UNIVERSITÉ
Grenoble
Alpes

Laboratoire
interdisciplinaire de Physique

Modelling eco-evolutionary dynamics to answer major questions in macroecology and historical biogeography

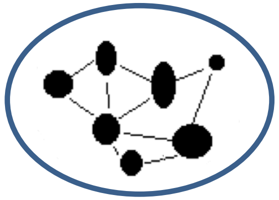
François Munoz

Ecole de recherche de la Chaire MMB

Friday June 14th 2024

Understanding and predicting the dynamics of biodiversity in environments that change in space and time

Local community dynamics

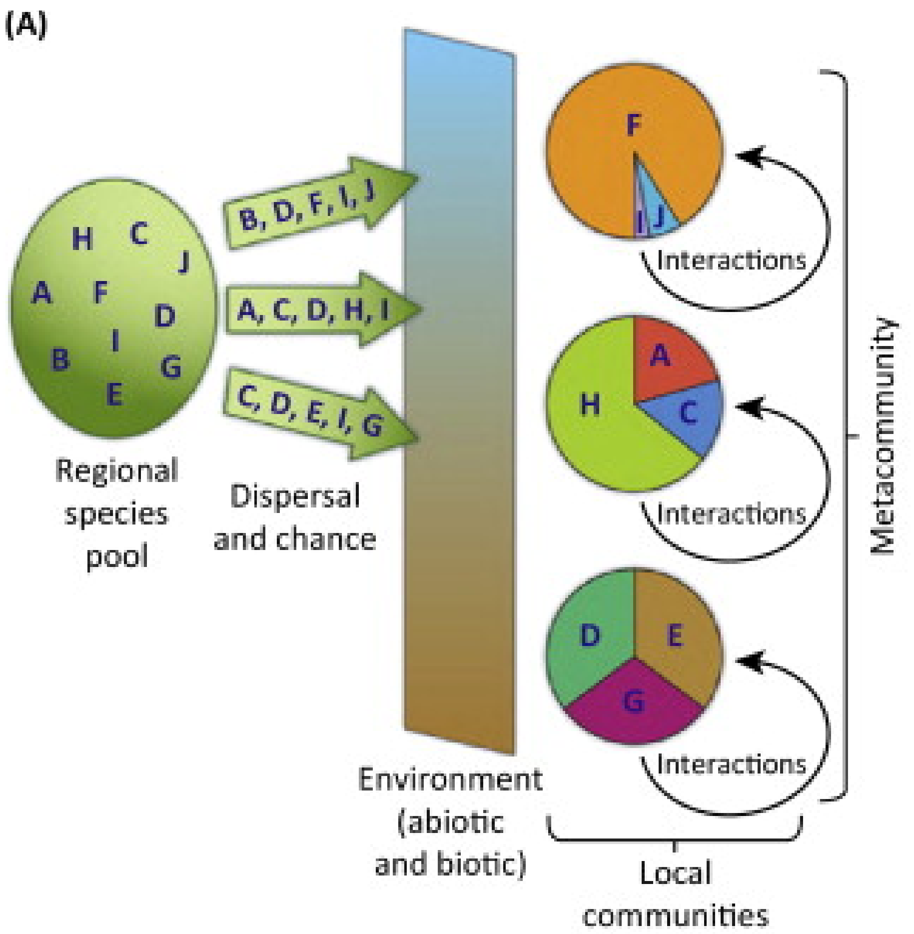


Metapopulation and metacommunity dynamics

Evolutionary and biogeographical dynamics

Hierarchy of processes

Community assembly



(Mittelbach and Schemske 2015)

Dispersal, establishment, reproduction and local survival
→ Neutral theory

+

Variable composition depending on environment and role of interactions
→ Niche theory

=

Patterns of diversity within and between communities

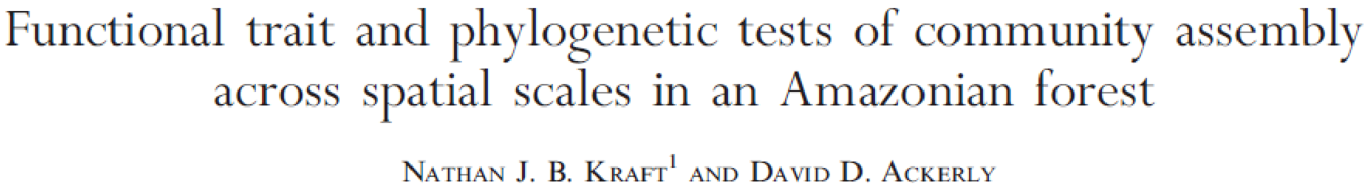
Neutral theory as a null model?

Heated discussion on the meanning of the good fits:

- Some think that the neutral theory is « true »
- Some other think it is bullshit, with a number of methodological issues

Hypothetico-deductive point of view: the theory is correct until we find predictions that do not fit correctly (falsification)

Many studies elaborate in this direction

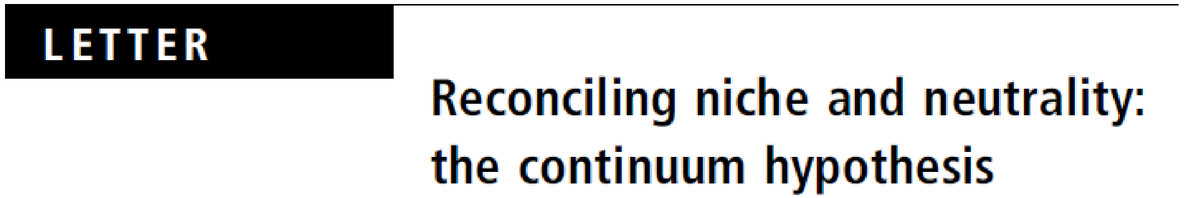


Functional trait and phylogenetic tests of community assembly across spatial scales in an Amazonian forest

Nathan J. B. Kraft[1] and David D. Ackerly

Integrative framework: the theory has to be integrated into a larger theoretical framework.

Stochastic neutral process are trivially occurring in any ecosystem, due to the limited number of individuals and the limited ability of dispersal



LETTER

Reconciling niche and neutrality: the continuum hypothesis

Back to the basics

Apart from discussing the role of neutral processes in community dynamics, a critical issue is still wether the « functional equivalence » assumption is clearly enough specified

This assumption is weak:

1. Species demographic and dispersal parameters are equal in probability. It is enough to consider **equivalence in expectation**, and thereby to allow some level of variation within and between species,

2. Equivalence in survival and establishment can arise in very different ways, and is fully **consistent with a variety of niches** in the community. The only constraint is that the success of the species is **equivalent over the long term**.

3. Even if differences of fitness exist, they can be overcome by regional processes (dispersal-competition trade-off) or by stabilizing mechanisms

Stabilizing processes

The Chesson (2000) framework (see also Adler et al 2007): **stabilizing processes** causes species to limit themselves more than they limit others. Then niches cause intraspecific effects to be more negative than interspecific effects.

$$\bar{r}_i = \underbrace{b_i\left(k_i - \bar{k}\right)}_{\text{Deviation from average fitness}} + \frac{b_i(1-\rho)D}{n-1}$$

Growth rate of invader

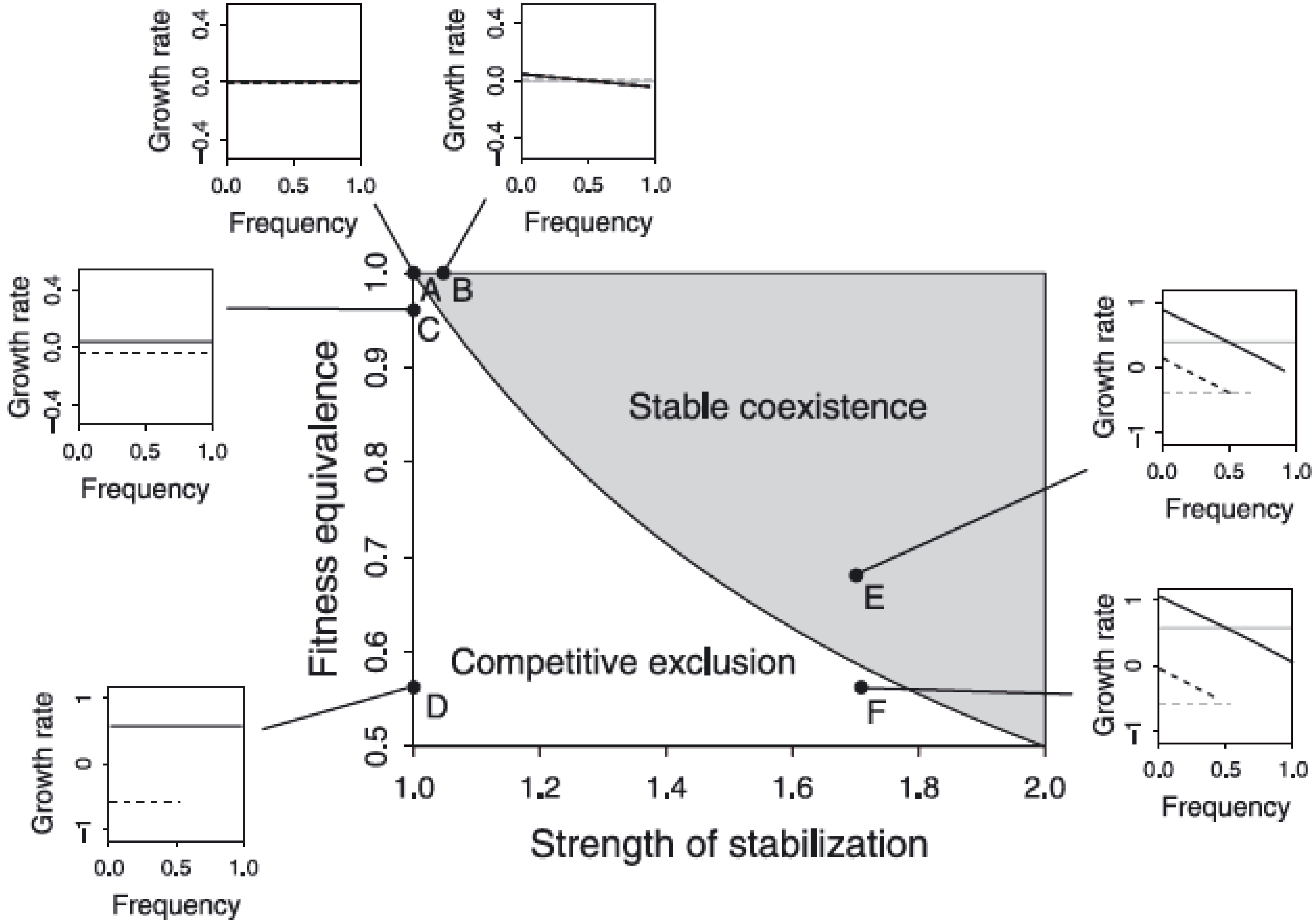
Stabilization due to niche difference

1. and 2. **Equalization or weak stabilization** of fitness differences sustains the equivalence assumption

3. **Strong stabilization:** in this case the equivalence assumption does not hold because fitness differences between species are too large, but the differences does not influence the resulting dynamics beyond the neutral model

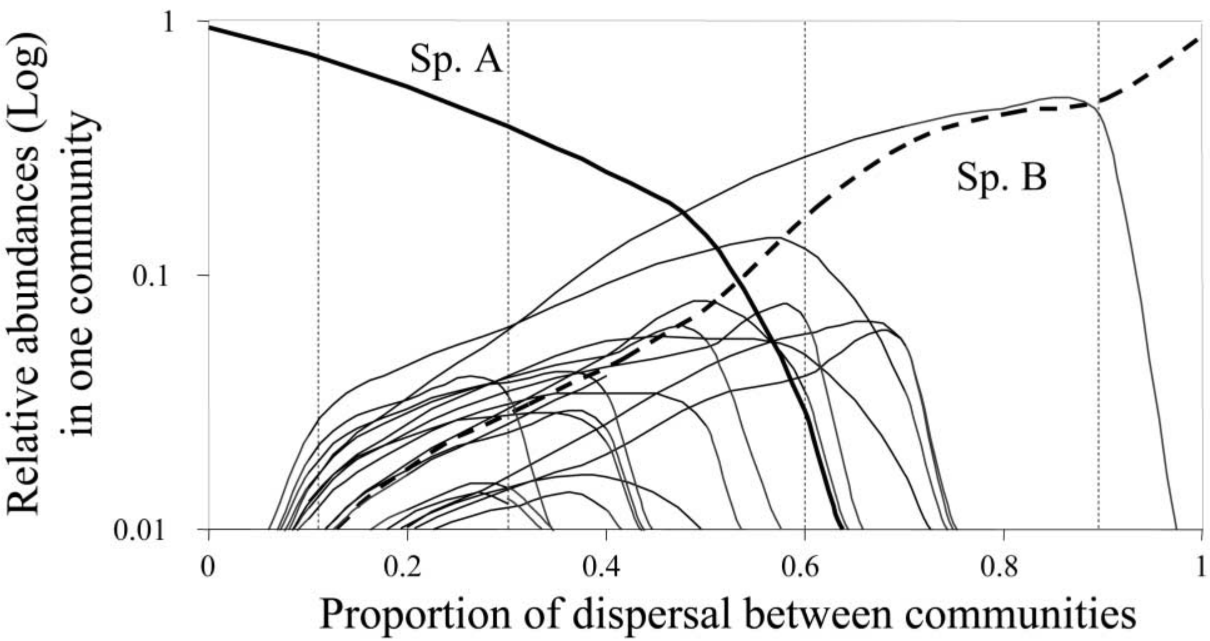
Stabilizing processes

A framework of stabilizing processes for two species (Adler et al 2007)



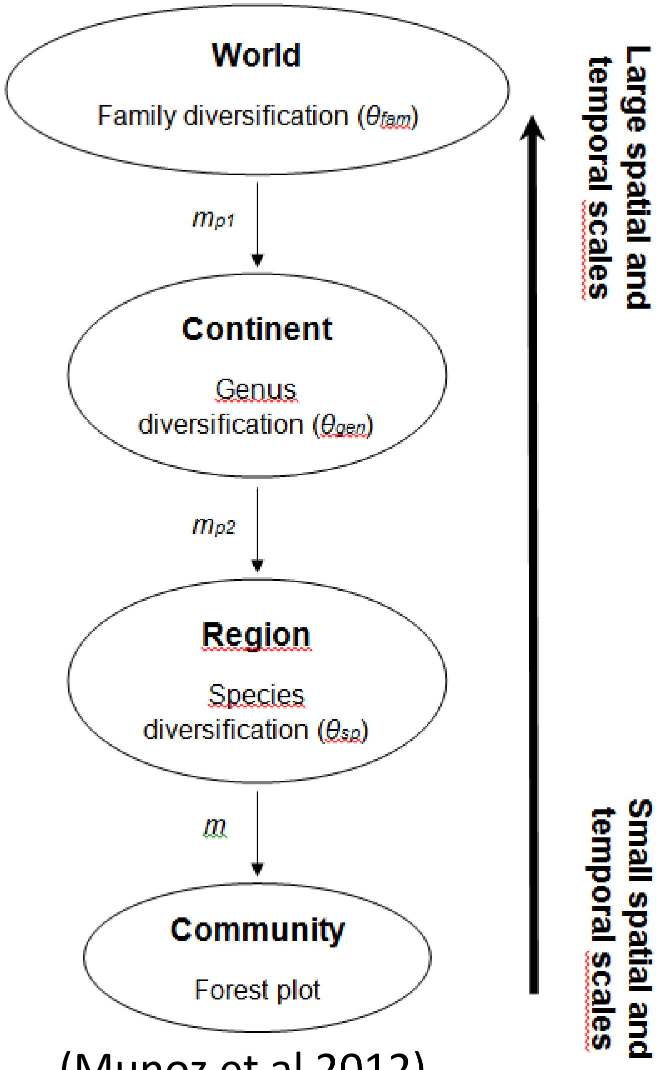
Stabilizing processes

A further critical aspect is the **coupling of local and regional scales**, based on migration and speciation processes



The intensity of migration and the variation in competitive abilities modulate the composition of communities

(Mouquet and Loreau 2003)



World

Family diversification (θ_{fam})

m_{p1}

Continent

Genus diversification (θ_{gen})

m_{p2}

Region

Species diversification (θ_{sp})

m

Community

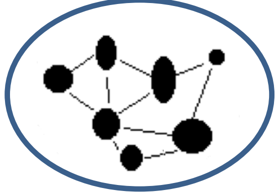
Forest plot

Large spatial and temporal scales

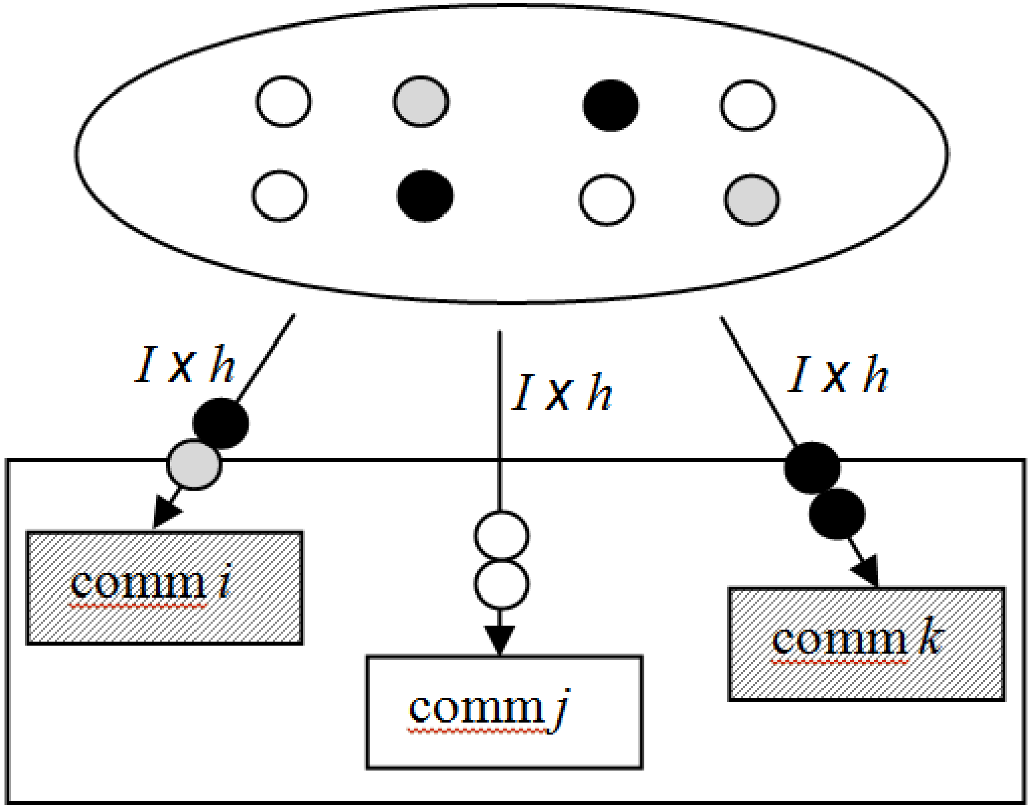
Small spatial and temporal scales

(Munoz et al 2012)

Scale-dependent neutrality



A departure from neutrality is expected during immigration over environmental gradients

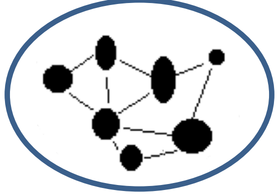


There are niche differences among species in the metacommunity

Habitat filtering during immigration

Their success of establishment depends on the local habitat

Still neutral local zero-sum game



Scale-dependent neutrality

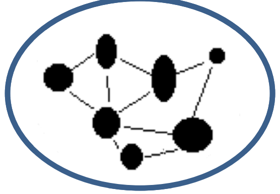
Weighted lottery of immigrants according to both dispersal limitation and habitat preferences

- For each individual the algorithm first decides whether the individual has a new immigrating ancestor (*i.e.* an ancestor that has no descendants among the previously considered individuals). The probability of this event is **governed by *I* and the fit of species niches to local environment** (weighted lottery).

