

# Étude de quelques hypothèses nouvelles pour la modélisation de la macroévolution

## Soutenance de thèse

Marc Manceau

21 Juin 2018

Composition du jury :

Nicolas Galtier Rapporteur

Lacey Knowles Rapporteuse (absente)

Amaury Lambert Directeur de thèse

Nicolas Lartillot Examineur

Hélène Morlon Directrice de thèse

Sandrine Pavoine Examinatrice

Stéphane Robin Examineur



# Macroévolution

## Échelle de temps longue



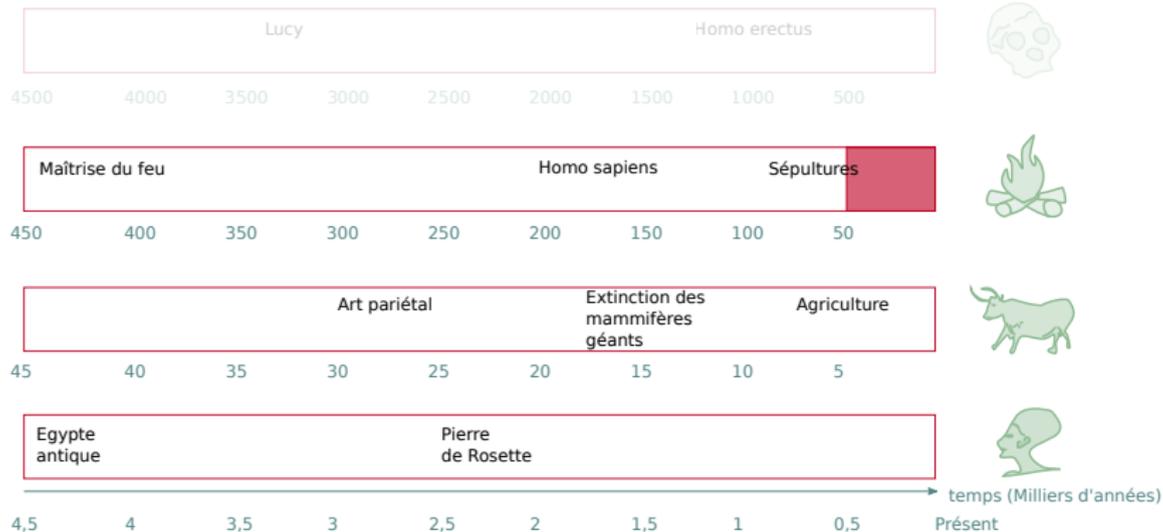
# Macroévolution

## Échelle de temps longue



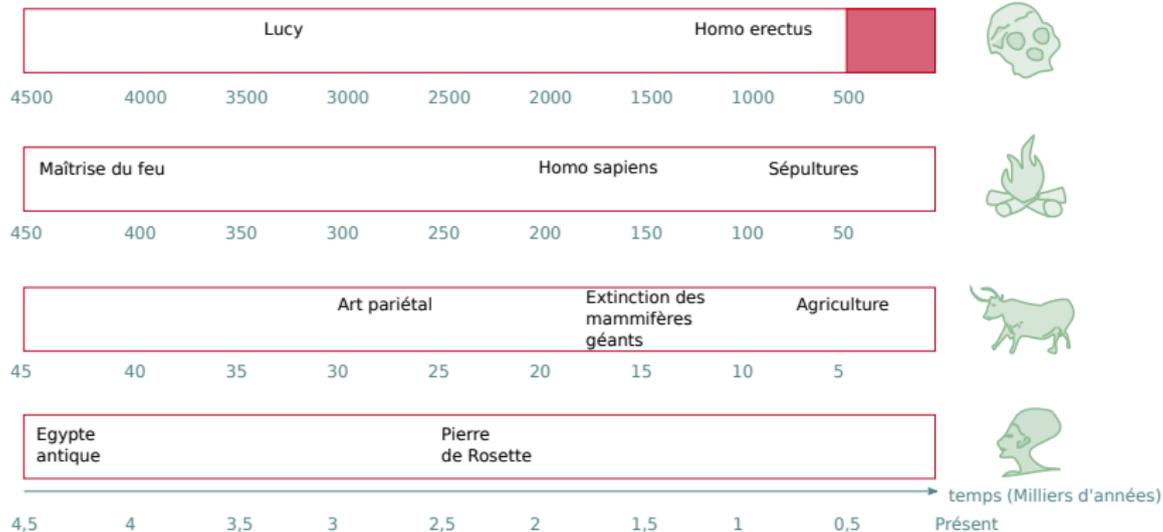
# Macroévolution

## Échelle de temps longue



# Macroévolution

## Échelle de temps longue



# Macroévolution

## Échelle de temps longue



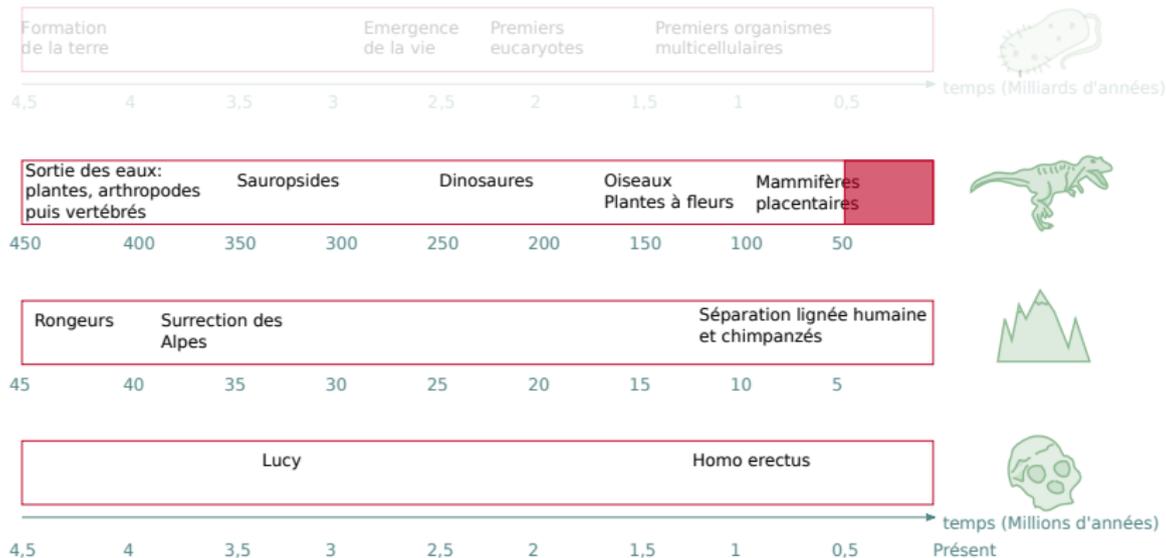
# Macroévolution

## Échelle de temps longue



# Macroévolution

## Échelle de temps longue



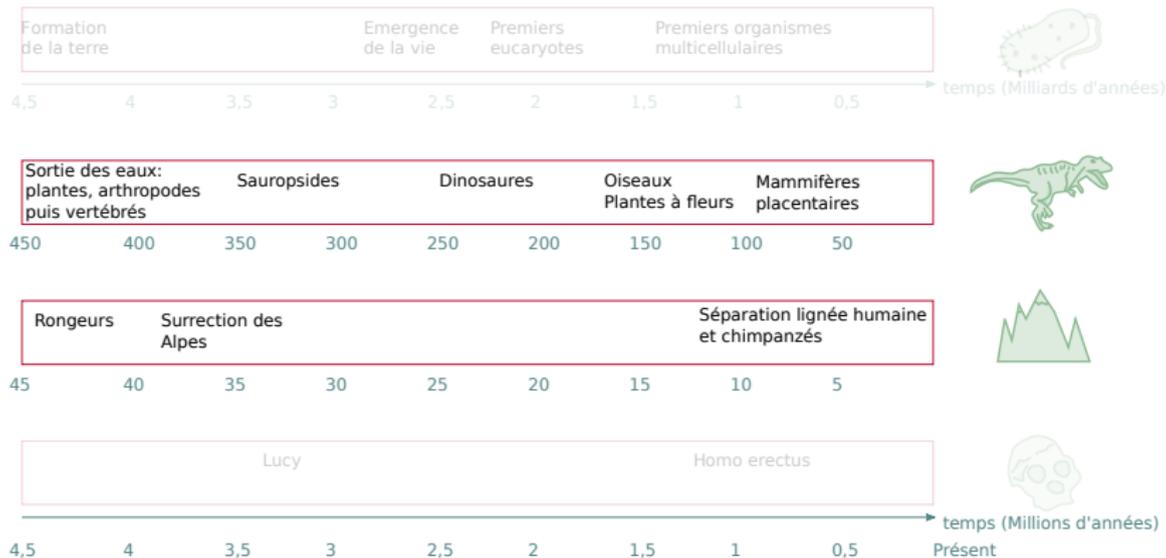
# Macroévolution

## Échelle de temps longue



# Macroévolution

## Échelle de temps longue



# Macroévolution

Accessible par l'étude de fossiles



~ 500 Millions d'années.



~ 300 Millions d'années.



~ 40 Millions d'années.

Crédits photo: Keith Schengili-Roberts, Woudloper, Didier Descouens.

# Macroévolution

Accessible par l'étude de la biodiversité actuelle

## ► Phénotypes des organismes



## ► Séquences génétiques

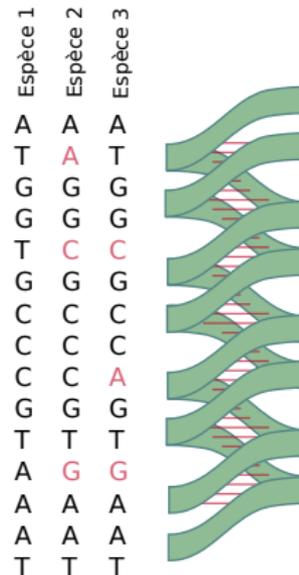
# Macroévolution

Accessible par l'étude de la biodiversité actuelle

## ► Phénotypes des organismes



## ► Séquences génétiques



# Macroévolution

## Approche comparative

Questions concrètes :

1. Qu'est-ce qui détermine leur couleur, leur taille, la forme des branchies ?
2. A quoi ressemblait l'ancêtre commun de tous ces nudibranches ?
3. Comment les classifie-t-on en espèces différentes ?
4. A quel rythme se produisent les événements de spéciation/extinction ?



Approche comparative :

1. Comparer les caractères observés dans la nature.
2. Proposer des scénarios d'évolution:
  - ▶ modèles probabilistes
  - ▶ origine mécaniste
3. Ajuster ces modèles aux données observées.

# Macroévolution

## Approche comparative

Questions concrètes :

1. Qu'est-ce qui détermine leur couleur, leur taille, la forme des branchies ?
2. A quoi ressemblait l'ancêtre commun de tous ces nudibranches ?
3. Comment les classifie-t-on en espèces différentes ?
4. A quel rythme se produisent les événements de spéciation/extinction ?



Approche comparative :

1. Comparer les caractères observés dans la nature.
2. Proposer des scénarios d'évolution:
  - ▶ modèles probabilistes
  - ▶ origine mécaniste
3. Ajuster ces modèles aux données observées.

# Macroévolution

## Approche comparative

Questions concrètes :

1. Qu'est-ce qui détermine leur couleur, leur taille, la forme des branchies ?
2. A quoi ressemblait l'ancêtre commun de tous ces nudibranches ?
3. Comment les classifie-t-on en espèces différentes ?
4. A quel rythme se produisent les événements de spéciation/extinction ?



Approche comparative :

1. Comparer les caractères observés dans la nature.
2. Proposer des scénarios d'évolution:
  - ▶ modèles probabilistes
  - ▶ origine mécaniste
3. Ajuster ces modèles aux données observées.

# Macroévolution

## Approche comparative

Questions concrètes :

1. Qu'est-ce qui détermine leur couleur, leur taille, la forme des branchies ?
2. A quoi ressemblait l'ancêtre commun de tous ces nudibranches ?
3. Comment les classifie-t-on en espèces différentes ?
4. A quel rythme se produisent les événements de spéciation/extinction ?



Approche comparative :

1. Comparer les caractères observés dans la nature.
2. Proposer des scénarios d'évolution:
  - ▶ modèles probabilistes
  - ▶ origine mécaniste
3. Ajuster ces modèles aux données observées.

# Macroévolution

## Approche comparative

Questions concrètes :

1. Qu'est-ce qui détermine leur couleur, leur taille, la forme des branchies ?
2. A quoi ressemblait l'ancêtre commun de tous ces nudibranches ?
3. Comment les classifie-t-on en espèces différentes ?
4. A quel rythme se produisent les événements de spéciation/extinction ?



Approche comparative :

1. Comparer les caractères observés dans la nature.
2. Proposer des scénarios d'évolution:
  - ▶ modèles probabilistes
  - ▶ origine mécaniste
3. Ajuster ces modèles aux données observées.

# Macroévolution

## Approche comparative

Questions concrètes :

1. Qu'est-ce qui détermine leur couleur, leur taille, la forme des branchies ?
2. A quoi ressemblait l'ancêtre commun de tous ces nudibranches ?
3. Comment les classifie-t-on en espèces différentes ?
4. A quel rythme se produisent les événements de spéciation/extinction ?



Approche comparative :

1. Comparer les caractères observés dans la nature.
2. Proposer des scénarios d'évolution:
  - ▶ modèles probabilistes
  - ▶ origine mécaniste
3. Ajuster ces modèles aux données observées.

# Macroévolution

## Approche comparative

Questions concrètes :

1. Qu'est-ce qui détermine leur couleur, leur taille, la forme des branchies ?
2. A quoi ressemblait l'ancêtre commun de tous ces nudibranches ?
3. Comment les classifie-t-on en espèces différentes ?
4. A quel rythme se produisent les événements de spéciation/extinction ?



Approche comparative :

1. Comparer les caractères observés dans la nature.
2. Proposer des scénarios d'évolution:
  - ▶ modèles probabilistes
  - ▶ origine mécaniste
3. Ajuster ces modèles aux données observées.

# Macroévolution

## Approche comparative

Questions concrètes :

1. Qu'est-ce qui détermine leur couleur, leur taille, la forme des branchies ?
2. A quoi ressemblait l'ancêtre commun de tous ces nudibranches ?
3. Comment les classifie-t-on en espèces différentes ?
4. A quel rythme se produisent les événements de spéciation/extinction ?

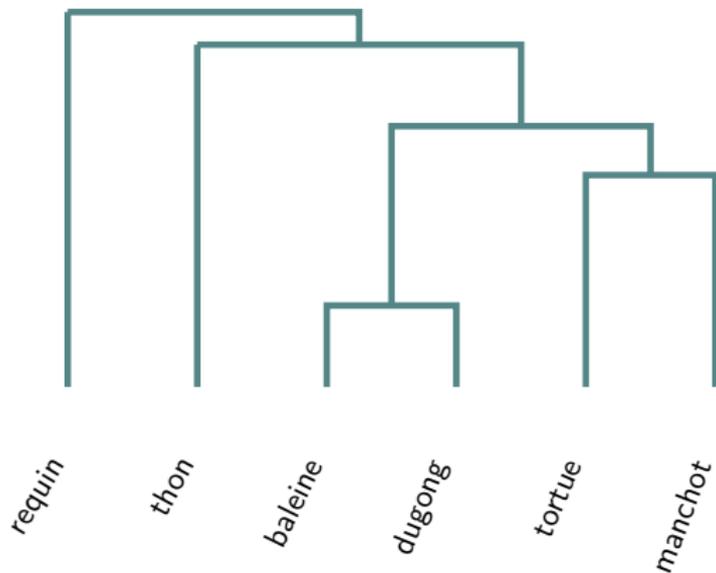


Approche comparative :

1. Comparer les caractères observés dans la nature.
2. Proposer des scénarios d'évolution:
  - ▶ modèles probabilistes
  - ▶ origine mécaniste
3. Ajuster ces modèles aux données observées.

# Arbre phylogénétique

Relations de parenté entre espèces



Feuille

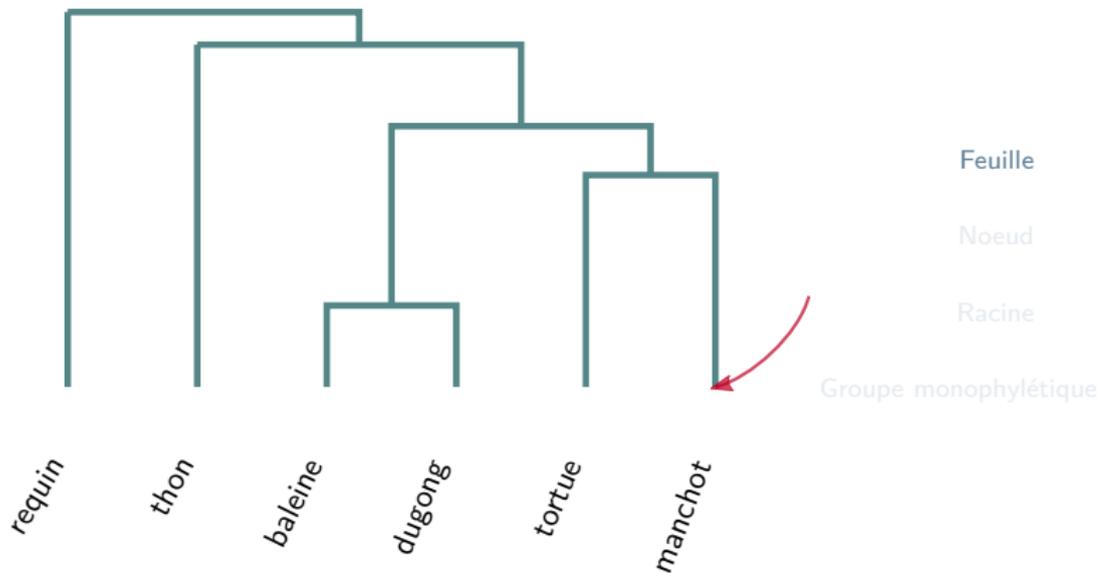
Noeud

Racine

Groupe monophylétique

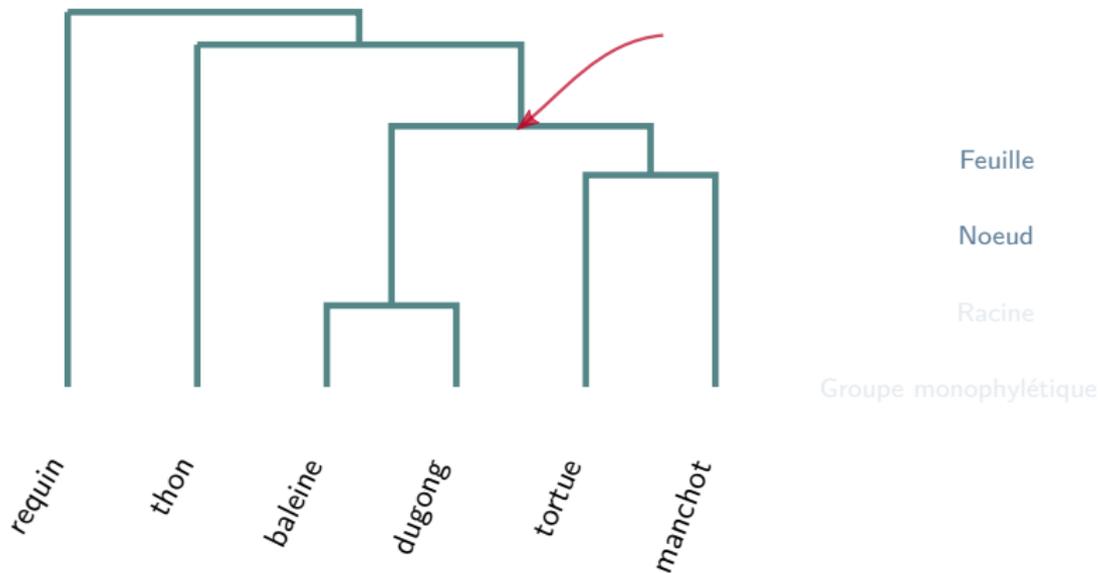
# Arbre phylogénétique

Relations de parenté entre espèces



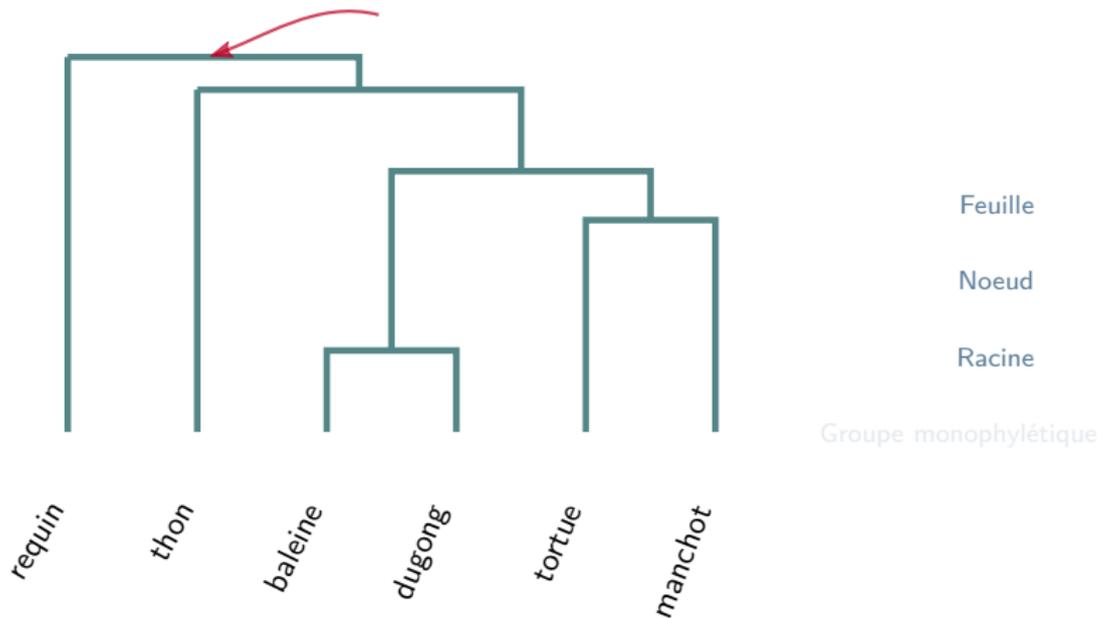
# Arbre phylogénétique

Relations de parenté entre espèces



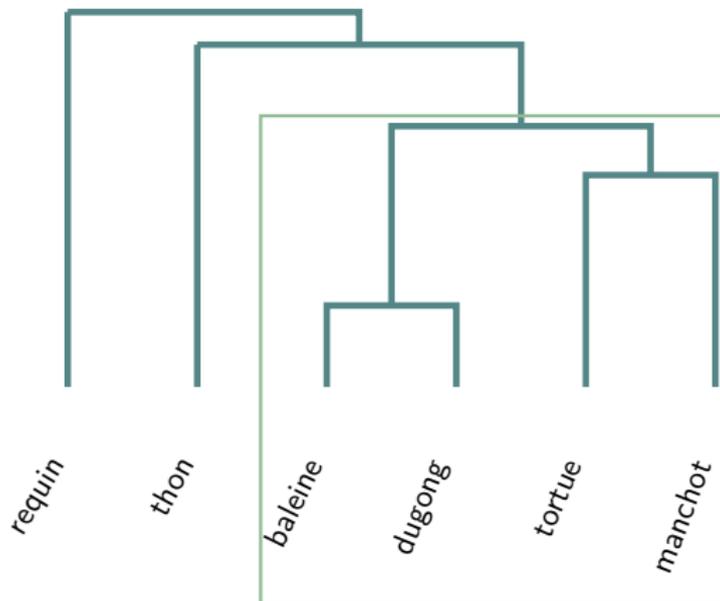
# Arbre phylogénétique

Relations de parenté entre espèces



# Arbre phylogénétique

Relations de parenté entre espèces



Feuille

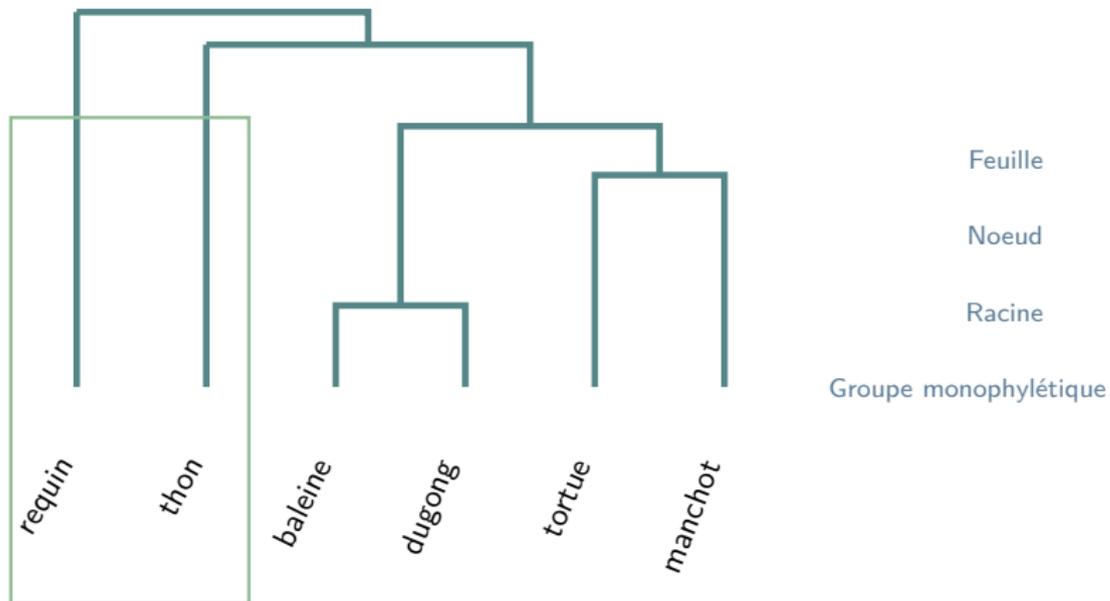
Noeud

Racine

Groupe monophylétique

# Arbre phylogénétique

Relations de parenté entre espèces



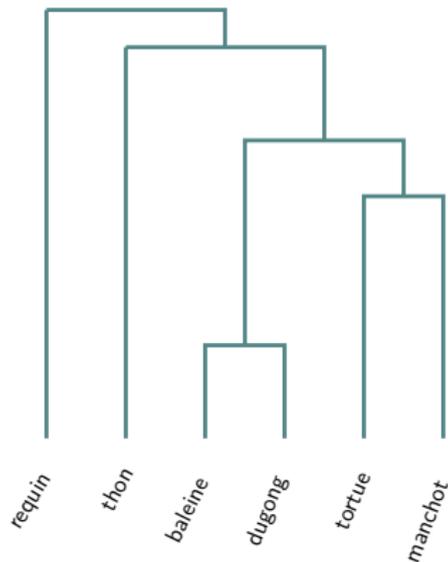
# Surcouche de modélisation probabiliste

Pour trois types d'observations différents

► Phylogénies

► Phénotypes

► Séquences génétiques

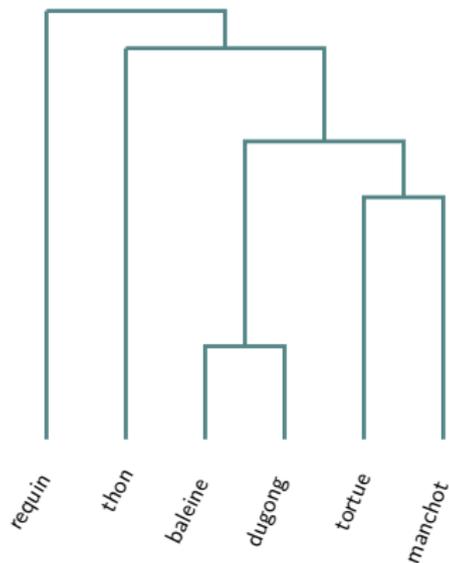


Modèle individu-centré minimaliste avec une dynamique de population hors-équilibre.

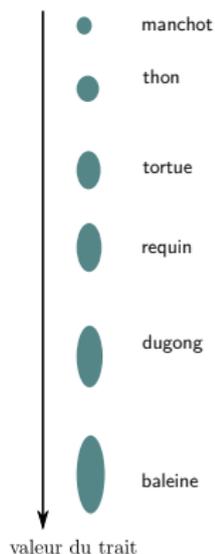
# Surcouche de modélisation probabiliste

Pour trois types d'observations différents

## ► Phylogénies



## ► Phénotypes



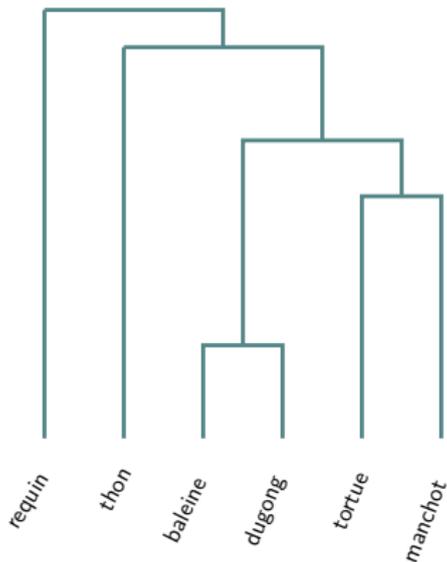
## ► Séquences génétiques

Caractères impliqués dans des interactions inter-spécifiques.

# Surcouche de modélisation probabiliste

Pour trois types d'observations différents

## ► Phylogénies



## ► Phénotypes

## ► Séquences génétiques

	requin	thon	baleine	dugong	tortue	manchot
A	'	'	'	'	'	'
T	'	A	A	'	'	'
C	G	G	G	G	G	G
G	'	'	'	'	'	'
A	C	'	'	'	'	'
A	T	'	'	'	'	'
G	'	'	'	A	A	'
C	'	'	'	'	'	'
:	:	:	:	:	:	:

Évolution ponctuelle de certaines séquences génétique.

## Introduction

Macroévolution

Arbre phylogénétique

Surcouche de modélisation probabiliste

## Diversification

État de l'art

La définition d'espèce en modélisation

Application à l'étude de la diversification



## Évolution phénotypique

État de l'art

Coévolution des caractères de différentes espèces

Proposition d'un cadre de modélisation

Exemples d'applications

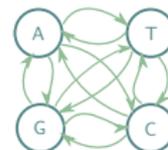


## Évolution moléculaire

Horloge moléculaire

Modélisation proposée

Méthode d'inférence



## Perspectives

Résumé

Liens entre les chapitres

## Introduction

- Macroévolution
- Arbre phylogénétique
- Surcouche de modélisation probabiliste

## Diversification

- État de l'art
- La définition d'espèce en modélisation
- Application à l'étude de la diversification



## Évolution phénotypique

- État de l'art
- Coévolution des caractères de différentes espèces
- Proposition d'un cadre de modélisation
- Exemples d'applications



## Évolution moléculaire

- Horloge moléculaire
- Modélisation proposée
- Méthode d'inférence

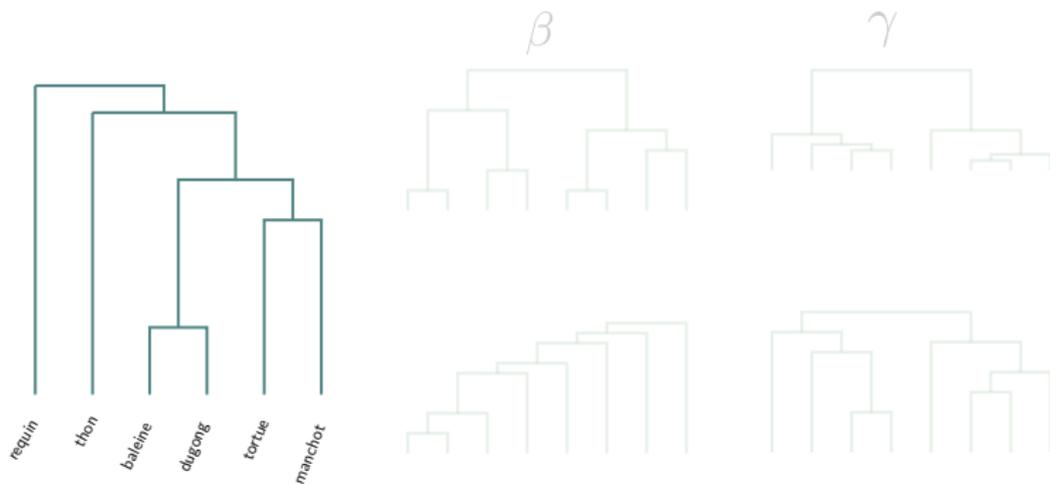


## Perspectives

- Résumé
- Liens entre les chapitres

# État de l'art

## Principes de l'approche de modélisation

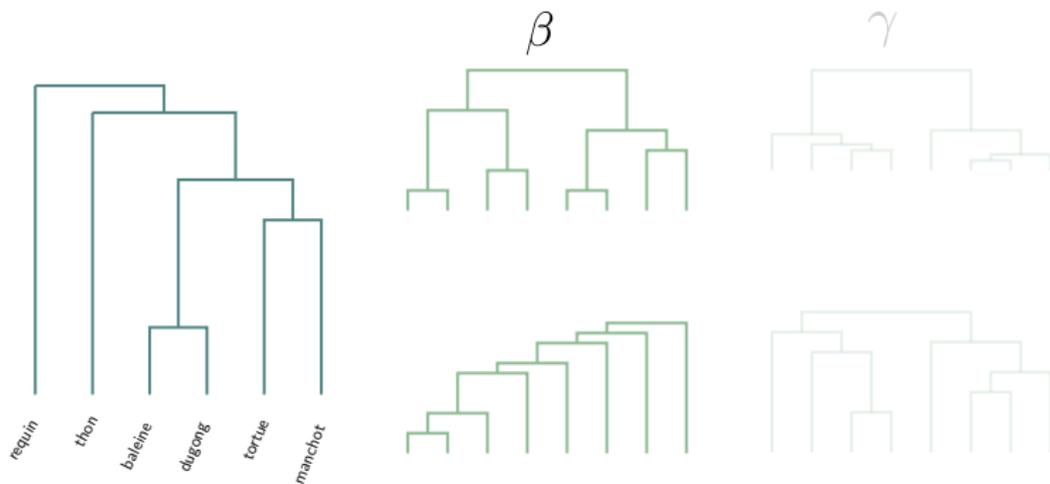


- ▶ Arbres plutôt déséquilibrés.
- ▶ Ayant des noeuds proches de la racine.
- ▶ Peut-on reproduire ces observations à l'aide d'un modèle individu-centré minimaliste ?

