb.2011.1439
Published online 12 October 2011

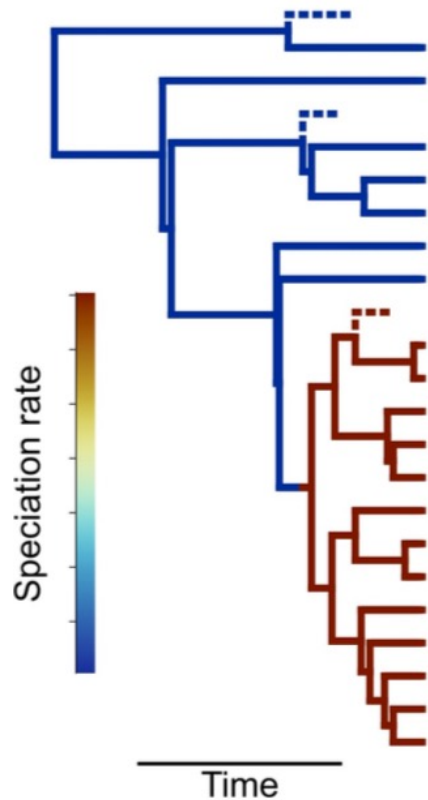
Diversity-dependence brings molecular phylogenies closer to agreement with the fossil record

**Rampal S. Etienne^{1,*}, Bart Haegeman², Tanja Stadler³, Tracy Aze⁴,
Paul N. Pearson⁴, Andy Purvis⁵ and Albert B. Phillimore⁵**

¹*Community and Conservation Ecology, Centre for Ecological and Evolutionary Studies,*

Les modèles de naissance-mort, combinés à une approche d'inférence statistique, permettent d'estimer les taux de spéciation et d'extinction et les variations passées de la biodiversité