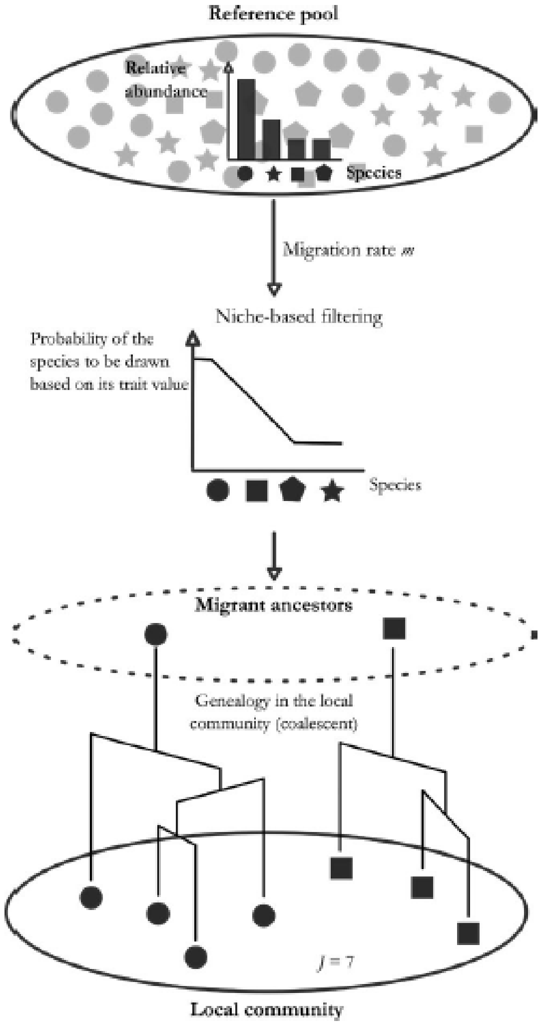
- If the individual has a new immigrating ancestor, the individual is assigned a new ancestry label and the algorithm continues to decide whether the new immigrating ancestor is of a new species.
  → The probability of this event is governed by θ. If the new immigrating ancestor is of a new species, the individual and its ancestor are assigned a new species label, otherwise they receive the species label of a randomly chosen already considered ancestor.

- If the individual does not have a new immigrating ancestor, the individual is assigned the ancestry and species labels of a randomly chosen already considered individual; it then has the same immigrating ancestor as this randomly chosen individual.

Scale-dependent neutrality

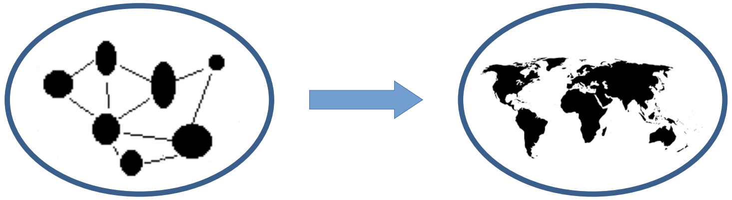


Weighted lottery of immigrants according to both dispersal limitation and habitat preferences



- Coalescent-based model of community assembly
- With environmental filtering depending on species trait values
- With reference to a given regional pool of species

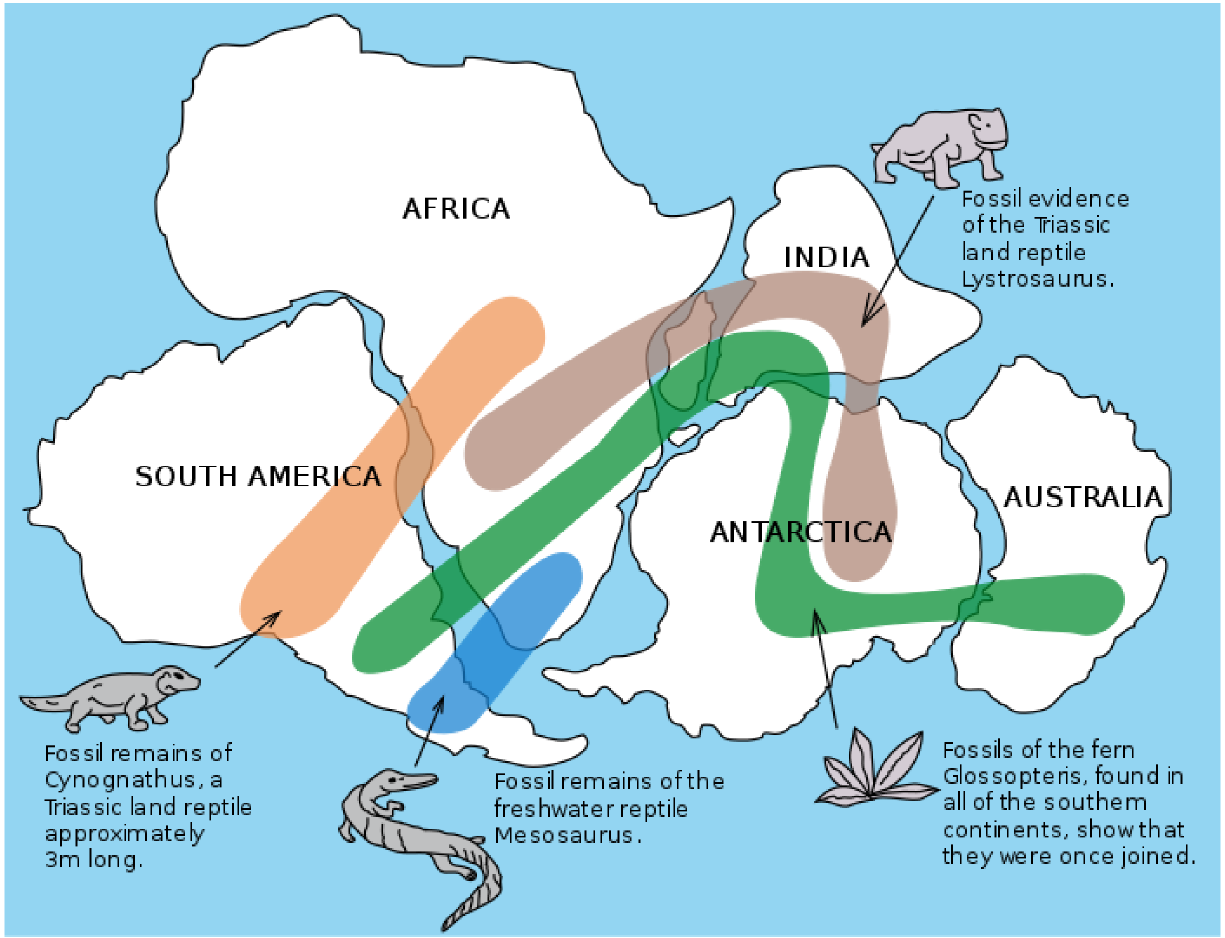
What about the composition and dynamics in the species pool?



(Munoz et al. 2018)

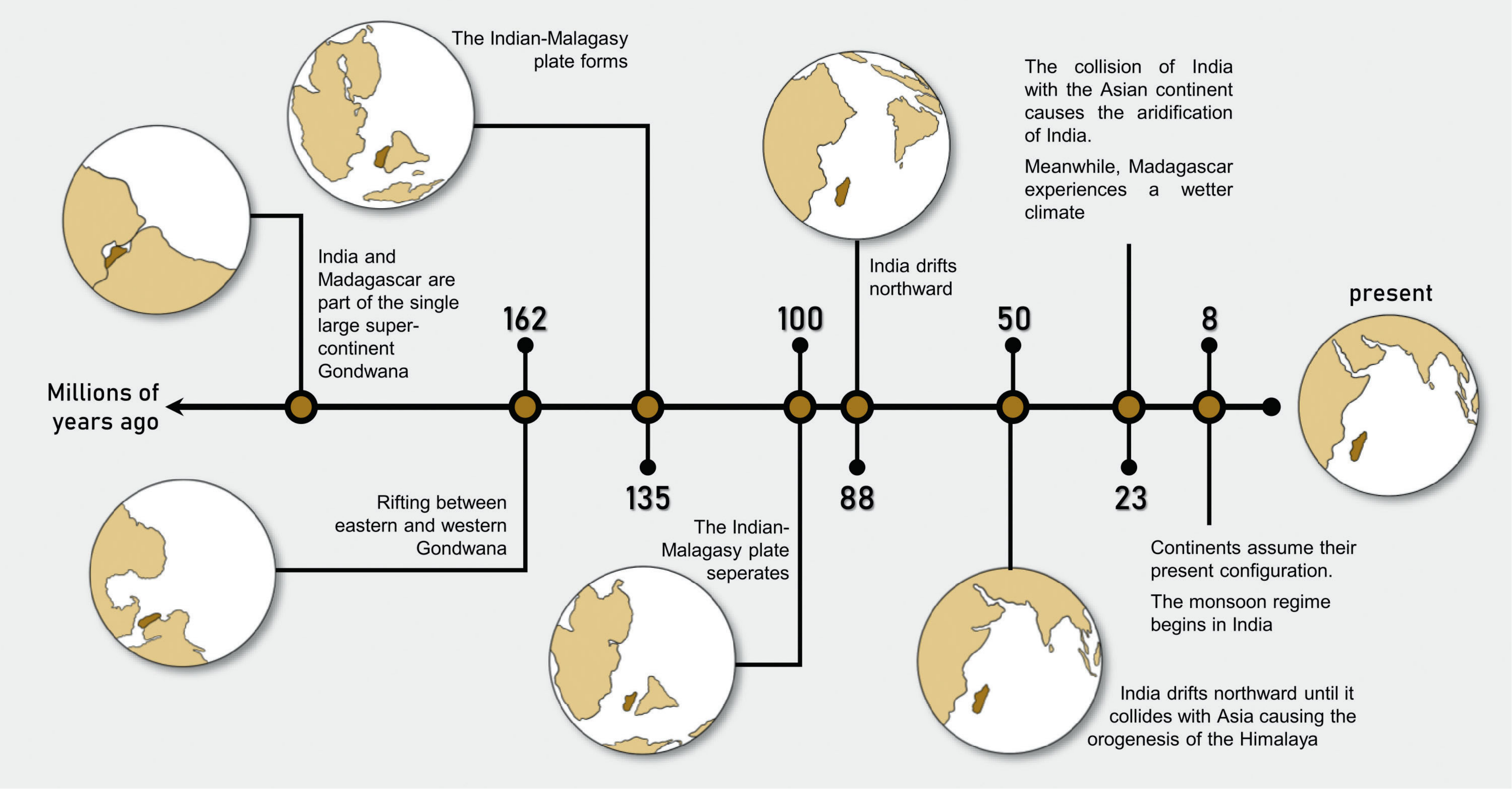
Biogeography and biodiversity dynamics

How are the spatial and temporal dynamics of taxa driven by environmental changes?



Biogeography and biodiversity dynamics

The journey of India and Madagascar



Millions of years ago

India and Madagascar are part of the single large super-continent Gondwana

162

Rifting between eastern and western Gondwana

135

The Indian-Malagasy plate forms

100

The Indian-Malagasy plate seperates

88

India drifts northward

50

India drifts northward until it collides with Asia causing the orogenesis of the Himalaya

23

The collision of India with the Asian continent causes the aridification of India.

Meanwhile, Madagascar experiences a wetter climate

8

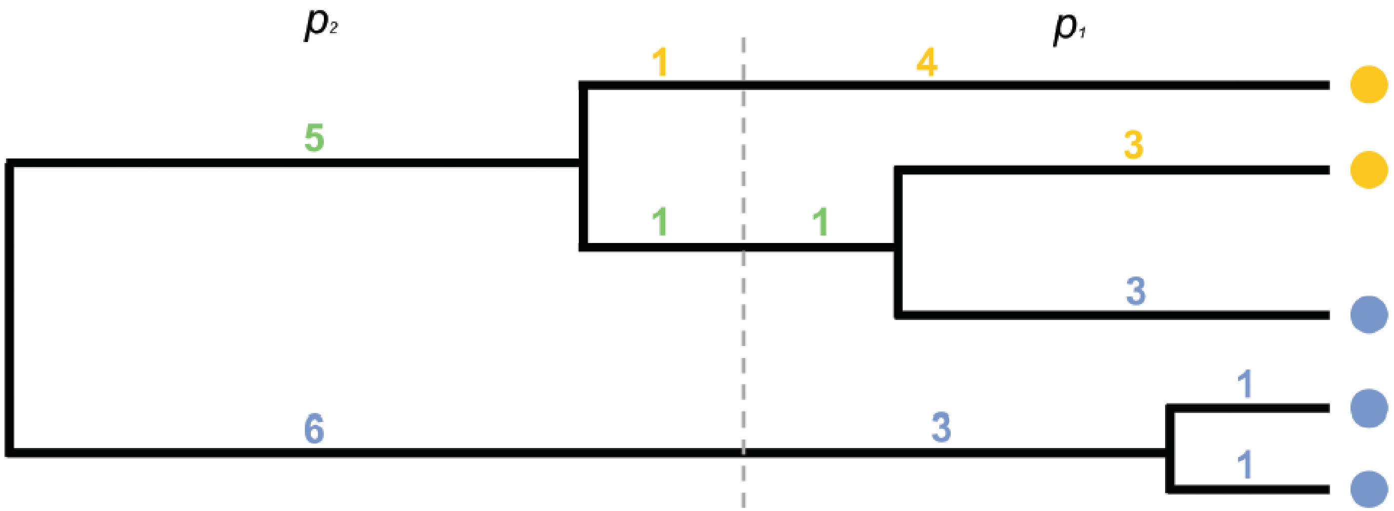
Continents assume their present configuration.