Aldous, 2001, *Stat. Sci.*  
Blum et François, 2006, *Syst. Biol.*

# État de l'art

## Principes de l'approche de modélisation

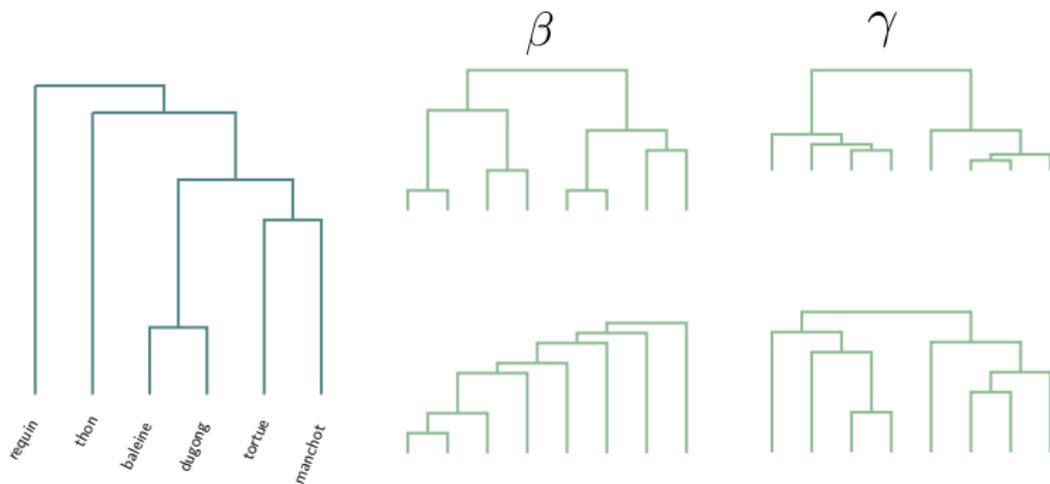


- ▶ Arbres plutôt déséquilibrés.
- ▶ Ayant des noeuds proches de la racine.
- ▶ Peut-on reproduire ces observations à l'aide d'un modèle individu-centré minimaliste ?

Aldous, 2001, *Stat. Sci.*  
Blum et François, 2006, *Syst. Biol.*

# État de l'art

## Principes de l'approche de modélisation

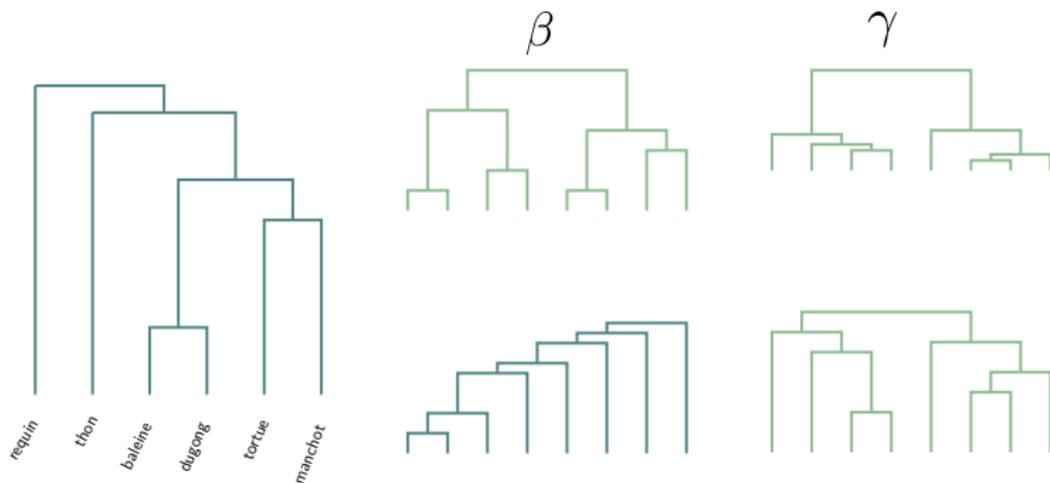


- ▶ Arbres plutôt déséquilibrés.
- ▶ Ayant des noeuds proches de la racine.
- ▶ Peut-on reproduire ces observations à l'aide d'un modèle individu-centré minimaliste ?

Aldous, 2001, *Stat. Sci.*  
Blum et François, 2006, *Syst. Biol.*

# État de l'art

## Principes de l'approche de modélisation

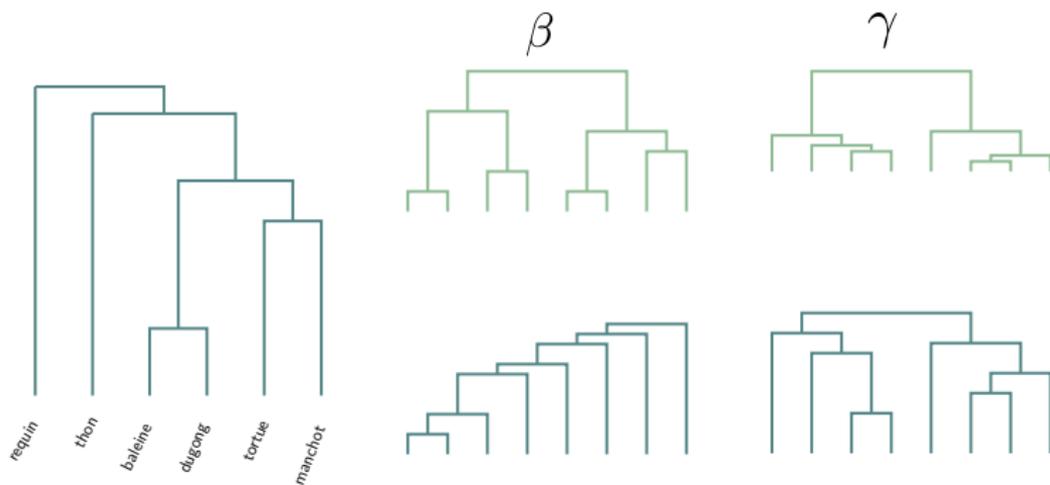


- ▶ Arbres plutôt déséquilibrés.
- ▶ Ayant des noeuds proches de la racine.
- ▶ Peut-on reproduire ces observations à l'aide d'un modèle individu-centré minimaliste ?

Aldous, 2001, *Stat. Sci.*  
Blum et François, 2006, *Syst. Biol.*

# État de l'art

## Principes de l'approche de modélisation

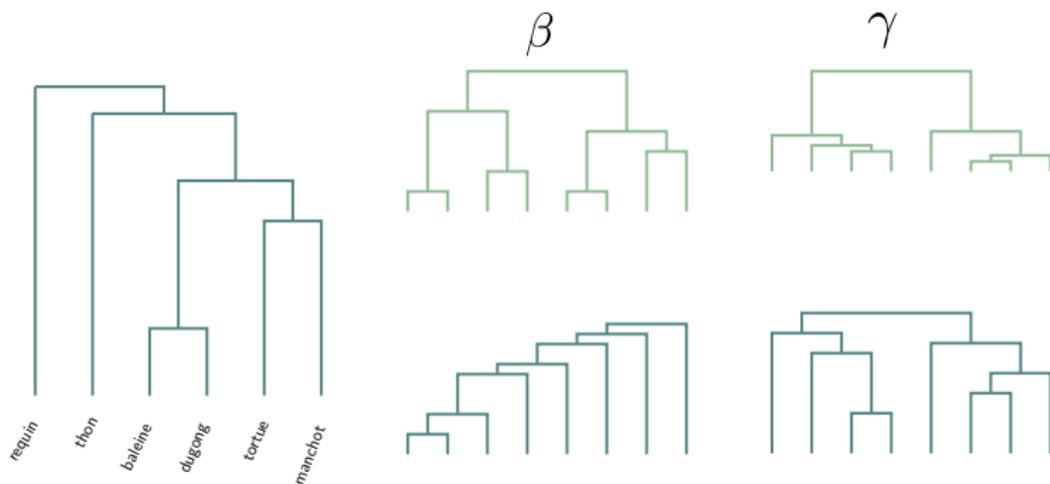


- ▶ Arbres plutôt déséquilibrés.
- ▶ Ayant des noeuds proches de la racine.
- ▶ Peut-on reproduire ces observations à l'aide d'un modèle individu-centré minimaliste ?

Aldous, 2001, *Stat. Sci.*  
Blum et François, 2006, *Syst. Biol.*

# État de l'art

## Principes de l'approche de modélisation



- ▶ Arbres plutôt déséquilibrés.
- ▶ Ayant des noeuds proches de la racine.
- ▶ Peut-on reproduire ces observations à l'aide d'un modèle individu-centré minimaliste ?

Aldous, 2001, *Stat. Sci.*  
Blum et François, 2006, *Syst. Biol.*

# État de l'art

## Modèles espèces-centrés



### Chaque espèce :

- ▶ donne naissance à taux  $b$ .
- ▶ est indépendante de toutes les autres.
- ▶ meurt à taux  $d$ .

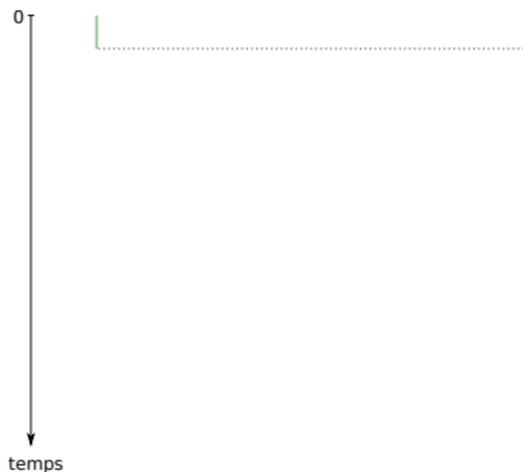
### Produit des arbres :

- ▶ trop équilibrés.
- ▶ ayant les noeuds trop proches du présent.

Nee, May et Harvey, 1994, *Phil. Trans. Biol. Sci.*

# État de l'art

## Modèles espèces-centrés



Chaque espèce :

- ▶ donne naissance à taux  $b$ .
- ▶ est indépendante de toutes les autres.
- ▶ meurt à taux  $d$ .

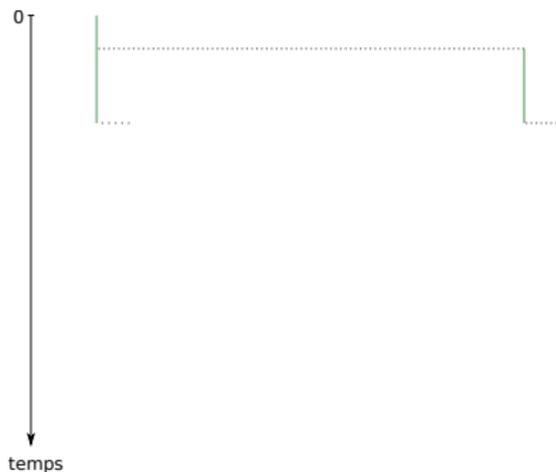
Produit des arbres :

- ▶ trop équilibrés.
- ▶ ayant les noeuds trop proches du présent.

Nee, May et Harvey, 1994, *Phil. Trans. Biol. Sci.*

# État de l'art

## Modèles espèces-centrés



Chaque espèce :

- ▶ donne naissance à taux  $b$ .
- ▶ est indépendante de toutes les autres.
- ▶ meurt à taux  $d$ .

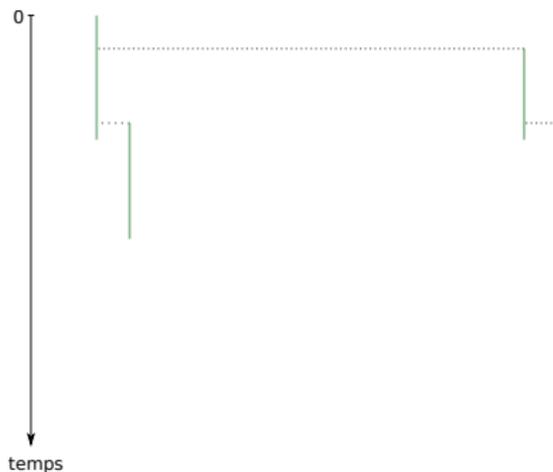
Produit des arbres :

- ▶ trop équilibrés.
- ▶ ayant les noeuds trop proches du présent.

Nee, May et Harvey, 1994, *Phil. Trans. Biol. Sci.*

# État de l'art

## Modèles espèces-centrés



Chaque espèce :

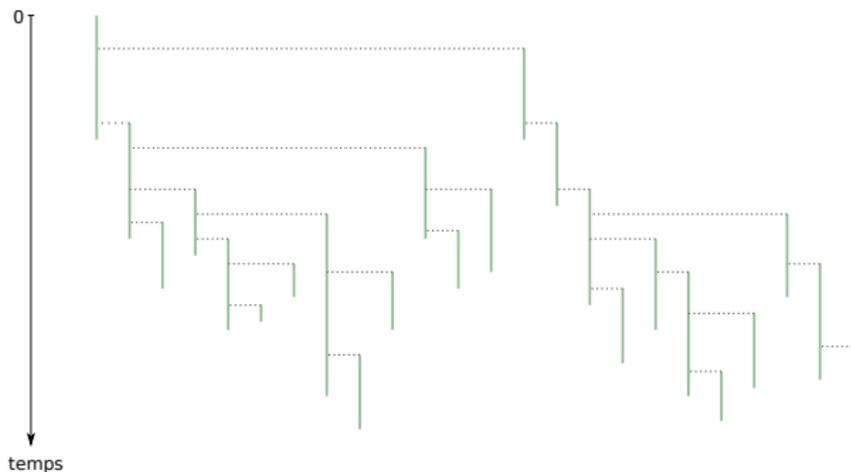
- ▶ donne naissance à taux  $b$ .
- ▶ est indépendante de toutes les autres.
- ▶ meurt à taux  $d$ .

Produit des arbres :

- ▶ trop équilibrés.
- ▶ ayant les noeuds trop proches du présent.

# État de l'art

## Modèles espèces-centrés



Chaque espèce :

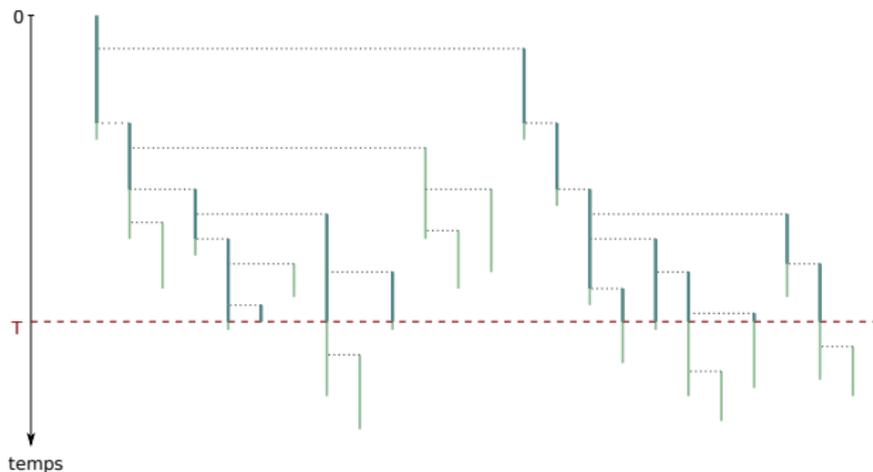
- ▶ donne naissance à taux  $b$ .
- ▶ est indépendante de toutes les autres.
- ▶ meurt à taux  $d$ .

Produit des arbres :

- ▶ trop équilibrés.
- ▶ ayant les noeuds trop proches du présent.

# État de l'art

## Modèles espèces-centrés



Chaque espèce :

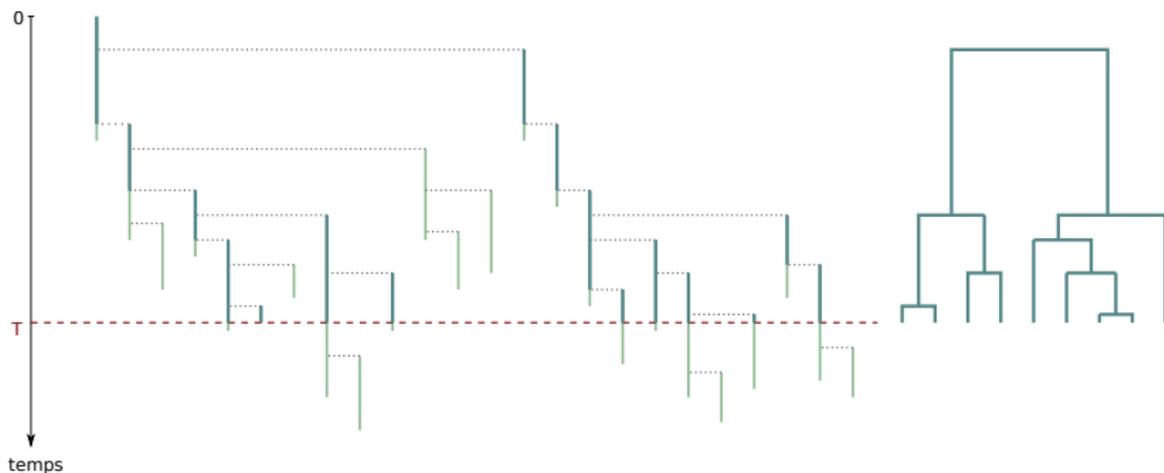
- ▶ donne naissance à taux  $b$ .
- ▶ est indépendante de toutes les autres.
- ▶ meurt à taux  $d$ .

Produit des arbres :

- ▶ trop équilibrés.
- ▶ ayant les noeuds trop proches du présent.

# État de l'art

## Modèles espèces-centrés



Chaque espèce :

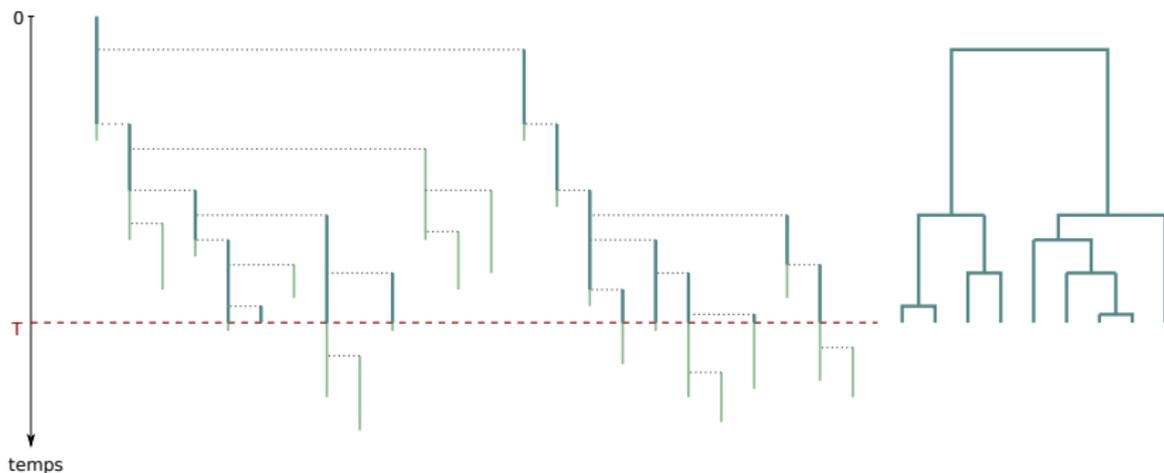
- ▶ donne naissance à taux  $b$ .
- ▶ est indépendante de toutes les autres.
- ▶ meurt à taux  $d$ .

Produit des arbres :

- ▶ trop équilibrés.
- ▶ ayant les noeuds trop proches du présent.

# État de l'art

## Modèles espèces-centrés



Chaque espèce :

- ▶ donne naissance à taux  $b$ .
- ▶ est indépendante de toutes les autres.
- ▶ meurt à taux  $d$ .

Produit des arbres :

- ▶ trop équilibrés.
- ▶ ayant les noeuds trop proches du présent.

# État de l'art

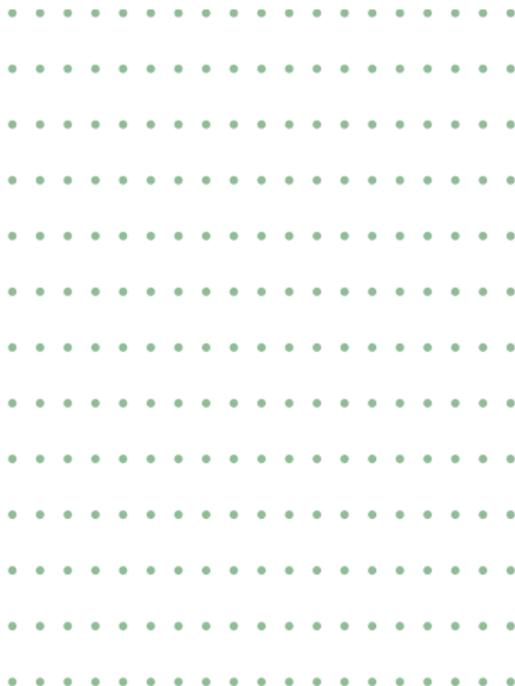
## Modèles individus-centrés

### Dynamique de population :

- ▶ taille fixée.
- ▶ générations discrètes.
- ▶ choix uniforme du parent dans la génération précédente.

### Superposition d'un modèle de spéciation.

Hubbell, 2001, *The Unified Neutral Theory of Biodiversity and Biogeography*.



# État de l'art

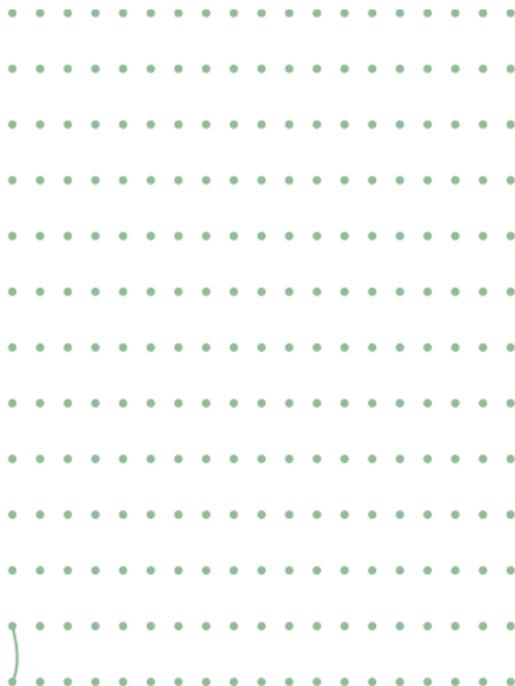
## Modèles individus-centrés

Dynamique de population :

- ▶ taille fixée.
- ▶ générations discrètes.
- ▶ choix uniforme du parent dans la génération précédente.

Superposition d'un modèle de spéciation.

Hubbell, 2001, *The Unified Neutral Theory of Biodiversity and Biogeography*.



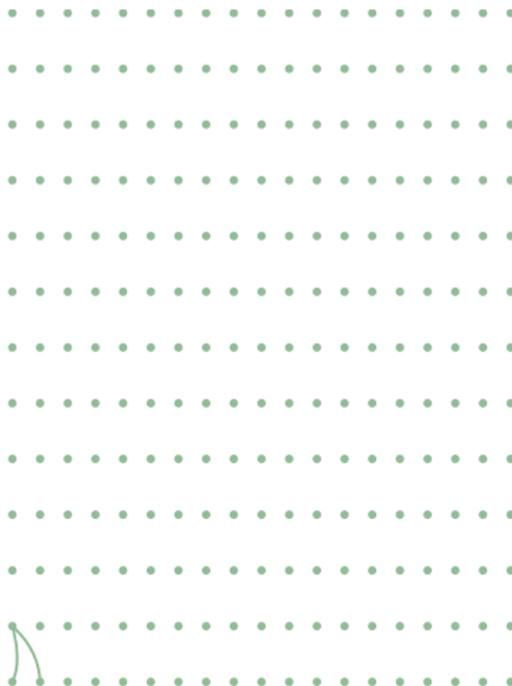
# État de l'art

## Modèles individus-centrés

Dynamique de population :

- ▶ taille fixée.
- ▶ générations discrètes.
- ▶ choix uniforme du parent dans la génération précédente.

Superposition d'un modèle de spéciation.



Hubbell, 2001, *The Unified Neutral Theory of Biodiversity and Biogeography*.

# État de l'art

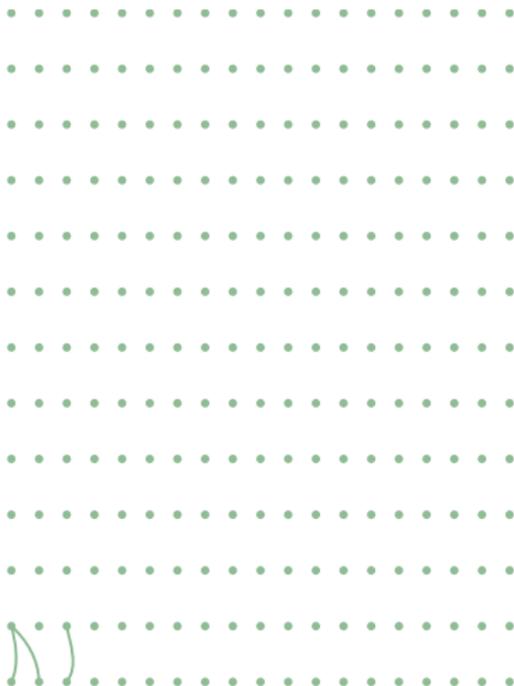
## Modèles individus-centrés

Dynamique de population :

- ▶ taille fixée.
- ▶ générations discrètes.
- ▶ choix uniforme du parent dans la génération précédente.

Superposition d'un modèle de spéciation.

Hubbell, 2001, *The Unified Neutral Theory of Biodiversity and Biogeography*.



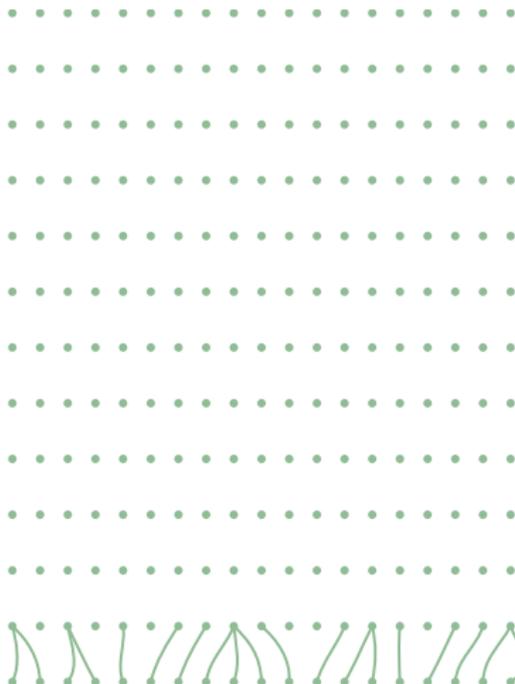
# État de l'art

## Modèles individus-centrés

Dynamique de population :

- ▶ taille fixée.
- ▶ générations discrètes.
- ▶ choix uniforme du parent dans la génération précédente.

Superposition d'un modèle de spéciation.



Hubbell, 2001, *The Unified Neutral Theory of Biodiversity and Biogeography*.

# État de l'art

## Modèles individus-centrés

Dynamique de population :

- ▶ taille fixée.
- ▶ générations discrètes.
- ▶ choix uniforme du parent dans la génération précédente.

Superposition d'un modèle de spéciation.



Hubbell, 2001, *The Unified Neutral Theory of Biodiversity and Biogeography*.

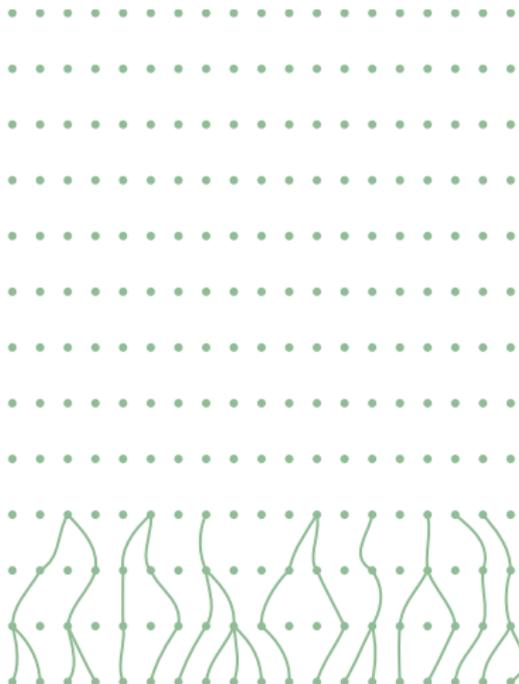
# État de l'art

## Modèles individus-centrés

Dynamique de population :

- ▶ taille fixée.
- ▶ générations discrètes.
- ▶ choix uniforme du parent dans la génération précédente.

Superposition d'un modèle de spéciation.



Hubbell, 2001, *The Unified Neutral Theory of Biodiversity and Biogeography*.

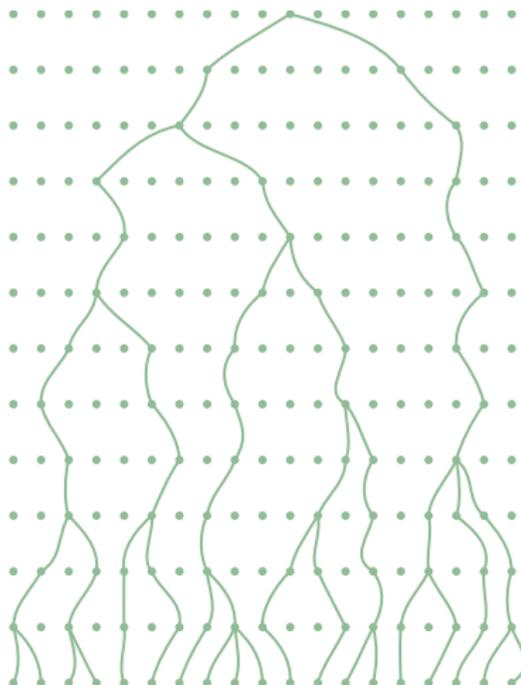
# État de l'art

## Modèles individus-centrés

Dynamique de population :

- ▶ taille fixée.
- ▶ générations discrètes.
- ▶ choix uniforme du parent dans la génération précédente.

Superposition d'un modèle de spéciation.



Hubbell, 2001, *The Unified Neutral Theory of Biodiversity and Biogeography*.

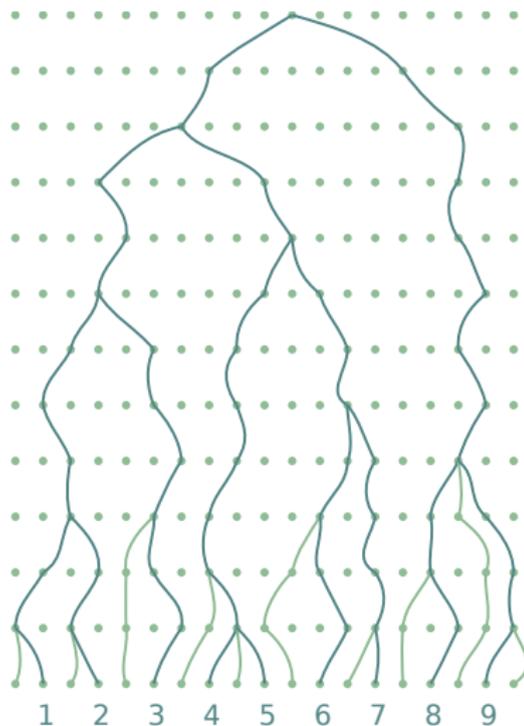
# État de l'art

## Modèles individus-centrés

Dynamique de population :

- ▶ taille fixée.
- ▶ générations discrètes.
- ▶ choix uniforme du parent dans la génération précédente.

Superposition d'un modèle de spéciation.



Hubbell, 2001, *The Unified Neutral Theory of Biodiversity and Biogeography*.

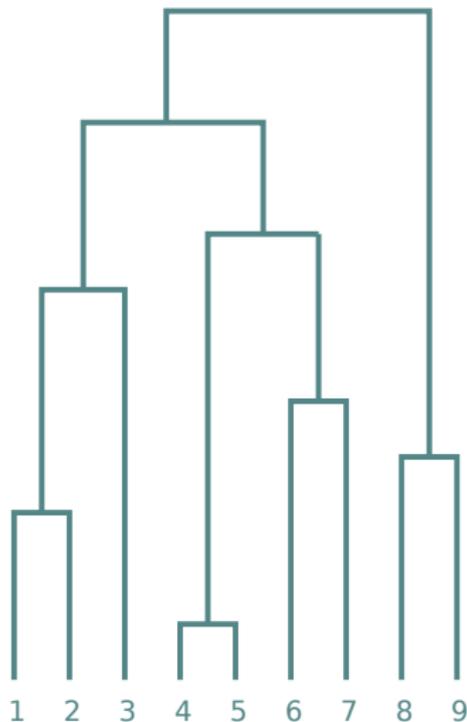
# État de l'art

## Modèles individus-centrés

Dynamique de population :

- ▶ taille fixée.
- ▶ générations discrètes.
- ▶ choix uniforme du parent dans la génération précédente.

Superposition d'un modèle de spéciation.