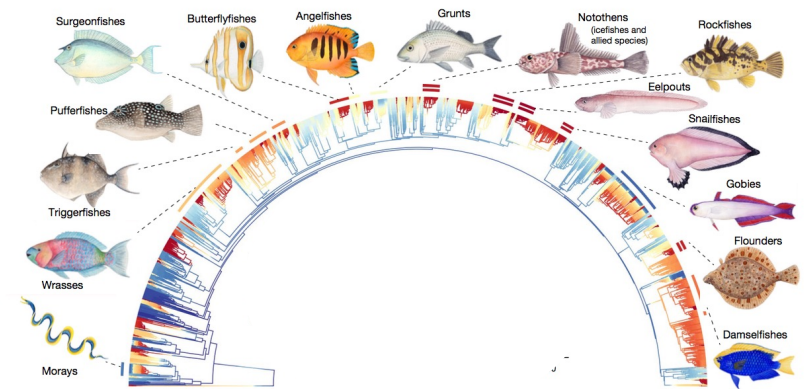
Modèle de diversification



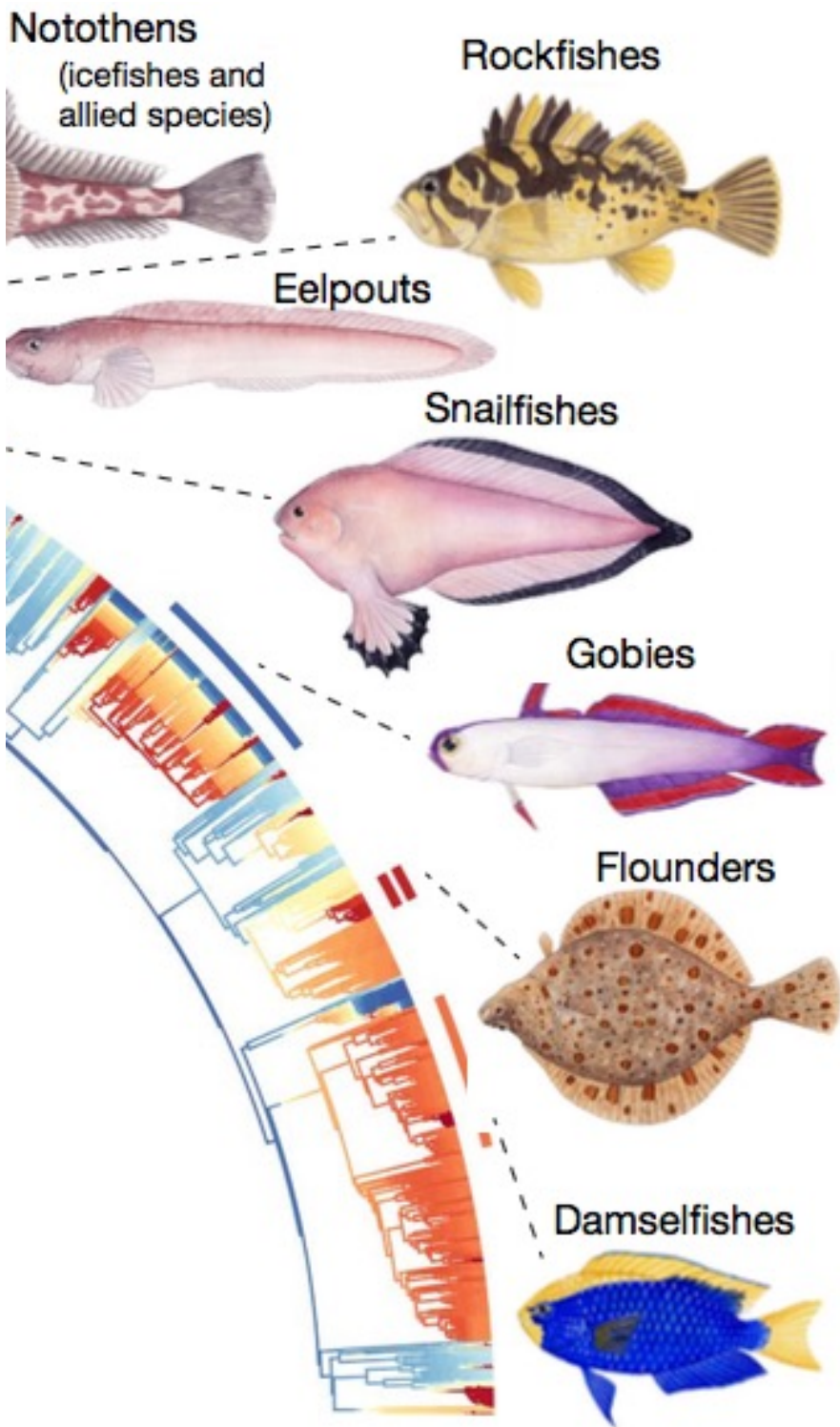
Inférence Statistique



Arbre phylogénétique



Estimation des taux de spéciation et extinction, et de la diversité passée



Dans ces approches, nous considérons l'arbre phylogénétique (daté) comme une donnée

L'arbre est en fait construit à partir d'un modèle d'évolution des séquences

Des les approches bayésiennes de reconstruction phylogénétique, le modèle de naissance-mort est utilisé comme un prior

Inférence bayésienne « complète »

Inférence
Statistique

Alignement de séquences

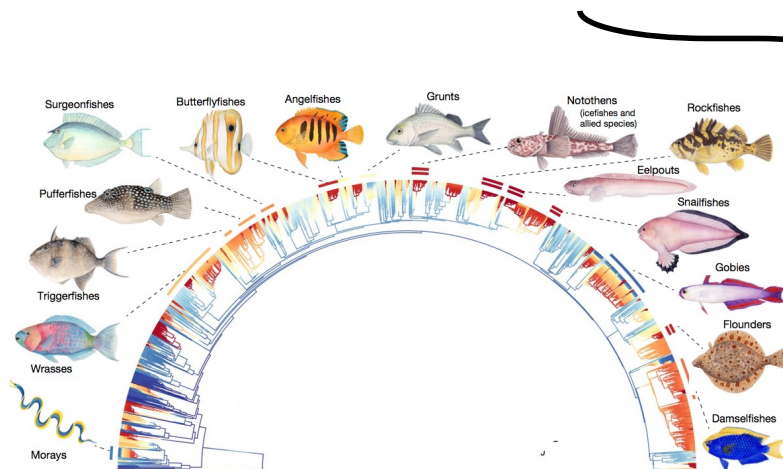
Modèle de diversification

+



Modèle d'évolution de séquence

T	-	G	C	G	C	A	A	T	A	T	T
T	T	G	C	G	G	C	T	A	C	A	A
C	A	G	G	T	A	C	A	C	C	T	C
C	G	G	G	G	A	C	T	G	C	T	T
C	T	G	G	G	-	C	G	G	T	T	A
C	A	G	T	T	A	C	G	A	T	A	G
C	C	G	C	A	A	C	A	C	C	A	C