The monsoon regime begins in India

present

Biogeography and biodiversity dynamics

Lineages Per Period (LPP) metric of phylogenetic diversity through time

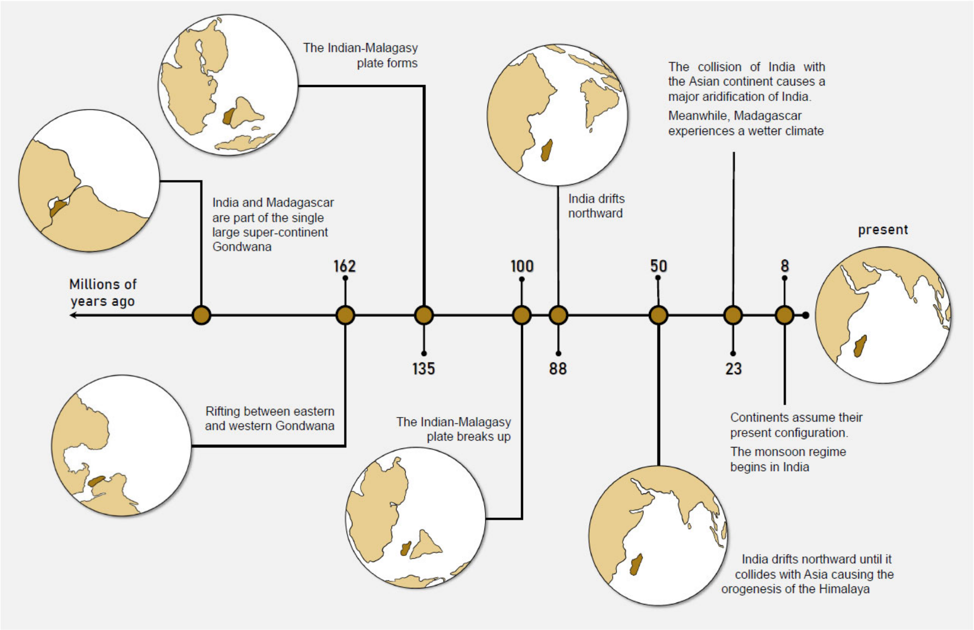


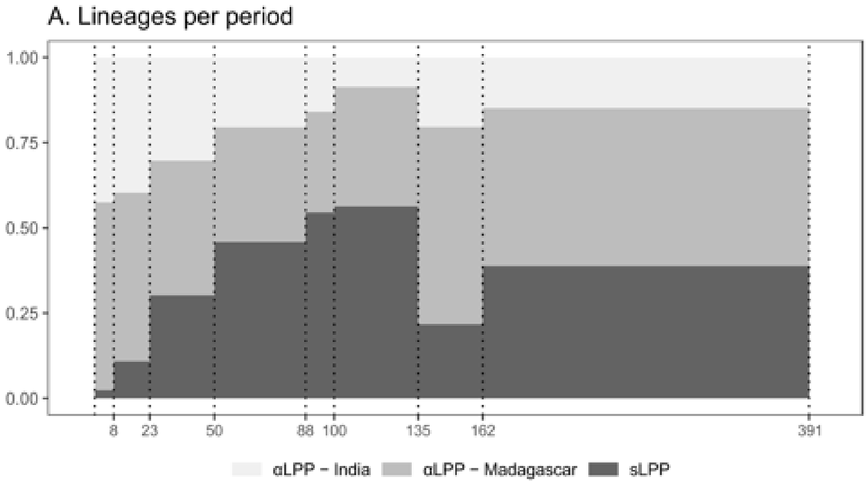
$\alpha_1 LPP_2 = 1 / 13 = 0{,}08$
$\alpha_2 LPP_2 = 6 / 13 = 0{,}46$
$sLPP_2 = (5+1) / 13 = 0{,}46$

$\alpha_1 LPP_1 = (3+4) / 16 = 0{,}44$
$\alpha_2 LPP_1 = (1+1+3+3) / 16 = 0{,}5$
$sLPP_1 = 1 / 16 = 0{,}06$

Biogeography and biodiversity dynamics

Museums and craddles

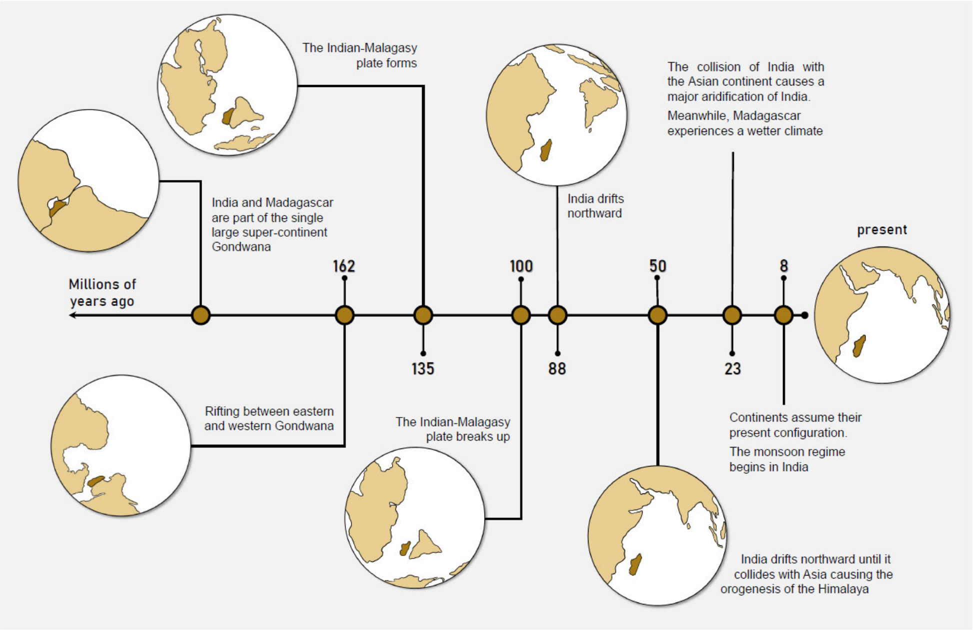


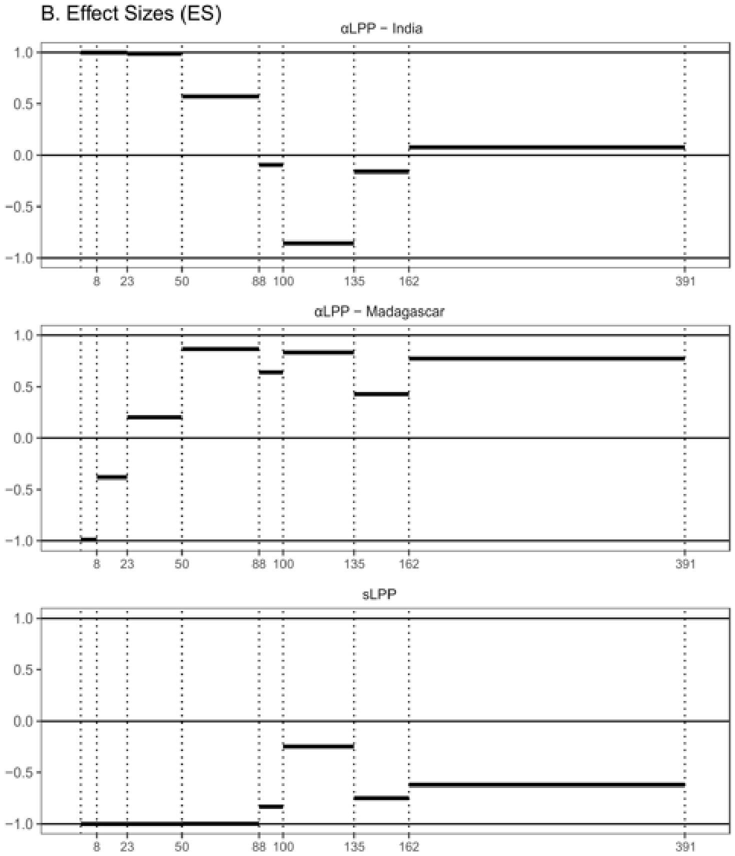

A. Lineages per period

(Barthélémy et al 2022)

Biogeography and biodiversity dynamics

Museums and craddles

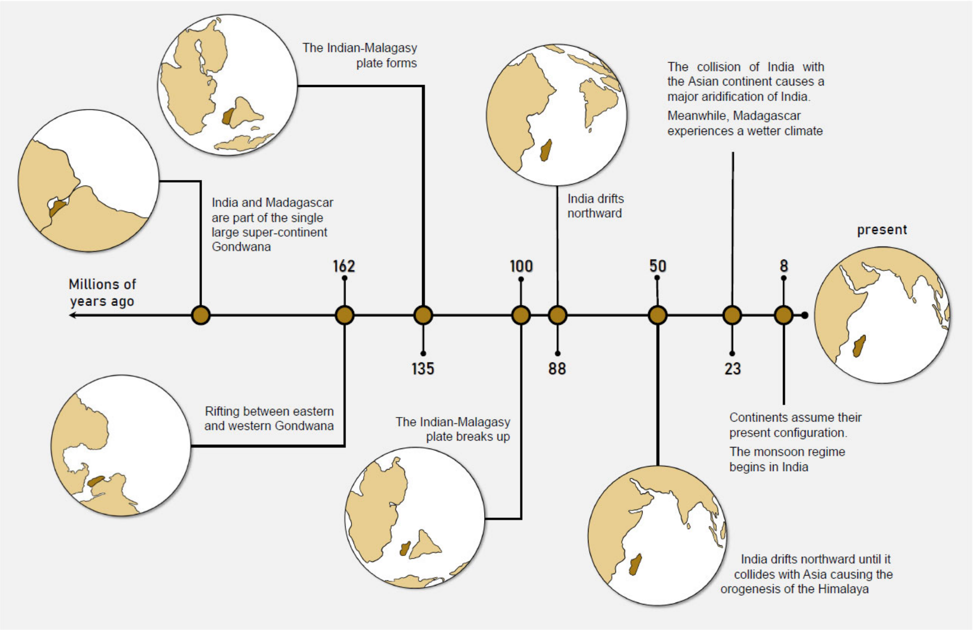


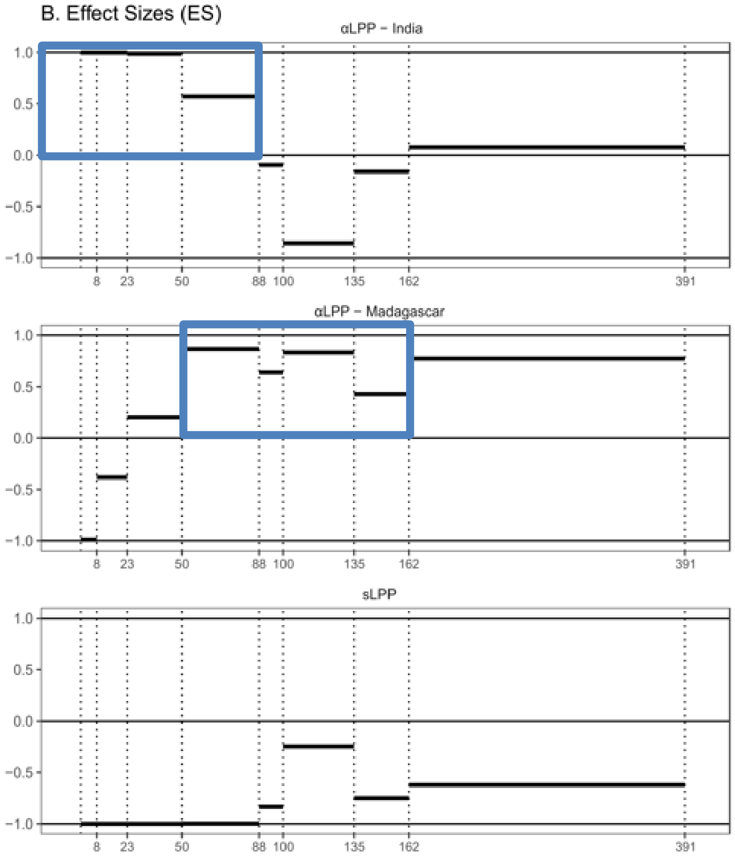


(Barthélémy et al 2022)

Biogeography and biodiversity dynamics

Museums and craddles

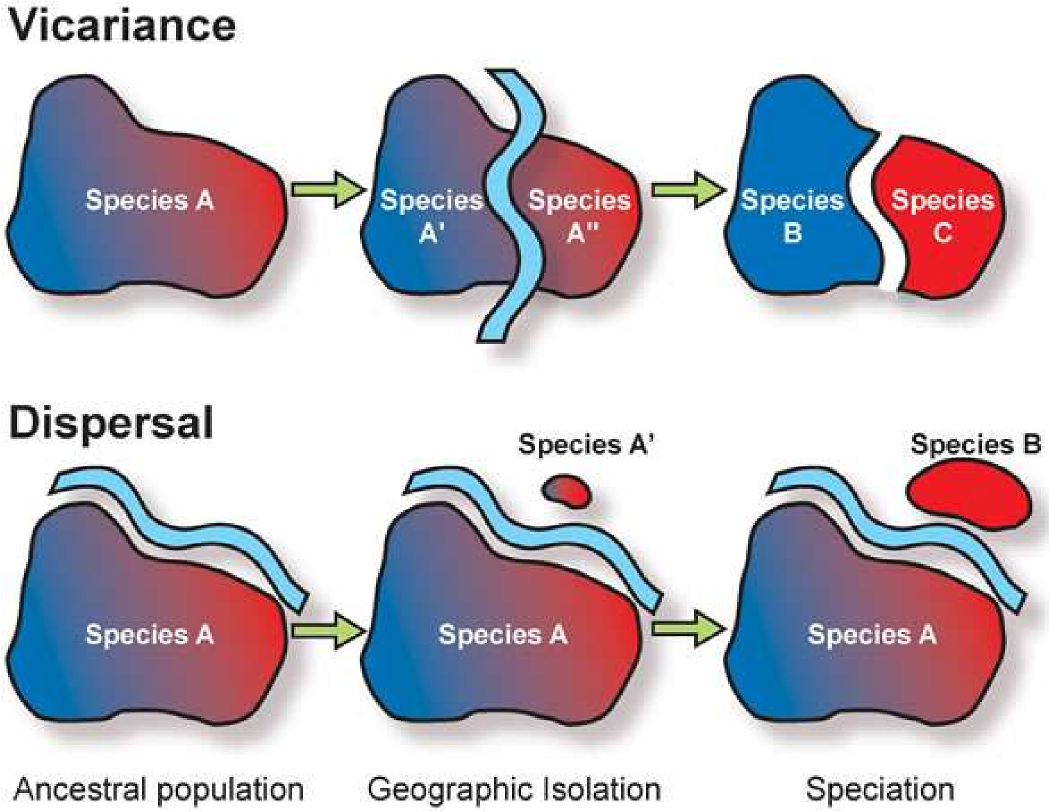




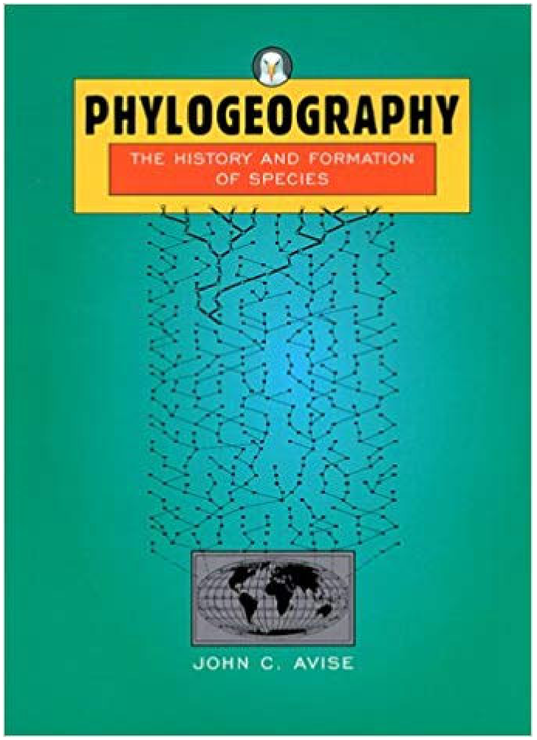
(Barthélémy et al 2022)

Biogeography and biodiversity dynamics

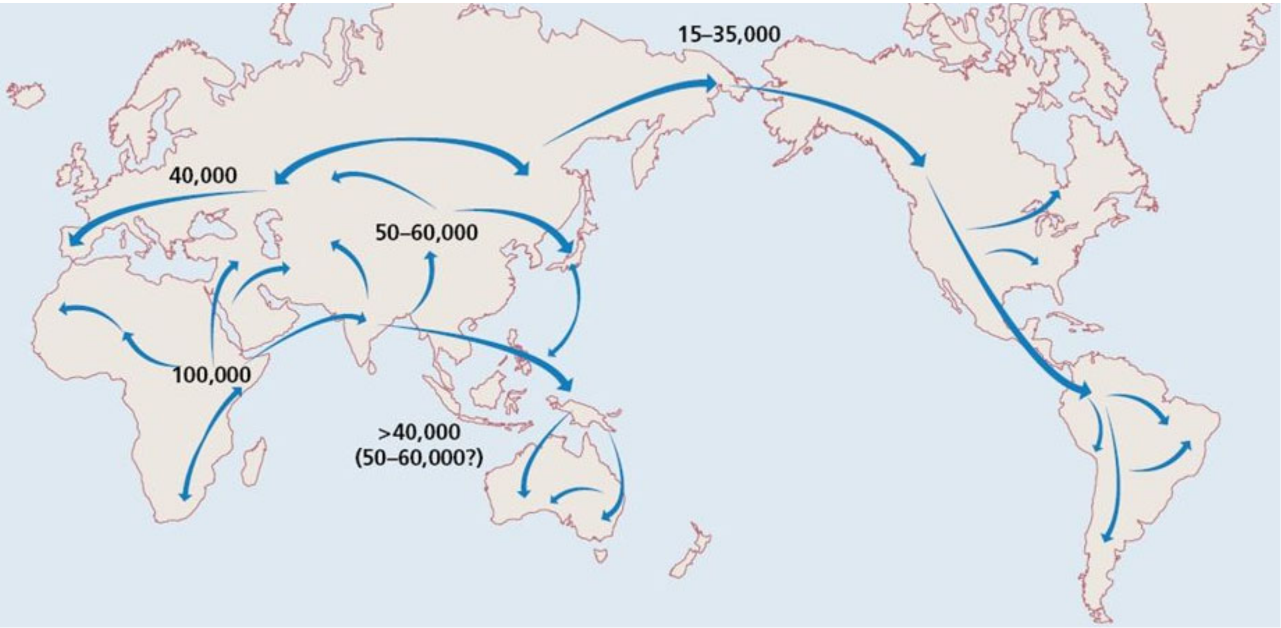
Historical biogeography: roles of dispersal and vicariance in relation to geodynamics and environmental changes



Biogeography and biodiversity dynamics



Objective : reconstruct fragmentation events, migration and population dynamics from the current genetic structure of populations



Human phylogeography (Cavalli-Sforza 2003)