Hubbell, 2001, *The Unified Neutral Theory of Biodiversity and Biogeography*.

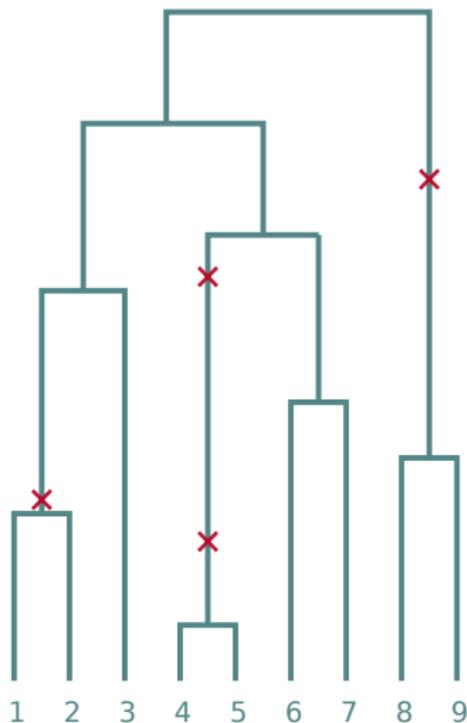
# État de l'art

## Modèles individus-centrés

Dynamique de population :

- ▶ taille fixée.
- ▶ générations discrètes.
- ▶ choix uniforme du parent dans la génération précédente.

Superposition d'un modèle de spéciation.



Hubbell, 2001, *The Unified Neutral Theory of Biodiversity and Biogeography*.

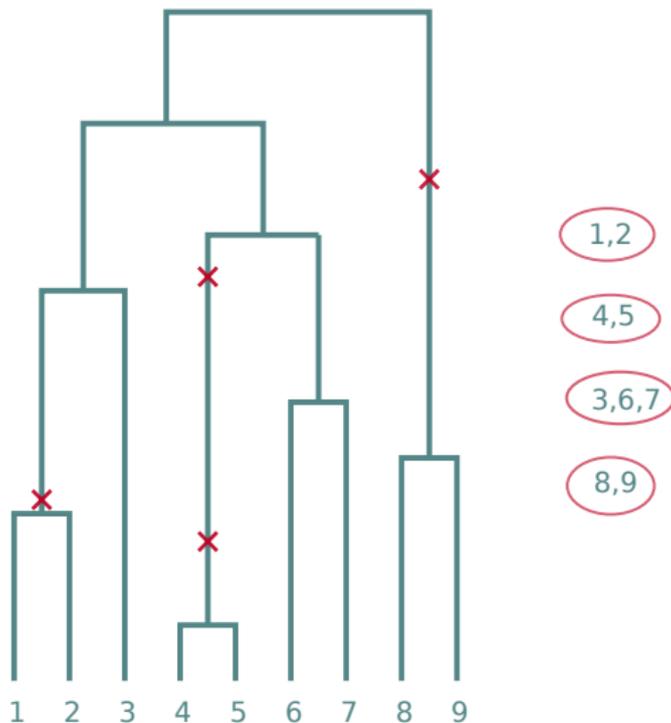
# État de l'art

## Modèles individus-centrés

Dynamique de population :

- ▶ taille fixée.
- ▶ générations discrètes.
- ▶ choix uniforme du parent dans la génération précédente.

Superposition d'un modèle de spéciation.

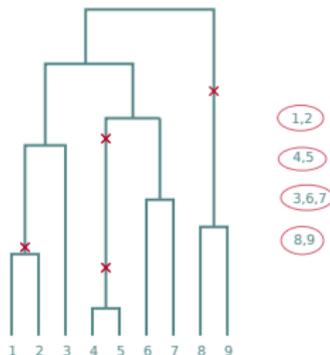


Hubbell, 2001, *The Unified Neutral Theory of Biodiversity and Biogeography*.

## La définition d'espèce en modélisation

### Passage d'une généalogie à une phylogénie

- ▶ Les modèles produisent des espèces qui sont des groupes non-monophylétiques d'individus.
- ▶ Difficulté à reconstruire la phylogénie à partir de la généalogie.
- ▶ Production d'arbres suffisamment déséquilibrés mais ayant des noeuds bien trop proches du présent.



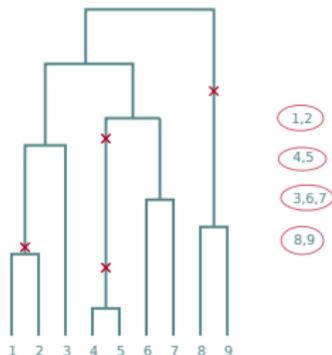
Jabot et Chave, 2009, *Ecol. Lett.*

Davies et al., 2011, *Evolution*.

## La définition d'espèce en modélisation

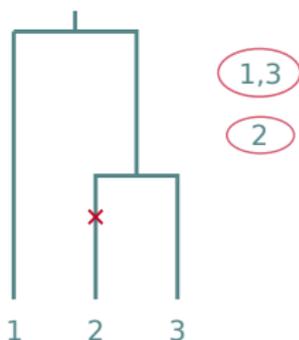
### Passage d'une généalogie à une phylogénie

- ▶ Les modèles produisent des espèces qui sont des groupes non-monophylétiques d'individus.
- ▶ Difficulté à reconstruire la phylogénie à partir de la généalogie.
- ▶ Production d'arbres suffisamment déséquilibrés mais ayant des noeuds bien trop proches du présent.



Jabot et Chave, 2009, *Ecol. Lett.*

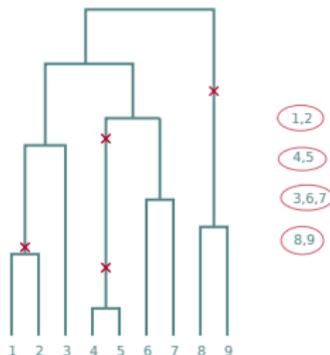
Davies et al., 2011, *Evolution*.



## La définition d'espèce en modélisation

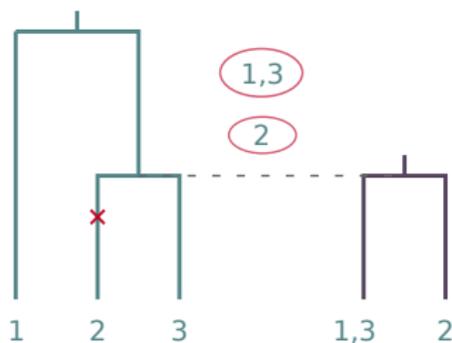
### Passage d'une généalogie à une phylogénie

- ▶ Les modèles produisent des espèces qui sont des groupes non-monophylétiques d'individus.
- ▶ Difficulté à reconstruire la phylogénie à partir de la généalogie.
- ▶ Production d'arbres suffisamment déséquilibrés mais ayant des noeuds bien trop proches du présent.



Jabot et Chave, 2009, *Ecol. Lett.*

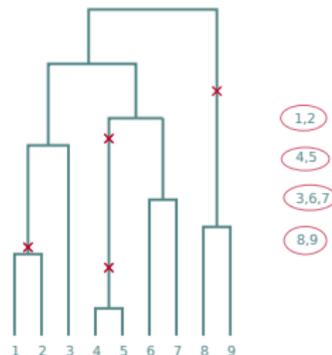
Davies et al., 2011, *Evolution*.



## La définition d'espèce en modélisation

### Passage d'une généalogie à une phylogénie

- ▶ Les modèles produisent des espèces qui sont des groupes non-monophylétiques d'individus.
- ▶ Difficulté à reconstruire la phylogénie à partir de la généalogie.
- ▶ Production d'arbres suffisamment déséquilibrés mais ayant des noeuds bien trop proches du présent.



Jabot et Chave, 2009, *Ecol. Lett.*

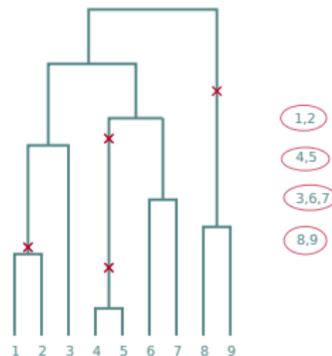
Davies et al., 2011, *Evolution.*



## La définition d'espèce en modélisation

### Passage d'une généalogie à une phylogénie

- ▶ Les modèles produisent des espèces qui sont des groupes non-monophylétiques d'individus.
- ▶ Difficulté à reconstruire la phylogénie à partir de la généalogie.
- ▶ Production d'arbres suffisamment déséquilibrés mais ayant des noeuds bien trop proches du présent.



Jabot et Chave, 2009, *Ecol. Lett.*

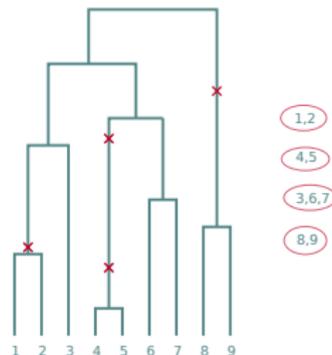
Davies et al., 2011, *Evolution*.



# La définition d'espèce en modélisation

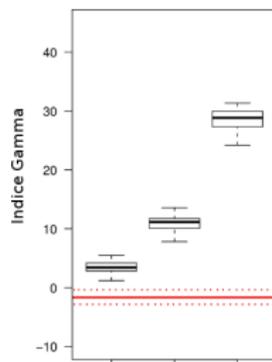
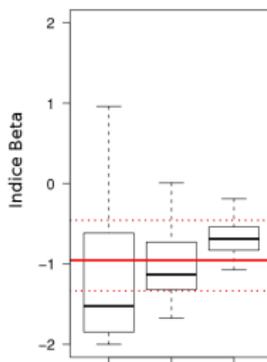
## Passage d'une généalogie à une phylogénie

- ▶ Les modèles produisent des espèces qui sont des groupes non-monophylétiques d'individus.
- ▶ Difficulté à reconstruire la phylogénie à partir de la généalogie.
- ▶ Production d'arbres suffisamment déséquilibrés mais ayant des noeuds bien trop proches du présent.



Jabot et Chave, 2009, *Ecol. Lett.*

Davies et al., 2011, *Evolution*.

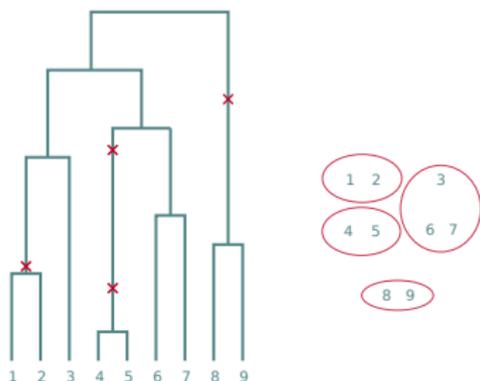


# La définition d'espèce en modélisation

## Proposition de deux nouvelles définitions

**Loose** Partition monophylétique la plus fine telle que deux individus de deux espèces différentes soient toujours différents.

**Lacy** Partition monophylétique la plus grossière telle que deux individus d'une même espèce soient toujours semblables.

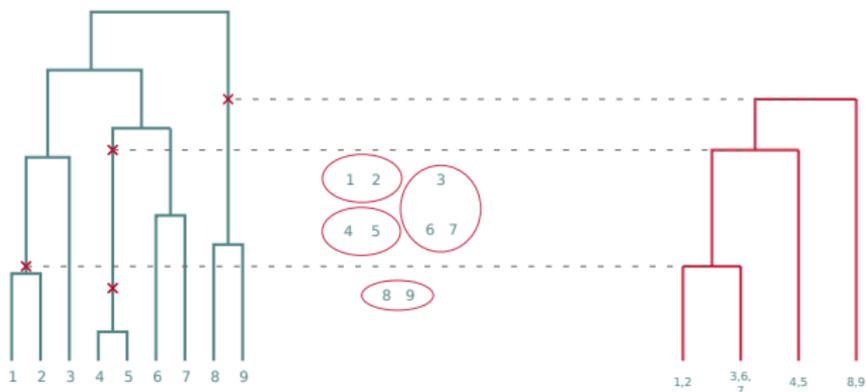


# La définition d'espèce en modélisation

## Proposition de deux nouvelles définitions

**Loose** Partition monophylétique la plus fine telle que deux individus de deux espèces différentes soient toujours différents.

**Lacy** Partition monophylétique la plus grossière telle que deux individus d'une même espèce soient toujours semblables.

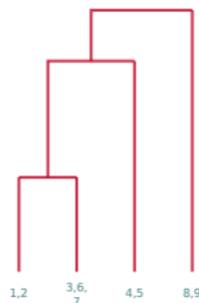
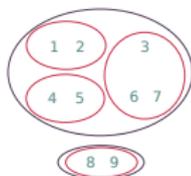
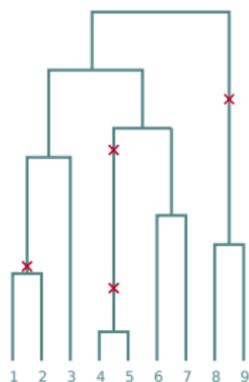


# La définition d'espèce en modélisation

## Proposition de deux nouvelles définitions

**Loose** Partition monophylétique la plus fine telle que deux individus de deux espèces différentes soient toujours différents.

**Lacy** Partition monophylétique la plus grossière telle que deux individus d'une même espèce soient toujours semblables.

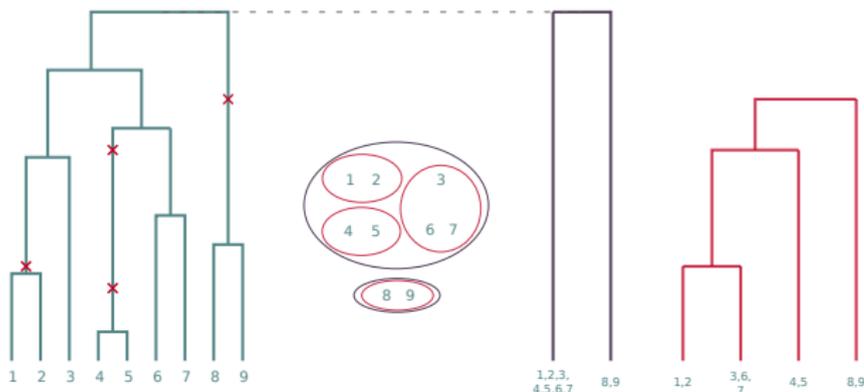


# La définition d'espèce en modélisation

## Proposition de deux nouvelles définitions

**Loose** Partition monophylétique la plus fine telle que deux individus de deux espèces différentes soient toujours différents.

**Lacy** Partition monophylétique la plus grossière telle que deux individus d'une même espèce soient toujours semblables.

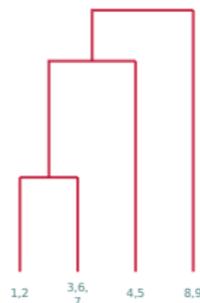
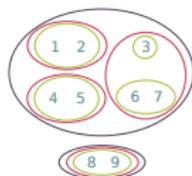
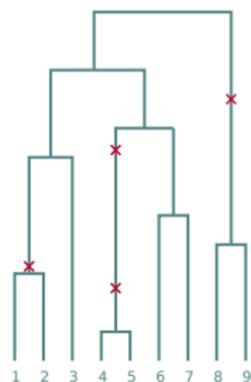


Manceau et Lambert, 2018, Bull. Math. Biol. (en révision).

# La définition d'espèce en modélisation

Proposition de deux nouvelles définitions

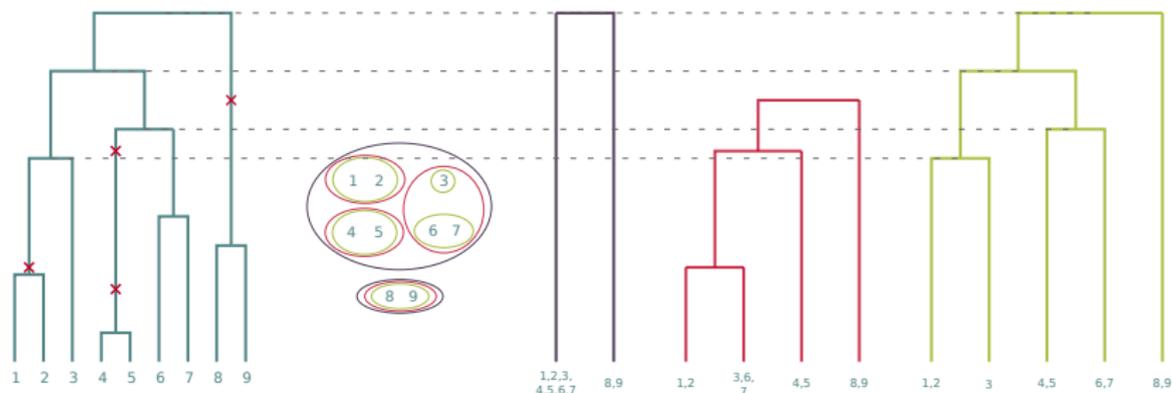
- Loose** Partition monophylétique la plus fine telle que deux individus de deux espèces différentes soient toujours différents.
- Lacy** Partition monophylétique la plus grossière telle que deux individus d'une même espèce soient toujours semblables.



# La définition d'espèce en modélisation

## Proposition de deux nouvelles définitions

- Loose** Partition monophylétique la plus fine telle que deux individus de deux espèces différentes soient toujours différents.
- Lacy** Partition monophylétique la plus grossière telle que deux individus d'une même espèce soient toujours semblables.

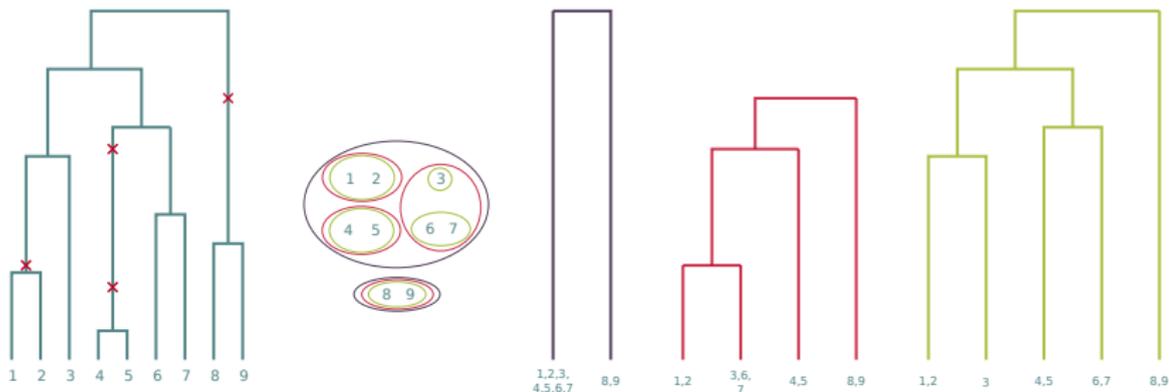


Manceau et Lambert, 2018, Bull. Math. Biol. (en révision).

# La définition d'espèce en modélisation

## Proposition de deux nouvelles définitions

- Loose** Partition monophylétique la plus fine telle que deux individus de deux espèces différentes soient toujours différents.
- Lacy** Partition monophylétique la plus grossière telle que deux individus d'une même espèce soient toujours semblables.



Manceau et Lambert, 2018, Bull. Math. Biol. (en révision).

# Application à l'étude de la diversification

## Deux hypothèses nouvelles

- ▶ **Modèle de naissance-mort individu-centré.**

- ▶ Définition d'espèce *loose*.



Manceau, Lambert et Morlon, 2015, *Ecol. lett.*

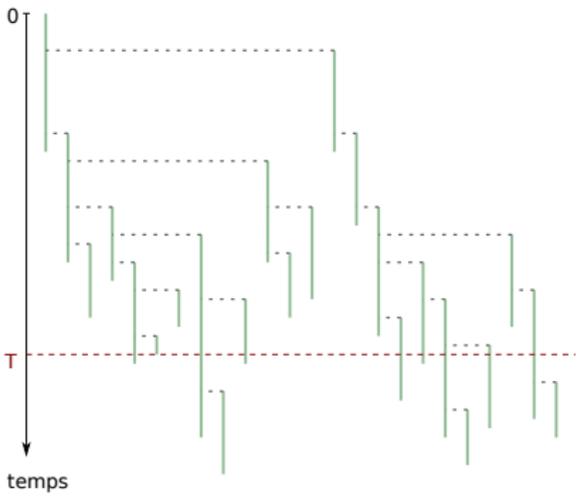


# Application à l'étude de la diversification

## Deux hypothèses nouvelles

► **Modèle de naissance-mort individu-centré.**

► Définition d'espèce *loose*.



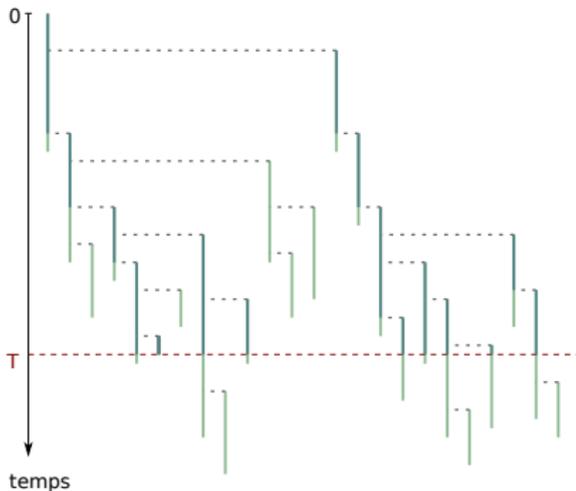
Manceau, Lambert et Morlon, 2015, *Ecol. lett.*

# Application à l'étude de la diversification

## Deux hypothèses nouvelles

- ▶ **Modèle de naissance-mort individu-centré.**

- ▶ Définition d'espèce *loose*.



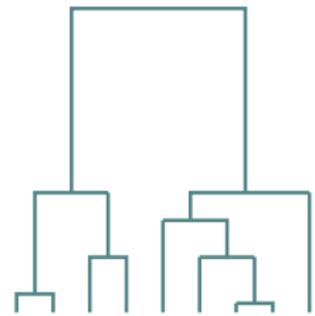
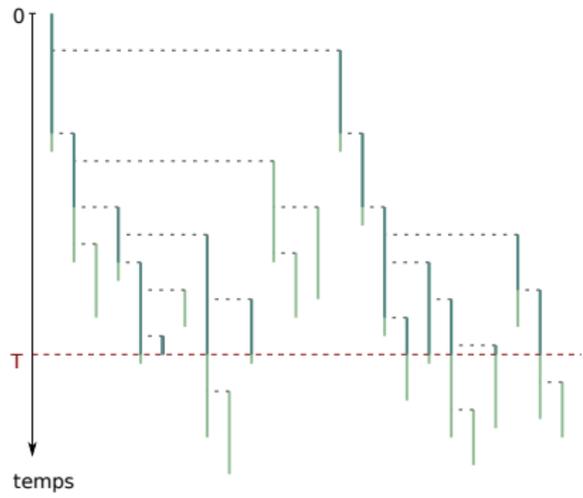
Manceau, Lambert et Morlon, 2015, *Ecol. lett.*

# Application à l'étude de la diversification

## Deux hypothèses nouvelles

► **Modèle de naissance-mort individu-centré.**

► **Définition d'espèce *loose*.**



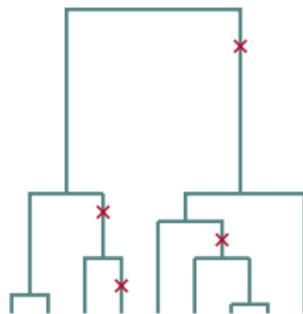
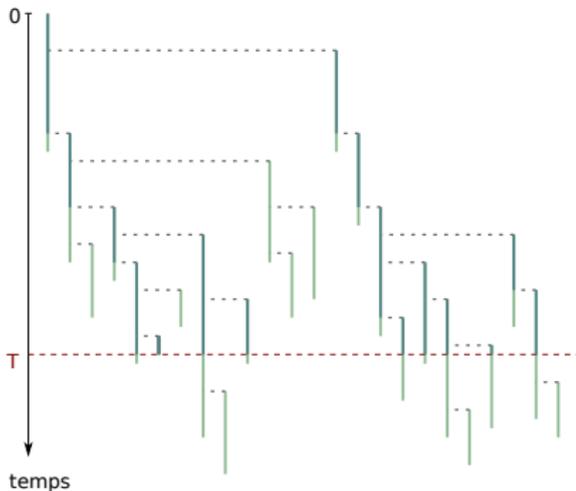
Manceau, Lambert et Morlon, 2015, *Ecol. lett.*

# Application à l'étude de la diversification

## Deux hypothèses nouvelles

- ▶ Modèle de naissance-mort individu-centré.

- ▶ Définition d'espèce *loose*.



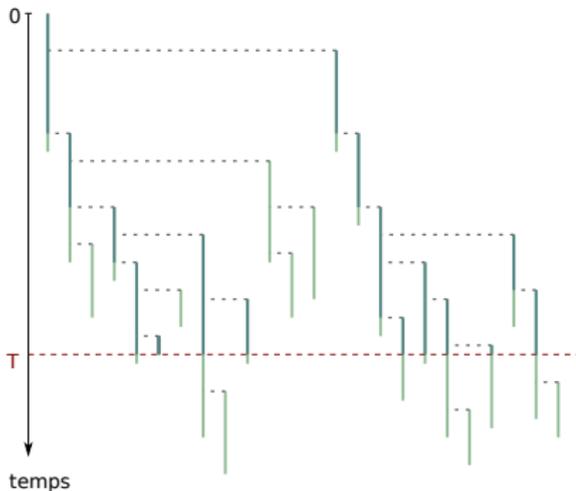
Manceau, Lambert et Morlon, 2015, *Ecol. lett.*

# Application à l'étude de la diversification

## Deux hypothèses nouvelles

- ▶ Modèle de naissance-mort individu-centré.

- ▶ Définition d'espèce *loose*.



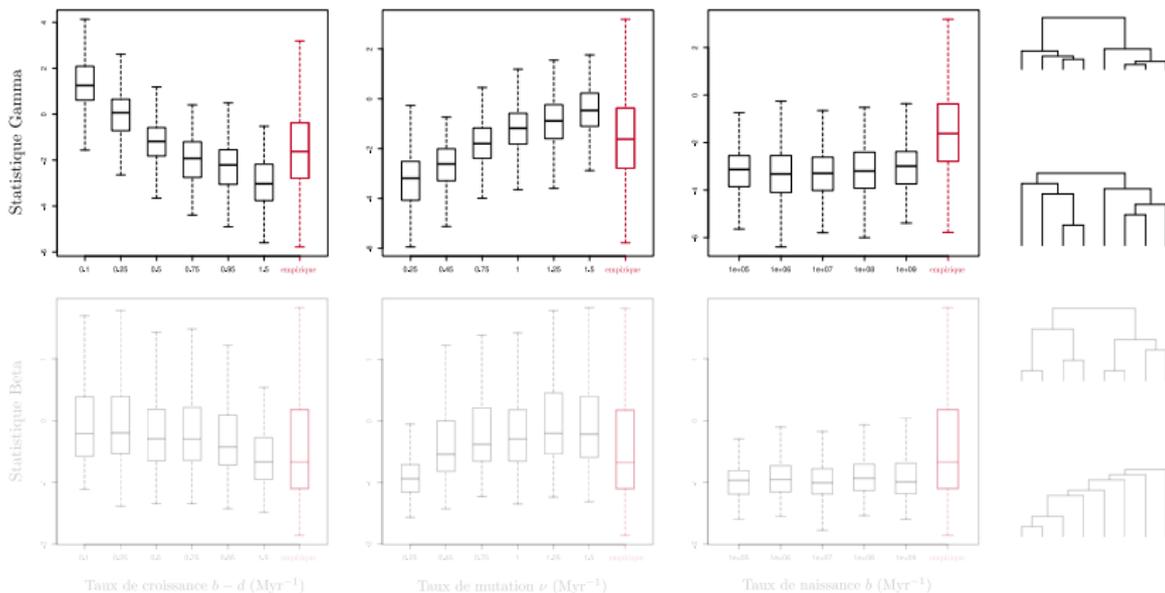
Manceau, Lambert et Morlon, 2015, *Ecol. lett.*

# Application à l'étude de la diversification

## Comparaison aux formes d'arbres empiriques

Le modèle permet de générer des arbres

- ▶ ayant la bonne distribution de temps de branchement.
- ▶ suffisamment déséquilibrés.

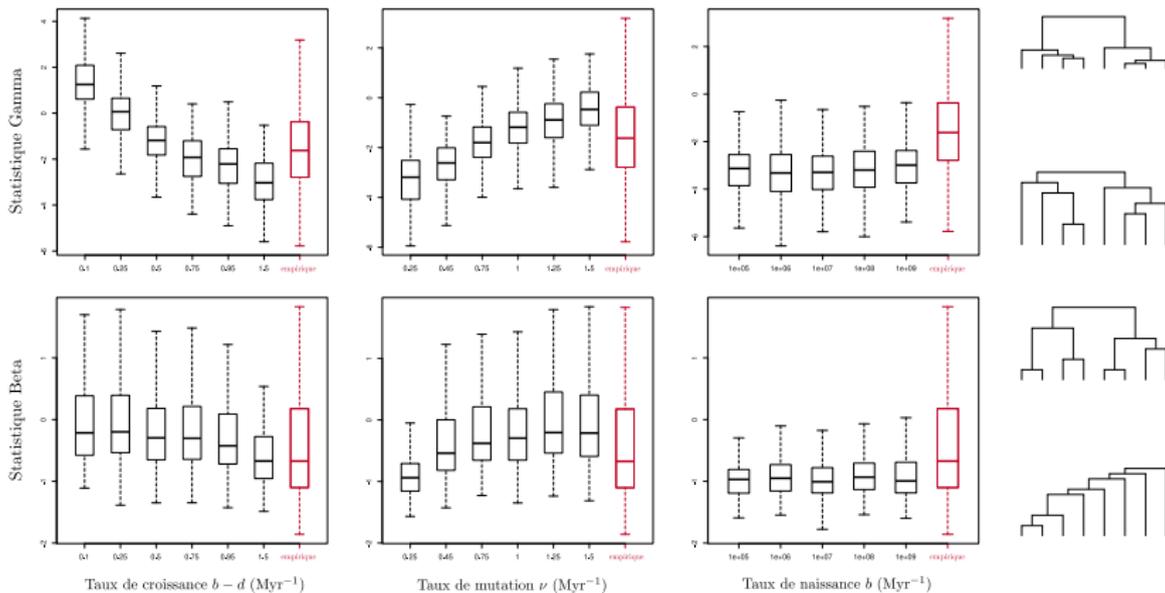


# Application à l'étude de la diversification

## Comparaison aux formes d'arbres empiriques

Le modèle permet de générer des arbres

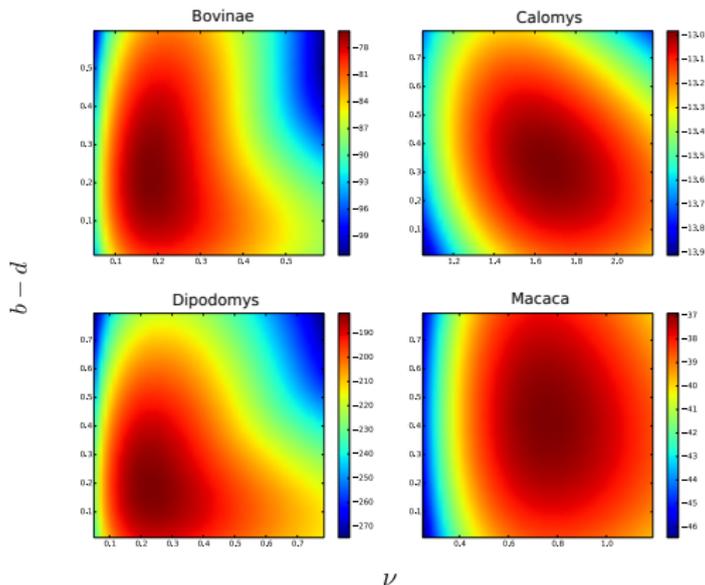
- ▶ ayant la bonne distribution de temps de branchement.
- ▶ suffisamment déséquilibrés.



# Application à l'étude de la diversification

## Méthode d'inférence statistique

- ▶ On peut calculer numériquement la densité de probabilité d'un arbre observé.

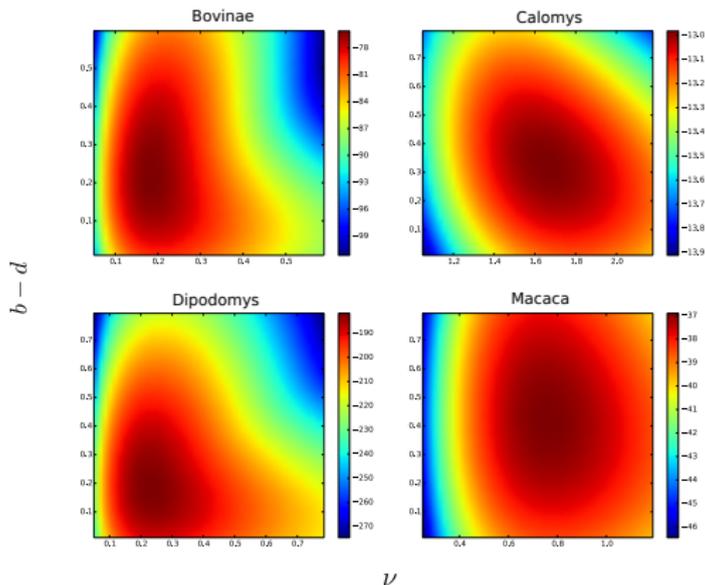


- ▶ On estime les paramètres par optimisation numérique de la vraisemblance.

# Application à l'étude de la diversification

## Méthode d'inférence statistique

- ▶ On peut calculer numériquement la densité de probabilité d'un arbre observé.