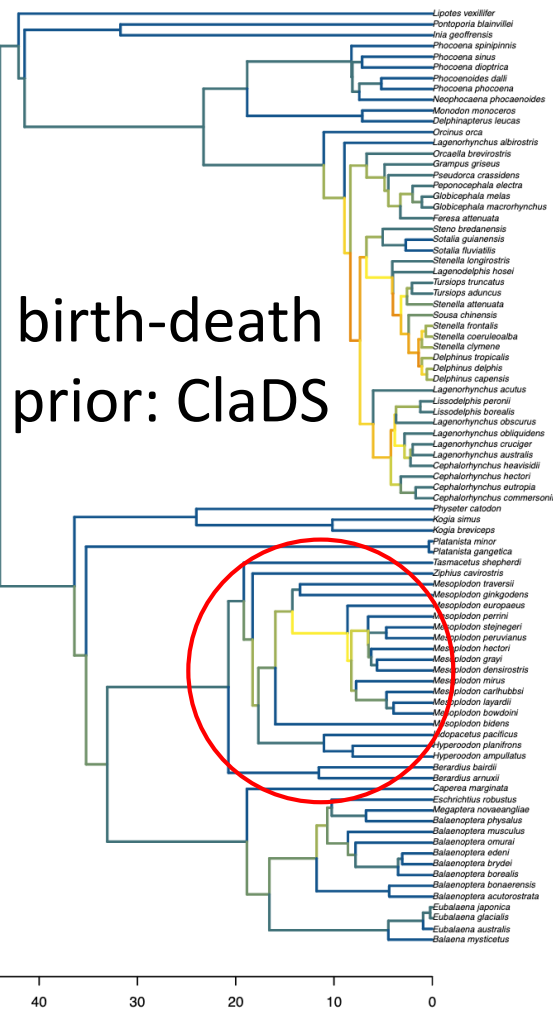


+

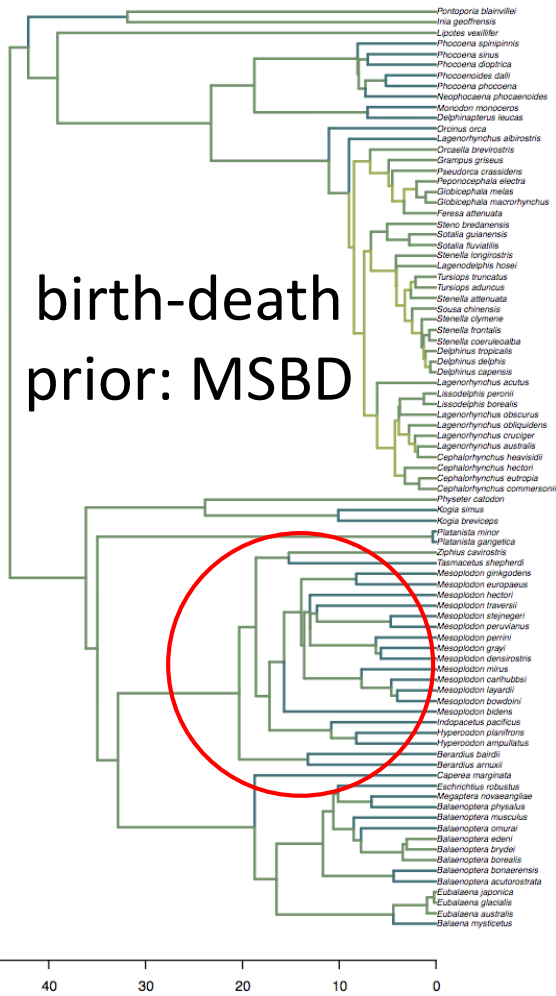
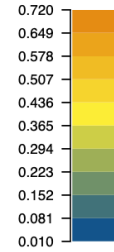
Estimation des taux de
spéciation et extinction



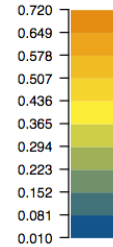
ClaDS now integrated in BEAST2 for full Bayesian phylogenetic inference



birth-death
prior: ClaDS



birth-death
prior: MSBD



Available for ClaDS₂ (constant turnover) and extinction also following a lognormal variation at speciation