Biogeography and biodiversity dynamics

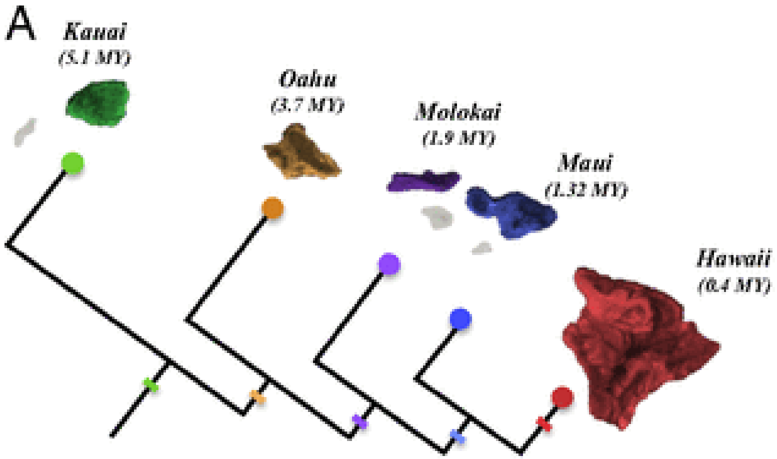
Eco-evolutionary perspective

Fundamental processes:

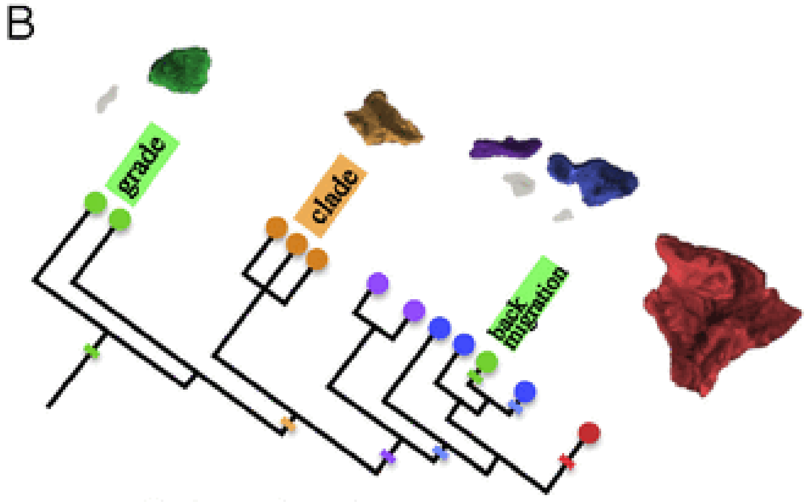
- Speciation
- Migration
- Drift
- Selection

Drive the patterning of organism lineages in space and time

(Shaw and Gillespie 2016)



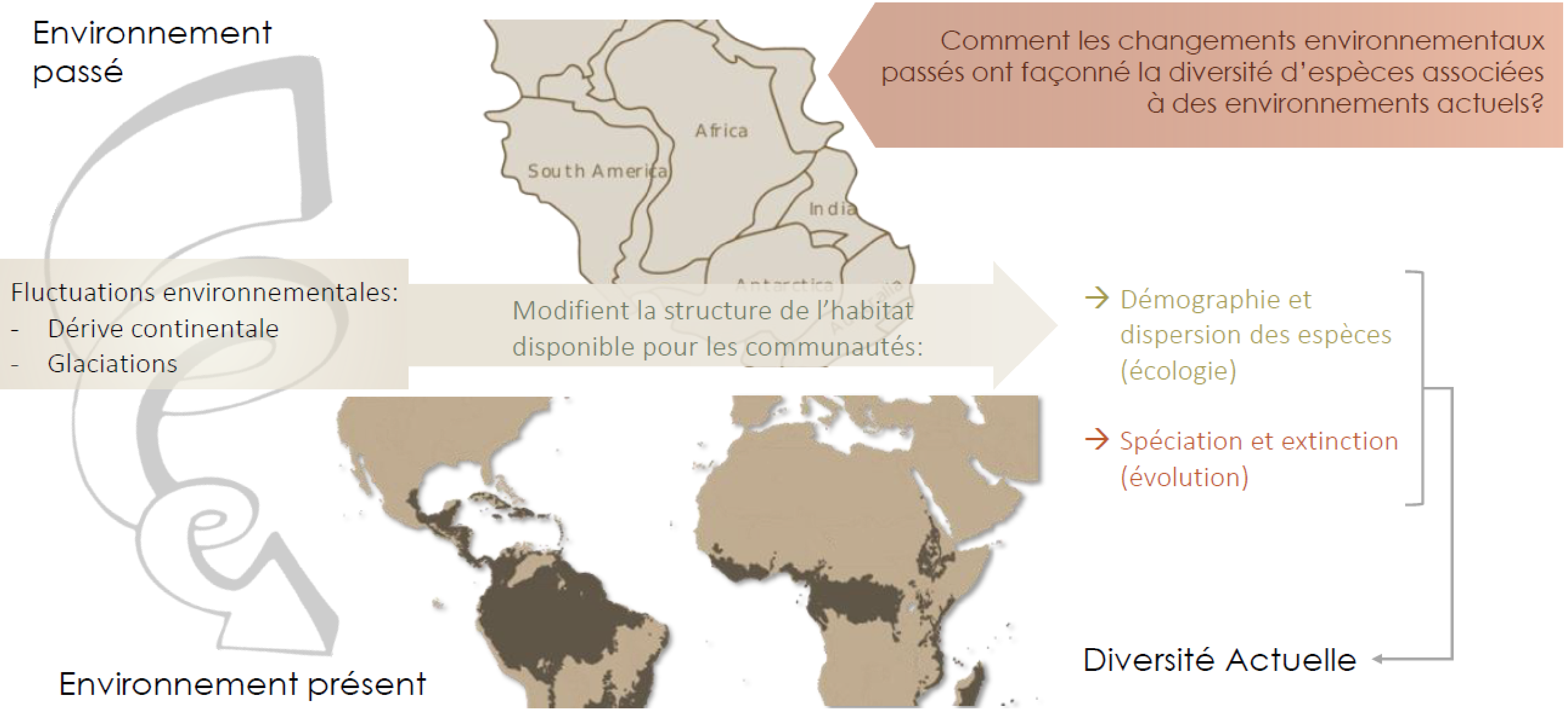
Simple progression



Clades and grades

Biogeography and biodiversity dynamics

Eco-evolutionary perspective



Environnement passé

Comment les changements environnementaux passés ont façonné la diversité d'espèces associées à des environnements actuels?

Fluctuations environnementales:
- Dérive continentale
- Glaciations

Modifient la structure de l'habitat disponible pour les communautés:

→ Démographie et dispersion des espèces (écologie)

→ Spéciation et extinction (évolution)

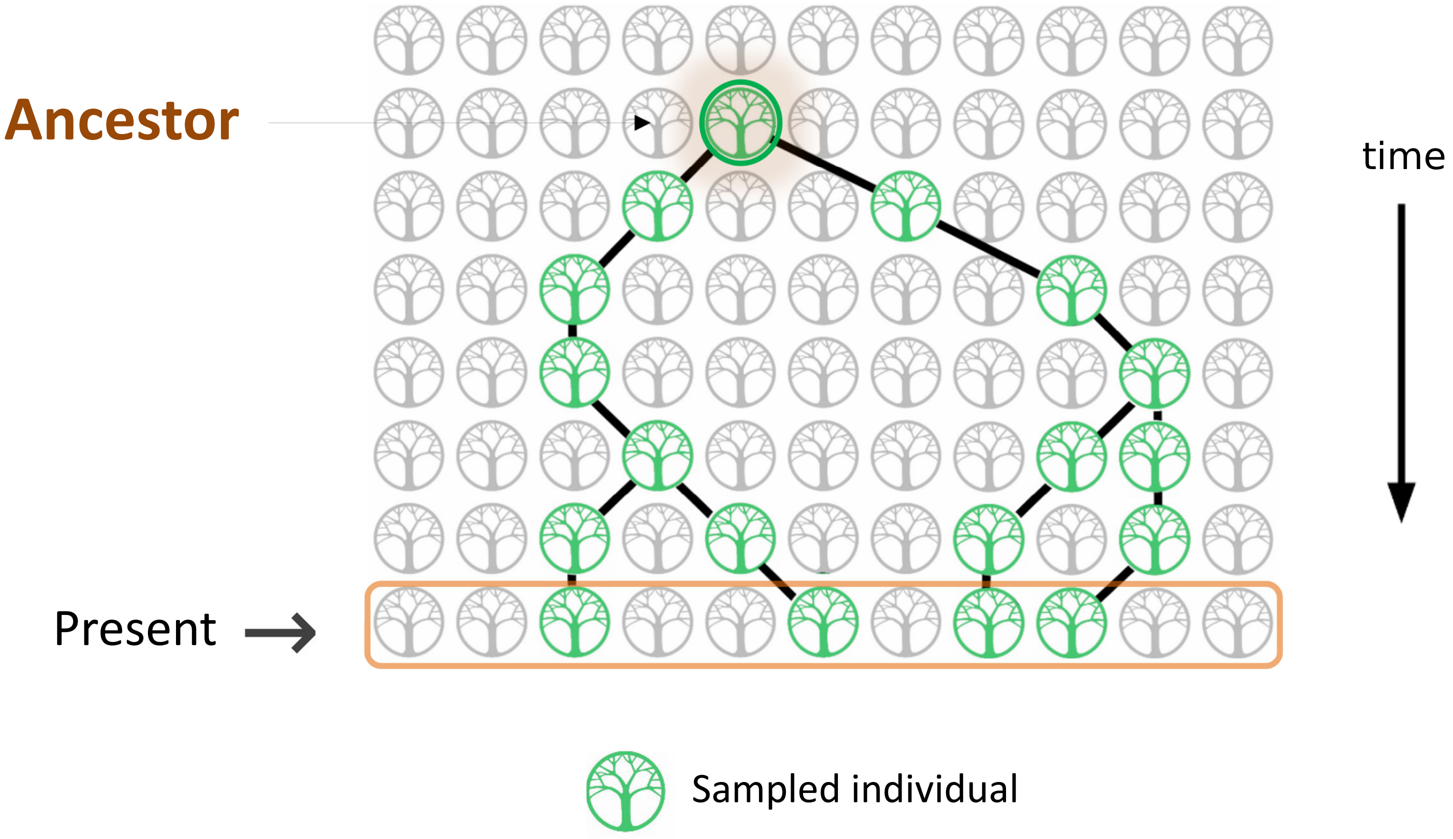
Diversité Actuelle

Environnement présent

Genealogical approach

Coalescent : rebuilding shared coancestry

In a population



Ancestor

time

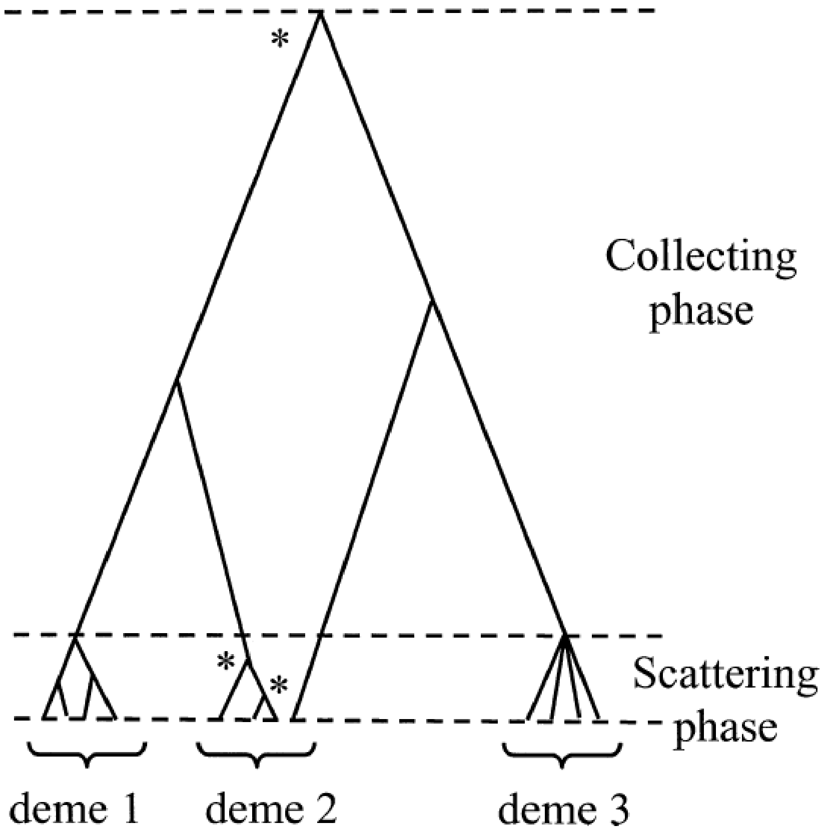
Present →

Sampled individual

Genealogical approach

Coalescent : rebuilding shared coancestry

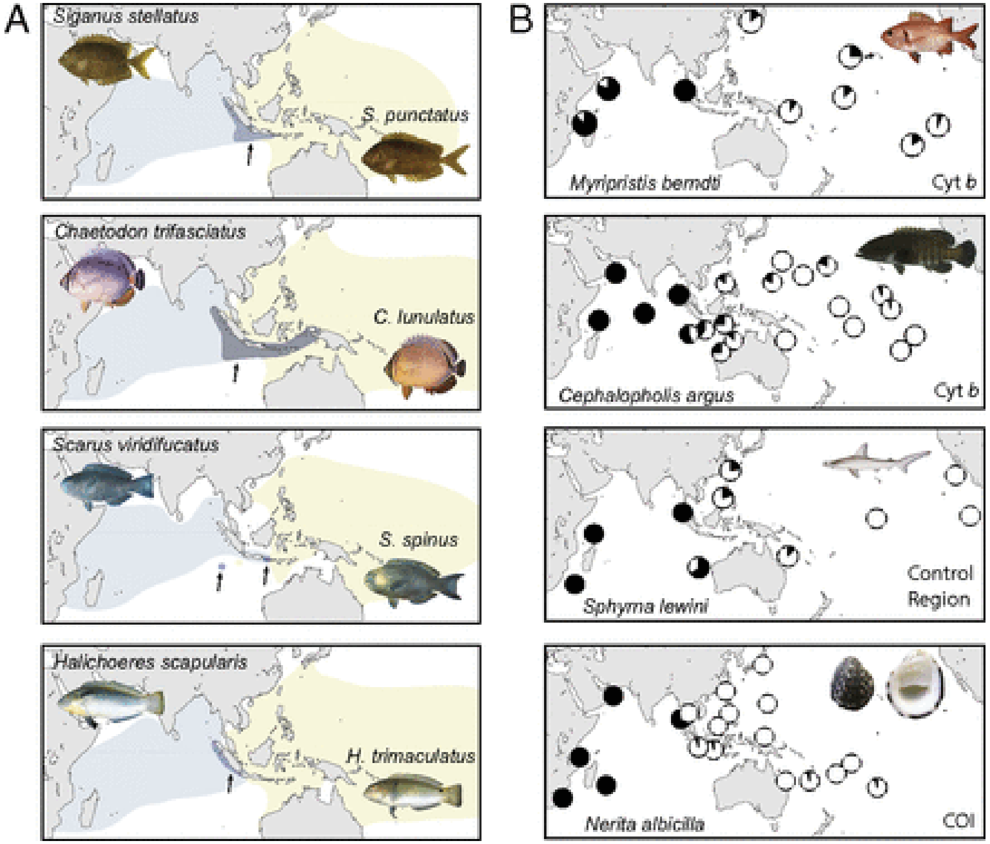
In a metapopulation



(Pannell 2003)

Biogeography and biodiversity dynamics

Comparative phylogeography: reconstruct fragmentation events, migration and population dynamics common to a set of co-distributed species

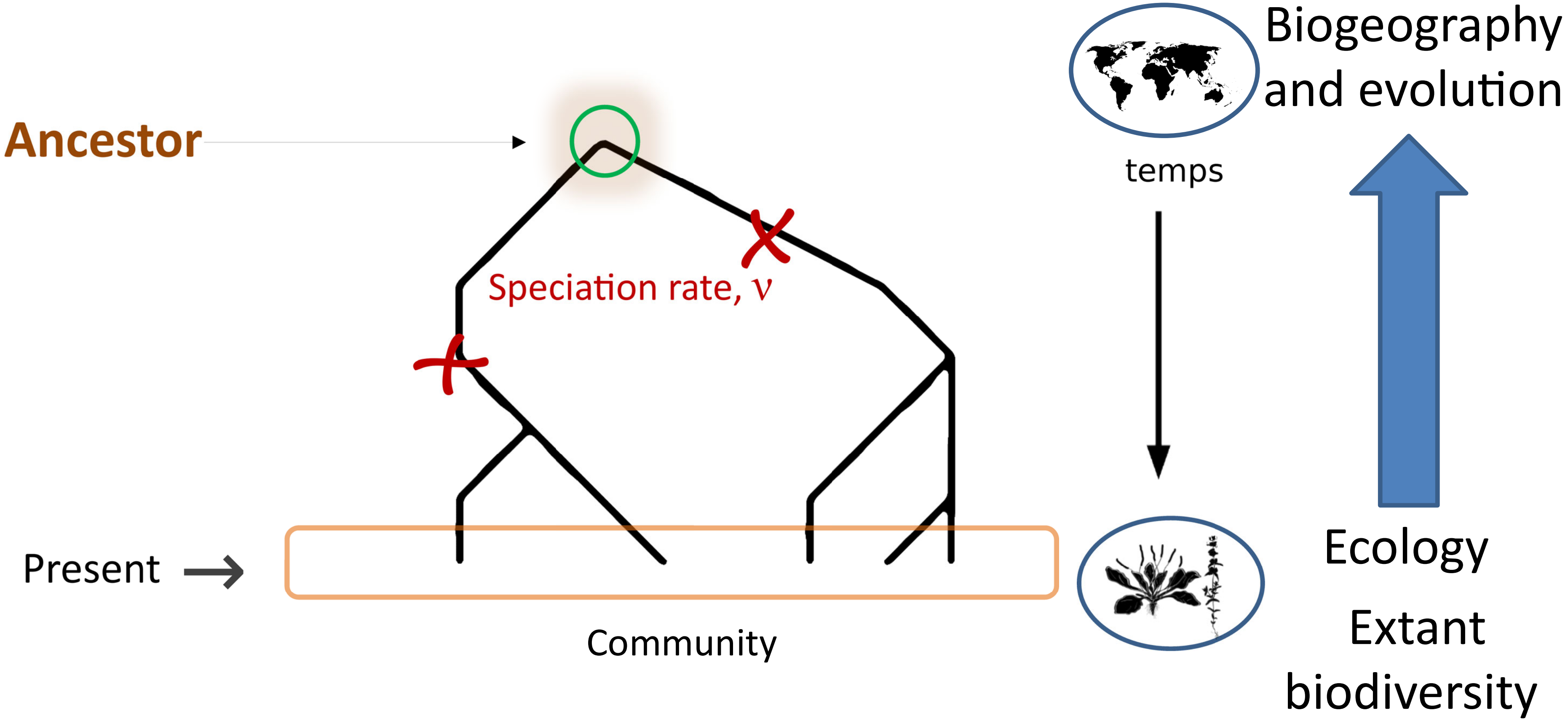


Recurrent impact of the Indo-specific barrier on interspecific distributions (left) and intraspecific genetic differentiation (right)