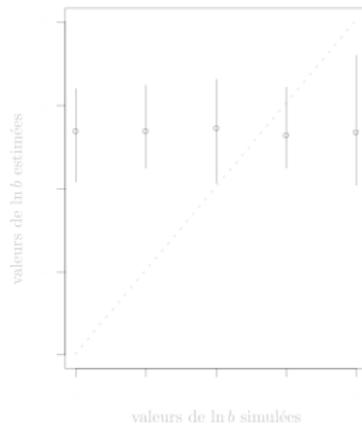
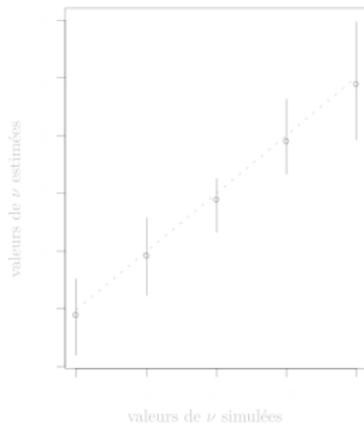
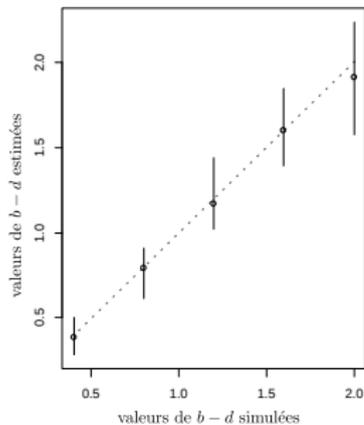


- ▶ On estime les paramètres par optimisation numérique de la vraisemblance.

# Application à l'étude de la diversification

## Méthode d'inférence statistique

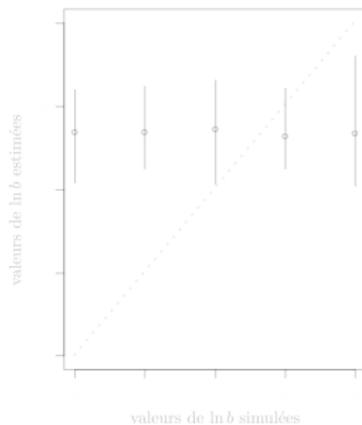
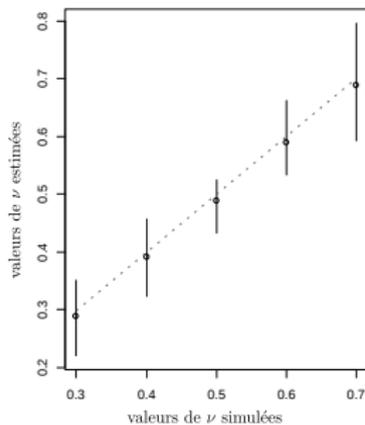
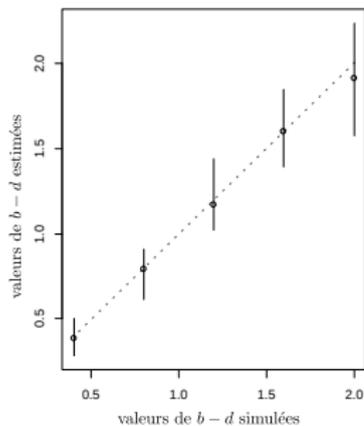
- ▶ Les estimateurs de  $b - d$  et  $\nu$  sont non biaisés.
- ▶ Mais on ne sait pas retrouver  $b$ .



# Application à l'étude de la diversification

## Méthode d'inférence statistique

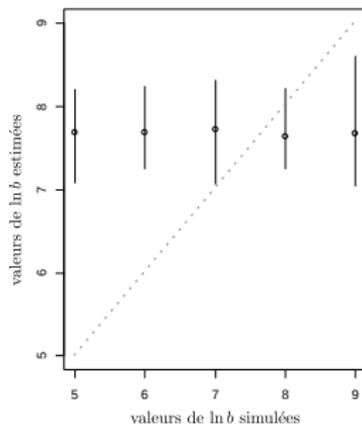
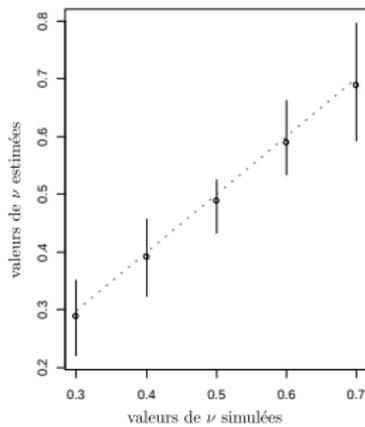
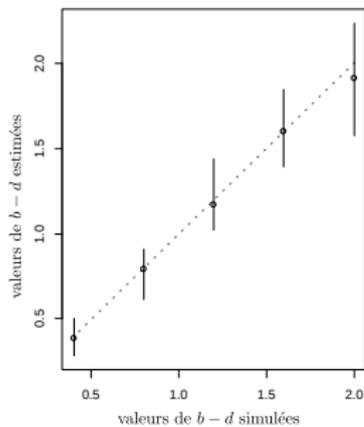
- ▶ Les estimateurs de  $b - d$  et  $\nu$  sont non biaisés.
- ▶ Mais on ne sait pas retrouver  $b$ .



# Application à l'étude de la diversification

## Méthode d'inférence statistique

- ▶ Les estimateurs de  $b - d$  et  $\nu$  sont non biaisés.
- ▶ Mais on ne sait pas retrouver  $b$ .

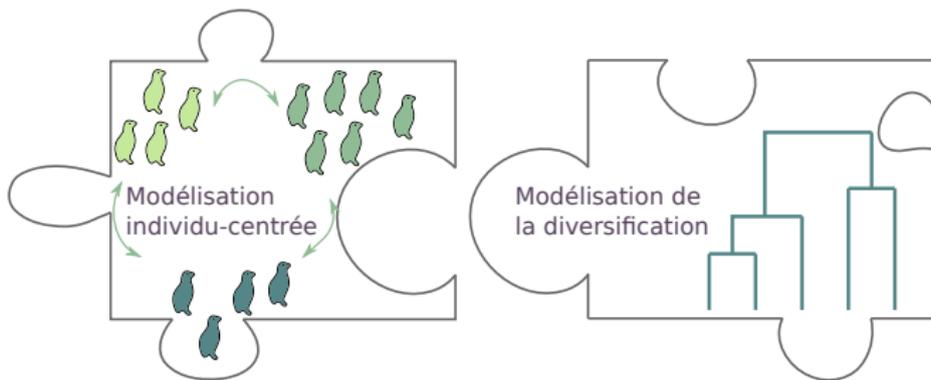


# Application à l'étude de la diversification

## Résumé du projet

Un modèle de diversification individu-centré proposant deux hypothèses inhabituelles :

1. dynamique de population donnée par un processus de naissance-mort,
2. définition d'espèce *loose*.



Conclusion et perspectives :

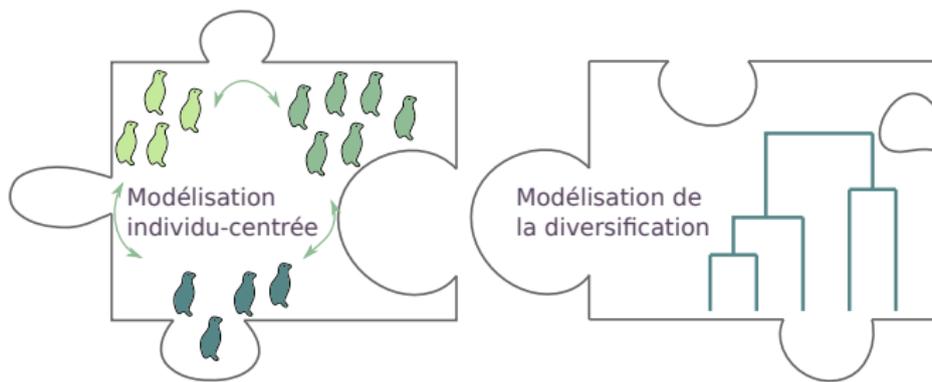
- ▶ Reproduit bien les formes d'arbres empiriques.
- ▶ Quelle est l'importance relative des deux hypothèses ?

# Application à l'étude de la diversification

## Résumé du projet

Un modèle de diversification individu-centré proposant deux hypothèses inhabituelles :

1. dynamique de population donnée par un processus de naissance-mort,
2. définition d'espèce *loose*.



Conclusion et perspectives :

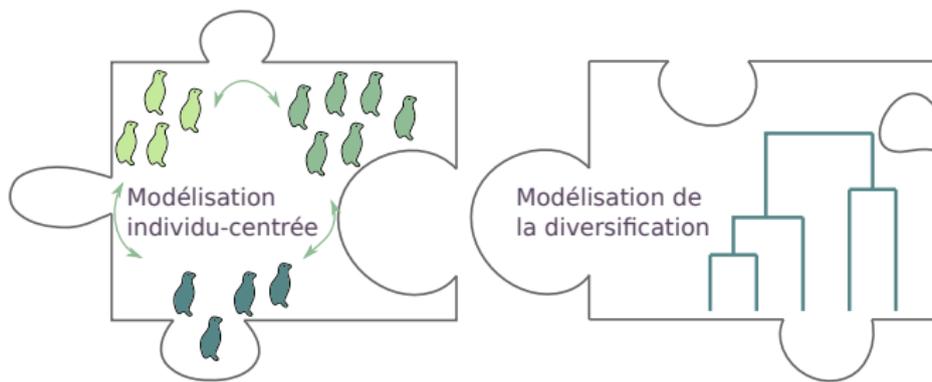
- ▶ Reproduit bien les formes d'arbres empiriques.
- ▶ Quelle est l'importance relative des deux hypothèses ?

# Application à l'étude de la diversification

## Résumé du projet

Un modèle de diversification individu-centré proposant deux hypothèses inhabituelles :

1. dynamique de population donnée par un processus de naissance-mort,
2. définition d'espèce *loose*.



Conclusion et perspectives :

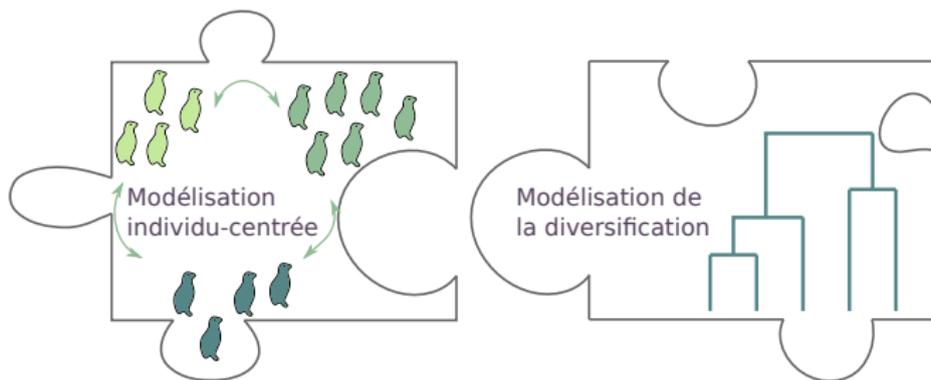
- ▶ Reproduit bien les formes d'arbres empiriques.
- ▶ Quelle est l'importance relative des deux hypothèses ?

# Application à l'étude de la diversification

## Résumé du projet

Un modèle de diversification individu-centré proposant deux hypothèses inhabituelles :

1. dynamique de population donnée par un processus de naissance-mort,
2. définition d'espèce *loose*.



Conclusion et perspectives :

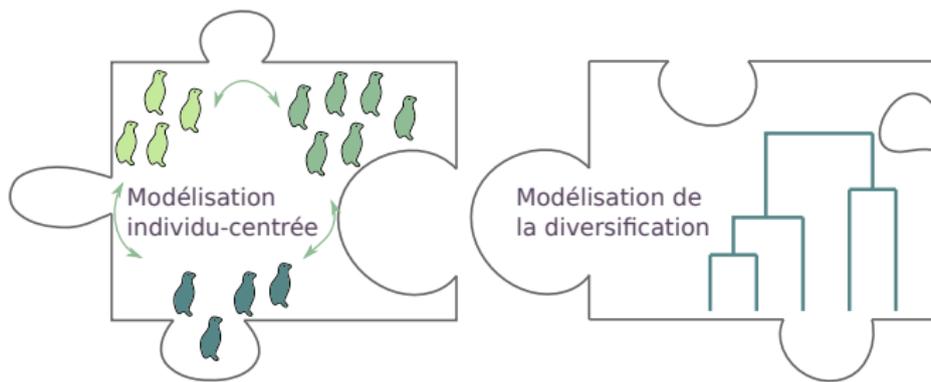
- ▶ Reproduit bien les formes d'arbres empiriques.
- ▶ Quelle est l'importance relative des deux hypothèses ?

# Application à l'étude de la diversification

## Résumé du projet

Un modèle de diversification individu-centré proposant deux hypothèses inhabituelles :

1. dynamique de population donnée par un processus de naissance-mort,
2. définition d'espèce *loose*.



Conclusion et perspectives :

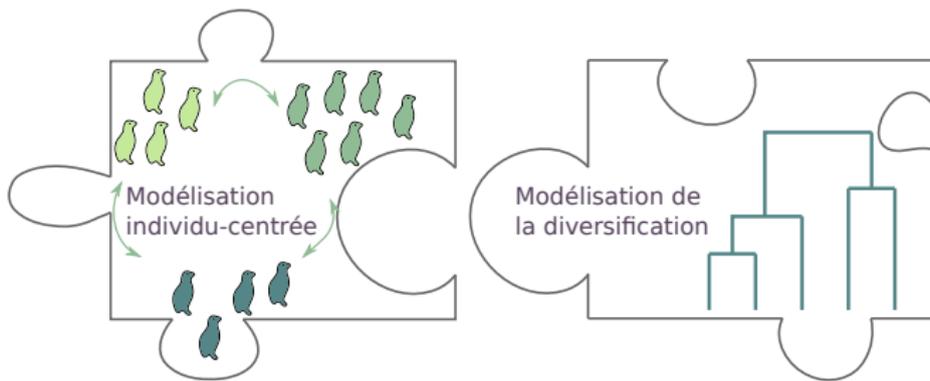
- ▶ Reproduit bien les formes d'arbres empiriques.
- ▶ Quelle est l'importance relative des deux hypothèses ?

# Application à l'étude de la diversification

## Résumé du projet

Un modèle de diversification individu-centré proposant deux hypothèses inhabituelles :

1. dynamique de population donnée par un processus de naissance-mort,
2. définition d'espèce *loose*.



Conclusion et perspectives :

- ▶ Reproduit bien les formes d'arbres empiriques.
- ▶ Quelle est l'importance relative des deux hypothèses ?

## Introduction

- Macroévolution
- Arbre phylogénétique
- Surcouche de modélisation probabiliste

## Diversification

- État de l'art
- La définition d'espèce en modélisation
- Application à l'étude de la diversification



## Évolution phénotypique

- État de l'art
- Coévolution des caractères de différentes espèces
- Proposition d'un cadre de modélisation
- Exemples d'applications



## Évolution moléculaire

- Horloge moléculaire
- Modélisation proposée
- Méthode d'inférence



## Perspectives

- Résumé
- Liens entre les chapitres

# État de l'art

Hypothèse centrale: indépendance des lignées

On compare :

- ▶ des caractères d'espèces différentes.
- ▶ dont on connaît la phylogénie.

Modélisation classique :

- ▶ le caractère à la racine est tiré dans une certaine loi.
- ▶ le long d'une branche il suit une certaine trajectoire aléatoire.
- ▶ il est copié en deux copies *indépendantes* à chaque branchement.



Felsenstein, 1973, *Am. J. Hum. Genet.*

# État de l'art

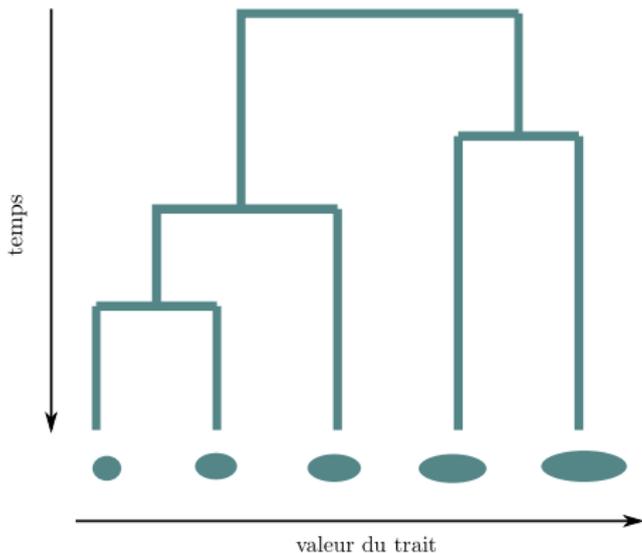
Hypothèse centrale: indépendance des lignées

On compare :

- ▶ des caractères d'espèces différentes.
- ▶ dont on connaît la phylogénie.

Modélisation classique :

- ▶ le caractère à la racine est tiré dans une certaine loi.
- ▶ le long d'une branche il suit une certaine trajectoire aléatoire.
- ▶ il est copié en deux copies *indépendantes* à chaque branchement.



Felsenstein, 1973, *Am. J. Hum. Genet.*

# État de l'art

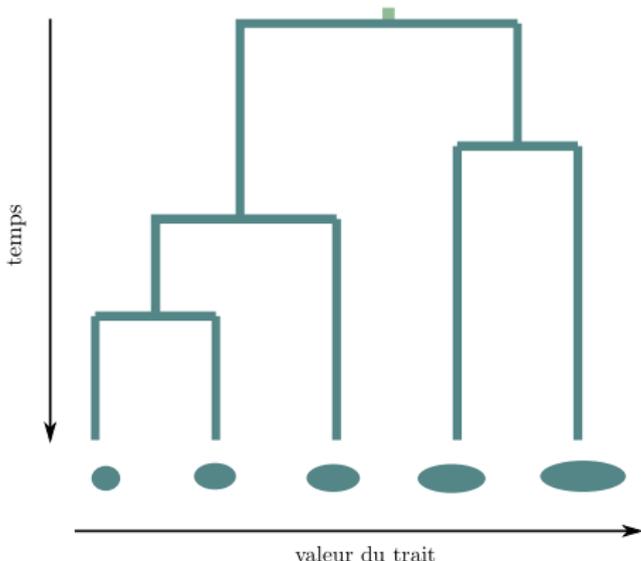
Hypothèse centrale: indépendance des lignées

On compare :

- ▶ des caractères d'espèces différentes.
- ▶ dont on connaît la phylogénie.

Modélisation classique :

- ▶ le caractère à la racine est tiré dans une certaine loi.
- ▶ le long d'une branche il suit une certaine trajectoire aléatoire.
- ▶ il est copié en deux copies *indépendantes* à chaque branchement.



Felsenstein, 1973, *Am. J. Hum. Genet.*

# État de l'art

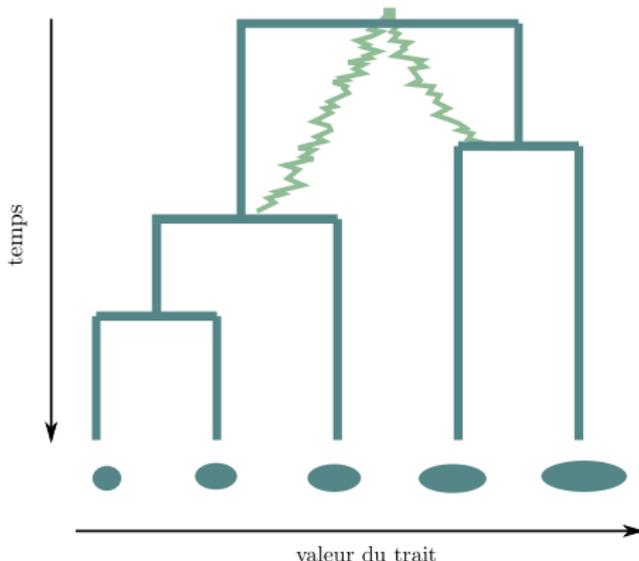
Hypothèse centrale: indépendance des lignées

On compare :

- ▶ des caractères d'espèces différentes.
- ▶ dont on connaît la phylogénie.

Modélisation classique :

- ▶ le caractère à la racine est tiré dans une certaine loi.
- ▶ le long d'une branche il suit une certaine trajectoire aléatoire.
- ▶ il est copié en deux copies *indépendantes* à chaque branchement.



Felsenstein, 1973, *Am. J. Hum. Genet.*

# État de l'art

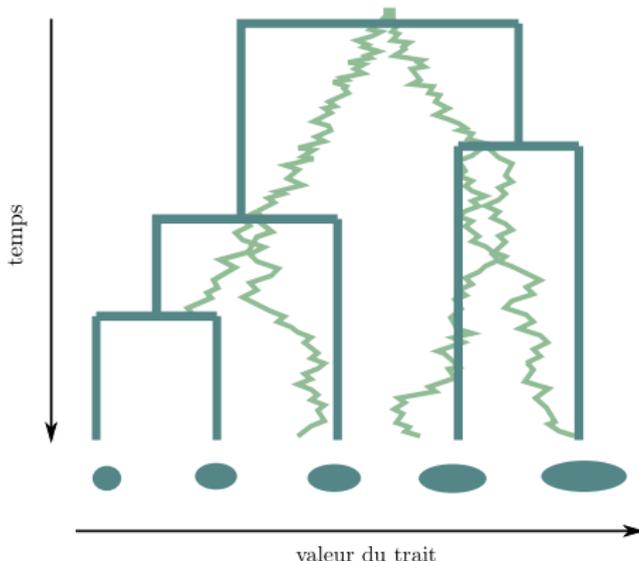
Hypothèse centrale: indépendance des lignées

On compare :

- ▶ des caractères d'espèces différentes.
- ▶ dont on connaît la phylogénie.

Modélisation classique :

- ▶ le caractère à la racine est tiré dans une certaine loi.
- ▶ le long d'une branche il suit une certaine trajectoire aléatoire.
- ▶ il est copié en deux copies *indépendantes* à chaque branchement.



Felsenstein, 1973, *Am. J. Hum. Genet.*

# État de l'art

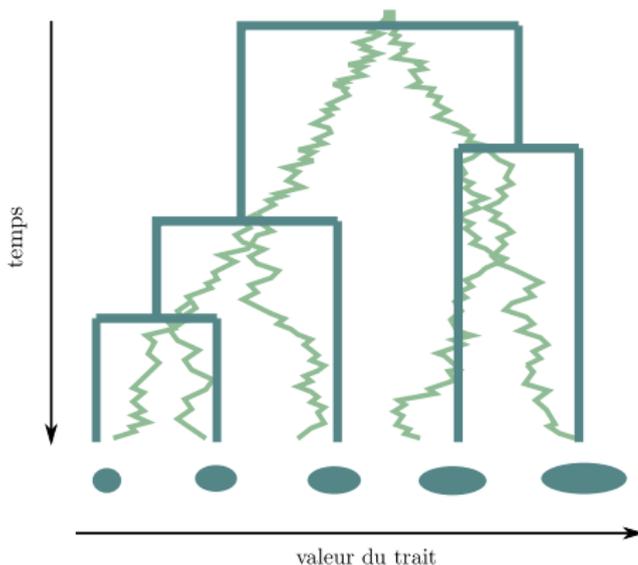
Hypothèse centrale: indépendance des lignées

On compare :

- ▶ des caractères d'espèces différentes.
- ▶ dont on connaît la phylogénie.

Modélisation classique :

- ▶ le caractère à la racine est tiré dans une certaine loi.
- ▶ le long d'une branche il suit une certaine trajectoire aléatoire.
- ▶ il est copié en deux copies *indépendantes* à chaque branchement.



Felsenstein, 1973, *Am. J. Hum. Genet.*

# État de l'art

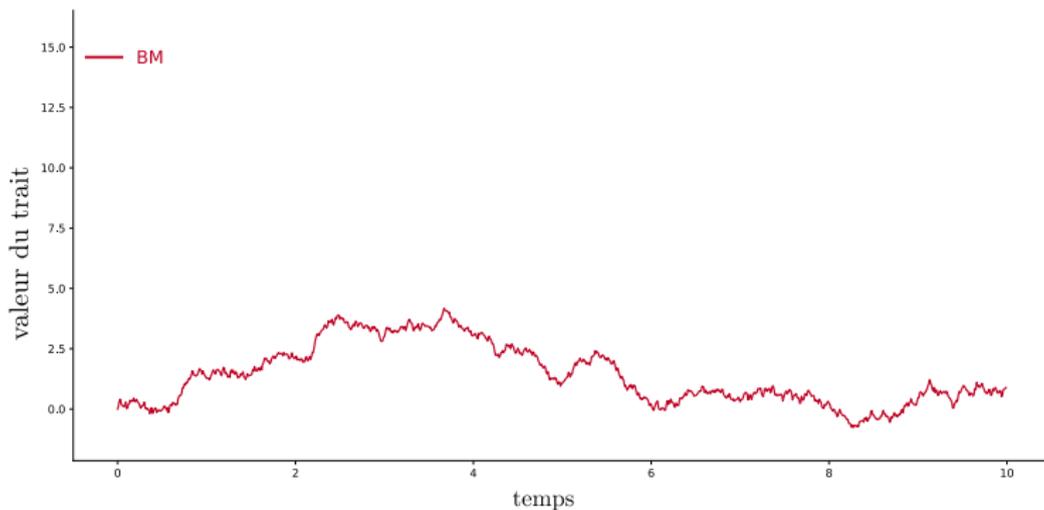
## Processus stochastique le long d'un arbre fixé

Aperçu de quelques modèles de trajectoires aléatoires :

Mouvement brownien  $dX_t = \sigma dB_t$

Ornstein-Uhlenbeck  $dX_t = \psi(\theta - X_t)dt + \sigma dB_t$

Brownien avec dérive  $dX_t = a dt + \sigma dB_t$



# État de l'art

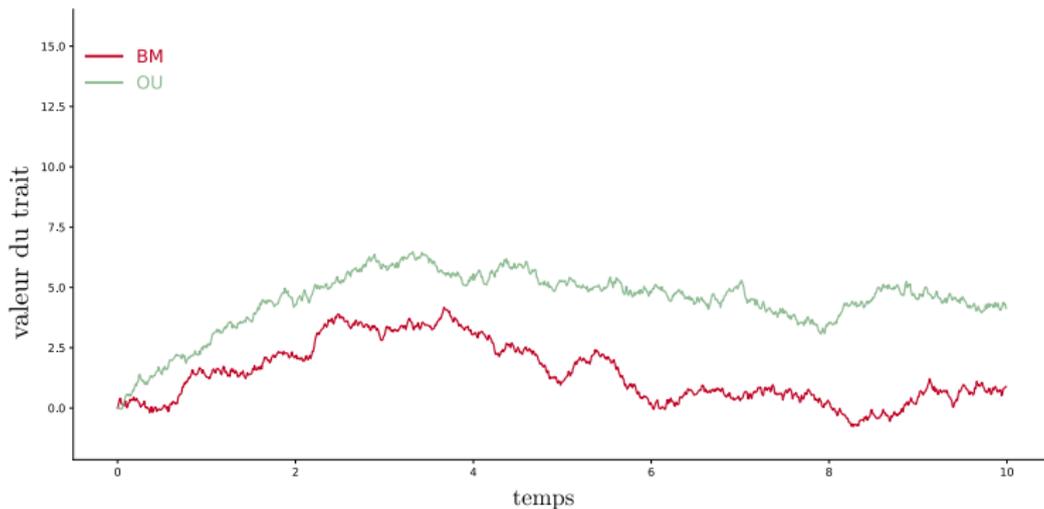
## Processus stochastique le long d'un arbre fixé

Aperçu de quelques modèles de trajectoires aléatoires :

Mouvement brownien  $dX_t = \sigma dB_t$

Ornstein-Uhlenbeck  $dX_t = \psi(\theta - X_t)dt + \sigma dB_t$

Brownien avec dérive  $dX_t = a dt + \sigma dB_t$



# État de l'art

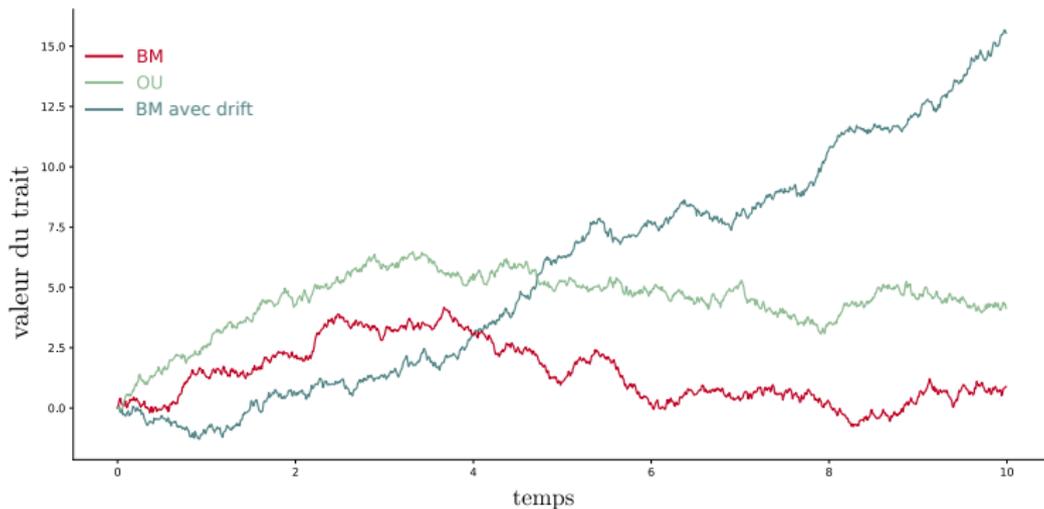
## Processus stochastique le long d'un arbre fixé

Aperçu de quelques modèles de trajectoires aléatoires :

Mouvement brownien  $dX_t = \sigma dB_t$

Ornstein-Uhlenbeck  $dX_t = \psi(\theta - X_t)dt + \sigma dB_t$

Brownien avec dérive  $dX_t = adt + \sigma dB_t$



## Coévolution des caractères de différentes espèces

### Attraction ou répulsion dans l'espace des caractères

- ▶ On possède de bons exemples de caractères non-indépendants.

**Imitation** attraction vers le caractère moyen dans une communauté.

**Mutualisme** coévolution de caractères dans deux clades.

- ▶ On souhaiterait étudier ces caractères dans un cadre similaire.
- ▶ Sans hypothèse d'indépendance entre caractères portés par lignées différentes.

# Coévolution des caractères de différentes espèces

## Attraction ou répulsion dans l'espace des caractères

- ▶ On possède de bons exemples de caractères non-indépendants.

**Imitation** attraction vers le caractère moyen dans une communauté.

**Mutualisme** coévolution de caractères dans deux clades.



- ▶ On souhaiterait étudier ces caractères dans un cadre similaire.
- ▶ Sans hypothèse d'indépendance entre caractères portés par lignées différentes.

# Coévolution des caractères de différentes espèces

## Attraction ou répulsion dans l'espace des caractères

- ▶ On possède de bons exemples de caractères non-indépendants.

**Imitation** attraction vers le caractère moyen dans une communauté.



**Mutualisme** coévolution de caractères dans deux clades.



- ▶ On souhaiterait étudier ces caractères dans un cadre similaire.
- ▶ Sans hypothèse d'indépendance entre caractères portés par lignées différentes.

# Coévolution des caractères de différentes espèces

## Attraction ou répulsion dans l'espace des caractères

- ▶ On possède de bons exemples de caractères non-indépendants.

**Imitation** attraction vers le caractère moyen dans une communauté.



**Mutualisme** coévolution de caractères dans deux clades.



- ▶ On souhaiterait étudier ces caractères dans un cadre similaire.
- ▶ Sans hypothèse d'indépendance entre caractères portés par lignées différentes.

# Proposition d'un cadre de modélisation

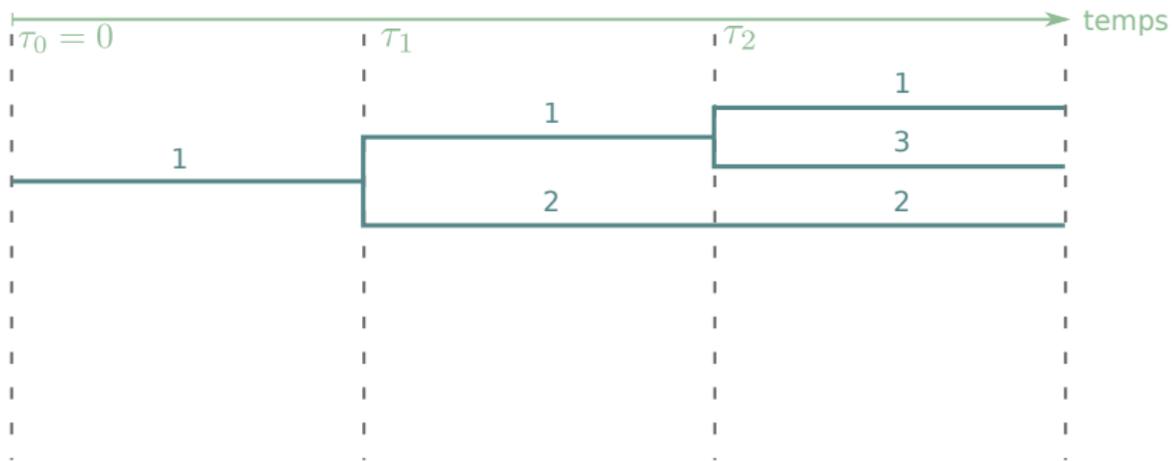
## Le principe



- ▶ La distribution des caractères au présent est gaussienne.
- ▶ On peut calculer numériquement sa moyenne et sa variance.