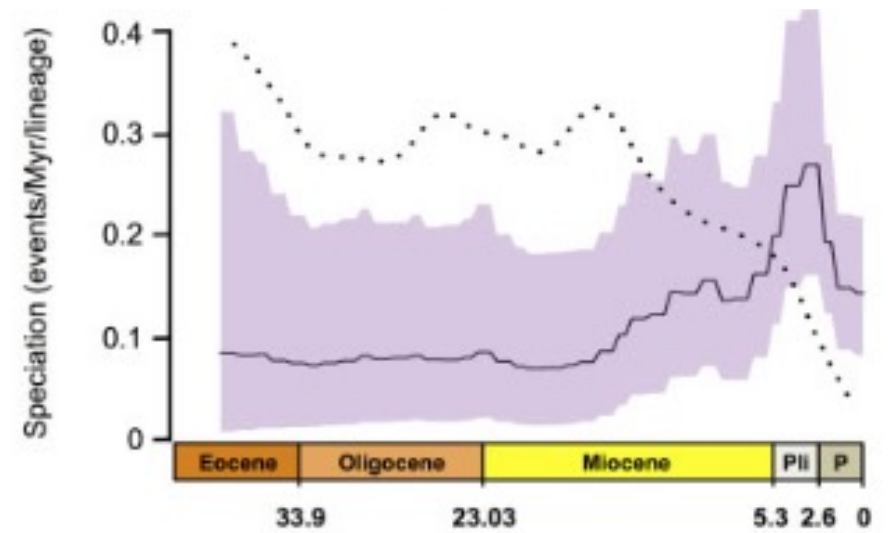
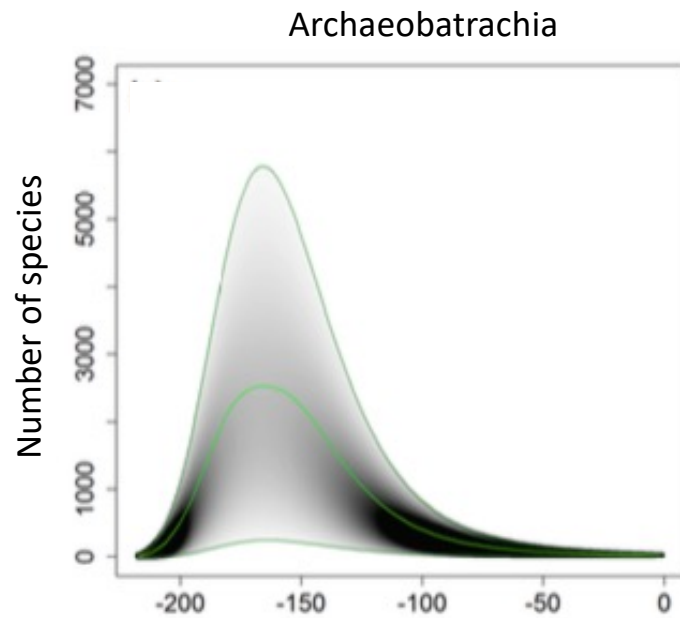


Beast2

Bayesian evolutionary analysis by sampling trees

Barido-Sottani & Morlon
Syst Bio (2023)