(Bowen et al. 2016)

Genealogical approach

Multispecies coalescent : integrating speciation events



Ancestor

Speciation rate, ν

Present

Community

temps

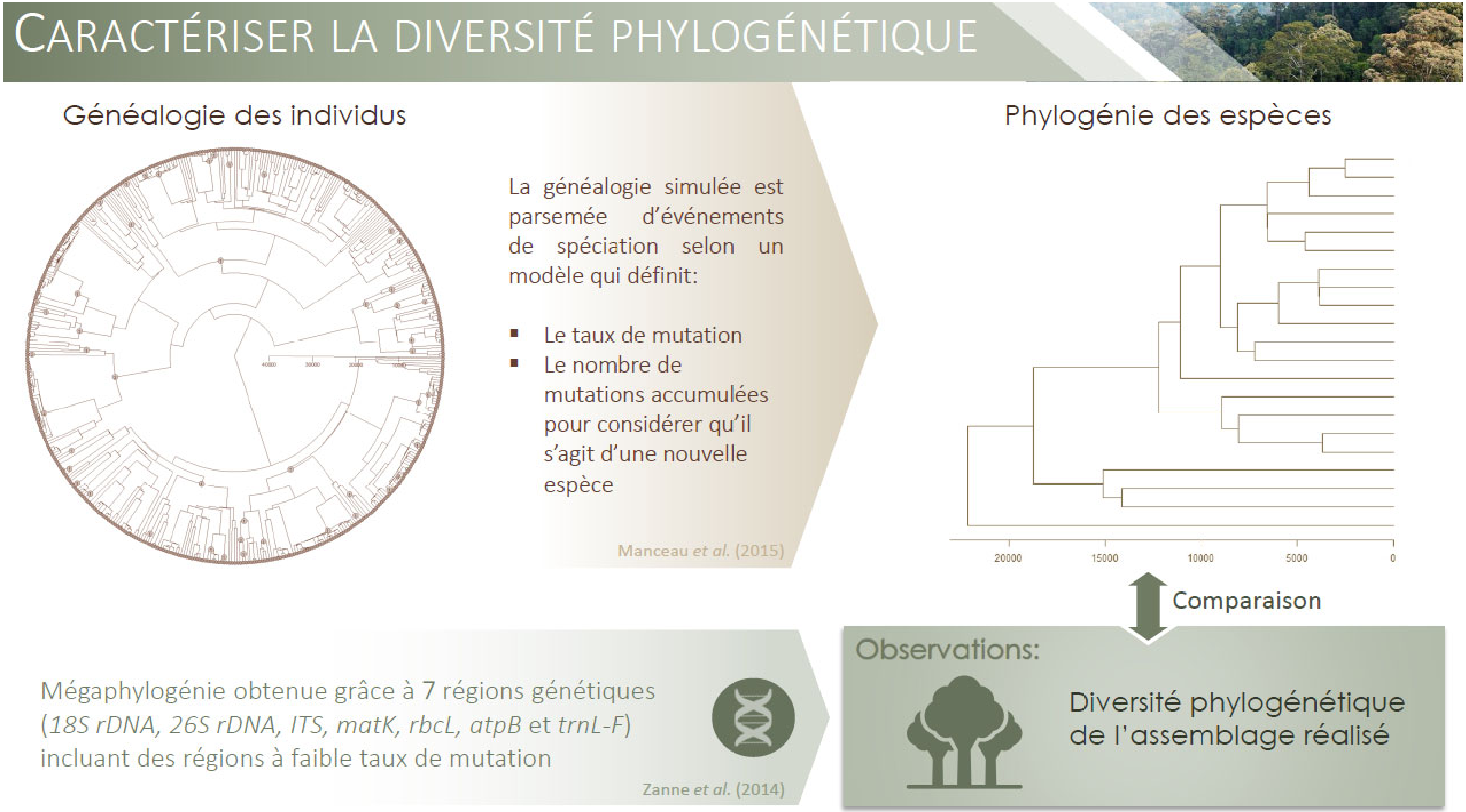
Biogeography and evolution

Ecology

Extant biodiversity

Genealogical approach

Past processes having determined extant diversity



CARACTÉRISER LA DIVERSITÉ PHYLOGÉNÉTIQUE

Généalogie des individus

La généalogie simulée est parsemée d'événements de spéciation selon un modèle qui définit:

- Le taux de mutation
- Le nombre de mutations accumulées pour considérer qu'il s'agit d'une nouvelle espèce

Manceau *et al.* (2015)

Phylogénie des espèces

Comparaison

Observations:

Diversité phylogénétique de l'assemblage réalisé

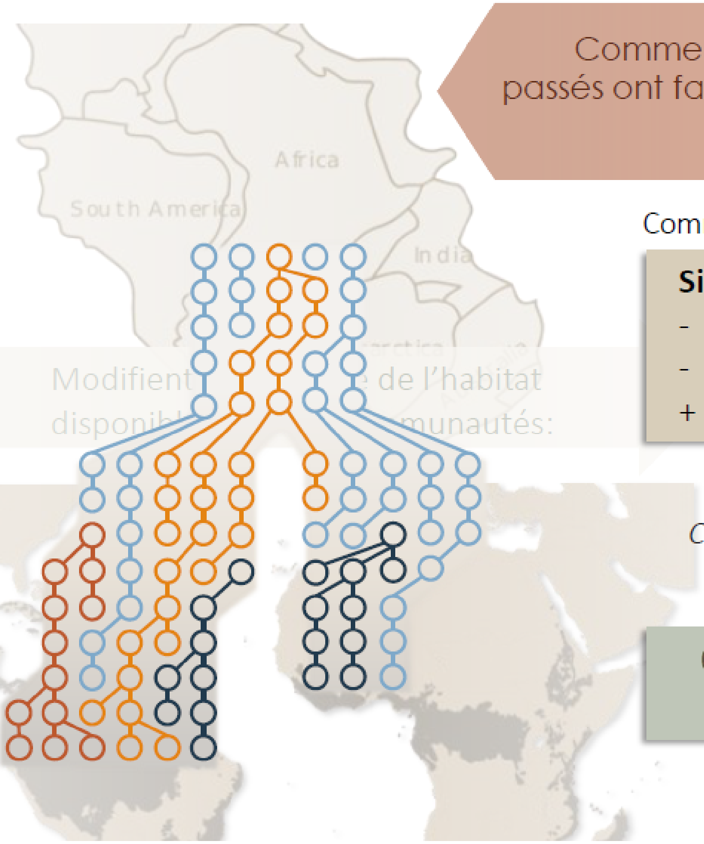
Mégaphylogénie obtenue grâce à 7 régions génétiques (*18S rDNA, 26S rDNA, ITS, matK, rbcL, atpB* et *trnL-F*) incluant des régions à faible taux de mutation

Zanne *et al.* (2014)

Biogeography and biodiversity dynamics

DÉVELOPPEMENT D'UNE APPROCHE ÉCO-ÉVOLUTIVE

Environnement passé



Comment les changements environnementaux passés ont façonné la diversité d'espèces associées à des environnements actuels?

Fluctuations environnementales:
- Dérive continentale
- Glaciations

Modifient ... de l'habitat disponib... munautés:

Comment évaluer l'impact de ses changements?

Simuler la généalogie des individus
- Démographie des espèces
- Evènements de migration
- \+ Spéciation

(écologie)

Cadre méthodologique développé en M2

(évolution)

Quelle dynamique éco-évolutive explique la diversité actuelle?

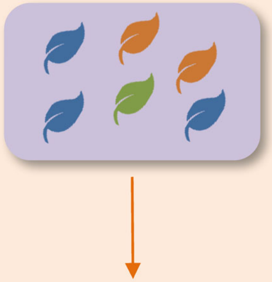
Diversité Actuelle

Environnement présent

Inference framework

Approximate Bayesian Computation

Observed community



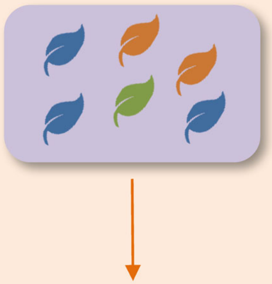
Calculation of observed *summary statistics*:

- richness
- diversity α, β …
- functional diversity

Inference framework

Approximate Bayesian Computation

Observed community



Calculation of observed *summary statistics*:
- richness
- diversity α, β …
- functional diversity

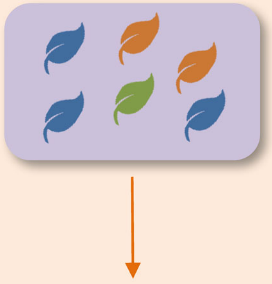
Formulation of an assembly model
- Definition of parameters
- And their *prior* distributions

Performing simulations for different parameter values

Inference framework

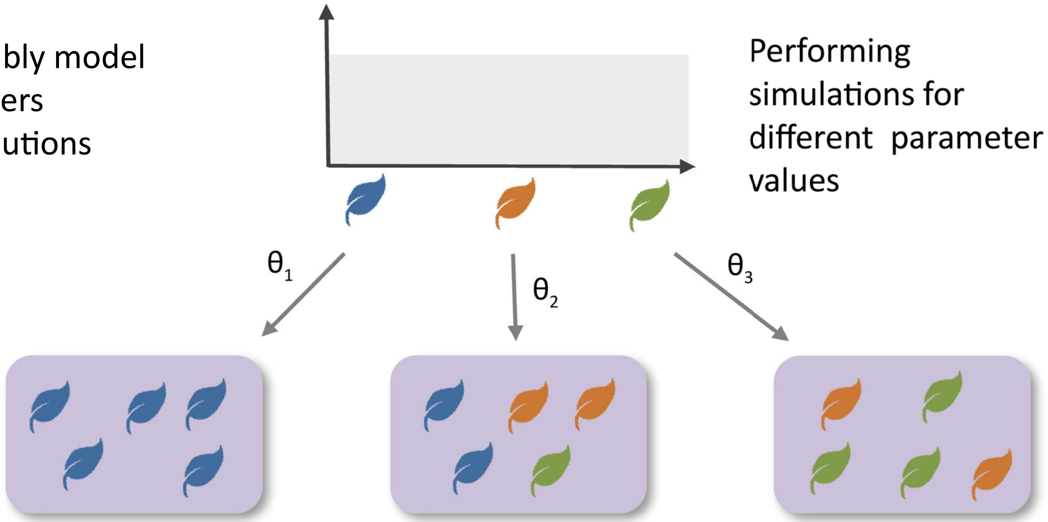
Approximate Bayesian Computation

Observed community



Calculation of observed *summary statistics*:
- richness
- diversity α, β …
- functional diversity

Formulation of an assembly model
- Definition of parameters
- And their *prior* distributions

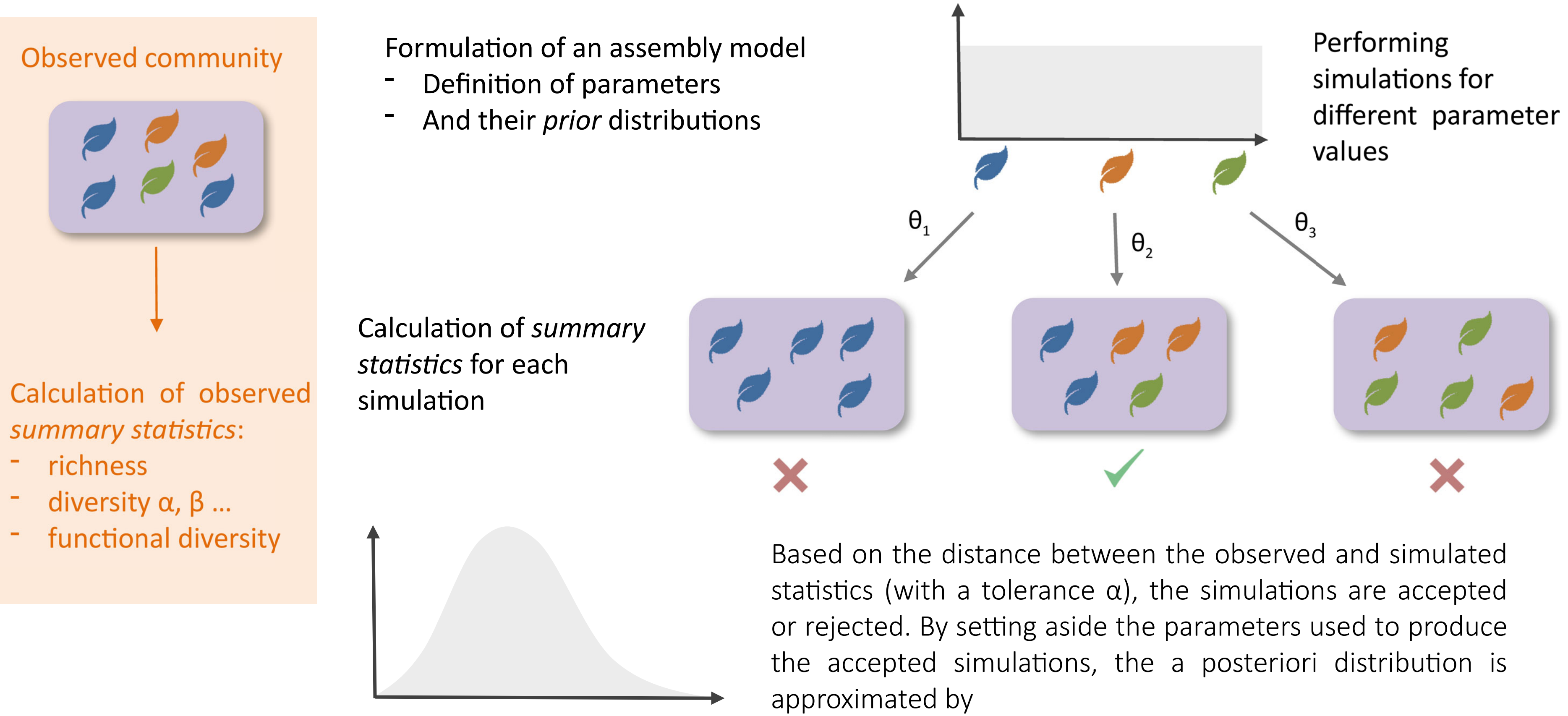


Performing simulations for different parameter values

Calculation of *summary statistics* for each simulation

Inference framework

Approximate Bayesian Computation



Observed community

Calculation of observed *summary statistics*:
- richness
- diversity α, β …
- functional diversity

Formulation of an assembly model
- Definition of parameters
- And their *prior* distributions

Performing simulations for different parameter values

Calculation of *summary statistics* for each simulation

Based on the distance between the observed and simulated statistics (with a tolerance α), the simulations are accepted or rejected. By setting aside the parameters used to produce the accepted simulations, the a posteriori distribution is approximated by

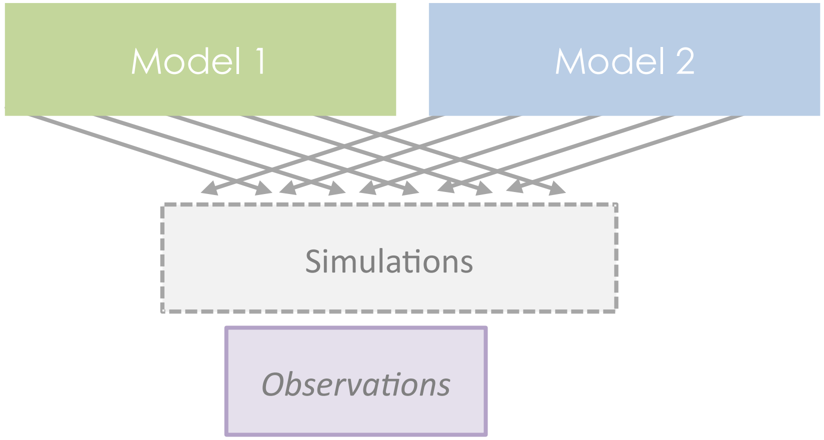
Inference framework

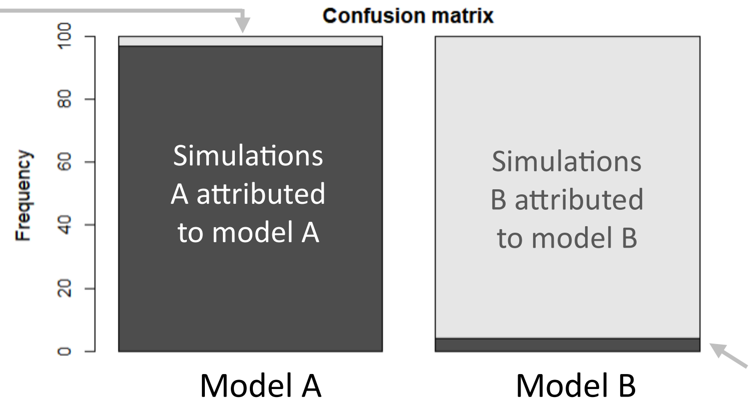
Approximate Bayesian Computation

✓ The way in which diversity patterns are simulated influences the result of the inference!