# Proposition d'un cadre de modélisation

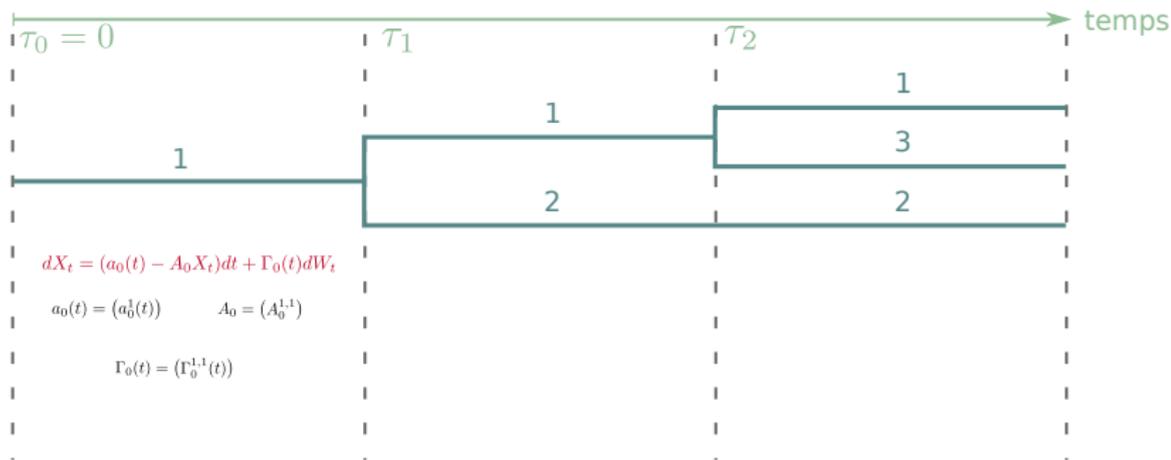
## Le principe



- ▶ La distribution des caractères au présent est gaussienne.
- ▶ On peut calculer numériquement sa moyenne et sa variance.

# Proposition d'un cadre de modélisation

## Le principe



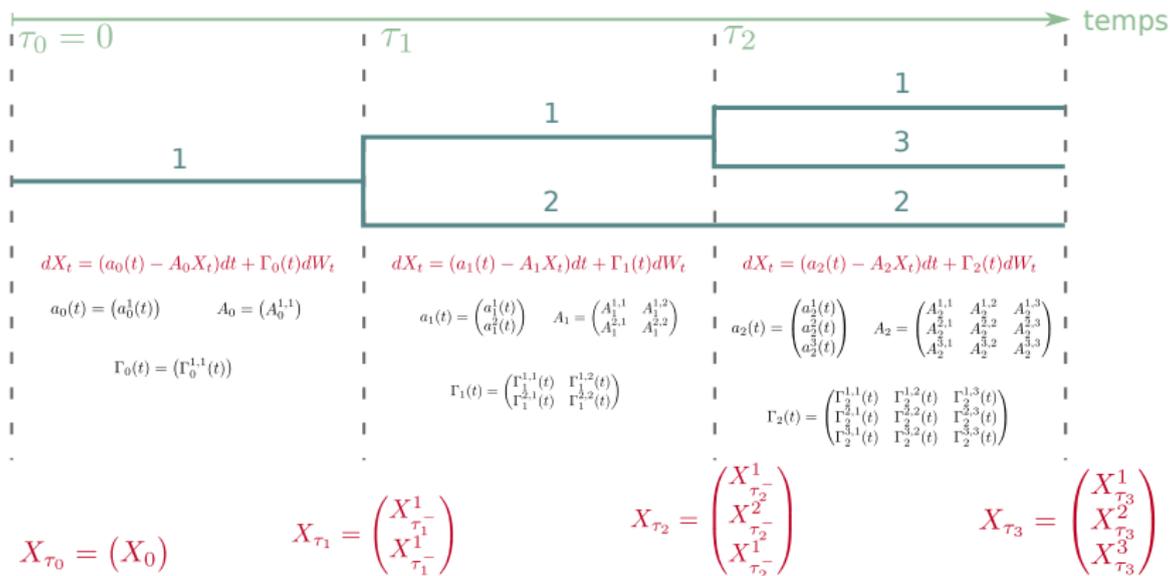
$$X_{\tau_0} = (X_0) \quad X_{\tau_1} = \begin{pmatrix} X_{\tau_1}^1 \\ X_{\tau_1}^1 \end{pmatrix}$$

- ▶ La distribution des caractères au présent est gaussienne.
- ▶ On peut calculer numériquement sa moyenne et sa variance.



# Proposition d'un cadre de modélisation

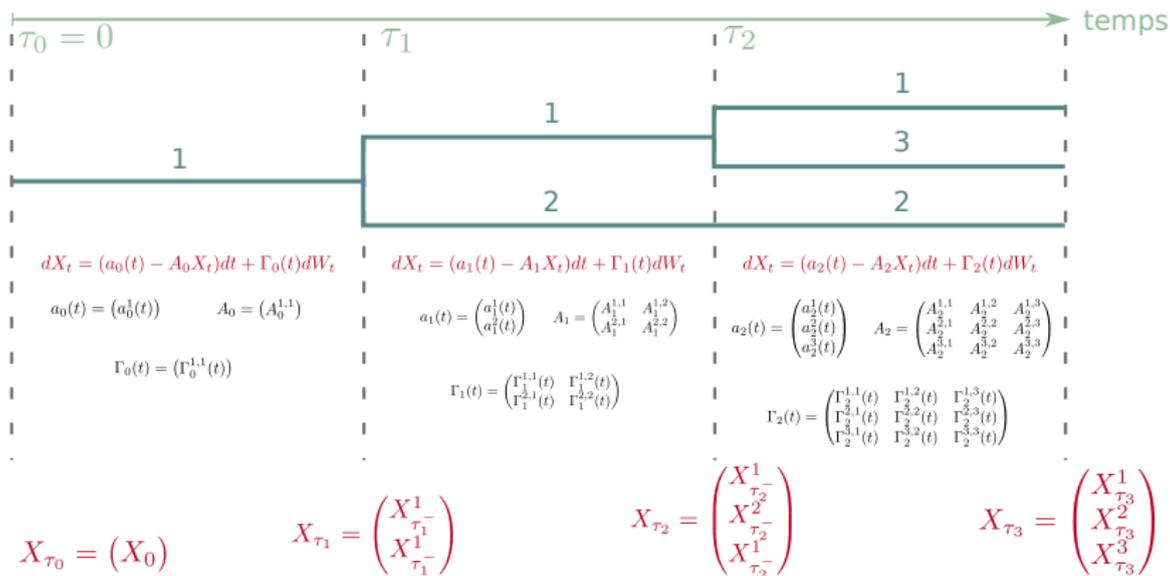
## Le principe



- ▶ La distribution des caractères au présent est gaussienne.
- ▶ On peut calculer numériquement sa moyenne et sa variance.

# Proposition d'un cadre de modélisation

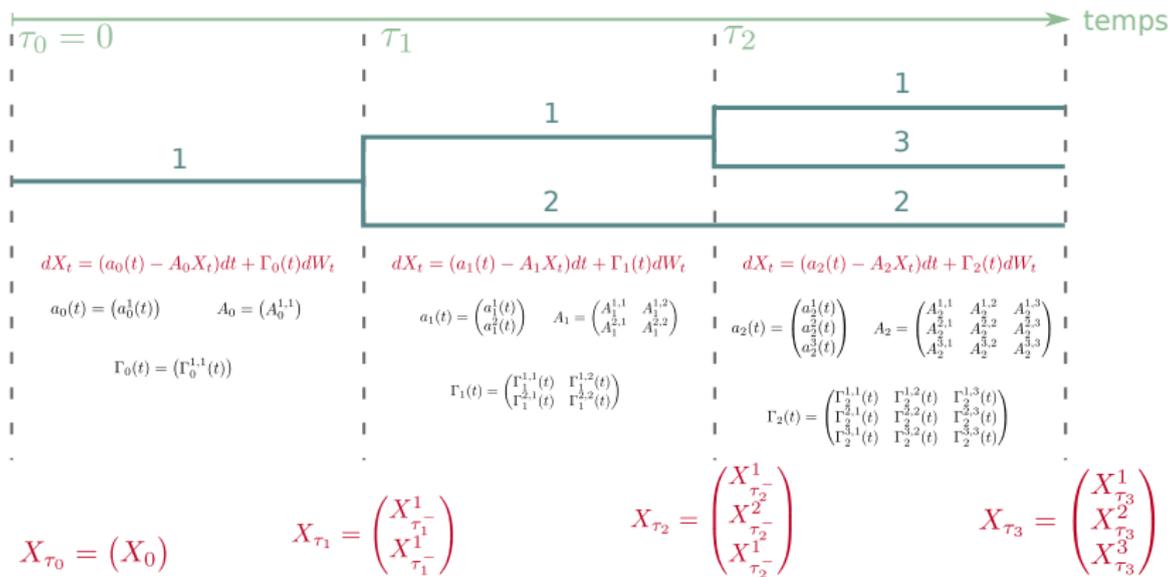
## Le principe



- La distribution des caractères au présent est gaussienne.
- On peut calculer numériquement sa moyenne et sa variance.

# Proposition d'un cadre de modélisation

## Le principe

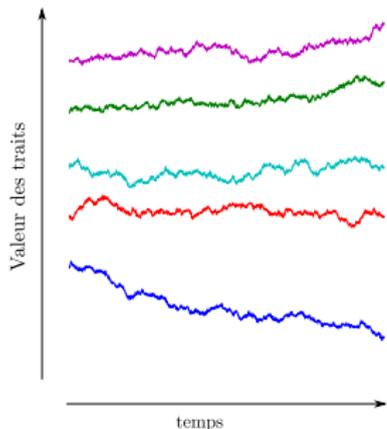


- ▶ La distribution des caractères au présent est gaussienne.
- ▶ On peut calculer numériquement sa moyenne et sa variance.

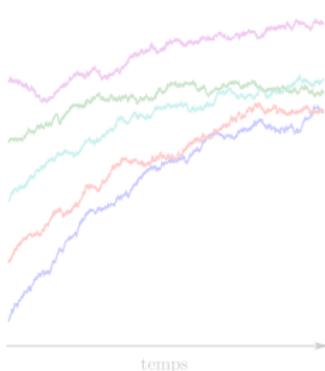
# Proposition d'un cadre de modélisation

Ce que l'on peut faire avec

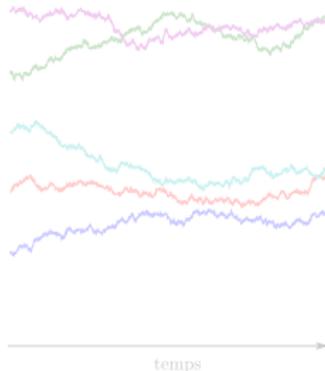
$$dX_t = (a - AX_t)dt + \Gamma dW_t$$



$$a = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \quad A = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix}$$



$$a = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \quad A = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}$$

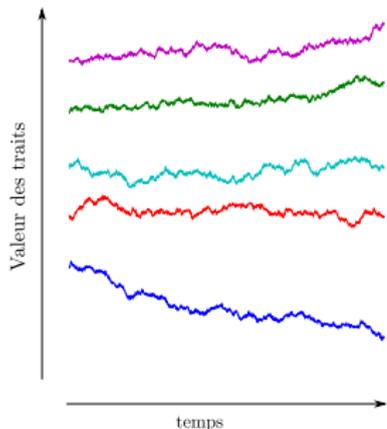


$$a = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \quad A = \begin{pmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ -\frac{1}{2} & 1 & -\frac{1}{2} & 0 & 0 \\ -\frac{1}{2} & -\frac{1}{2} & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}$$

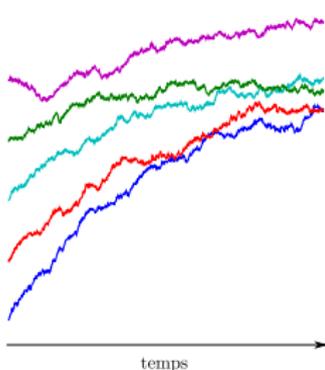
# Proposition d'un cadre de modélisation

Ce que l'on peut faire avec

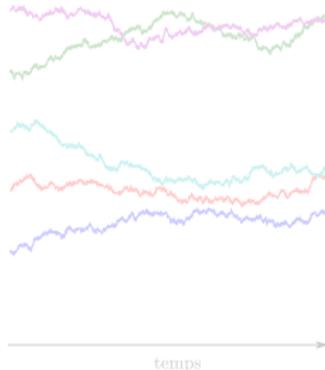
$$dX_t = (a - AX_t)dt + \Gamma dW_t$$



$$a = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \quad A = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix}$$



$$a = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \quad A = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}$$

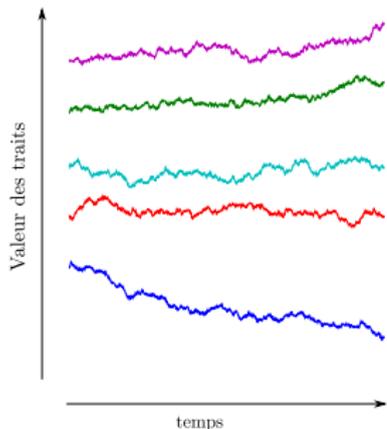


$$a = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \quad A = \begin{pmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ -\frac{1}{2} & 1 & -\frac{1}{2} & 0 & 0 \\ -\frac{1}{2} & -\frac{1}{2} & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}$$

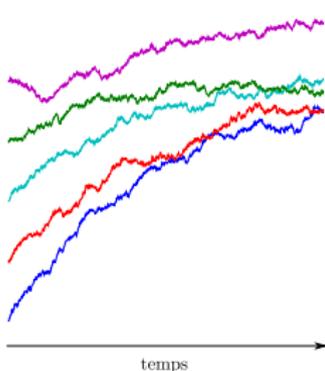
# Proposition d'un cadre de modélisation

Ce que l'on peut faire avec

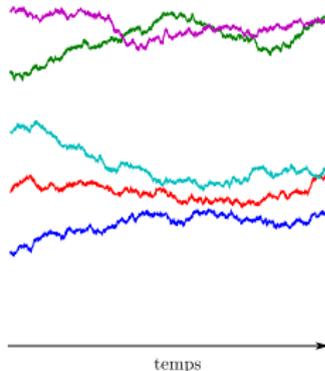
$$dX_t = (a - AX_t)dt + \Gamma dW_t$$



$$a = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \quad A = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix}$$



$$a = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \quad A = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}$$



$$a = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \quad A = \begin{pmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ -\frac{1}{2} & 1 & -\frac{1}{2} & 0 & 0 \\ -\frac{1}{2} & -\frac{1}{2} & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}$$

# Exemples d'applications

## Déplacement de caractères par compétition



Lézards *Anolis* dans les Antilles.

Traits étudiés :

- ▶ crâne, mâchoires.
- ▶ fémur, tibia, humerus, radius, ...
- ▶ pelvis, queue.

## Exemples d'applications

### Déplacement de caractères par compétition



Lézards *Anolis* dans les Antilles.

Traits étudiés :

- ▶ crâne, mâchoires.
- ▶ fémur, tibia, humerus, radius, ...
- ▶ pelvis, queue.



## Exemples d'applications

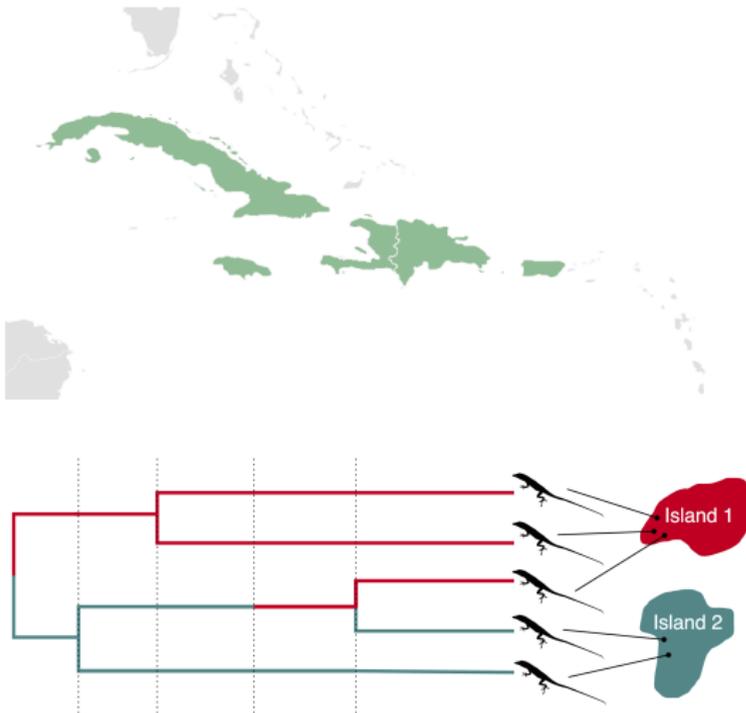
### Déplacement de caractères par compétition



Lézards *Anolis* dans les Antilles.

Traits étudiés :

- ▶ crâne, mâchoires.
- ▶ fémur, tibia, humerus, radius, ...
- ▶ pelvis, queue.

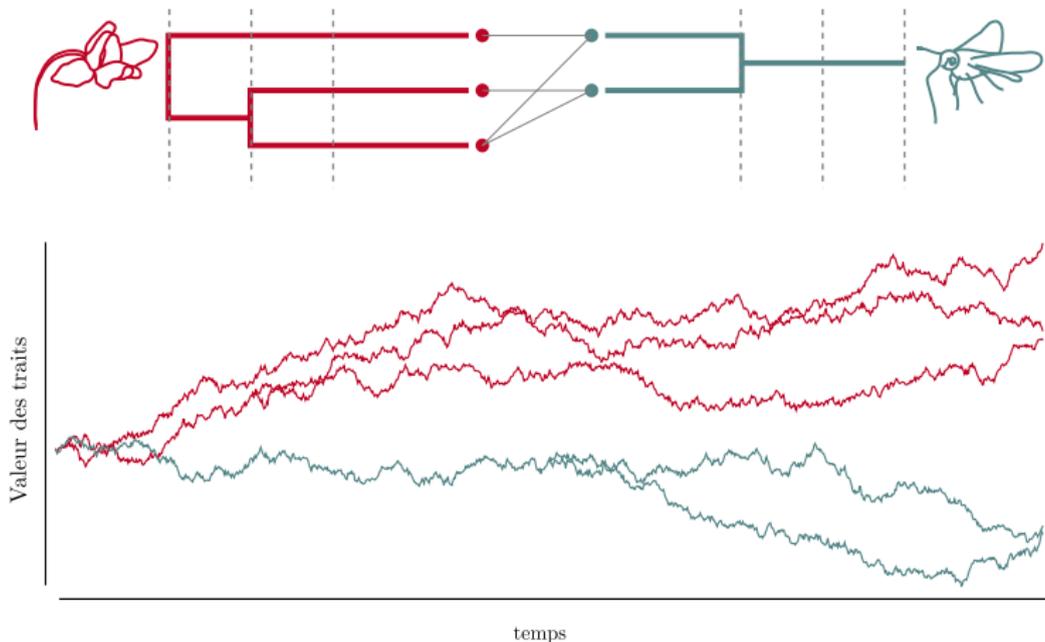


Drury et al., 2016, *Syst. Biol.*

## Exemples d'applications

### Coévolution de caractères sur différents clades

- ▶ Coévolution de caractères impliqués dans des interactions mutualistes.
- ▶ Pourrait être couplé à un modèle d'évolution du réseau d'interaction.



## Introduction

- Macroévolution
- Arbre phylogénétique
- Surcouche de modélisation probabiliste

## Diversification

- État de l'art
- La définition d'espèce en modélisation
- Application à l'étude de la diversification



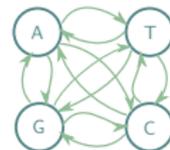
## Évolution phénotypique

- État de l'art
- Coévolution des caractères de différentes espèces
- Proposition d'un cadre de modélisation
- Exemples d'applications



## Évolution moléculaire

- Horloge moléculaire
- Modélisation proposée
- Méthode d'inférence



## Perspectives

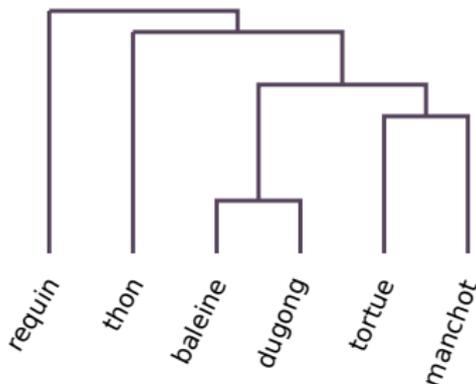
- Résumé
- Liens entre les chapitres

# Horloge moléculaire

Années 60, Zuckerkandl et Pauling

- Les différences deux à deux sont compatibles avec les longueurs de branches d'un arbre ultramétrique.

	<i>requin</i>	<i>thon</i>	<i>baleine</i>	<i>dugong</i>	<i>tortue</i>	<i>manchot</i>
<i>requin</i>	0	46	47	45	48	44
<i>thon</i>		0	43	42	39	41
<i>baleine</i>			0	9	30	33
<i>dugong</i>				0	32	34
<i>tortue</i>					0	27
<i>manchot</i>						0



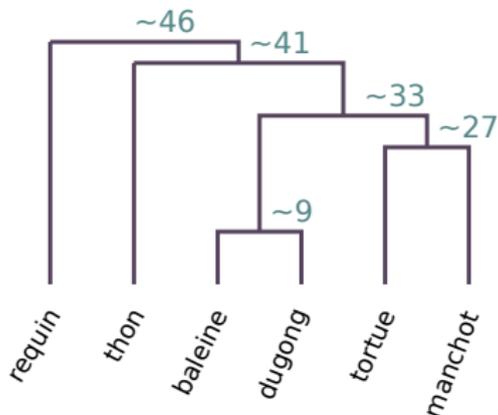
- On peut imaginer des mutations survenant selon un processus de Poisson à taux constant.

# Horloge moléculaire

Années 60, Zuckerkandl et Pauling

- ▶ Les différences deux à deux sont compatibles avec les longueurs de branches d'un arbre ultramétrique.

	requin	thon	baleine	dugong	tortue	manchot
requin	0	46	47	45	48	44
thon		0	43	42	39	41
baleine			0	9	30	33
dugong				0	32	34
tortue					0	27
manchot						0

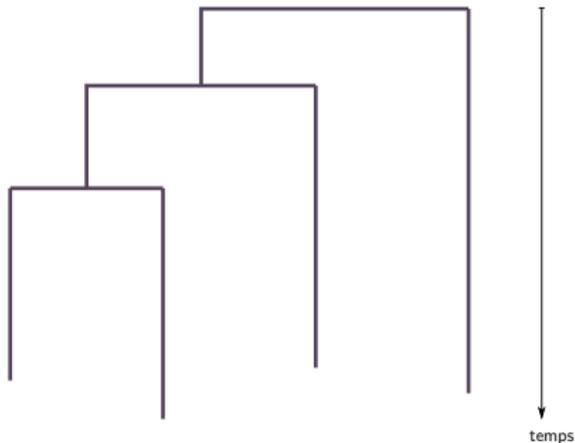
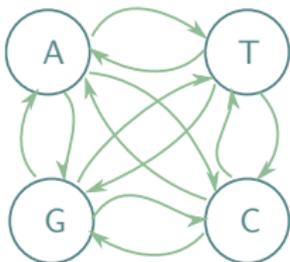


- ▶ On peut imaginer des mutations survenant selon un processus de Poisson à taux constant.

# Horloge moléculaire

## Modèle markovien d'évolution moléculaire

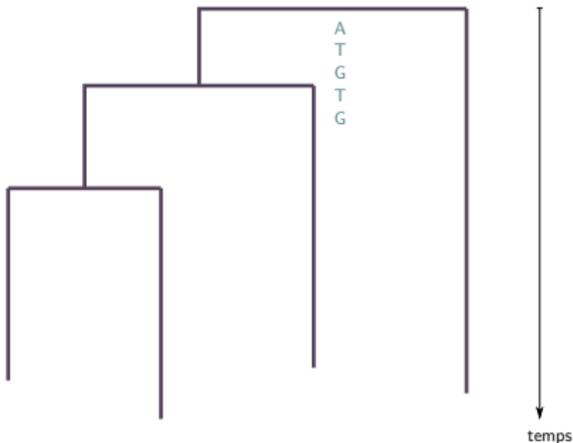
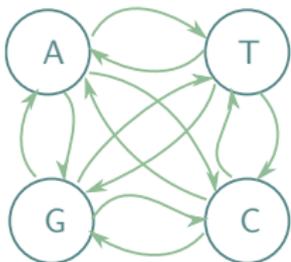
- ▶ On se donne une chaîne de Markov stationnaire.
- ▶ Les nucléotides à la racine sont tirés indépendamment dans la loi stationnaire.
- ▶ Chaque nucléotide évolue indépendamment des autres sur chaque branche.
- ▶ Le processus est copié en deux processus indépendants à chaque branchement.



# Horloge moléculaire

## Modèle markovien d'évolution moléculaire

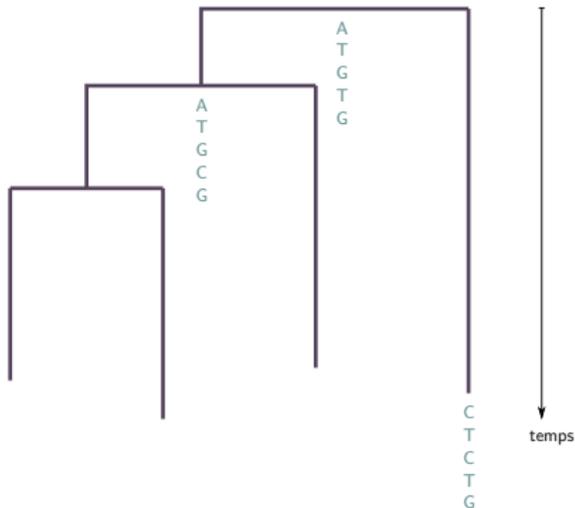
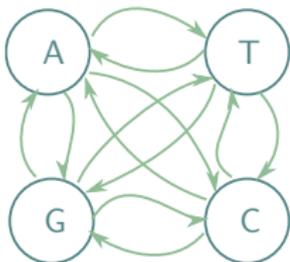
- ▶ On se donne une chaîne de Markov stationnaire.
- ▶ Les nucléotides à la racine sont tirés indépendamment dans la loi stationnaire.
- ▶ Chaque nucléotide évolue indépendamment des autres sur chaque branche.
- ▶ Le processus est copié en deux processus indépendants à chaque branchement.



# Horloge moléculaire

## Modèle markovien d'évolution moléculaire

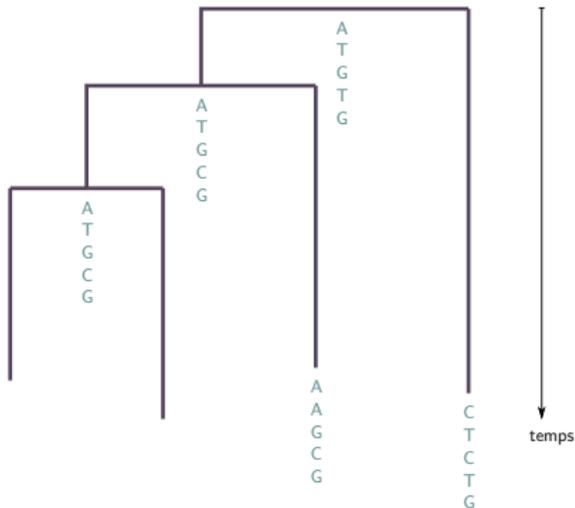
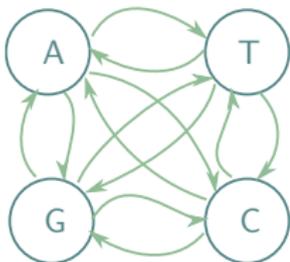
- ▶ On se donne une chaîne de Markov stationnaire.
- ▶ Les nucléotides à la racine sont tirés indépendamment dans la loi stationnaire.
- ▶ Chaque nucléotide évolue indépendamment des autres sur chaque branche.
- ▶ Le processus est copié en deux processus indépendants à chaque branchement.



# Horloge moléculaire

## Modèle markovien d'évolution moléculaire

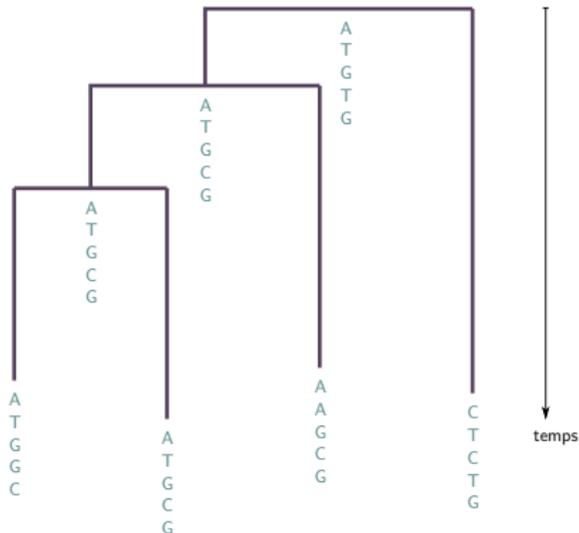
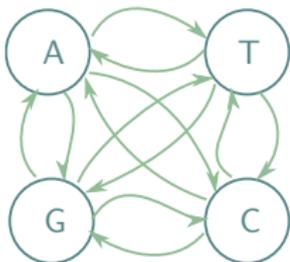
- ▶ On se donne une chaîne de Markov stationnaire.
- ▶ Les nucléotides à la racine sont tirés indépendamment dans la loi stationnaire.
- ▶ Chaque nucléotide évolue indépendamment des autres sur chaque branche.
- ▶ Le processus est copié en deux processus indépendants à chaque branchement.



# Horloge moléculaire

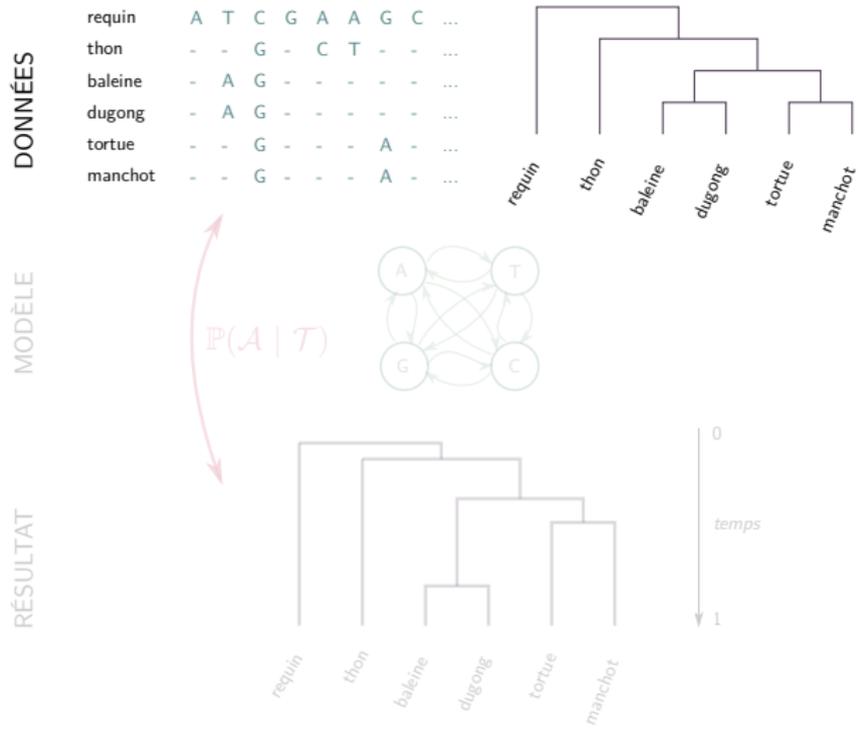
## Modèle markovien d'évolution moléculaire

- ▶ On se donne une chaîne de Markov stationnaire.
- ▶ Les nucléotides à la racine sont tirés indépendamment dans la loi stationnaire.
- ▶ Chaque nucléotide évolue indépendamment des autres sur chaque branche.
- ▶ Le processus est copié en deux processus indépendants à chaque branchement.



# Horloge moléculaire

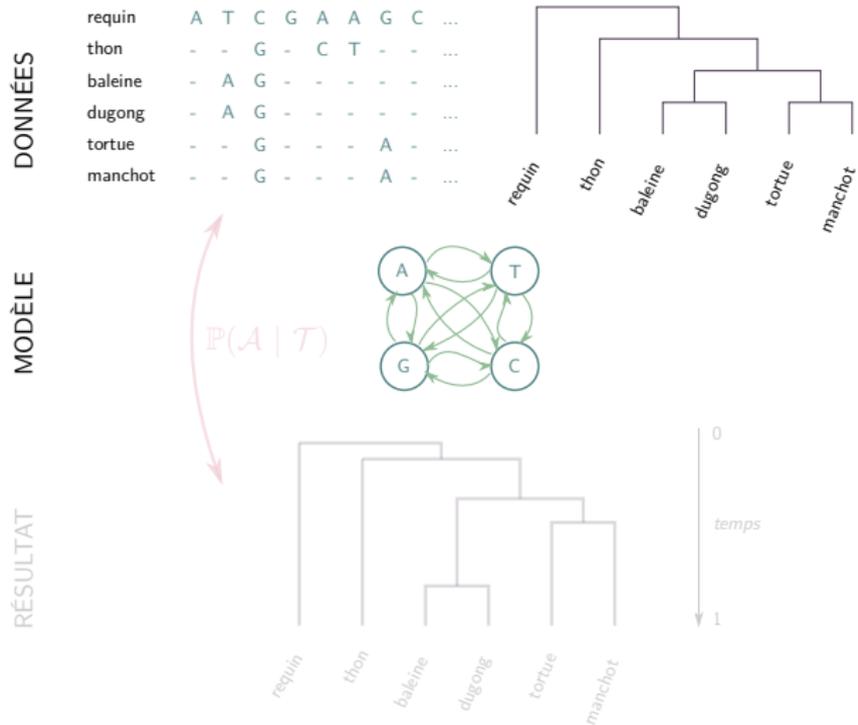
La version stricte ou totalement relâchée



Perrière et Brochier-Armanet, 2010, *Concepts et méthodes en phylogénie moléculaire*.