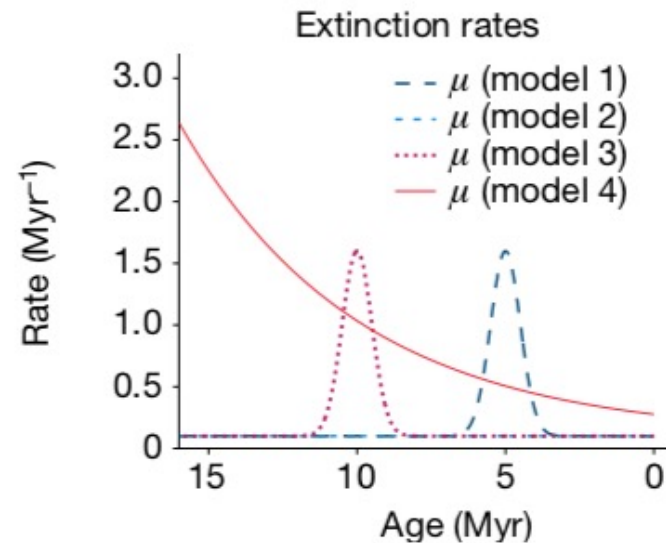
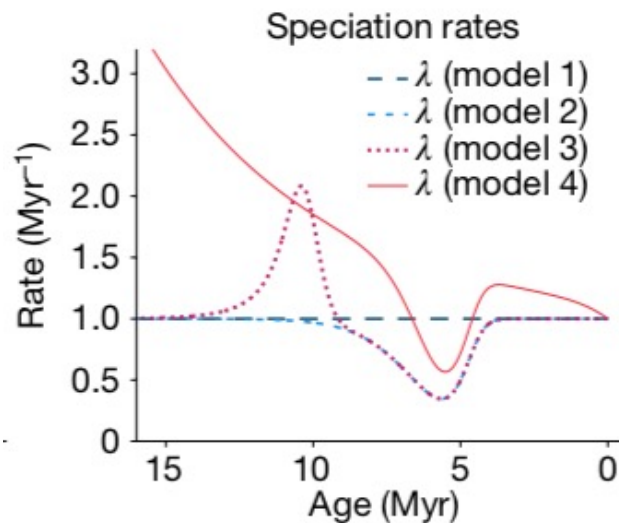
Quelle confiance dans ces inférences?



Problèmes d'identifiabilité?