- **How can we ensure that conditionality does not bias our conclusions?**

    1. Misclassification analysis (verification of the ability to select a model)





Simulations A attributed to model B

Simulations B attributed to model A

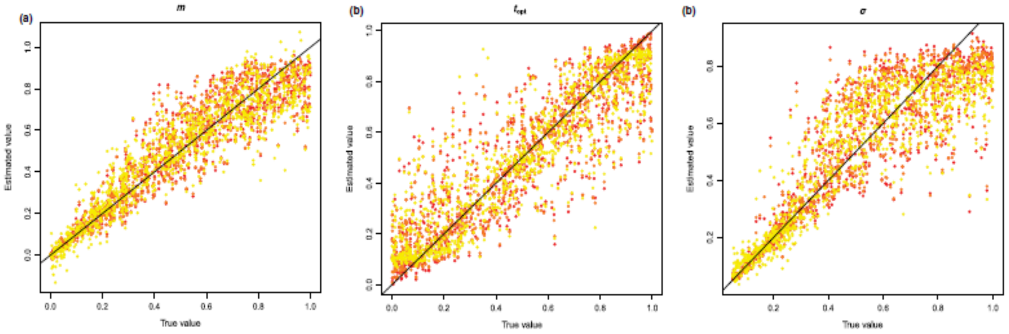
Inference framework

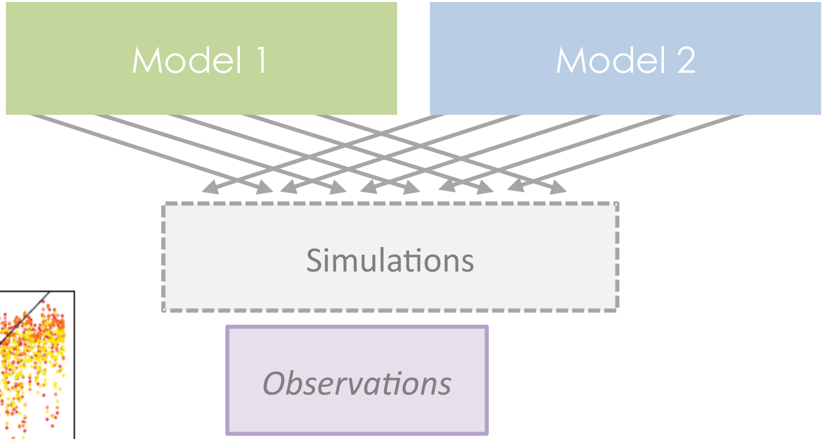
Approximate Bayesian Computation

✓ The way in which diversity patterns are simulated influences the result of the inference!

- **How can we ensure that conditionality does not bias our conclusions?**

  2. Cross-validation analysis (verification of the ability to infer the parameters of a model)





Model 1

Model 2

Simulations

Observations

Inference framework

Approximate Bayesian Computation

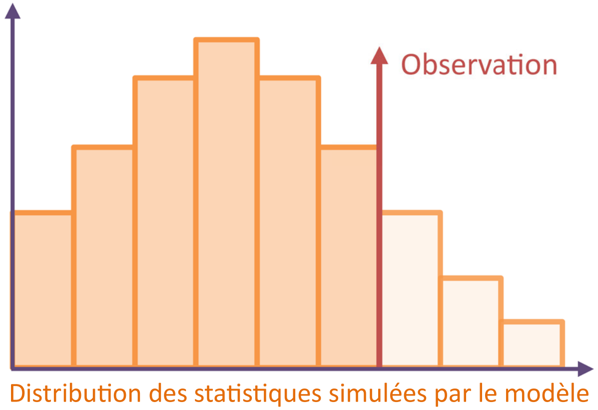
- **Posterior predictive checks**

  Is the selected model capable of reproducing observed biodiversity patterns?

Focus model
- ✓ A posteriori distribution of parameters that can be approximated by an estimator (mean, median, mode, etc.)

Simulation of *n* communities
- ✓ Calculating statistics (*summary statistics* or others)

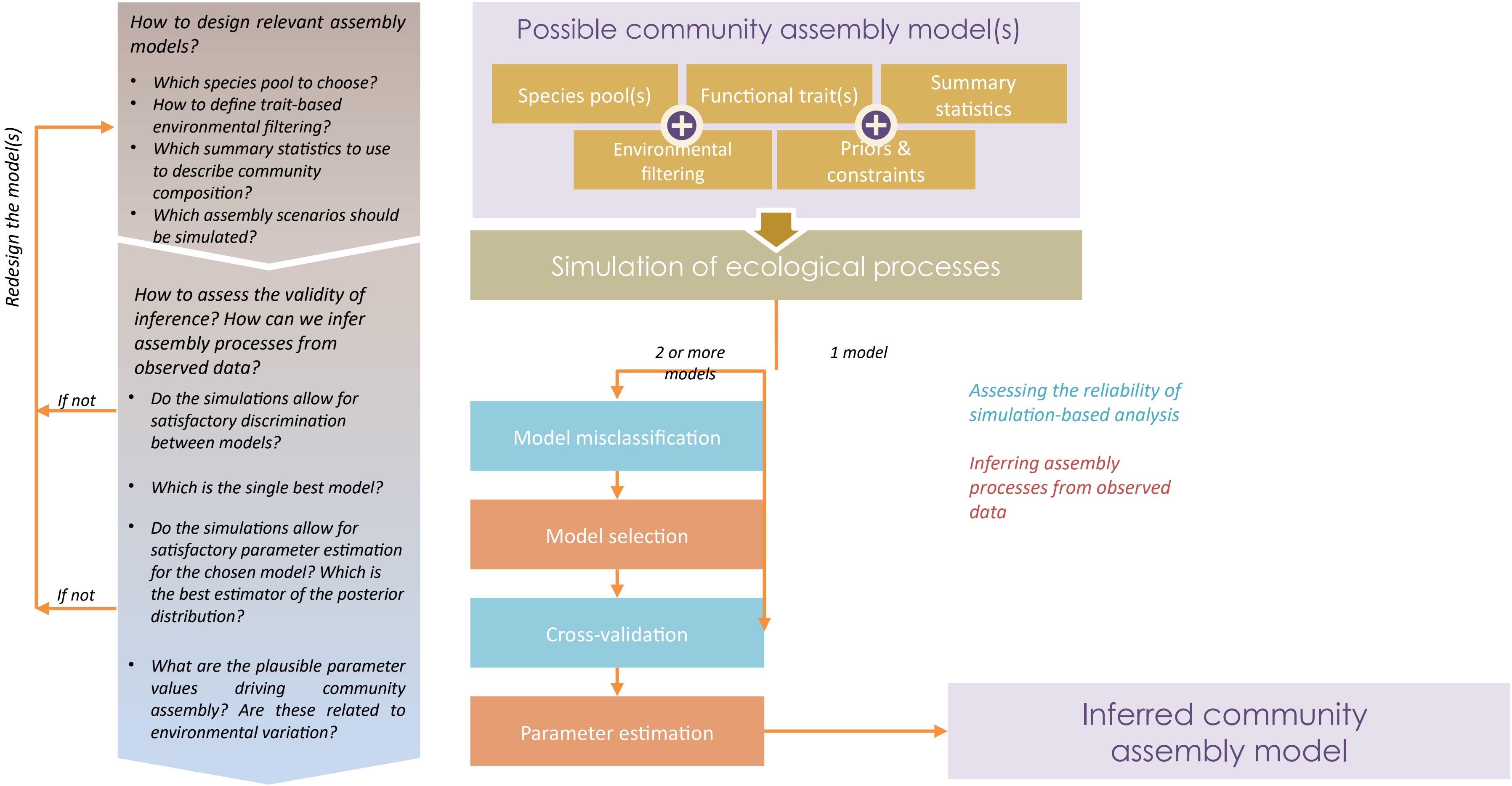
Comparison with the real community
- ✓ Probability that the observed statistic is different from the statsitics simulated with the model



Note: it is also possible to see whether certain species are systematically under/overestimated by the model, in which case another trait is involved

Inference framework

Approximate Bayesian Computation



Redesign the model(s)

How to design relevant assembly models?

- *Which species pool to choose?*
- *How to define trait-based environmental filtering?*
- *Which summary statistics to use to describe community composition?*
- *Which assembly scenarios should be simulated?*

How to assess the validity of inference? How can we infer assembly processes from observed data?

- *Do the simulations allow for satisfactory discrimination between models?*
- *Which is the single best model?*
- *Do the simulations allow for satisfactory parameter estimation for the chosen model? Which is the best estimator of the posterior distribution?*
- *What are the plausible parameter values driving community assembly? Are these related to environmental variation?*

If not

If not

Possible community assembly model(s)

Species pool(s)

Functional trait(s)

Summary statistics

Environmental filtering

Priors & constraints

Simulation of ecological processes

2 or more models

1 model

Model misclassification

Model selection

Cross-validation

Parameter estimation

Inferred community assembly model

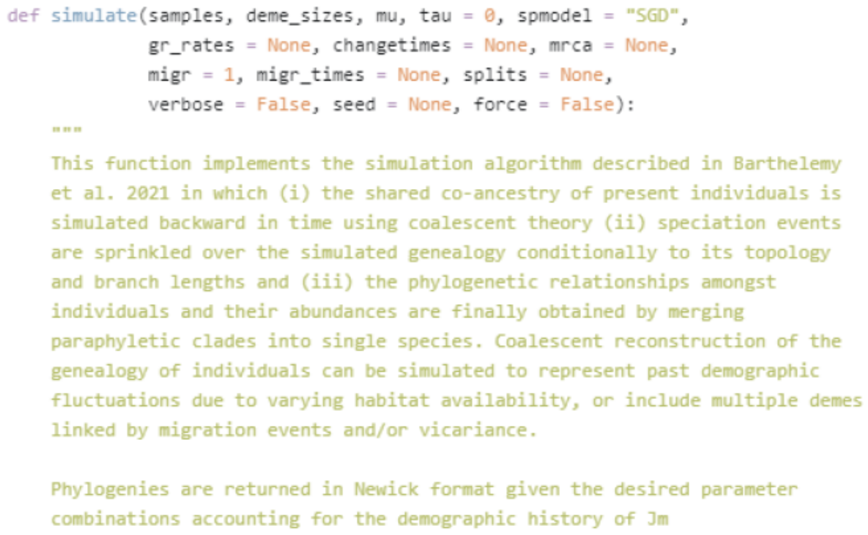
Assessing the reliability of simulation-based analysis

Inferring assembly processes from observed data

ecophylo project

Python package developed to model multispecies coalescence, performing inferences and hypothesis testing

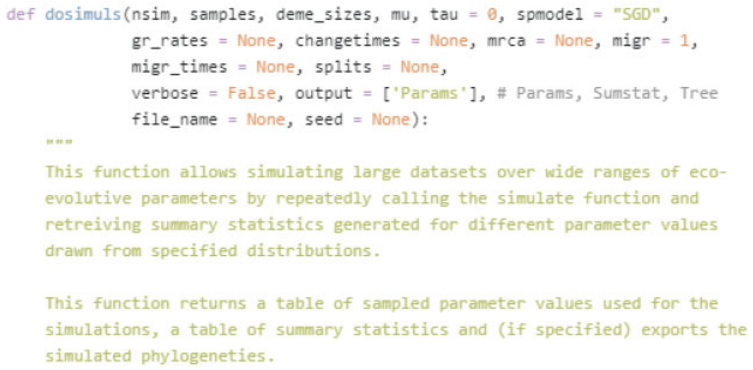
Simulate phylogeny according to a desired biogeographical scenario (specifying the various parameters of the eco-evolutionary model):



```
def simulate(samples, deme_sizes, mu, tau = 0, spmodel = "SGD",
             gr_rates = None, changetimes = None, mrca = None,
             migr = 1, migr_times = None, splits = None,
             verbose = False, seed = None, force = False):
    """
    This function implements the simulation algorithm described in Barthelemy
    et al. 2021 in which (i) the shared co-ancestry of present individuals is
    simulated backward in time using coalescent theory (ii) speciation events
    are sprinkled over the simulated genealogy conditionally to its topology
    and branch lengths and (iii) the phylogenetic relationships amongst
    individuals and their abundances are finally obtained by merging
    paraphyletic clades into single species. Coalescent reconstruction of the
    genealogy of individuals can be simulated to represent past demographic
    fluctuations due to varying habitat availability, or include multiple demes
    linked by migration events and/or vicariance.

    Phylogenies are returned in Newick format given the desired parameter
    combinations accounting for the demographic history of Jm
```

Simulate a large number of phylogenies according to a desired biogeographical scenario, but whose parameters vary according to prior distributions specified by the user:



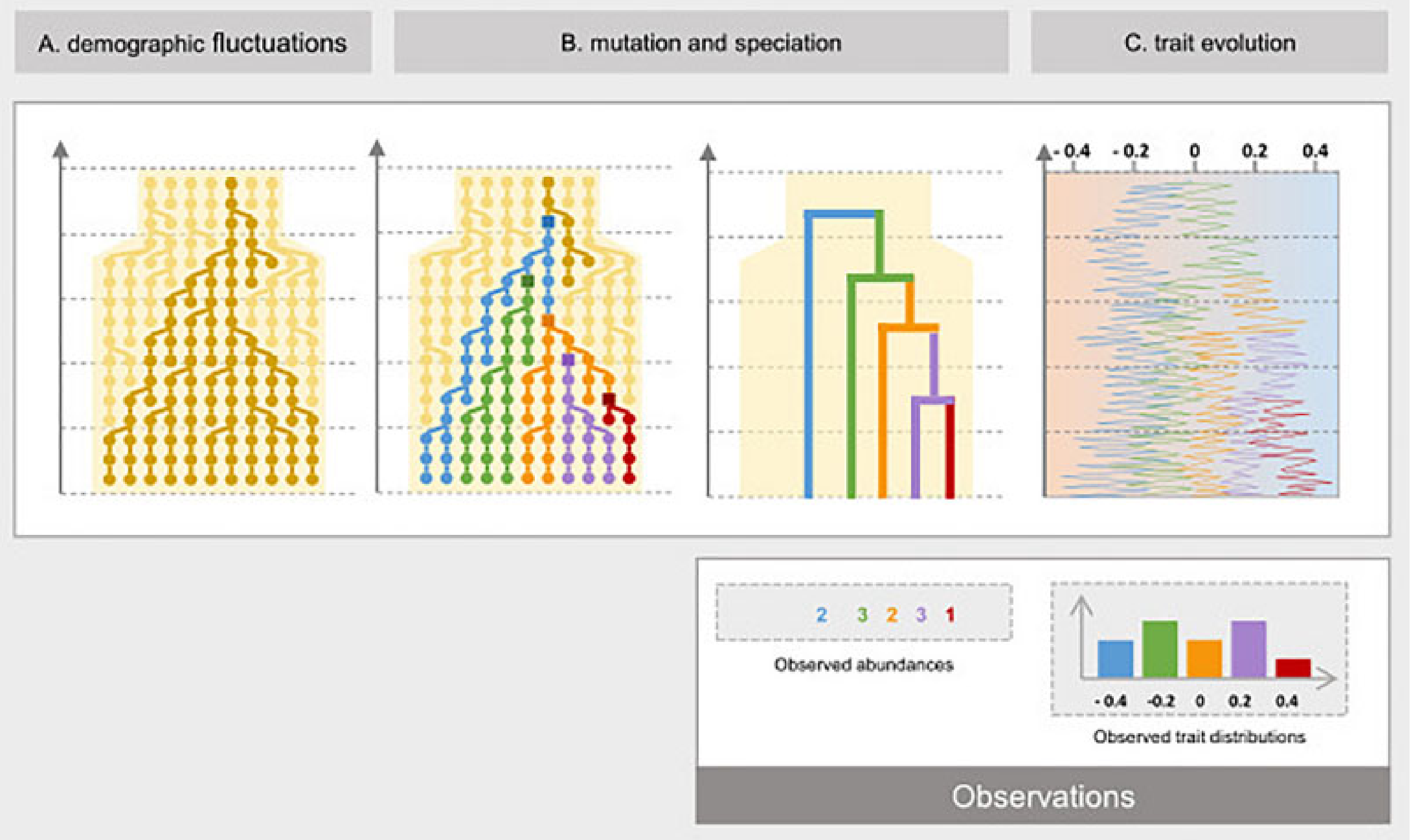
```
def dosimuls(nsim, samples, deme_sizes, mu, tau = 0, spmodel = "SGD",
             gr_rates = None, changetimes = None, mrca = None, migr = 1,
             migr_times = None, splits = None,
             verbose = False, output = ['Params'], # Params, Sumstat, Tree
             file_name = None, seed = None):
    """
    This function allows simulating large datasets over wide ranges of eco-
    evolutive parameters by repeatedly calling the simulate function and
    retreiving summary statistics generated for different parameter values
    drawn from specified distributions.

    This function returns a table of sampled parameter values used for the
    simulations, a table of summary statistics and (if specified) exports the
    simulated phylogeneties.
```

ecophylo project

Impact of habitat reduction (refugia)