# Horloge moléculaire

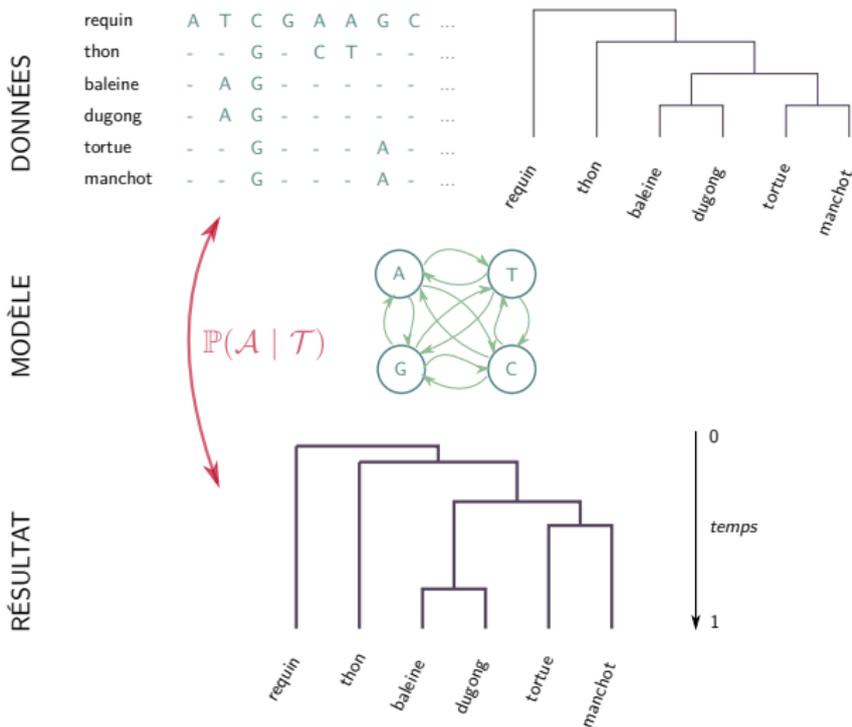
La version stricte ou totalement relâchée



Perrière et Brochier-Armanet, 2010, *Concepts et méthodes en phylogénie moléculaire*.

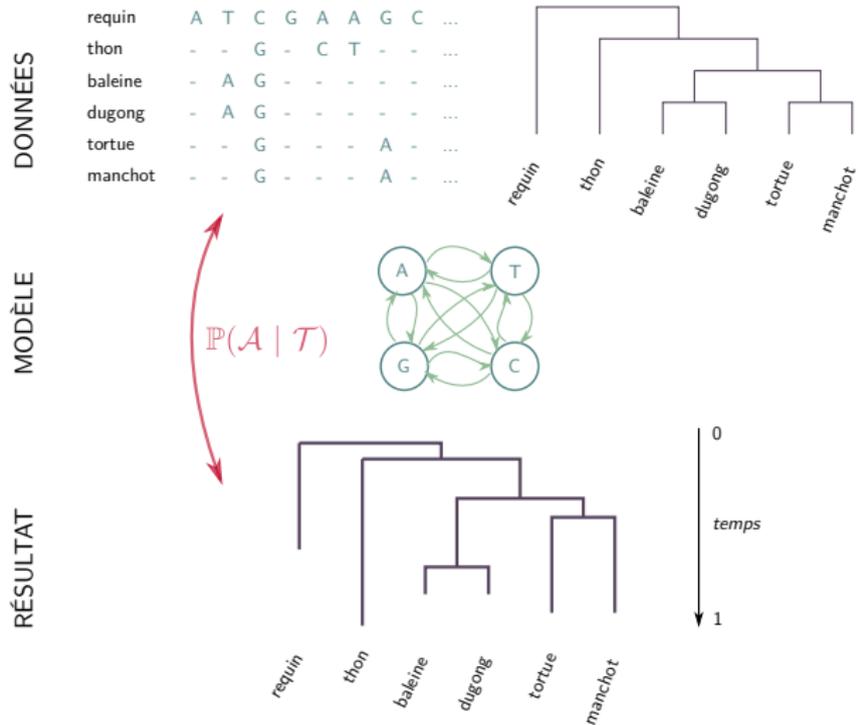
# Horloge moléculaire

La version stricte ou totalement relâchée



# Horloge moléculaire

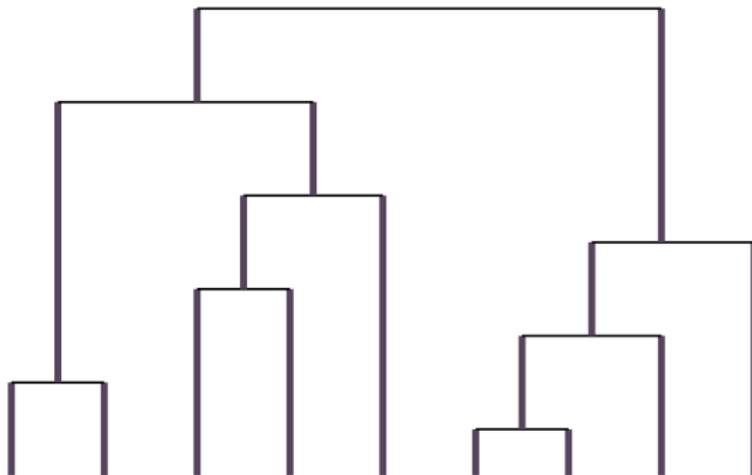
La version stricte ou totalement relâchée



Perrière et Brochier-Armanet, 2010, *Concepts et méthodes en phylogénie moléculaire*.

# Horloge moléculaire

Toute une gamme de relaxations

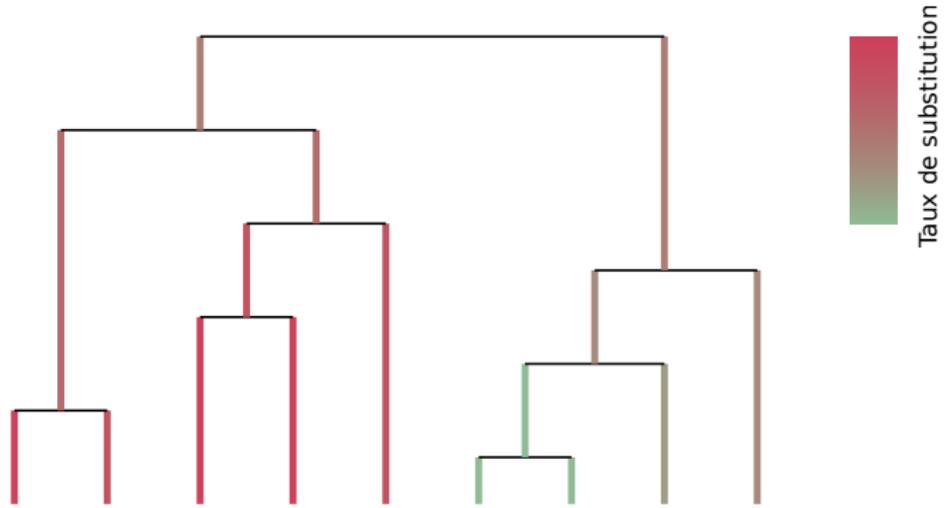


- ▶ Les longueurs de branches suivent une loi donnée.
- ▶ Le taux de substitution peut varier selon les branches, de façon auto-corrélée.
- ▶ ... ou non-auto-corrélée.

Lepage et al., 2007, *Mol. Biol. Evol.*

# Horloge moléculaire

Toute une gamme de relaxations

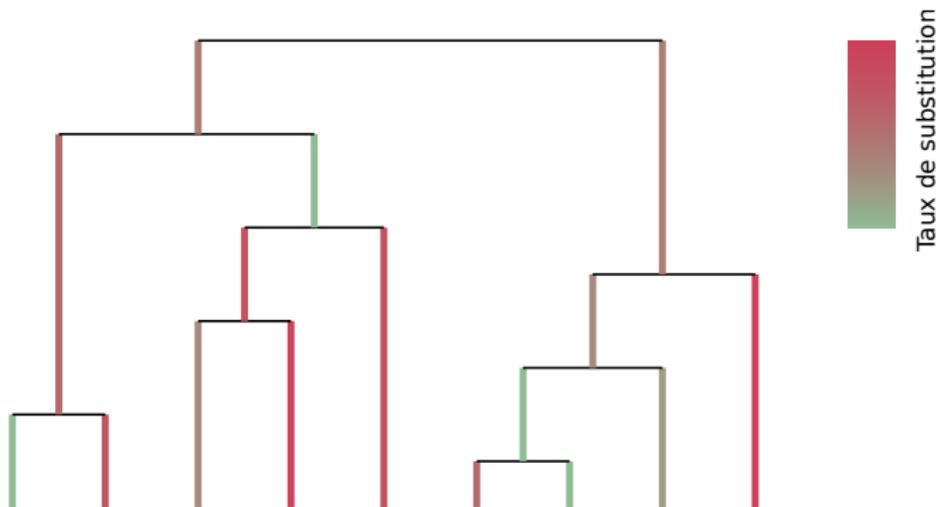


- ▶ Les longueurs de branches suivent une loi donnée.
- ▶ Le taux de substitution peut varier selon les branches, de façon auto-corrélée.
- ▶ ... ou non-auto-corrélée.

Lepage et al., 2007, *Mol. Biol. Evol.*

# Horloge moléculaire

Toute une gamme de relaxations



- ▶ Les longueurs de branches suivent une loi donnée.
- ▶ Le taux de substitution peut varier selon les branches, de façon auto-corrélée.
- ▶ ... ou non-auto-corrélée.

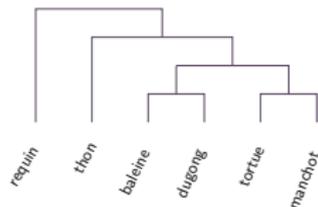
Lepage et al., 2007, *Mol. Biol. Evol.*

# Horloge moléculaire

## Principe de la datation

DONNÉES

requin	A	T	C	G	A	A	G	C	...
thon	-	-	G	-	C	T	-	-	...
baleine	-	A	G	-	-	-	-	-	...
dugong	-	A	G	-	-	-	-	-	...
tortue	-	-	G	-	-	-	A	-	...
manchot	-	-	G	-	-	-	A	-	...



MODÈLE



$\mathcal{A} :=$  Alignement  
 $P(\mathcal{A} | \mathcal{T}, \mathcal{R})$

$\mathcal{R} :=$  Taux de substitution  
 $P(\mathcal{R} | \mathcal{T})$

$\mathcal{T} :=$  Arbre  
 $P(\mathcal{T})$

$$P(\mathcal{T}, \mathcal{R} | \mathcal{A}) \propto P(\mathcal{A} | \mathcal{R}, \mathcal{T}) P(\mathcal{R} | \mathcal{T}) P(\mathcal{T})$$

RÉSULTAT

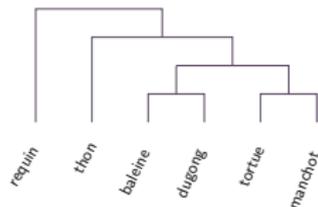


# Horloge moléculaire

## Principe de la datation

DONNÉES

requin	A	T	C	G	A	A	G	C	...
thon	-	-	G	-	C	T	-	-	...
baleine	-	A	G	-	-	-	-	-	...
dugong	-	A	G	-	-	-	-	-	...
tortue	-	-	G	-	-	-	A	-	...
manchot	-	-	G	-	-	-	A	-	...



MODÈLE



$\mathcal{A} :=$  Alignement  
 $\mathbb{P}(\mathcal{A} | \mathcal{T}, \mathcal{R})$

$\mathcal{R} :=$  Taux de substitution  
 $\mathbb{P}(\mathcal{R} | \mathcal{T})$

$\mathcal{T} :=$  Arbre  
 $\mathbb{P}(\mathcal{T})$

$$\mathbb{P}(\mathcal{T}, \mathcal{R} | \mathcal{A}) \propto \mathbb{P}(\mathcal{A} | \mathcal{R}, \mathcal{T}) \mathbb{P}(\mathcal{R} | \mathcal{T}) \mathbb{P}(\mathcal{T})$$

RÉSULTAT

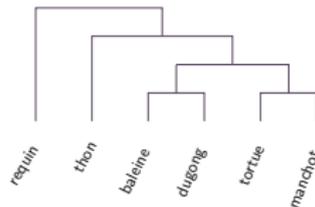


# Horloge moléculaire

## Principe de la datation

DONNÉES

requin	A	T	C	G	A	A	G	C	...
thon	-	-	G	-	C	T	-	-	...
baleine	-	A	G	-	-	-	-	-	...
dugong	-	A	G	-	-	-	-	-	...
tortue	-	-	G	-	-	-	A	-	...
manchot	-	-	G	-	-	-	A	-	...



MODÈLE



$\mathcal{A} :=$  Alignement  
 $\mathbb{P}(\mathcal{A} | \mathcal{T}, \mathcal{R})$



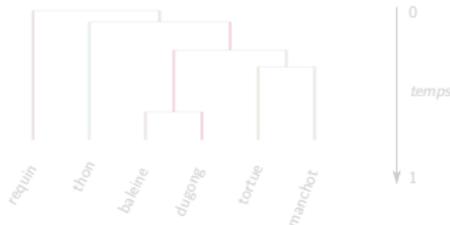
$\mathcal{R} :=$  Taux de substitution  
 $\mathbb{P}(\mathcal{R} | \mathcal{T})$



$\mathcal{T} :=$  Arbre  
 $\mathbb{P}(\mathcal{T})$

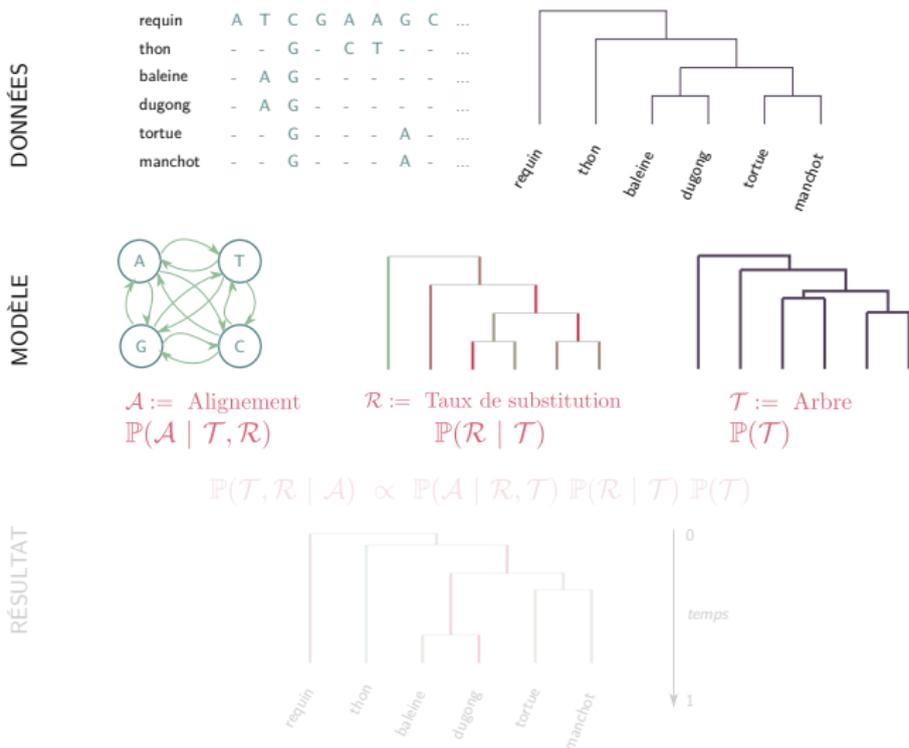
$$\mathbb{P}(\mathcal{T}, \mathcal{R} | \mathcal{A}) \propto \mathbb{P}(\mathcal{A} | \mathcal{R}, \mathcal{T}) \mathbb{P}(\mathcal{R} | \mathcal{T}) \mathbb{P}(\mathcal{T})$$

RÉSULTAT



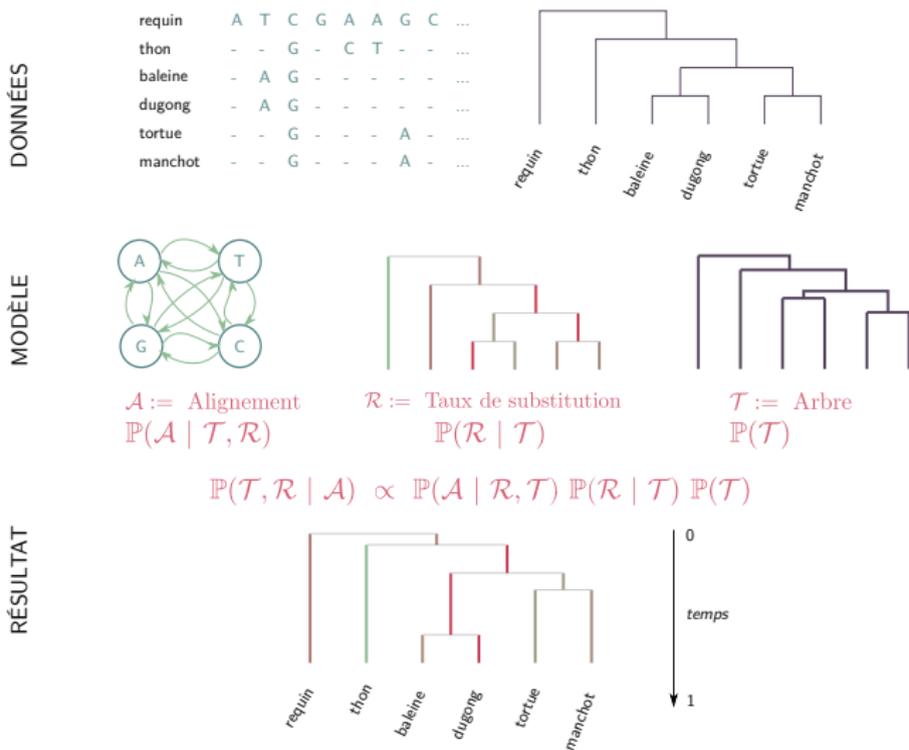
# Horloge moléculaire

## Principe de la datation



# Horloge moléculaire

## Principe de la datation



## Modélisation proposée

Motivation biologique : génomique de la spéciation

Nouvelle hypothèse : et si les déviations à l'horloge moléculaire étaient ponctuelles ?

Sélection divergente potentiellement forte à la spéciation :

- ▶ utilisation de ressources différentes,
- ▶ adaptation à un habitat différent,
- ▶ reconnaissance de partenaires sexuels.

Pourrait se traduire génétiquement :

- ▶ îlots génomiques de spéciation,
- ▶ accélération brève du taux de substitution.

## Modélisation proposée

Motivation biologique : génomique de la spéciation

Nouvelle hypothèse : et si les déviations à l'horloge moléculaire étaient ponctuelles ?

Sélection divergente potentiellement forte à la spéciation :

- ▶ utilisation de ressources différentes,
- ▶ adaptation à un habitat différent,
- ▶ reconnaissance de partenaires sexuels.

Pourrait se traduire génétiquement :

- ▶ îlots génomiques de spéciation,
- ▶ accélération brève du taux de substitution.

## Modélisation proposée

Motivation biologique : génomique de la spéciation

Nouvelle hypothèse : et si les déviations à l'horloge moléculaire étaient ponctuelles ?

Sélection divergente potentiellement forte à la spéciation :

- ▶ utilisation de ressources différentes,
- ▶ adaptation à un habitat différent,
- ▶ reconnaissance de partenaires sexuels.

Pourrait se traduire génétiquement :

- ▶ îlots génomiques de spéciation,
- ▶ accélération brève du taux de substitution.

## Modélisation proposée

Motivation biologique : génomique de la spéciation

Nouvelle hypothèse : et si les déviations à l'horloge moléculaire étaient ponctuelles ?

Sélection divergente potentiellement forte à la spéciation :

- ▶ utilisation de ressources différentes,
- ▶ adaptation à un habitat différent,
- ▶ reconnaissance de partenaires sexuels.

Pourrait se traduire génétiquement :

- ▶ îlots génomiques de spéciation,
- ▶ accélération brève du taux de substitution.

# Modélisation proposée

Motivation biologique : génomique de la spéciation

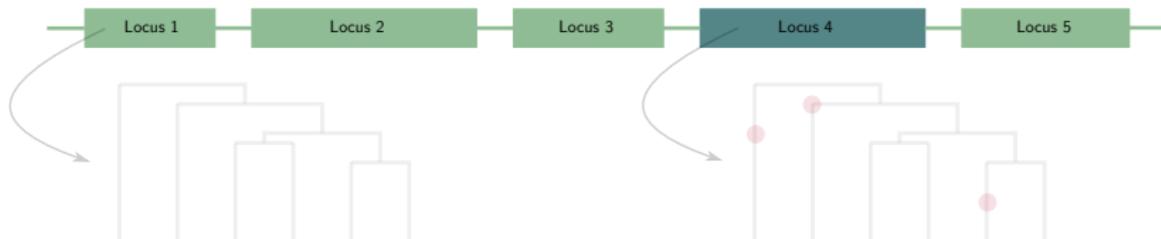
Nouvelle hypothèse : et si les déviations à l'horloge moléculaire étaient ponctuelles ?

Sélection divergente potentiellement forte à la spéciation :

- ▶ utilisation de ressources différentes,
- ▶ adaptation à un habitat différent,
- ▶ reconnaissance de partenaires sexuels.

Pourrait se traduire génétiquement :

- ▶ îlots génomiques de spéciation,
- ▶ accélération brève du taux de substitution.



## Modélisation proposée

Motivation biologique : génomique de la spéciation

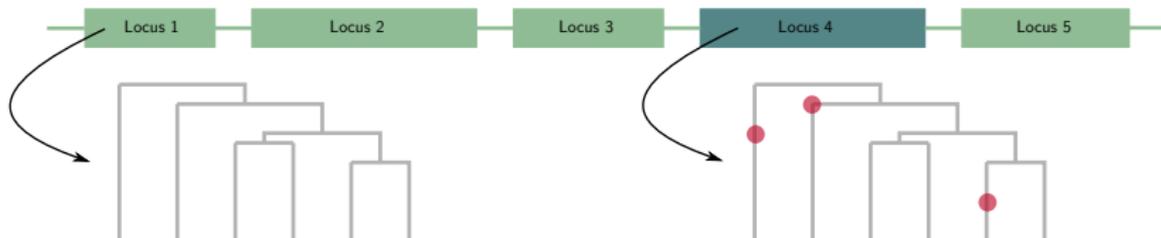
Nouvelle hypothèse : et si les déviations à l'horloge moléculaire étaient ponctuelles ?

Sélection divergente potentiellement forte à la spéciation :

- ▶ utilisation de ressources différentes,
- ▶ adaptation à un habitat différent,
- ▶ reconnaissance de partenaires sexuels.

Pourrait se traduire génétiquement :

- ▶ îlots génomiques de spéciation,
- ▶ accélération brève du taux de substitution.



# Modélisation proposée

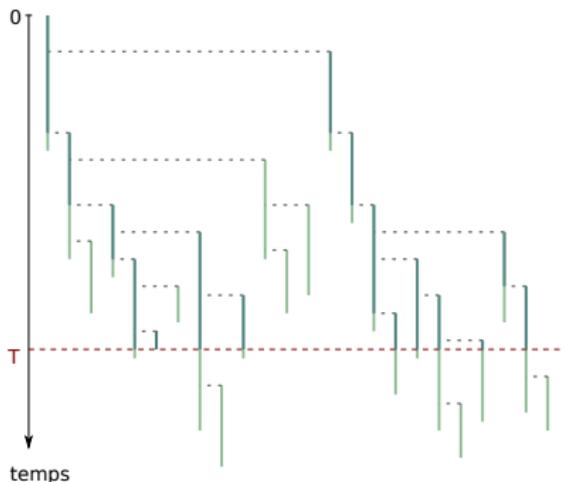
## Introduction de pics de mutation

Du processus complet ...

- ▶ Processus de naissance-mort,
- ▶ probabilité  $\nu$  de *pics* à la naissance.

... à l'arbre reconstruit.

- ▶ probabilité  $\nu$  de *pics* à un branchement,
- ▶ processus de Poisson inhomogène le long des branches.



- ▶ Évolution markovienne le long des branches,
- ▶ Sur un pic, chaque nucléotide de la séquence mute avec probabilité  $\kappa$ .

# Modélisation proposée

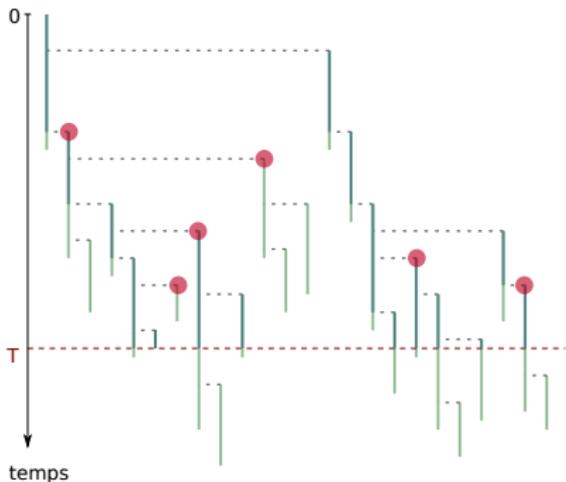
## Introduction de pics de mutation

Du processus complet ...

- ▶ Processus de naissance-mort,
- ▶ probabilité  $\nu$  de *pics* à la naissance.

... à l'arbre reconstruit.

- ▶ probabilité  $\nu$  de *pics* à un branchement,
- ▶ processus de Poisson inhomogène le long des branches.



- ▶ Évolution markovienne le long des branches,
- ▶ Sur un pic, chaque nucléotide de la séquence mute avec probabilité  $\kappa$ .

# Modélisation proposée

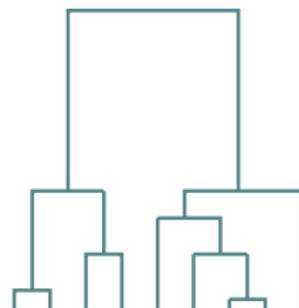
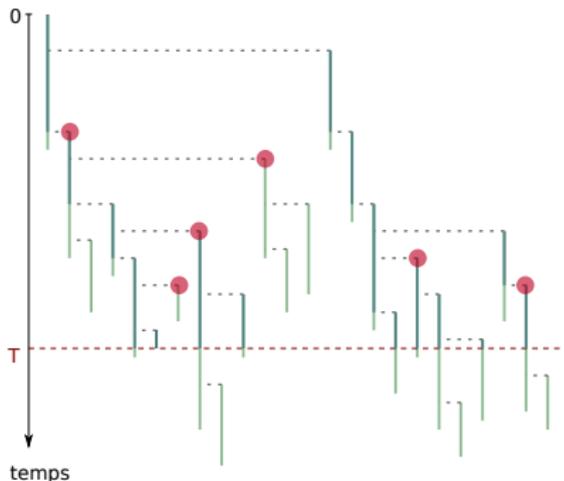
## Introduction de pics de mutation

Du processus complet ...

- ▶ Processus de naissance-mort,
- ▶ probabilité  $\nu$  de *pics* à la naissance.

... à l'arbre reconstruit.

- ▶ probabilité  $\nu$  de *pics* à un branchement,
- ▶ processus de Poisson inhomogène le long des branches.



- ▶ Évolution markovienne le long des branches,
- ▶ Sur un pic, chaque nucléotide de la séquence mute avec probabilité  $\kappa$ .

# Modélisation proposée

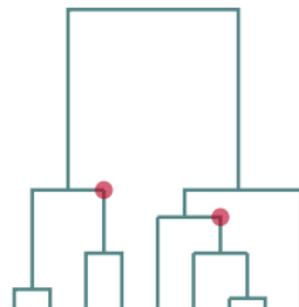
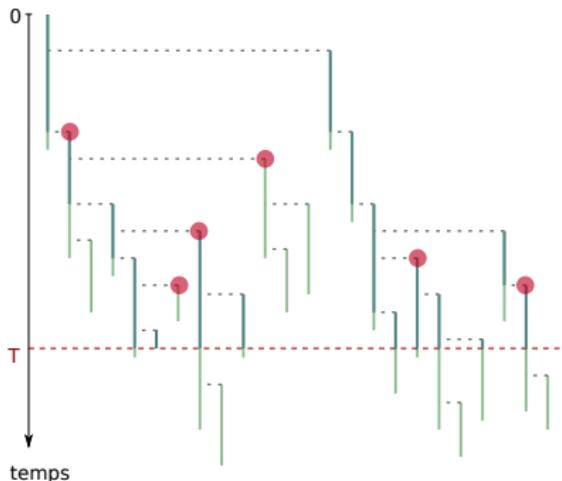
## Introduction de pics de mutation

Du processus complet ...

- ▶ Processus de naissance-mort,
- ▶ probabilité  $\nu$  de *pics* à la naissance.

... à l'arbre reconstruit.

- ▶ probabilité  $\nu$  de *pics* à un branchement,
- ▶ processus de Poisson inhomogène le long des branches.



- ▶ Évolution markovienne le long des branches,
- ▶ Sur un pic, chaque nucléotide de la séquence mute avec probabilité  $\kappa$ .

# Modélisation proposée

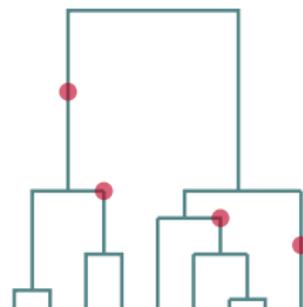
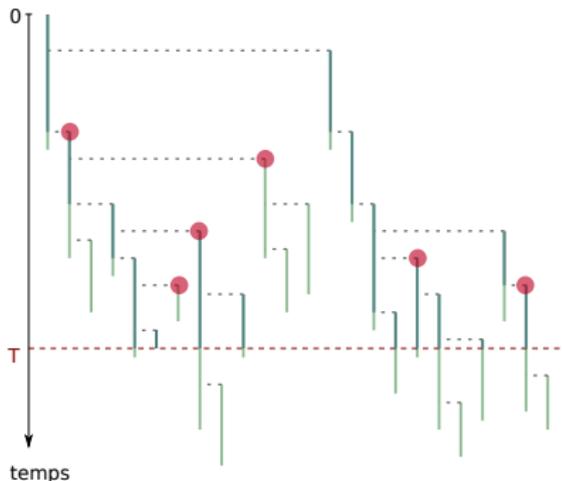
## Introduction de pics de mutation

Du processus complet ...

- ▶ Processus de naissance-mort,
- ▶ probabilité  $\nu$  de *pics* à la naissance.

... à l'arbre reconstruit.

- ▶ probabilité  $\nu$  de *pics* à un branchement,
- ▶ processus de Poisson inhomogène le long des branches.



- ▶ Évolution markovienne le long des branches,
- ▶ Sur un pic, chaque nucléotide de la séquence mute avec probabilité  $\kappa$ .

# Modélisation proposée

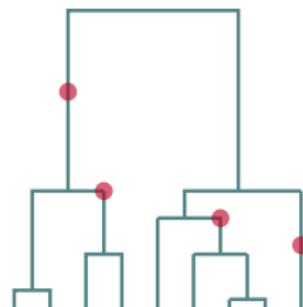
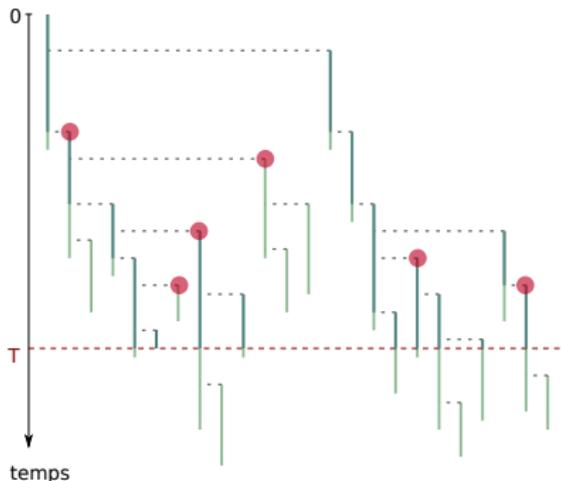
## Introduction de pics de mutation

Du processus complet ...

- ▶ Processus de naissance-mort,
- ▶ probabilité  $\nu$  de *pics* à la naissance.

... à l'arbre reconstruit.

- ▶ probabilité  $\nu$  de *pics* à un branchement,
- ▶ processus de Poisson inhomogène le long des branches.



- ▶ Évolution markovienne le long des branches,
- ▶ Sur un pic, chaque nucléotide de la séquence mute avec probabilité  $\kappa$ .

# Modélisation proposée

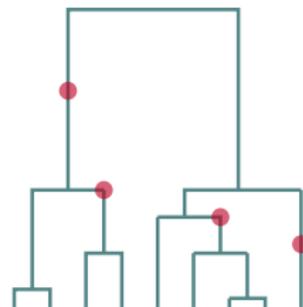
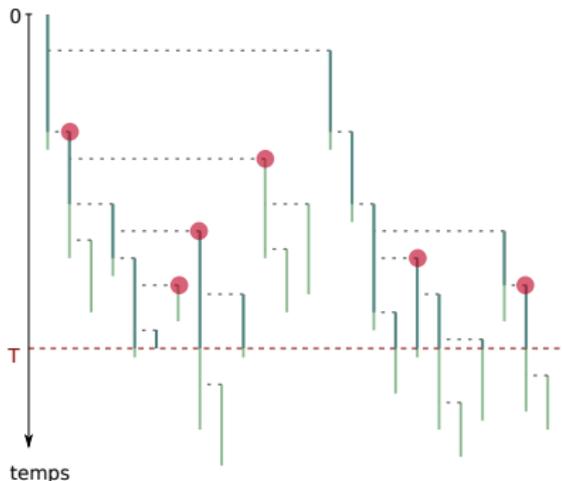
## Introduction de pics de mutation

Du processus complet ...