Extant timetrees are consistent with a myriad of diversification histories

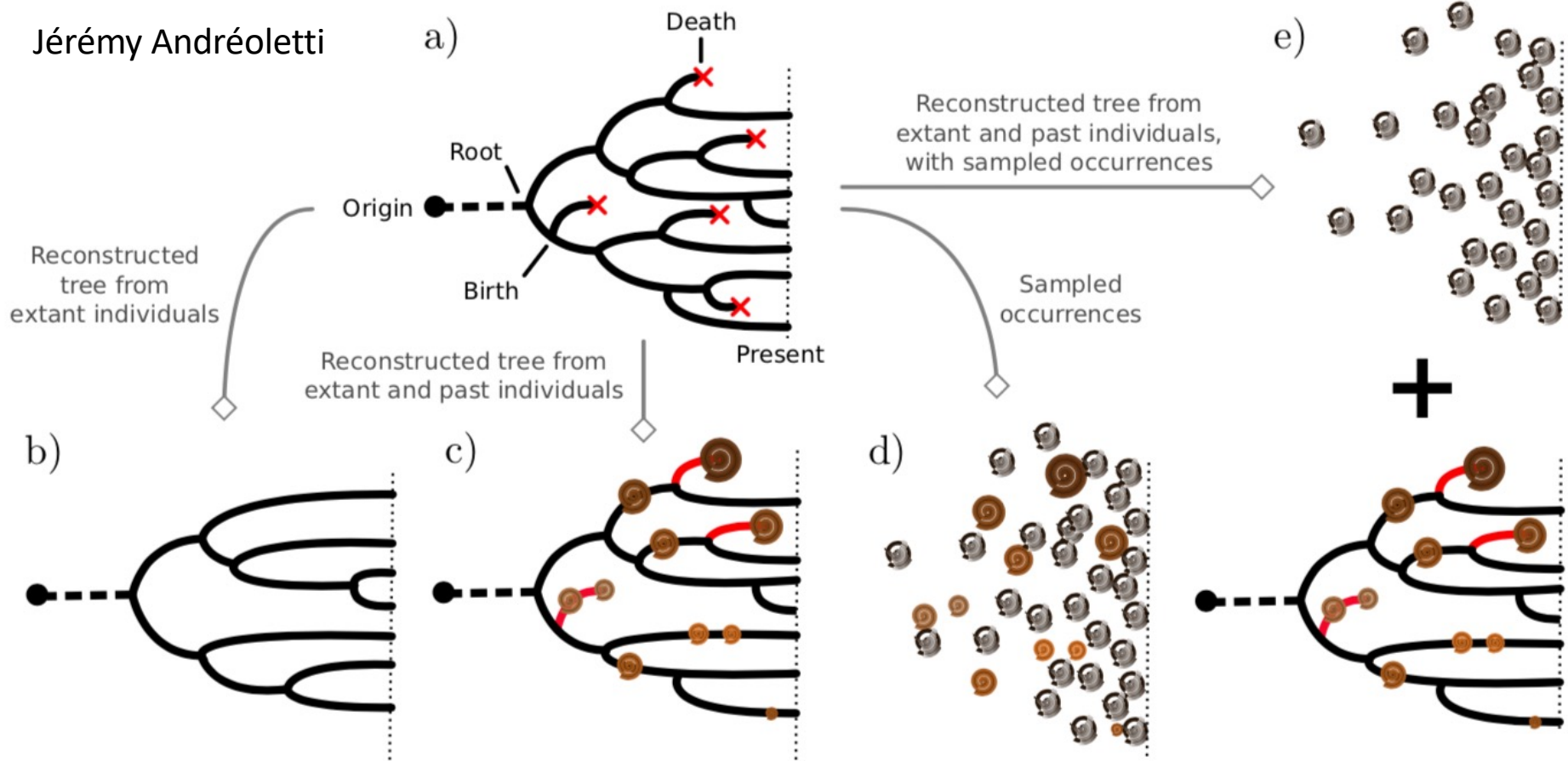
Louca & Pennell Nature 2020





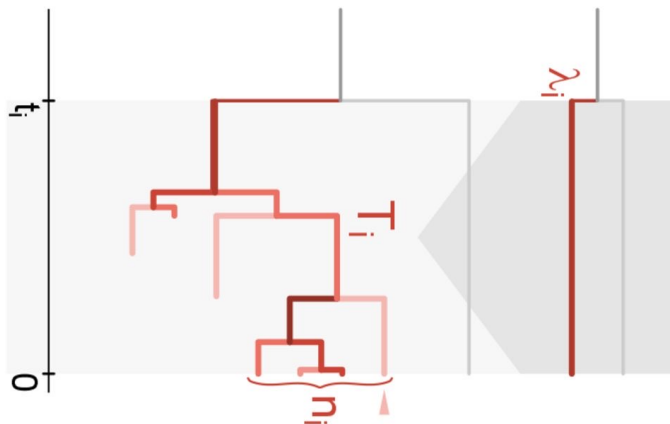
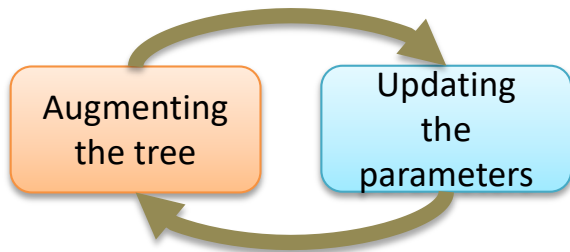
Intégrer données moléculaires et fossiles dans les inférences de la diversification

Jérémie Andréoletti

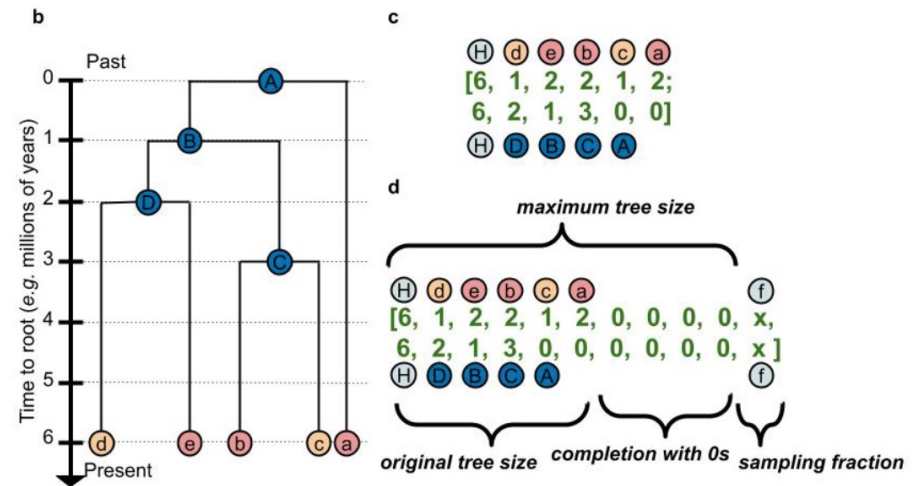


Les limites de l'inférence statistique « classique » et les approches alternatives