Model components

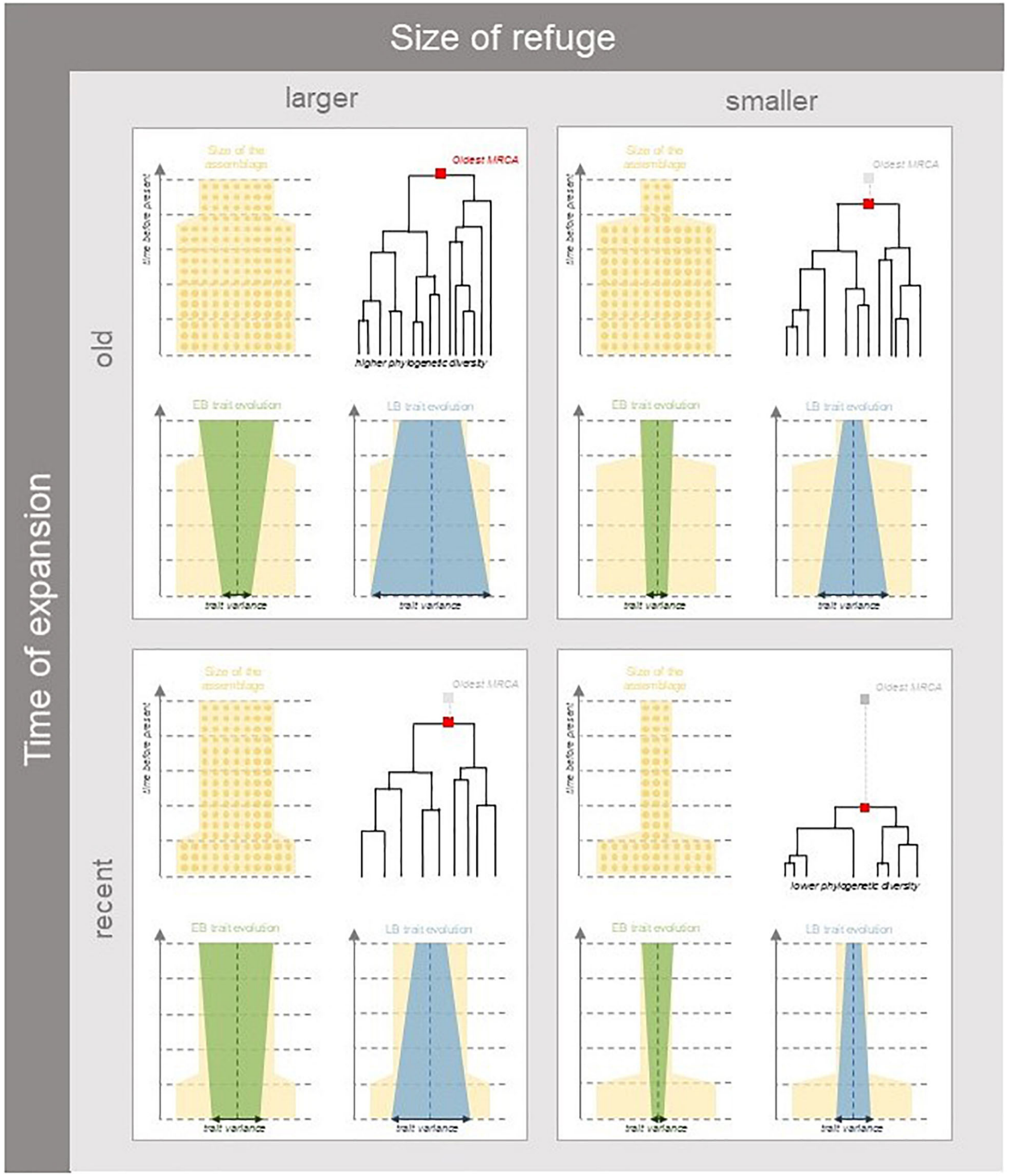


ecophylo project

Impact of habitat reduction (refugia)

Two parameters:

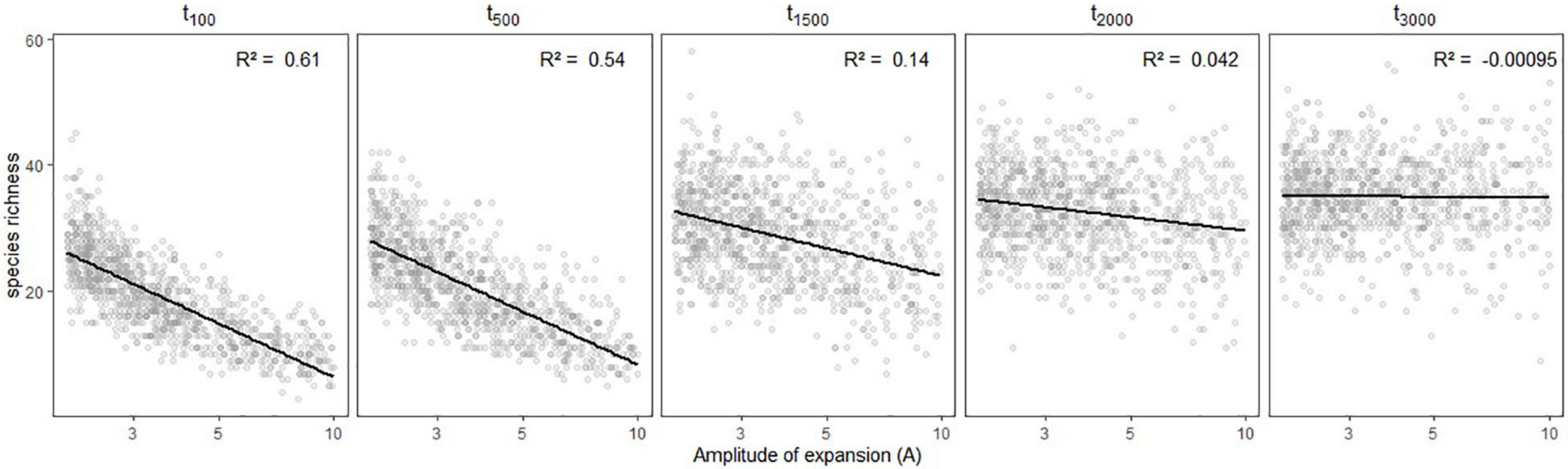
- Size of refuge
- Time of expansion



ecophylo project

Impact of habitat reduction (refugia)

Patterns of taxonomic diversity

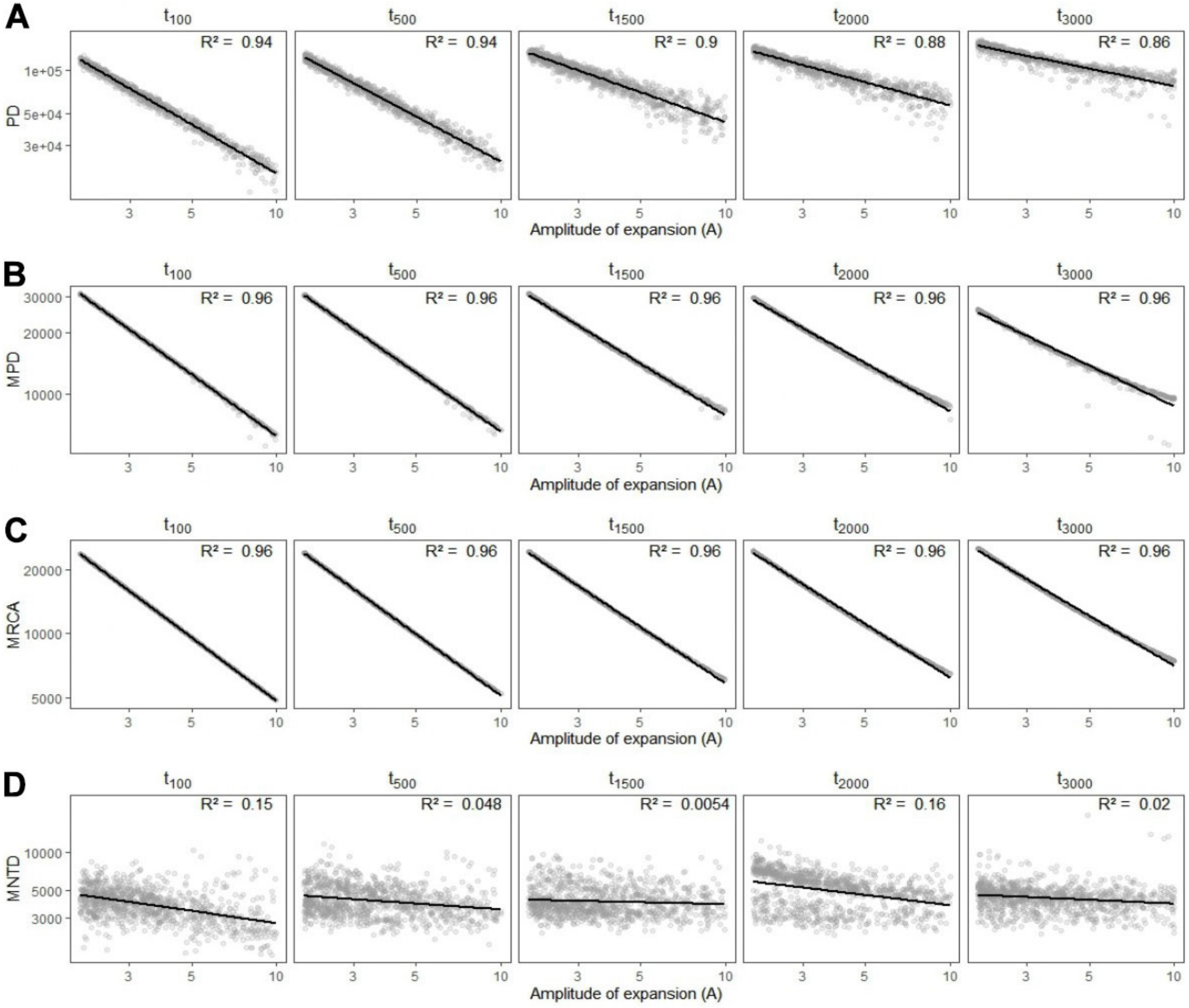


(Barthélémy et al. 2021)

ecophylo project

Impact of habitat reduction (refugia)

Patterns of phylogenetic diversity

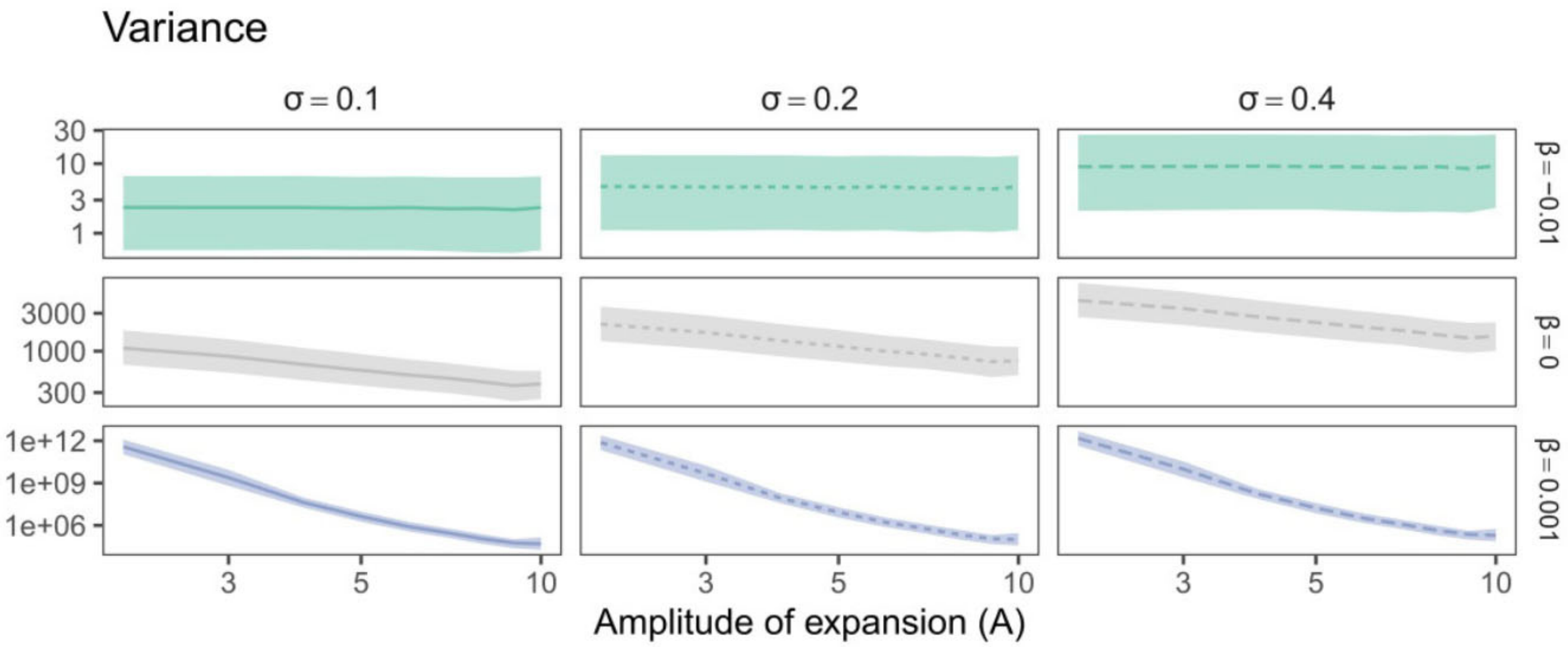


(Barthélémy et al. 2021)

ecophylo project

Impact of habitat reduction (refugia)

Patterns of functional diversity

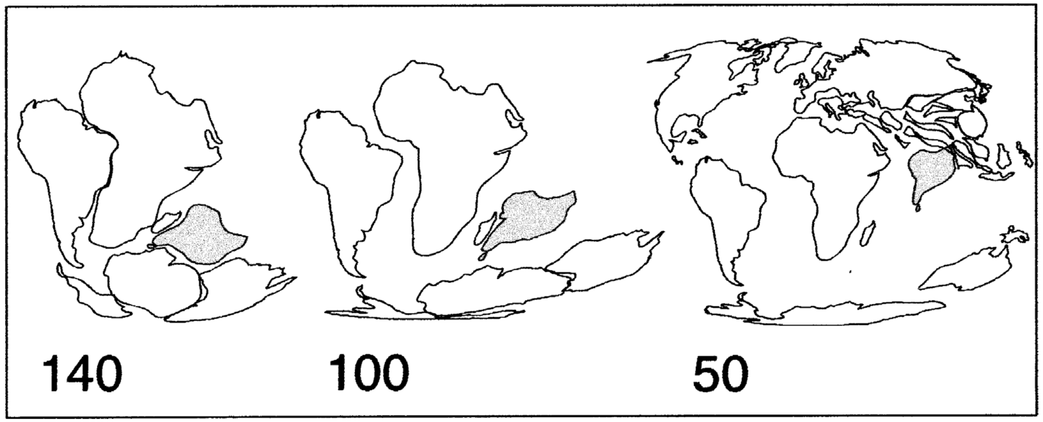


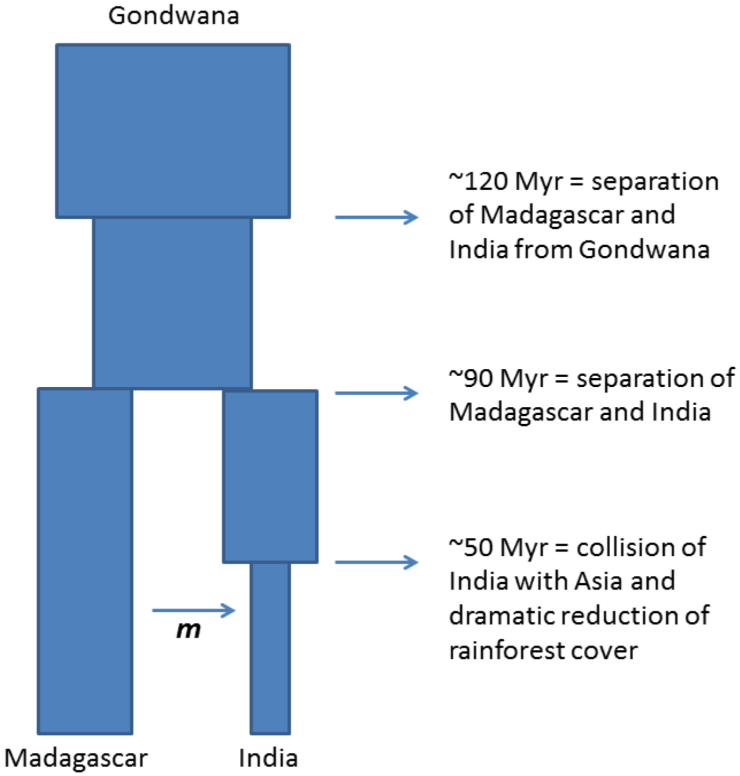
(Barthélémy et al. 2021)