- ▶ Processus de naissance-mort,
- ▶ probabilité  $\nu$  de *pics* à la naissance.

... à l'arbre reconstruit.

- ▶ probabilité  $\nu$  de *pics* à un branchement,
- ▶ processus de Poisson inhomogène le long des branches.



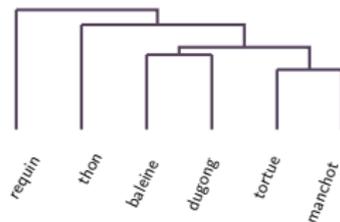
- ▶ Évolution markovienne le long des branches,
- ▶ Sur un pic, chaque nucléotide de la séquence mute avec probabilité  $\kappa$ .

# Méthode d'inférence

## Spikes sur l'arbre reconstruit

DONNÉES

requin	A	T	C	G	A	A	G	C	...
thon	-	-	G	-	C	T	-	-	...
baleine	-	A	G	-	-	-	-	-	...
dugong	-	A	G	-	-	-	-	-	...
tortue	-	-	G	-	-	-	A	-	...
manchot	-	-	G	-	-	-	A	-	...



MODÈLE



$\mathcal{A} :=$  Alignement  
 $\mathbb{P}(\mathcal{A} | \mathcal{S}, \mathcal{T})$



$\mathcal{S} :=$  Pics  
 $\mathbb{P}(\mathcal{S} | \mathcal{T})$



$\mathcal{T} :=$  Arbre  
 $\mathbb{P}(\mathcal{T})$

RÉSULTAT

$$\mathbb{P}(\mathcal{S} | \mathcal{A}, \mathcal{T}) \propto \mathbb{P}(\mathcal{A} | \mathcal{S}, \mathcal{T}) \mathbb{P}(\mathcal{S} | \mathcal{T}) \mathbb{P}(\mathcal{T})$$

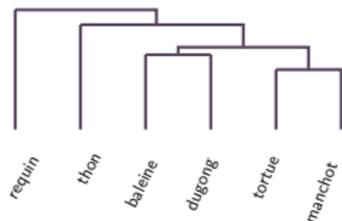


# Méthode d'inférence

## Spikes sur l'arbre reconstruit

DONNÉES

requin	A	T	C	G	A	A	G	C	...
thon	-	-	G	-	C	T	-	-	...
baleine	-	A	G	-	-	-	-	-	...
dugong	-	A	G	-	-	-	-	-	...
tortue	-	-	G	-	-	-	A	-	...
manchot	-	-	G	-	-	-	A	-	...



MODÈLE



$\mathcal{A} :=$  Alignement  
 $\mathbb{P}(\mathcal{A} | \mathcal{S}, \mathcal{T})$



$\mathcal{S} :=$  Pics  
 $\mathbb{P}(\mathcal{S} | \mathcal{T})$



$\mathcal{T} :=$  Arbre  
 $\mathbb{P}(\mathcal{T})$

$$\mathbb{P}(\mathcal{S} | \mathcal{A}, \mathcal{T}) \propto \mathbb{P}(\mathcal{A} | \mathcal{S}, \mathcal{T}) \mathbb{P}(\mathcal{S} | \mathcal{T}) \mathbb{P}(\mathcal{T})$$

RÉSULTAT

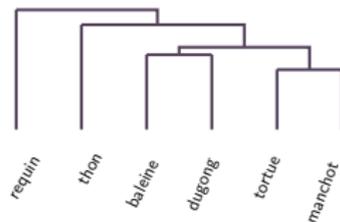


# Méthode d'inférence

## Spikes sur l'arbre reconstruit

DONNÉES

requin	A	T	C	G	A	A	G	C	...
thon	-	-	G	-	C	T	-	-	...
baleine	-	A	G	-	-	-	-	-	...
dugong	-	A	G	-	-	-	-	-	...
tortue	-	-	G	-	-	-	A	-	...
manchot	-	-	G	-	-	-	A	-	...



MODÈLE



$\mathcal{A} :=$  Alignement  
 $\mathbb{P}(\mathcal{A} | \mathcal{S}, \mathcal{T})$



$\mathcal{S} :=$  Pics  
 $\mathbb{P}(\mathcal{S} | \mathcal{T})$



$\mathcal{T} :=$  Arbre  
 $\mathbb{P}(\mathcal{T})$

RÉSULTAT

$$\mathbb{P}(\mathcal{S} | \mathcal{A}, \mathcal{T}) \propto \mathbb{P}(\mathcal{A} | \mathcal{S}, \mathcal{T}) \mathbb{P}(\mathcal{S} | \mathcal{T}) \mathbb{P}(\mathcal{T})$$

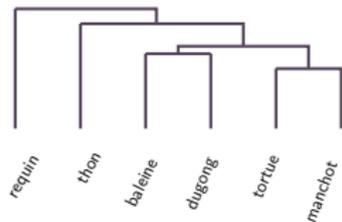


## Méthode d'inférence

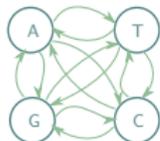
Spikes sur l'arbre reconstruit

DONNÉES

requin	A	T	C	G	A	A	G	C	...
thon	-	-	G	-	C	T	-	-	...
baleine	-	A	G	-	-	-	-	-	...
dugong	-	A	G	-	-	-	-	-	...
tortue	-	-	G	-	-	-	A	-	...
manchot	-	-	G	-	-	-	A	-	...



MODÈLE



$\mathcal{A} :=$  Alignement  
 $\mathbb{P}(\mathcal{A} | \mathcal{S}, \mathcal{T})$



$\mathcal{S} :=$  Pics  
 $\mathbb{P}(\mathcal{S} | \mathcal{T})$



$\mathcal{T} :=$  Arbre  
 $\mathbb{P}(\mathcal{T})$

RÉSULTAT

$$\mathbb{P}(\mathcal{S} | \mathcal{A}, \mathcal{T}) \propto \mathbb{P}(\mathcal{A} | \mathcal{S}, \mathcal{T}) \mathbb{P}(\mathcal{S} | \mathcal{T}) \mathbb{P}(\mathcal{T})$$

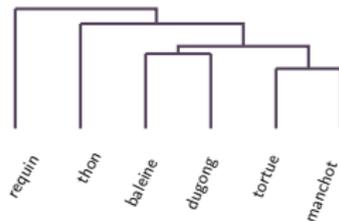


# Méthode d'inférence

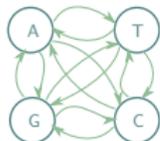
## Spikes sur l'arbre reconstruit

DONNÉES

requin	A	T	C	G	A	A	G	C	...
thon	-	-	G	-	C	T	-	-	...
baleine	-	A	G	-	-	-	-	-	...
dugong	-	A	G	-	-	-	-	-	...
tortue	-	-	G	-	-	-	A	-	...
manchot	-	-	G	-	-	-	A	-	...



MODÈLE



$\mathcal{A} :=$  Alignement  
 $\mathbb{P}(\mathcal{A} | \mathcal{S}, \mathcal{T})$



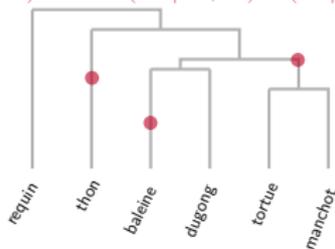
$\mathcal{S} :=$  Pics  
 $\mathbb{P}(\mathcal{S} | \mathcal{T})$



$\mathcal{T} :=$  Arbre  
 $\mathbb{P}(\mathcal{T})$

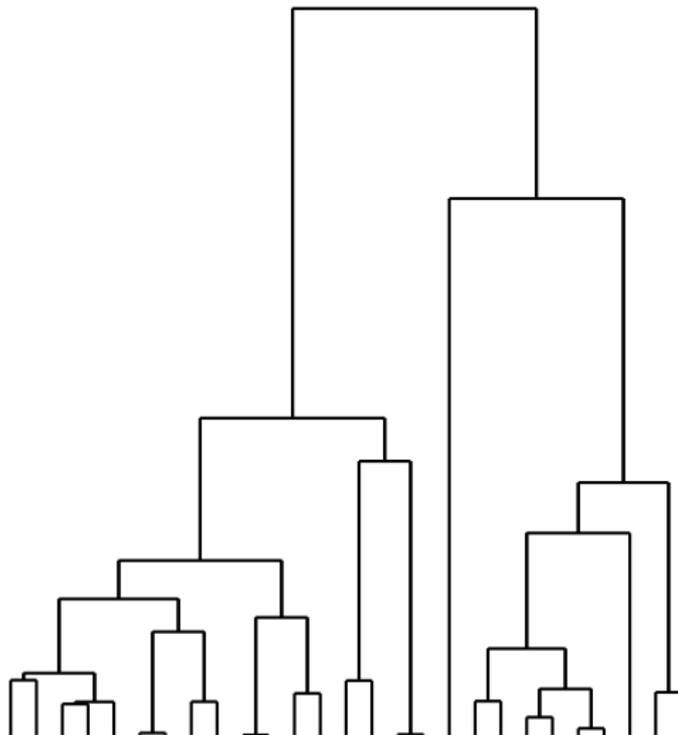
$$\mathbb{P}(\mathcal{S} | \mathcal{A}, \mathcal{T}) \propto \mathbb{P}(\mathcal{A} | \mathcal{S}, \mathcal{T}) \mathbb{P}(\mathcal{S} | \mathcal{T}) \mathbb{P}(\mathcal{T})$$

RÉSULTAT



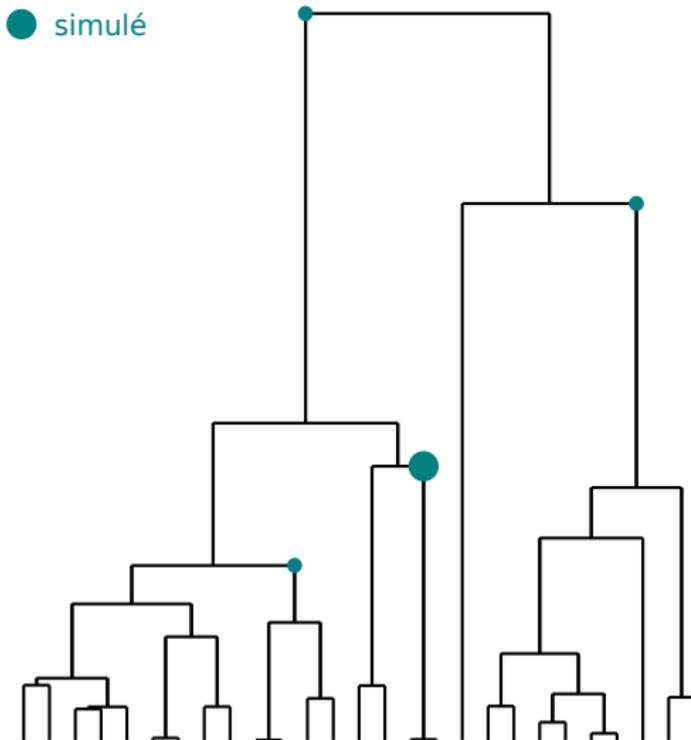
# Méthode d'inférence

Loi des pics sachant les longueurs de branche et l'alignement



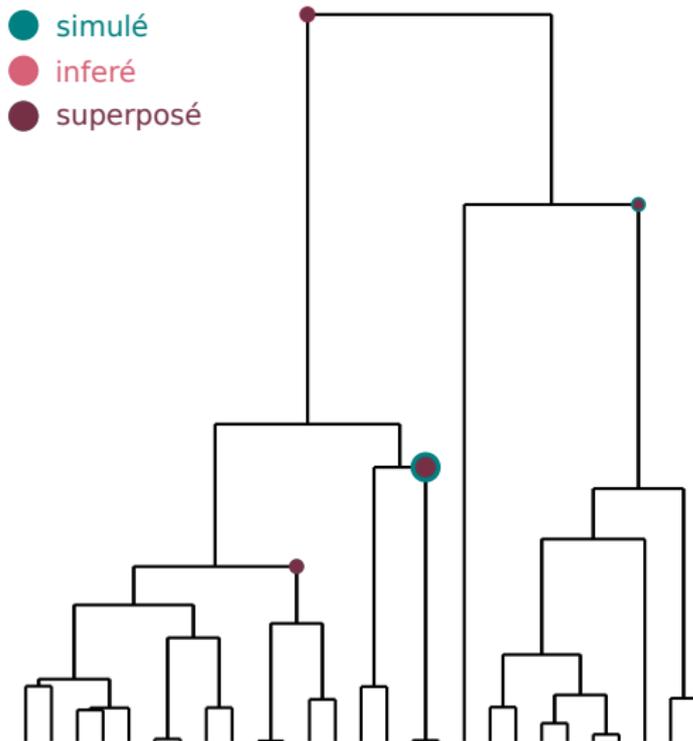
# Méthode d'inférence

Loi des pics sachant les longueurs de branche et l'alignement



# Méthode d'inférence

Loi des pics sachant les longueurs de branche et l'alignement



# Méthode d'inférence

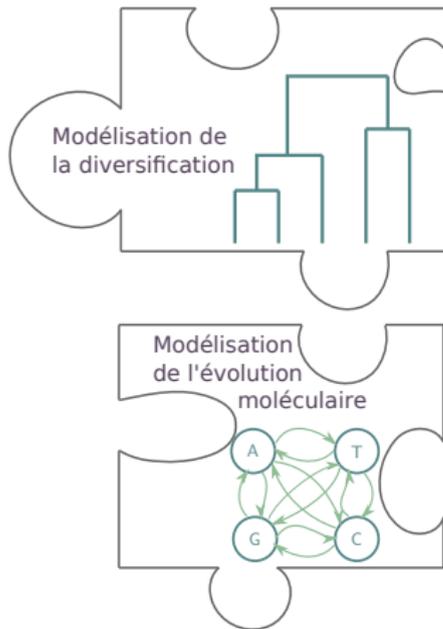
## Résumé

### Modèle et développements :

- ▶ Lien fort entre processus de spéciation et substitutions.
- ▶ Méthode d'inférence préliminaire.

### Perspectives :

- ▶ Tester le support relatif de modèles graduels / ponctués.
- ▶ Scanner des séquences pour chercher des gènes se comportant de façon "ponctuée".



# Méthode d'inférence

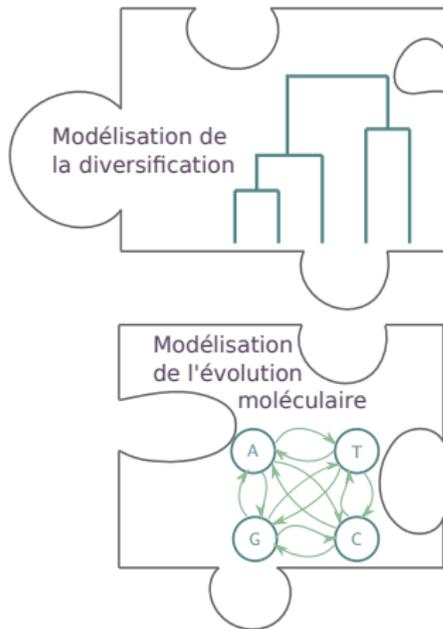
## Résumé

### Modèle et développements :

- ▶ Lien fort entre processus de spéciation et substitutions.
- ▶ Méthode d'inférence préliminaire.

### Perspectives :

- ▶ Tester le support relatif de modèles graduels / ponctués.
- ▶ Scanner des séquences pour chercher des gènes se comportant de façon "ponctuée".



# Méthode d'inférence

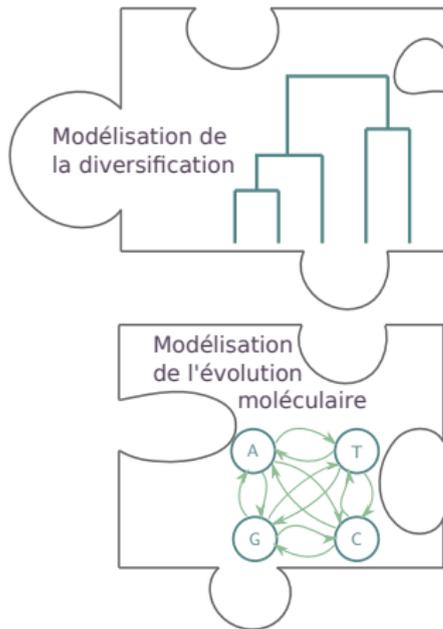
## Résumé

### Modèle et développements :

- ▶ Lien fort entre processus de spéciation et substitutions.
- ▶ Méthode d'inférence préliminaire.

### Perspectives :

- ▶ Tester le support relatif de modèles graduels / ponctués.
- ▶ Scanner des séquences pour chercher des gènes se comportant de façon "ponctuée".



# Méthode d'inférence

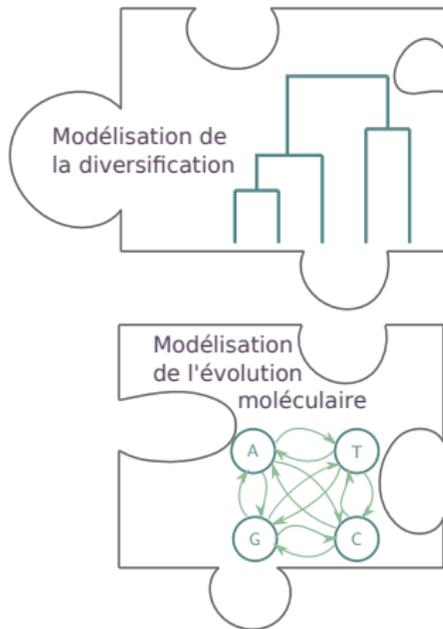
## Résumé

### Modèle et développements :

- ▶ Lien fort entre processus de spéciation et substitutions.
- ▶ Méthode d'inférence préliminaire.

### Perspectives :

- ▶ Tester le support relatif de modèles graduels / ponctués.
- ▶ Scanner des séquences pour chercher des gènes se comportant de façon "ponctuée".



## Introduction

- Macroévolution
- Arbre phylogénétique
- Surcouche de modélisation probabiliste

## Diversification

- État de l'art
- La définition d'espèce en modélisation
- Application à l'étude de la diversification



## Évolution phénotypique

- État de l'art
- Coévolution des caractères de différentes espèces
- Proposition d'un cadre de modélisation
- Exemples d'applications



## Évolution moléculaire

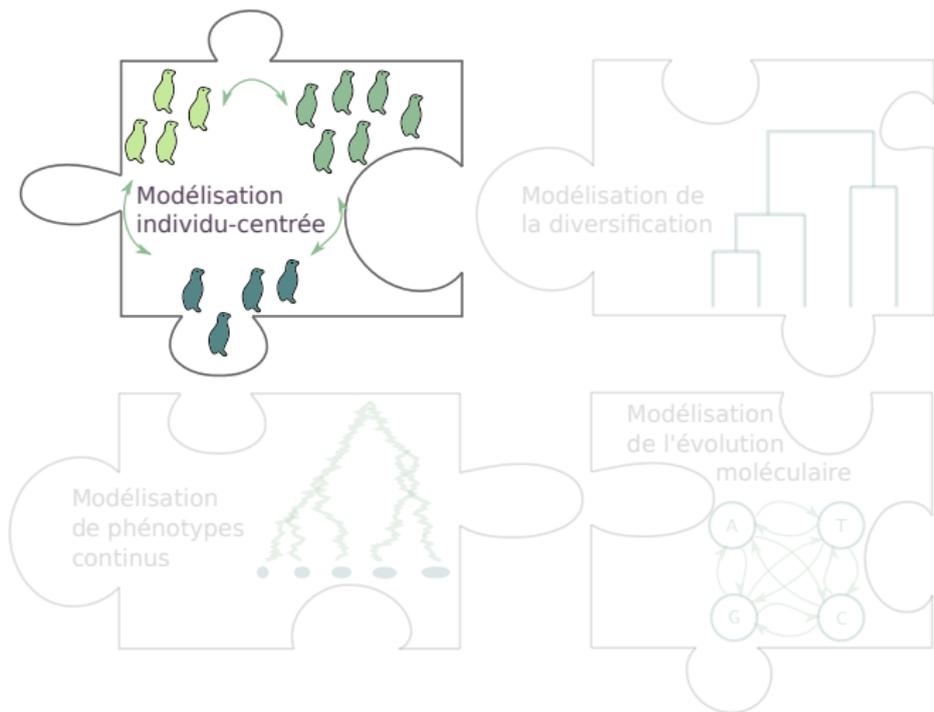
- Horloge moléculaire
- Modélisation proposée
- Méthode d'inférence



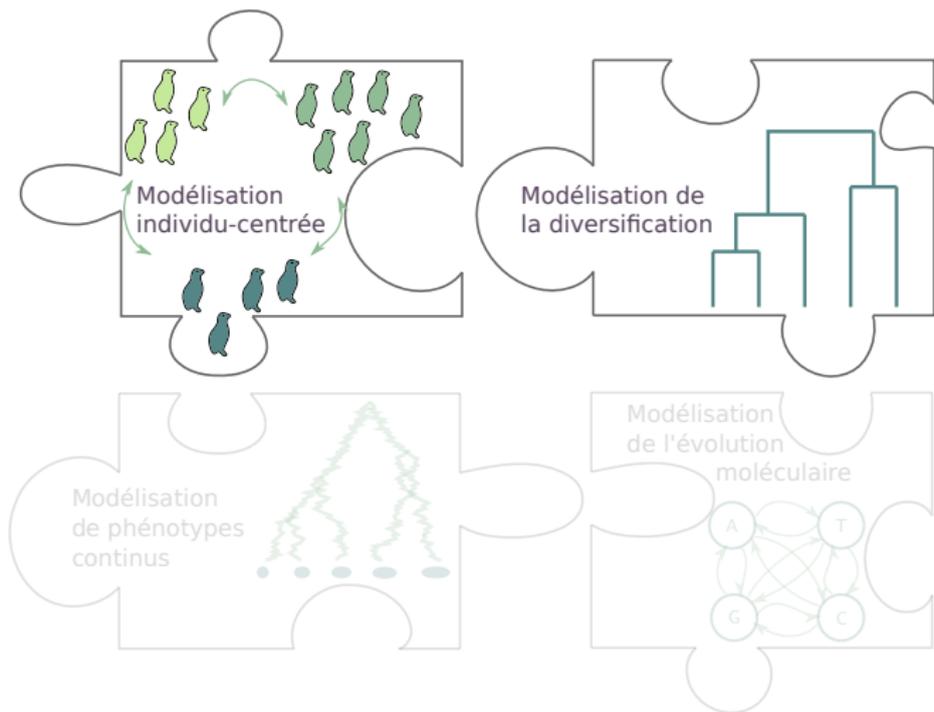
## Perspectives

- Résumé
- Liens entre les chapitres

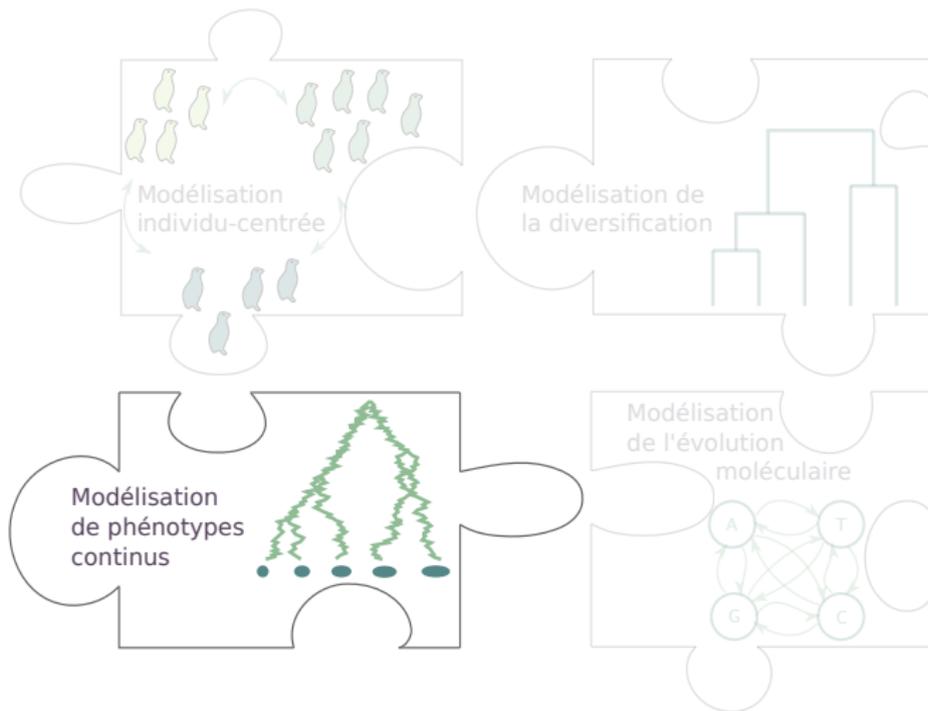
# Résumé



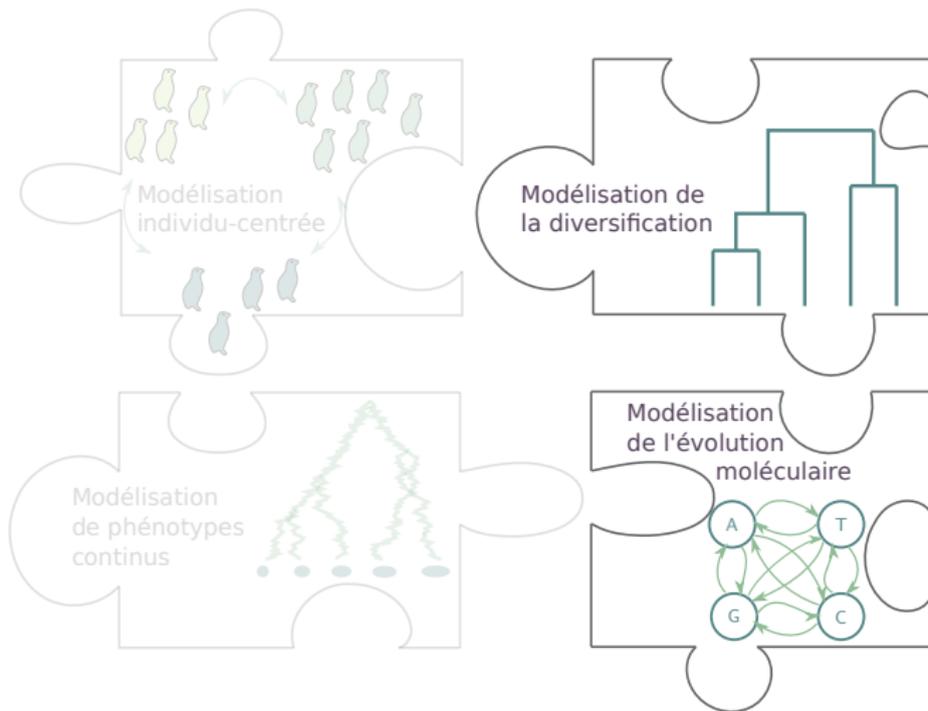
# Résumé



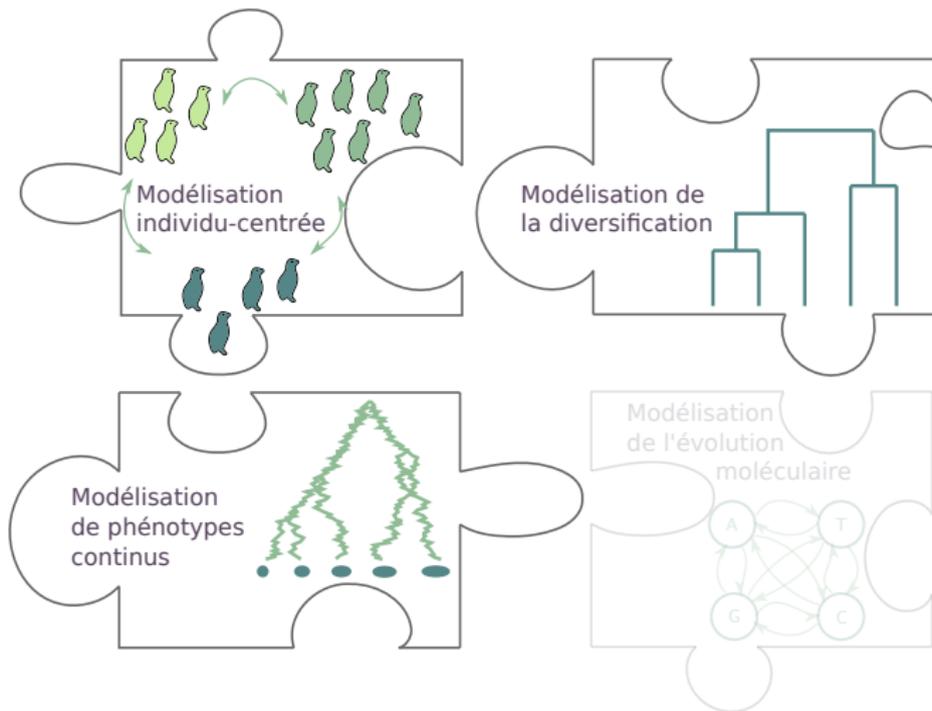
# Résumé



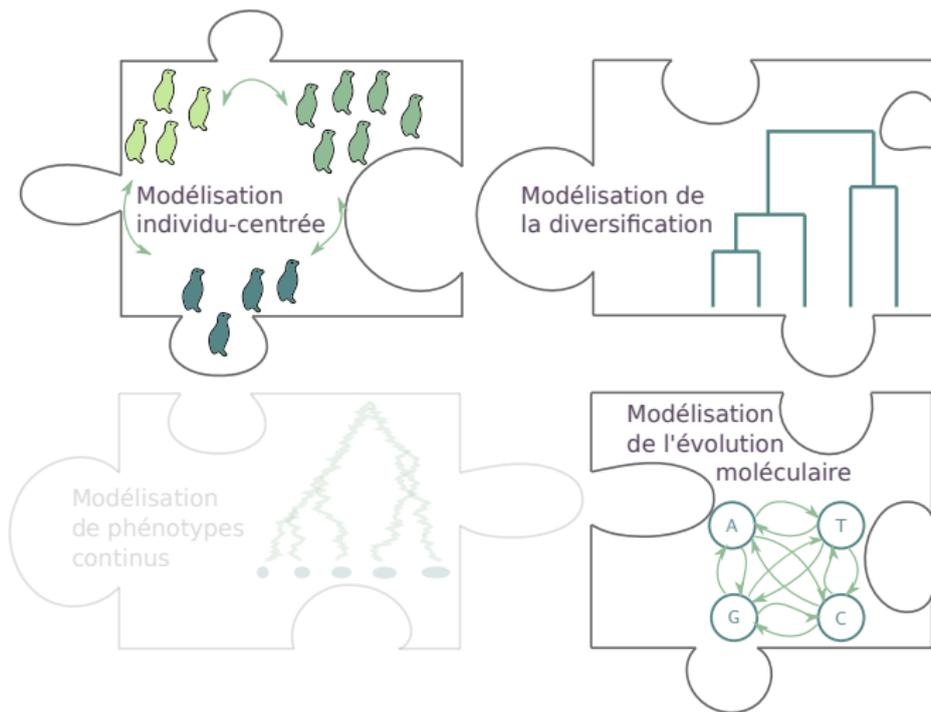
# Résumé



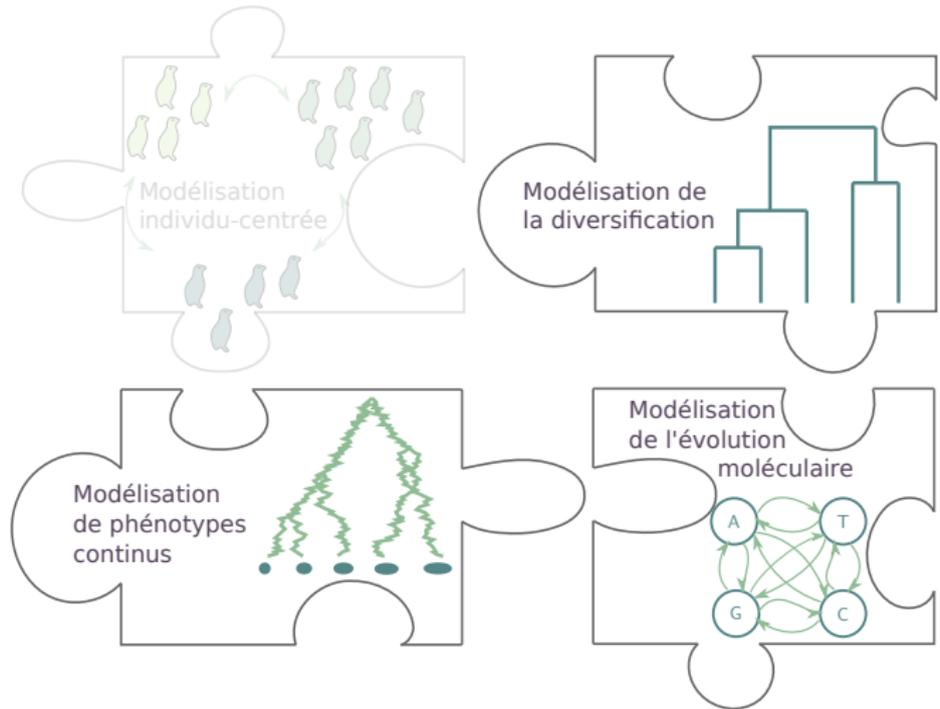
## Liens entre les chapitres



## Liens entre les chapitres



# Liens entre les chapitres



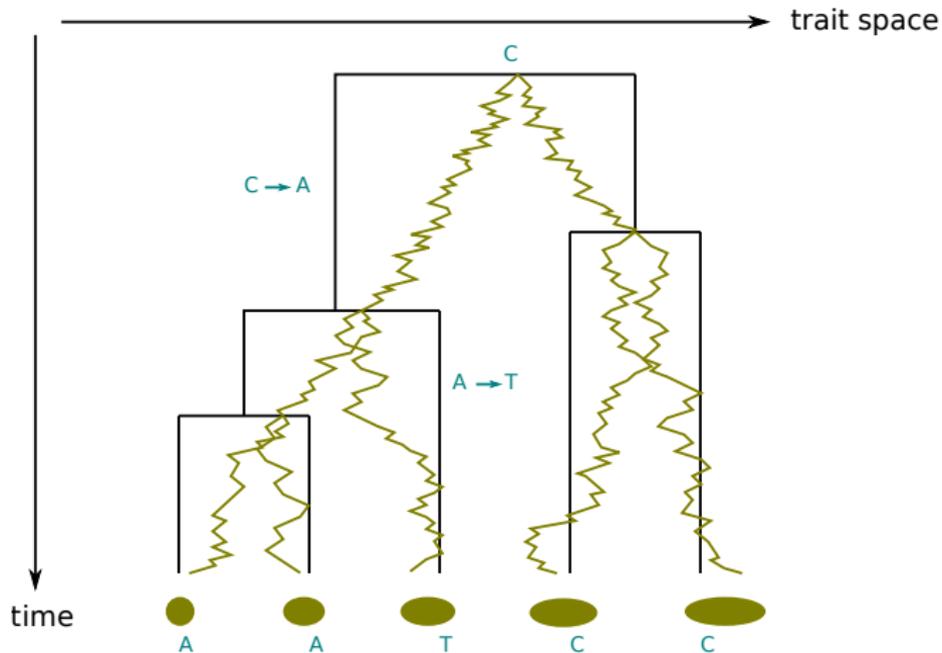
Lartillot et Poujol, 2011, Mol. Biol. Evol.