Inférence par augmentation de données



Apprentissage profond



Lambert et al.
Syst Bio 2023

These tools are implemented in well-documented, user-friendly packages

Methods in Ecology and Evolution



Methods in Ecology and Evolution 2016, 7, 589–597

doi: 10.1111/2041-210X.12526



APPLICATION

RPANDA: an R package for macroevolutionary analyses on phylogenetic trees

Hélène Morlon^{1*}, Eric Lewitus¹, Fabien L. Condamine², Marc Manceau¹, Julien Clavel¹ and Jonathan Drury¹



Beast2

Bayesian evolutionary analysis by sampling trees





UNDERSTANDING MACROEVOLUTIONARY DYNAMICS USING RPANDA AND JPANDA

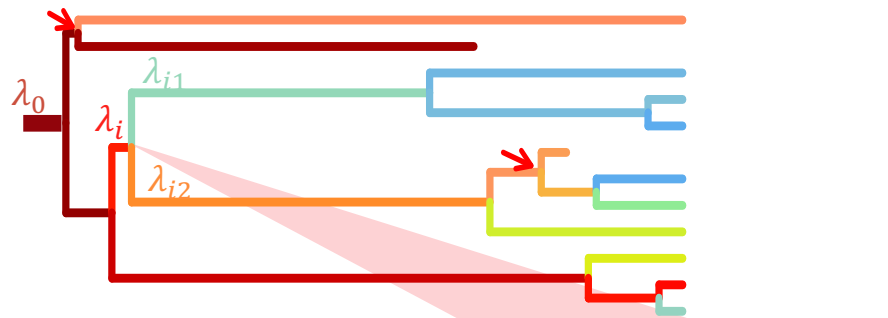


Beast2

Bayesian evolutionary analysis by sampling trees

BAYESIAN PHYLOGENETIC INFERENCE WITH BEAST2

Towards macroevolutionary models accounting for the interplay between speciation and demography / intraspecific genetic differentiation

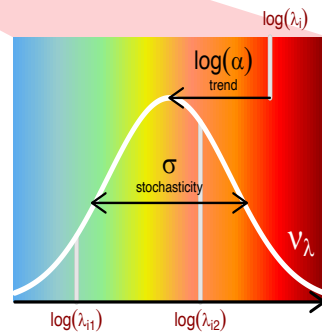


Maliet et al.
Nature Ecol Evol 2019

λ_0 initial speciation rate

α deterministic trend

σ^2 stochastic variation



+ Demographic process

each species follows density-dependent population dynamics

growth rate evolves as a Brownian

random split of individuals at speciation

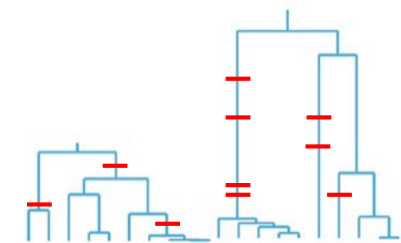
extinction naturally proceeds from the death of all individuals in a given species

+ Population genetics

demography controls N_e

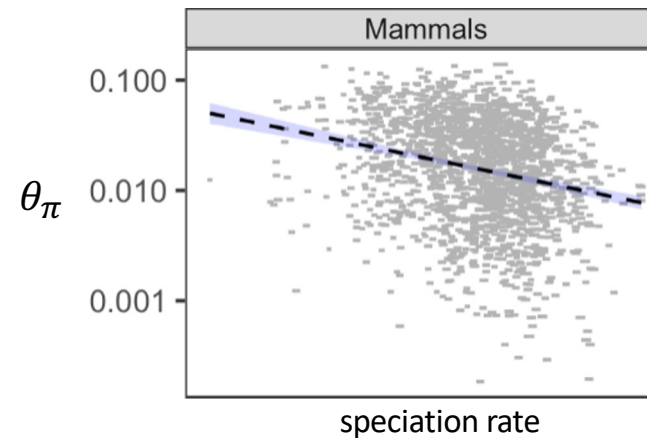


Overcast et al. in prep.



The model predicts either a positive or negative association between speciation rate and genetic diversity depending on the relative pace of speciation and accumulation of genetic diversity

Rapid speciation can limit the accumulation of genetic diversity



The model can be fitted to data using machine learning techniques