ecophylo project

Tropical rainforests: fragmentation scenarios, area changes and migration

Exemple : rainforests of India and Madagascar





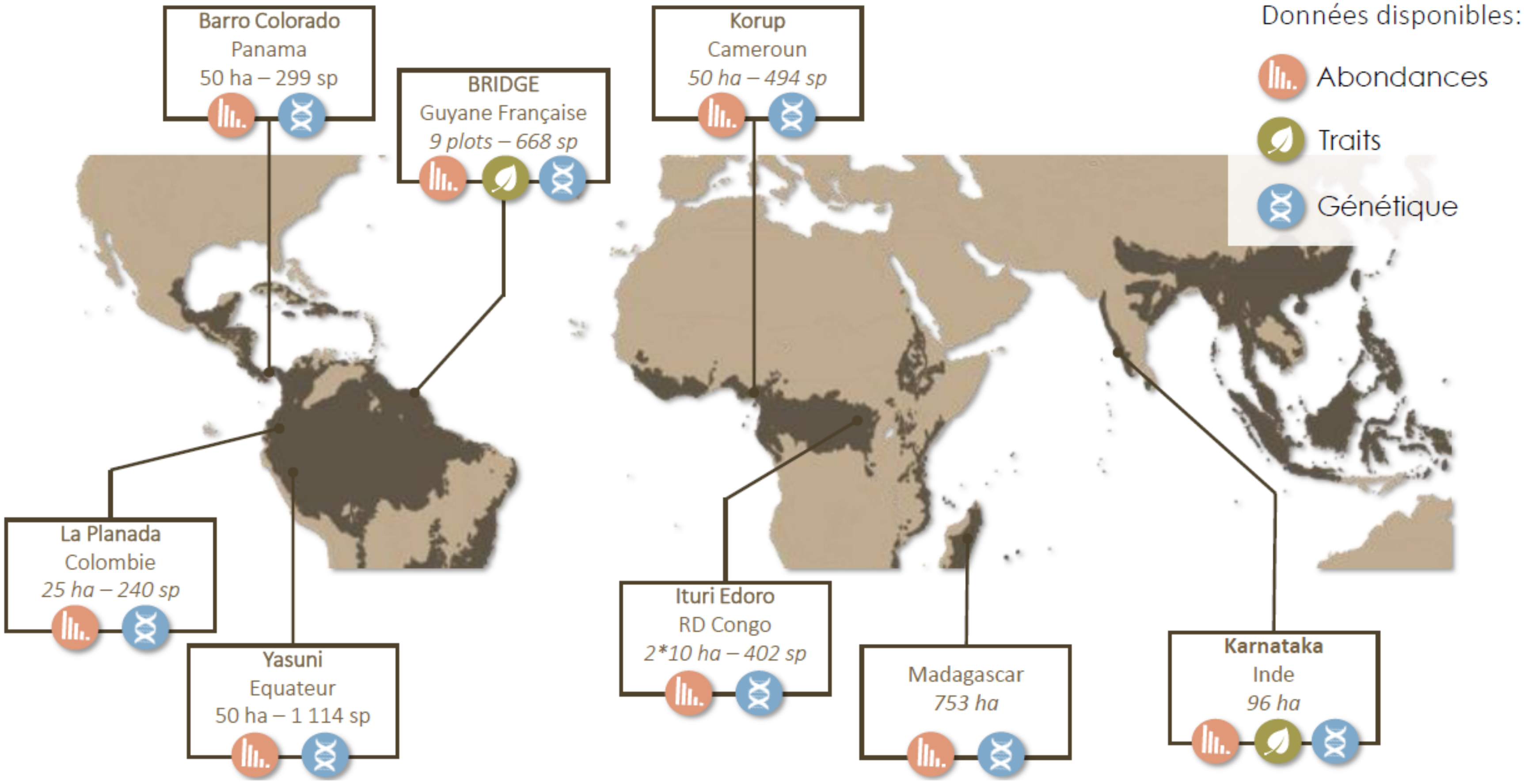
~120 Myr = separation of Madagascar and India from Gondwana

~90 Myr = separation of Madagascar and India

~50 Myr = collision of India with Asia and dramatic reduction of rainforest cover

ecophylo project

Tropical rainforests: forest inventories, genetic measurements and functional traits



Barro Colorado
Panama
50 ha – 299 sp

BRIDGE
Guyane Française
9 plots – 668 sp

Korup
Cameroun
50 ha – 494 sp

Données disponibles:
Abondances
Traits
Génétique

La Planada
Colombie
25 ha – 240 sp

Yasuni
Equateur
50 ha – 1 114 sp

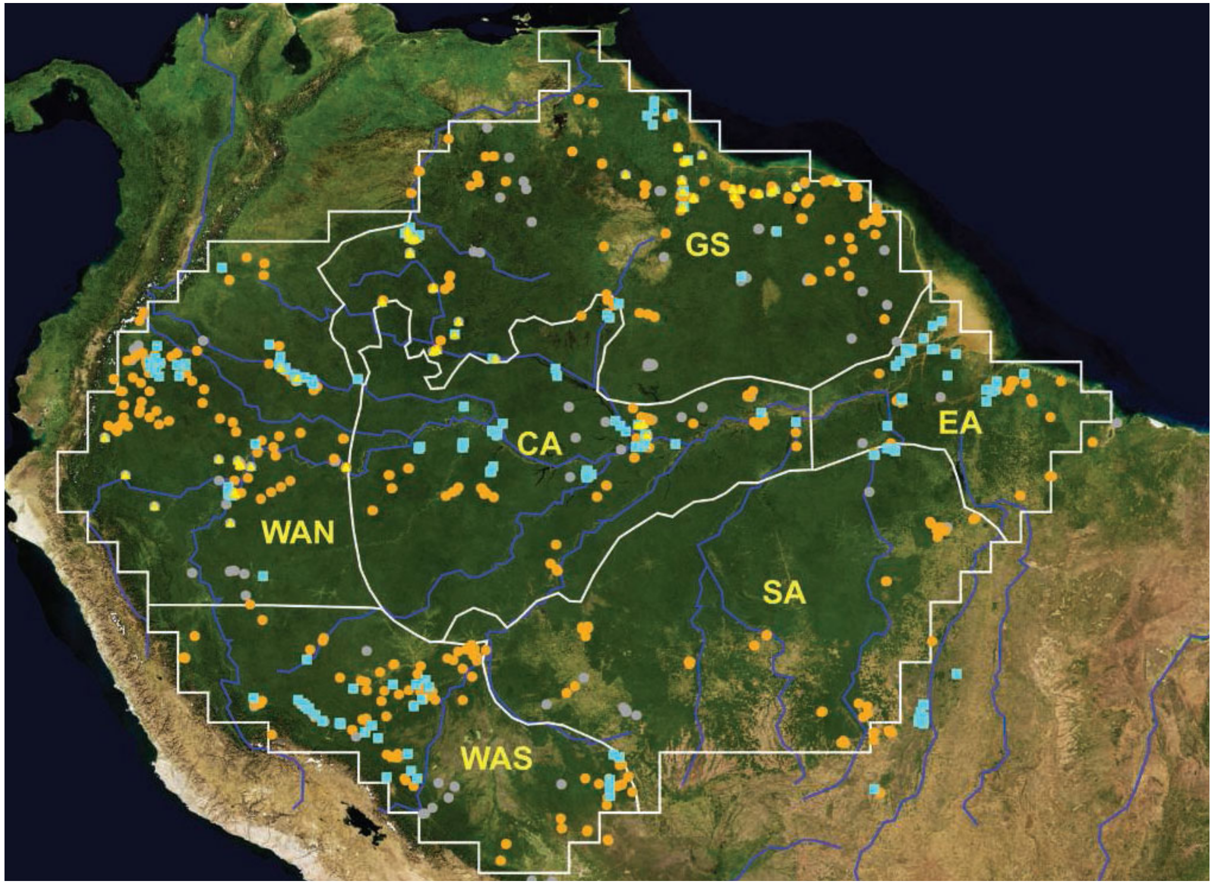
Ituri Edoro
RD Congo
*2*10 ha – 402 sp*

Madagascar
753 ha

Karnataka
Inde
96 ha

ecophylo project

Tropical rainforests: forest inventories, genetic measurements and functional traits



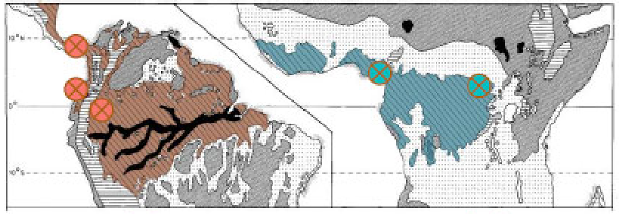
ATDN network within Amazonia

(ter Steege et al. 2013)

First results

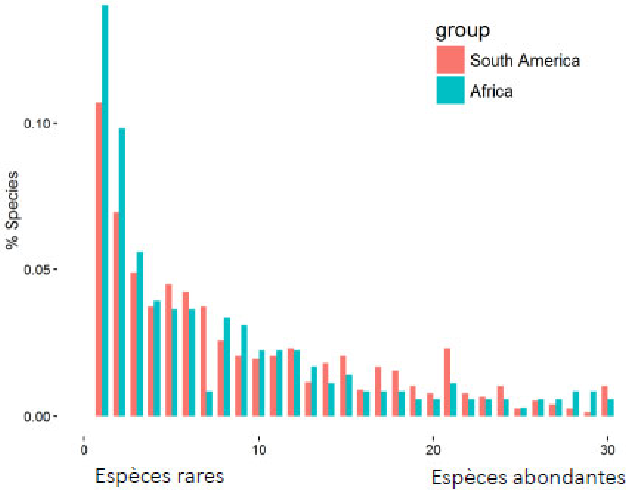
Imprint of forest habitat reduction during Last Glacial Maximum

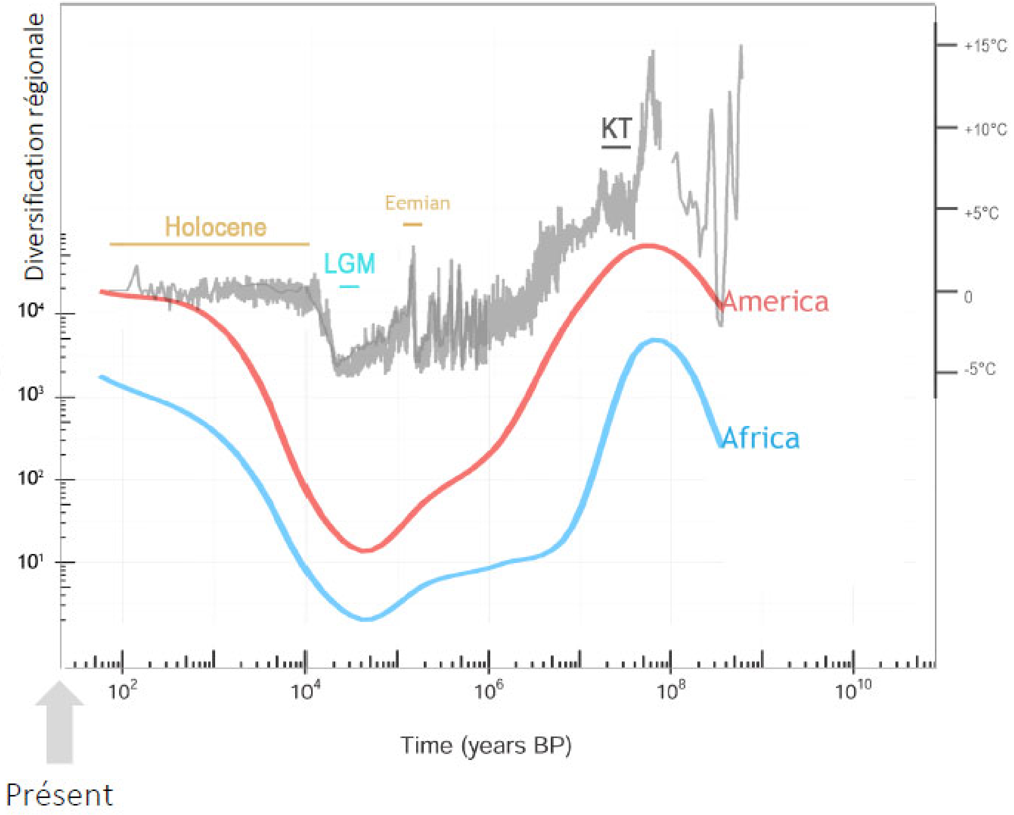
Les données:



~ 47,000 ind – 777 sp.

~ 30,000 ind – 357 sp.





(Munoz & Tournebize, in prep)

ecophylo project

Methodological challenges

- Spatially-explicit coalescent
- Coalescent with selection
- Influence of biotic interactions

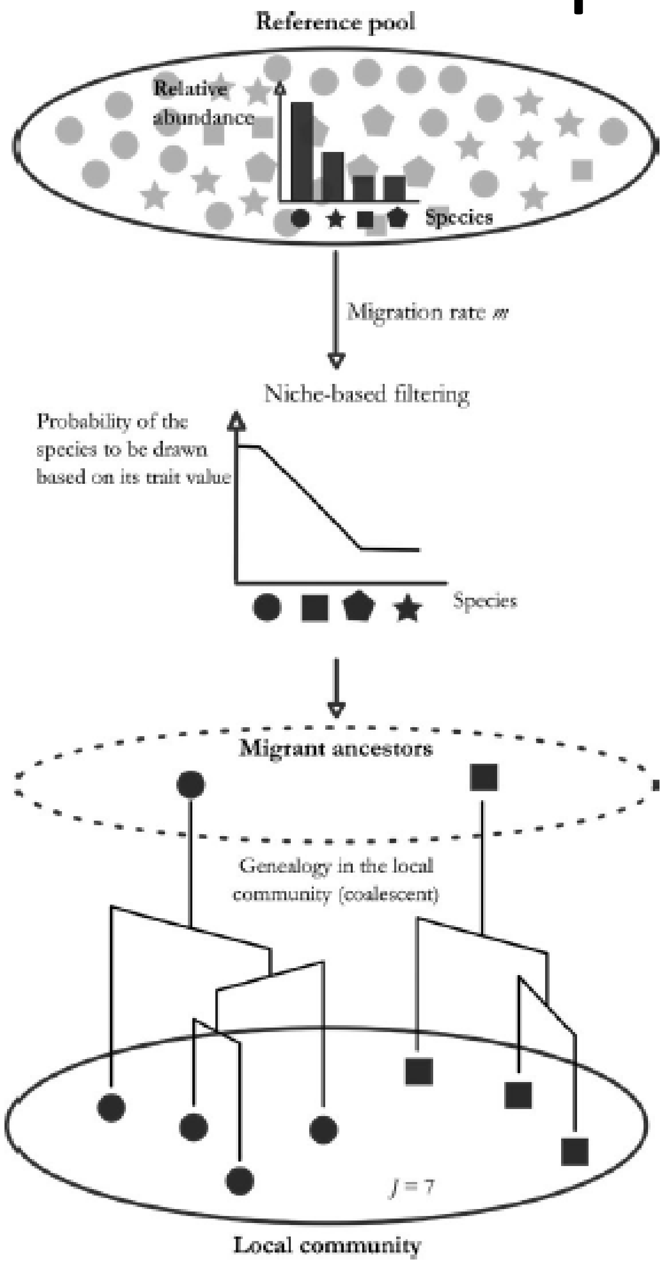
Evaluating inference and test capacities

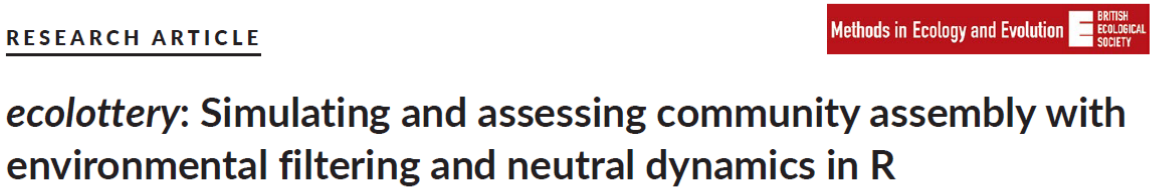
- Misclassification of alternative models
- Cross-validation for parameter estimation
- Predictive checks for assessing model consistency

Which data

- Different facets of biological diversity
- Phylogenetic + intraspecific genetic diversity
- Integrating fossil data

ecophylo vs. ecolottery





RESEARCH ARTICLE

Methods in Ecology and Evolution | BRITISH ECOLOGICAL SOCIETY

ecolottery: Simulating and assessing community assembly with environmental filtering and neutral dynamics in R

- ➢ Coalescent-based model of community assembly
- ➢ With environmental filtering depending on species trait values
- ➢ With reference to a given regional pool of species

Perspective to integrate the eco-evolutionary and community assembly models in a single framework

(Munoz et al. 2018)

UNIVERSITÉ Grenoble Alpes

LiPhy
Laboratoire interdisciplinaire de Physique

Modelling eco-evolutionary dynamics to answer major questions in macroecology and historical biogeography

François Munoz

Ecole de recherche de la Chaire MMB

Friday June 14th 2024

Cypripedium calceolus (Orchidacées)

Clematis alpina
(Renonculacées)

Orchis pallens (Orchidacées)

Dactylorhiza latifolia (Orchidacées)

Pelouse steppique

Stipa pennata (Poacées)

Pétroglyphe

Iphiclides podalirius, Flambé

Libelloides coccajus, Ascalaphe soufré

Neotinea ustulata Orchidacées)

Achillea tomentosa (Astéracées)

Anthericum liliago (Liliacées)

Matthiola fruticulosa ssp *valesiaca*
(Brassicacées)