# Remerciements



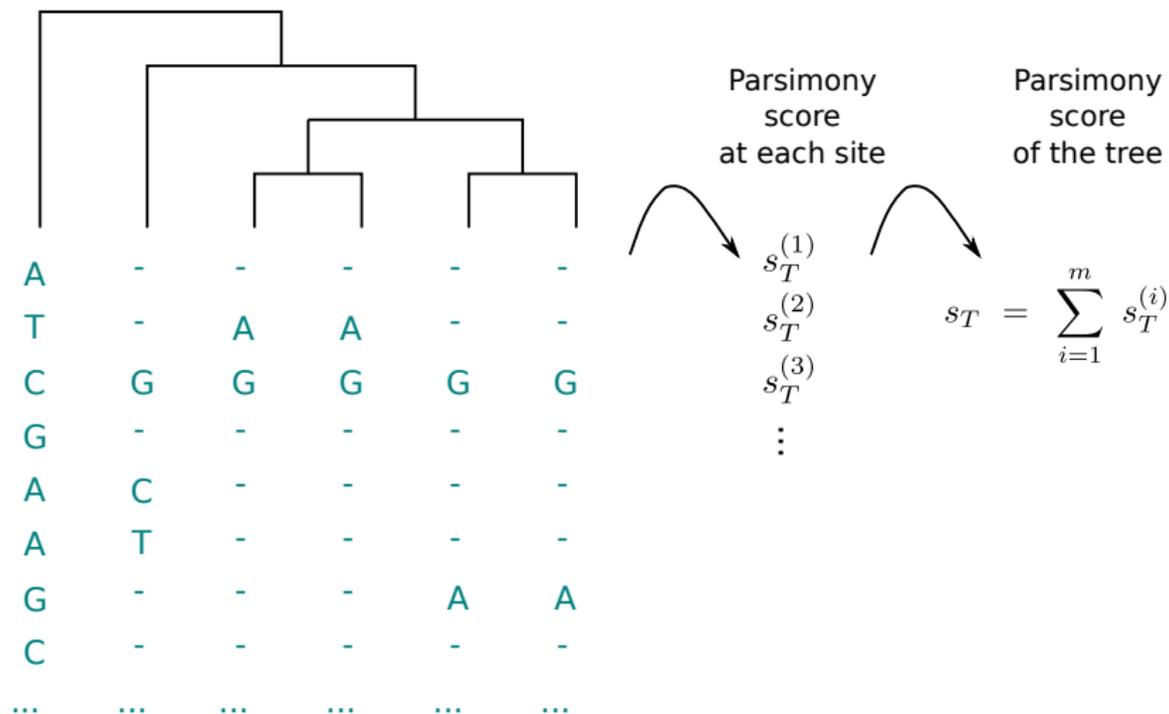
Merci à tous de votre présence !

# Évolution de traits sur un arbre



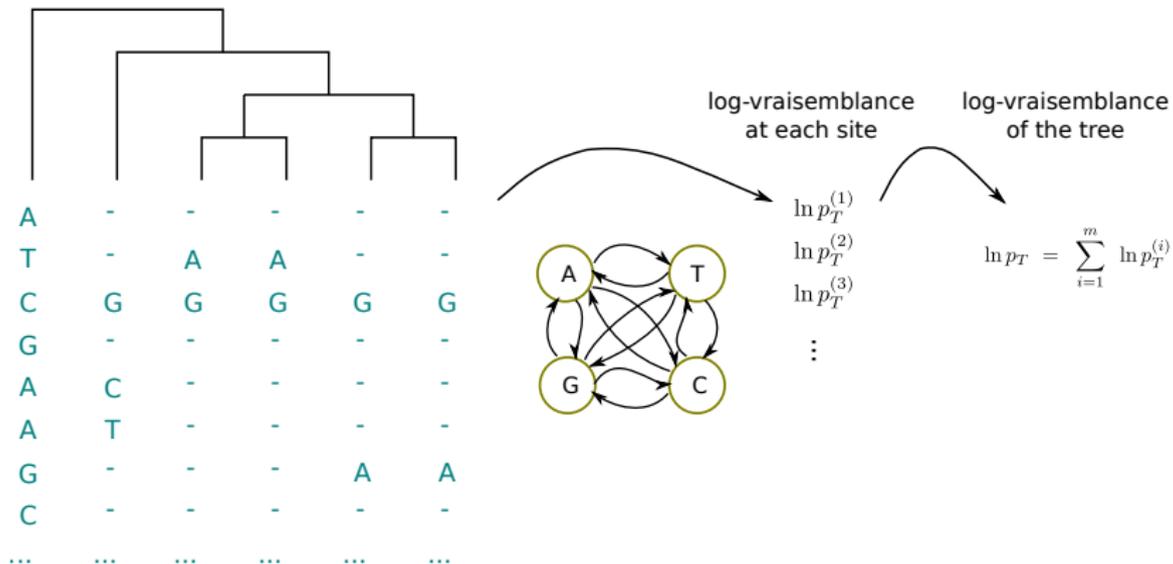
## Reconstruction d'une phylogénie

Par parcimonie



# Reconstruction d'une phylogénie

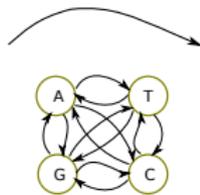
Par maximum de vraisemblance



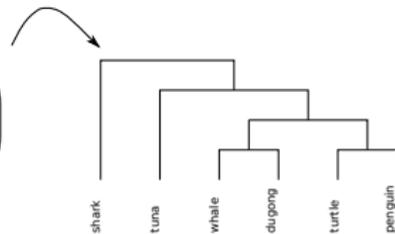
# Reconstruction d'une phylogénie

Par maximum de vraisemblance, méthode "des distances"

shark	A	T	C	G	A	A	G	C	...
tuna	-	-	G	-	C	T	-	-	...
whale	-	A	G	-	-	-	-	-	...
dugong	-	A	G	-	-	-	-	-	...
turtle	-	-	G	-	-	-	A	-	...
penguin	-	-	G	-	-	-	A	-	...

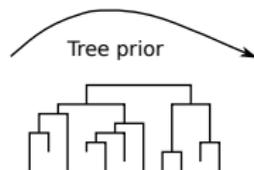
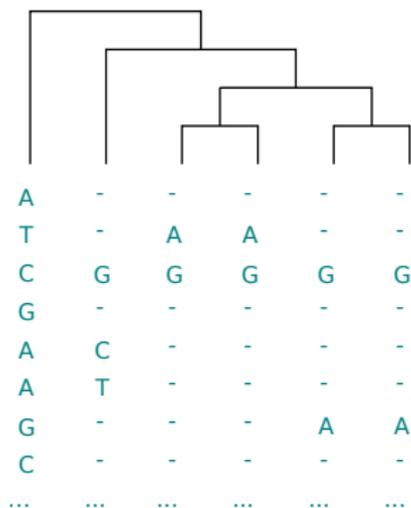
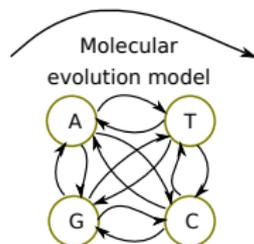


	shark	tuna	whale	...
shark	0	$d_{1,2}$	$d_{1,3}$	...
tuna	$d_{2,1}$	0	$d_{2,3}$	...
whale	$d_{3,1}$	$d_{3,2}$	0	...
...	...	...	...	...



# Reconstruction d'une phylogénie

Dans un cadre bayésien


 $p(T)$ 

 $\ln p_T^{(1)}$ 
 $\ln p_T^{(2)}$ 
 $\ln p_T^{(3)}$ 
 $\vdots$ 

$$\ln p(X = x | T) = \sum_{i=1}^m \ln p_T^{(i)}$$

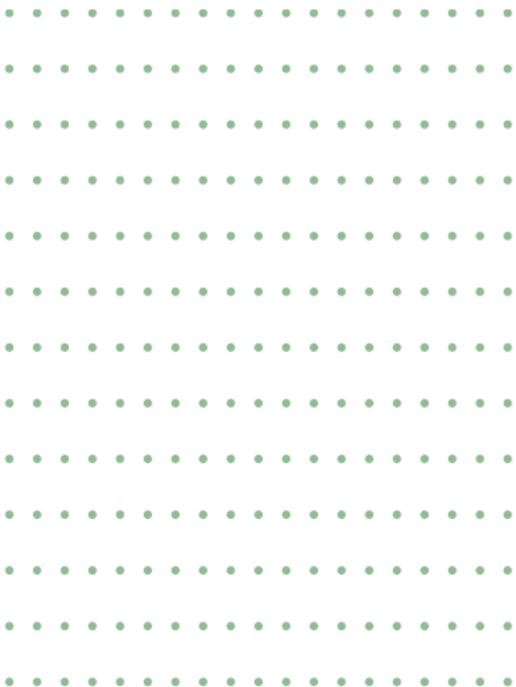
# Bestiaire de modèles de spéciation

## Modèles individus-centrés

Dynamique de population :

- ▶ taille fixée.
- ▶ générations discrètes.
- ▶ choix uniforme du parent dans la génération précédente.

Superposition d'un modèle de spéciation.



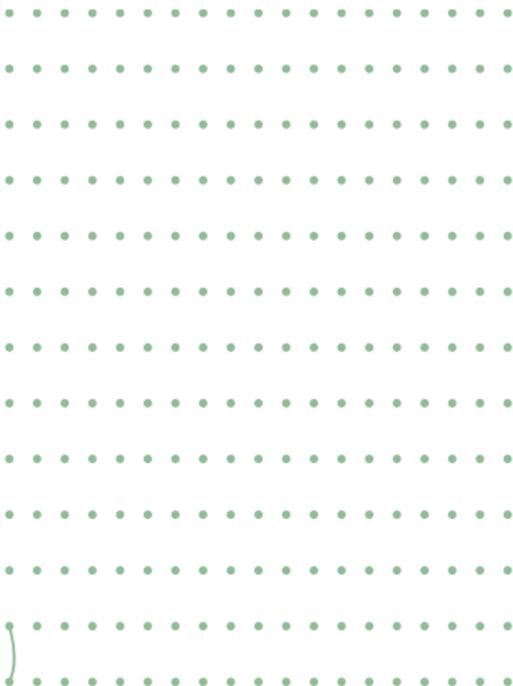
# Bestiaire de modèles de spéciation

## Modèles individus-centrés

Dynamique de population :

- ▶ taille fixée.
- ▶ générations discrètes.
- ▶ choix uniforme du parent dans la génération précédente.

Superposition d'un modèle de spéciation.



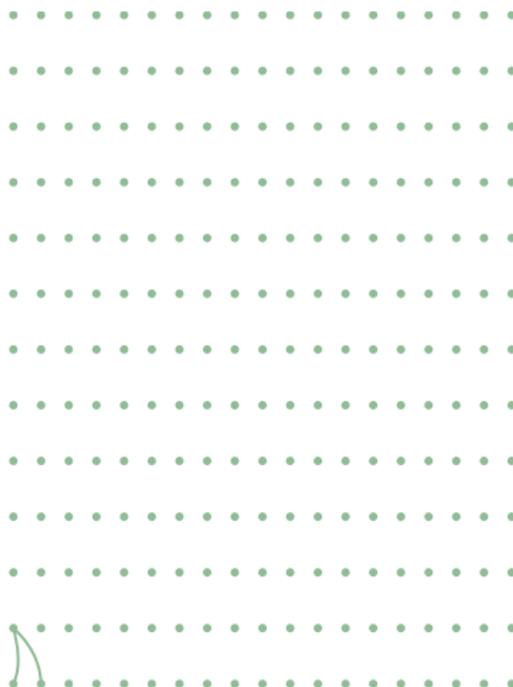
# Bestiaire de modèles de spéciation

## Modèles individus-centrés

Dynamique de population :

- ▶ taille fixée.
- ▶ générations discrètes.
- ▶ choix uniforme du parent dans la génération précédente.

Superposition d'un modèle de spéciation.



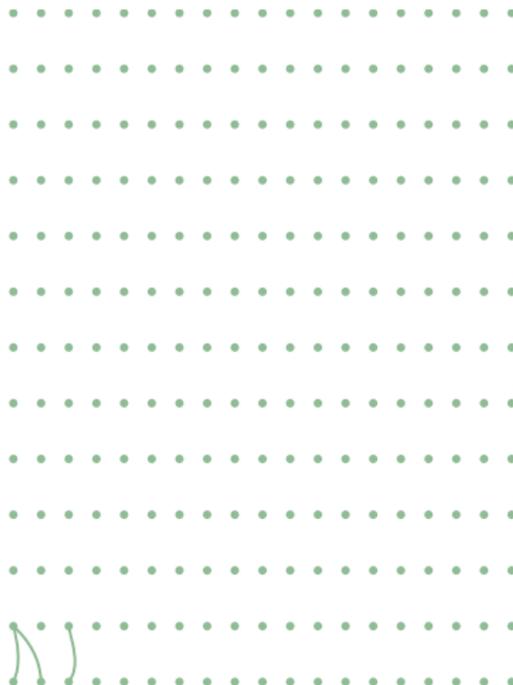
# Bestiaire de modèles de spéciation

## Modèles individus-centrés

Dynamique de population :

- ▶ taille fixée.
- ▶ générations discrètes.
- ▶ choix uniforme du parent dans la génération précédente.

Superposition d'un modèle de spéciation.



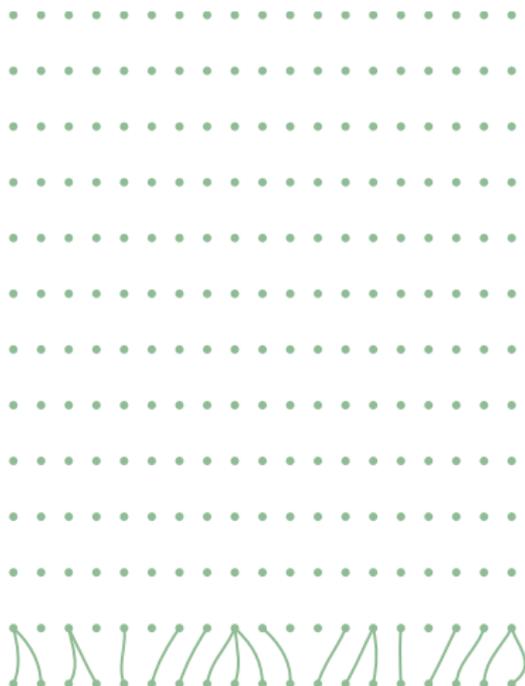
# Bestiaire de modèles de spéciation

## Modèles individus-centrés

Dynamique de population :

- ▶ taille fixée.
- ▶ générations discrètes.
- ▶ choix uniforme du parent dans la génération précédente.

Superposition d'un modèle de spéciation.



# Bestiaire de modèles de spéciation

## Modèles individus-centrés

Dynamique de population :

- ▶ taille fixée.
- ▶ générations discrètes.
- ▶ choix uniforme du parent dans la génération précédente.

Superposition d'un modèle de spéciation.



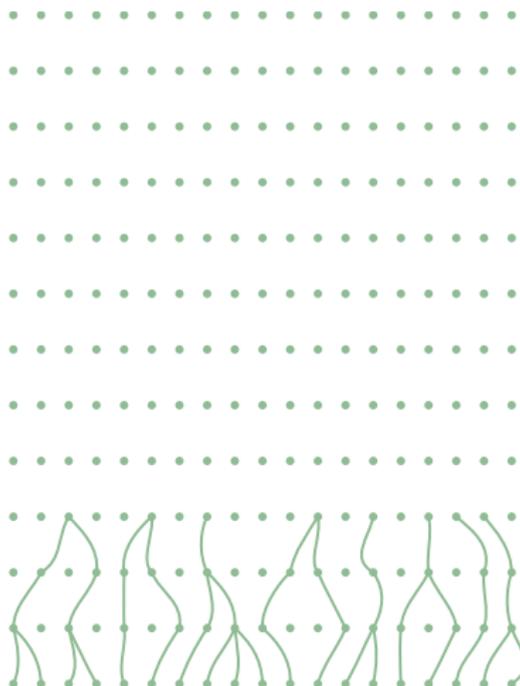
# Bestiaire de modèles de spéciation

## Modèles individus-centrés

Dynamique de population :

- ▶ taille fixée.
- ▶ générations discrètes.
- ▶ choix uniforme du parent dans la génération précédente.

Superposition d'un modèle de spéciation.



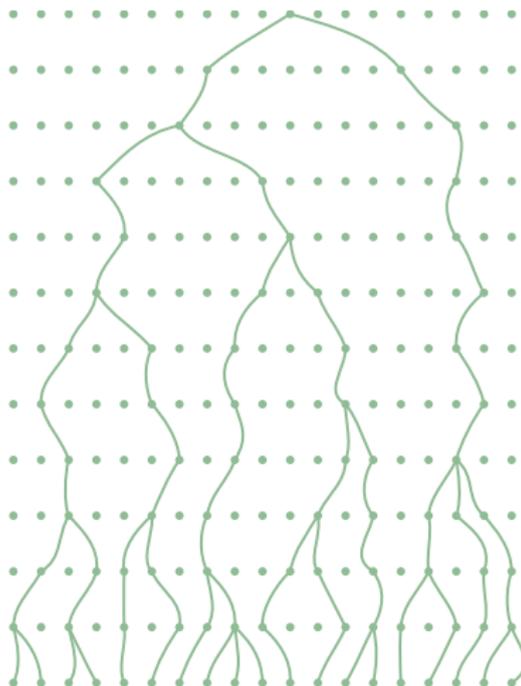
# Bestiaire de modèles de spéciation

## Modèles individus-centrés

Dynamique de population :

- ▶ taille fixée.
- ▶ générations discrètes.
- ▶ choix uniforme du parent dans la génération précédente.

Superposition d'un modèle de spéciation.



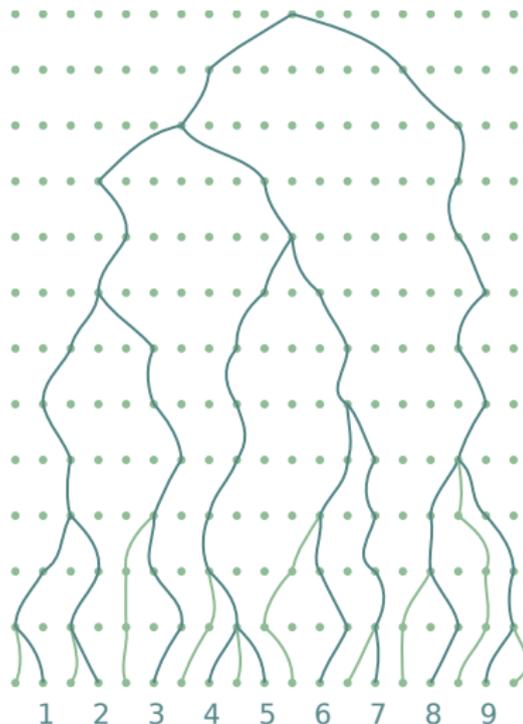
# Bestiaire de modèles de spéciation

## Modèles individus-centrés

Dynamique de population :

- ▶ taille fixée.
- ▶ générations discrètes.
- ▶ choix uniforme du parent dans la génération précédente.

Superposition d'un modèle de spéciation.



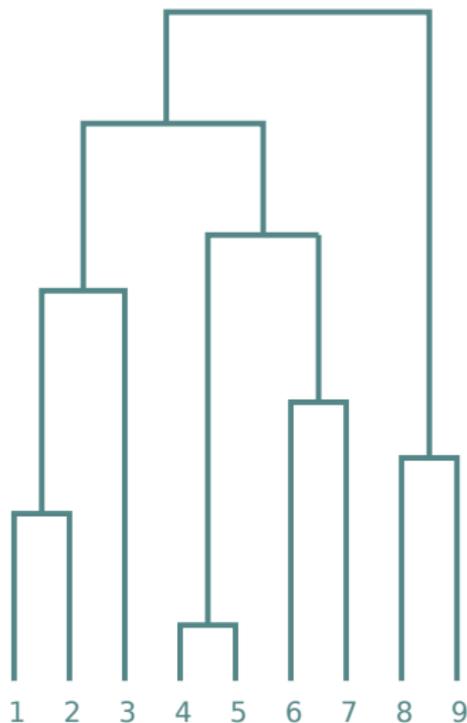
# Bestiaire de modèles de spéciation

## Modèles individus-centrés

Dynamique de population :

- ▶ taille fixée.
- ▶ générations discrètes.
- ▶ choix uniforme du parent dans la génération précédente.

Superposition d'un modèle de spéciation.



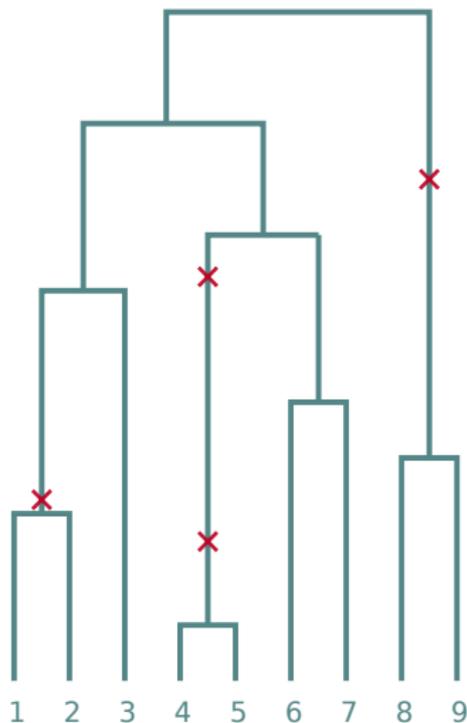
# Bestiaire de modèles de spéciation

## Modèles individus-centrés

Dynamique de population :

- ▶ taille fixée.
- ▶ générations discrètes.
- ▶ choix uniforme du parent dans la génération précédente.

Superposition d'un modèle de spéciation.



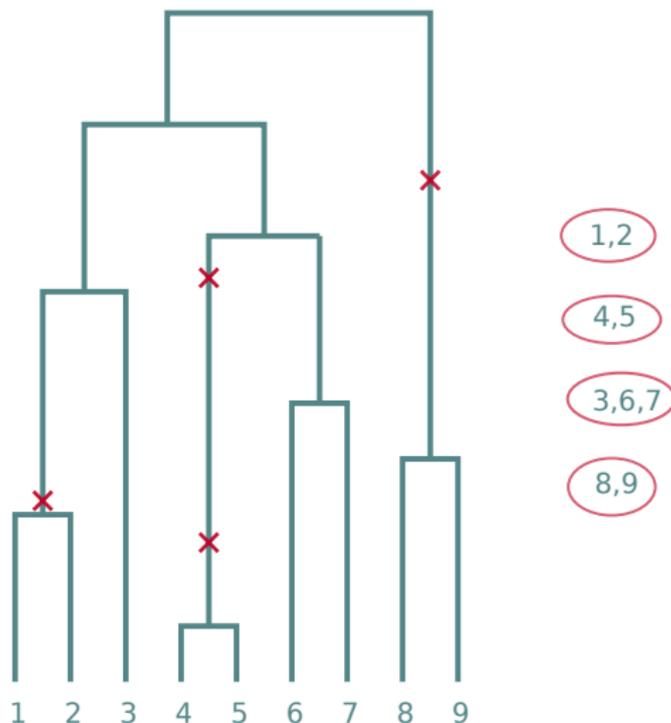
# Bestiaire de modèles de spéciation

## Modèles individus-centrés

Dynamique de population :

- ▶ taille fixée.
- ▶ générations discrètes.
- ▶ choix uniforme du parent dans la génération précédente.

Superposition d'un modèle de spéciation.



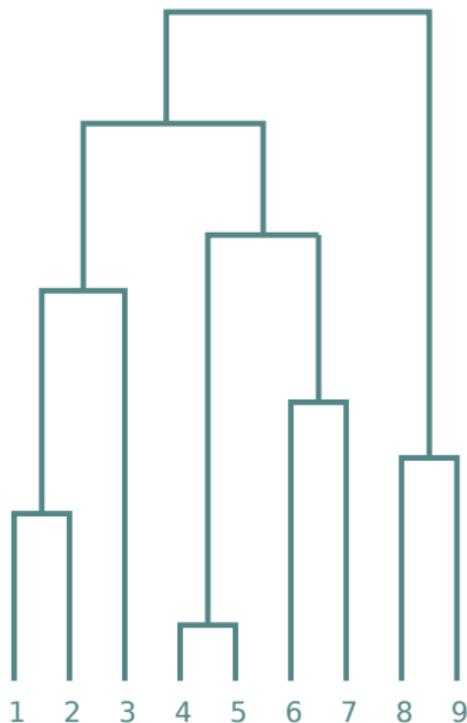
# Bestiaire de modèles de spéciation

## Modèles individus-centrés

Dynamique de population :

- ▶ taille fixée.
- ▶ générations discrètes.
- ▶ choix uniforme du parent dans la génération précédente.

Superposition d'un modèle de spéciation.



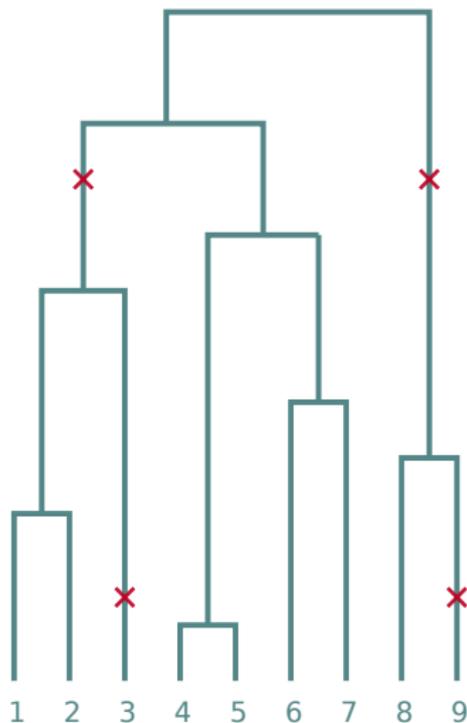
# Bestiaire de modèles de spéciation

## Modèles individus-centrés

Dynamique de population :

- ▶ taille fixée.
- ▶ générations discrètes.
- ▶ choix uniforme du parent dans la génération précédente.

Superposition d'un modèle de spéciation.



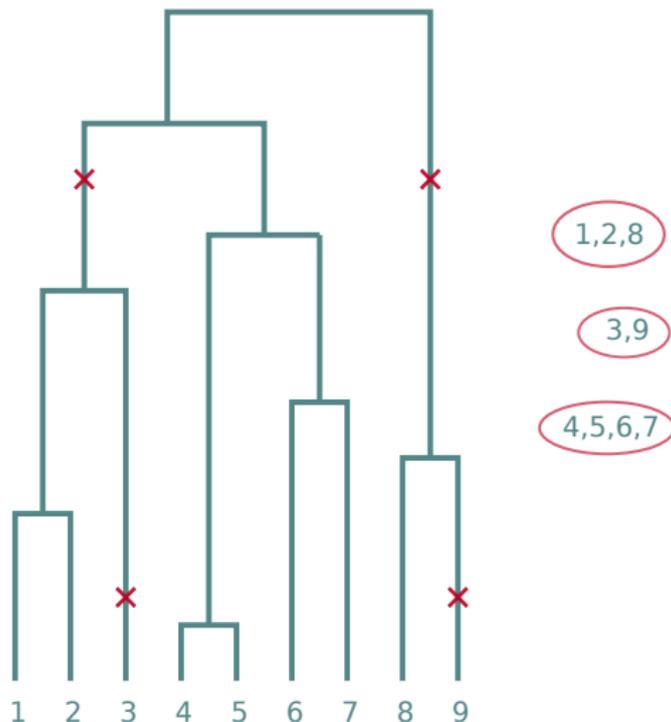
# Bestiaire de modèles de spéciation

## Modèles individus-centrés

Dynamique de population :

- ▶ taille fixée.
- ▶ générations discrètes.
- ▶ choix uniforme du parent dans la génération précédente.

Superposition d'un modèle de spéciation.



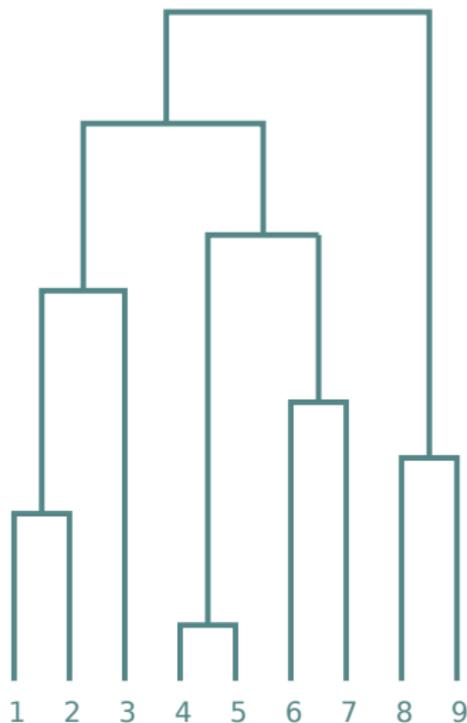
# Bestiaire de modèles de spéciation

## Modèles individus-centrés

Dynamique de population :

- ▶ taille fixée.
- ▶ générations discrètes.
- ▶ choix uniforme du parent dans la génération précédente.

Superposition d'un modèle de spéciation.



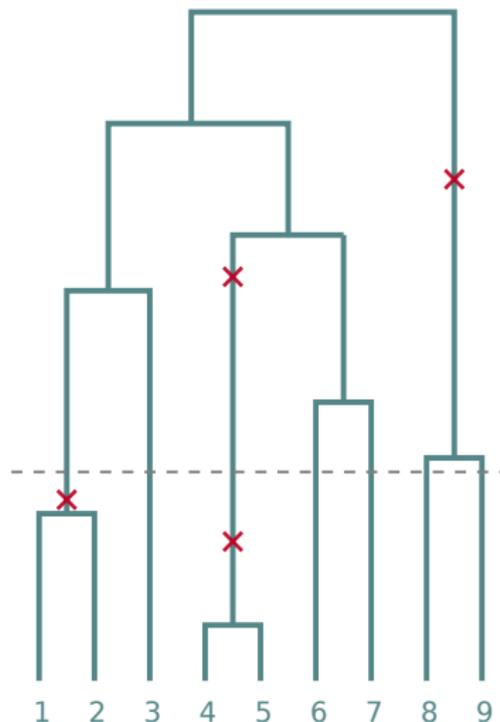
# Bestiaire de modèles de spéciation

## Modèles individus-centrés

Dynamique de population :

- ▶ taille fixée.
- ▶ générations discrètes.
- ▶ choix uniforme du parent dans la génération précédente.

Superposition d'un modèle de spéciation.



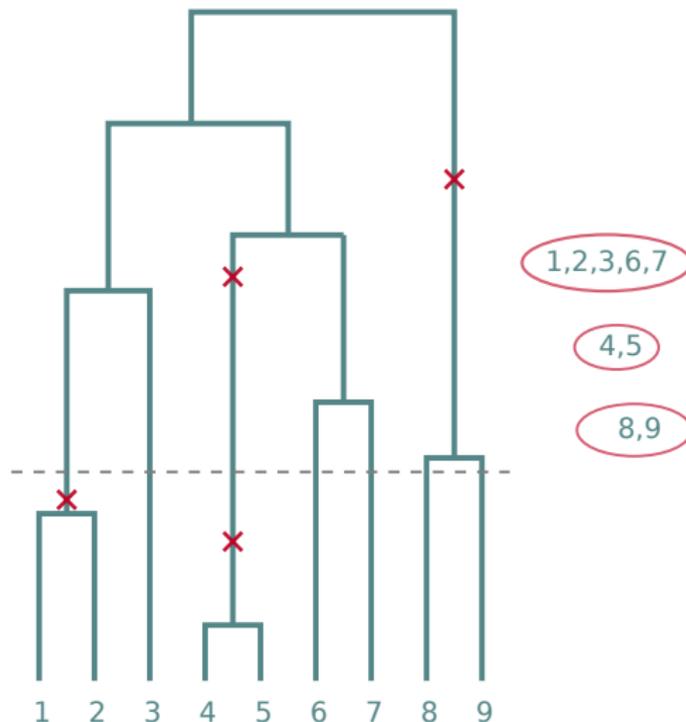
# Bestiaire de modèles de spéciation

## Modèles individus-centrés

Dynamique de population :

- ▶ taille fixée.
- ▶ générations discrètes.
- ▶ choix uniforme du parent dans la génération précédente.

Superposition d'un modèle de spéciation.



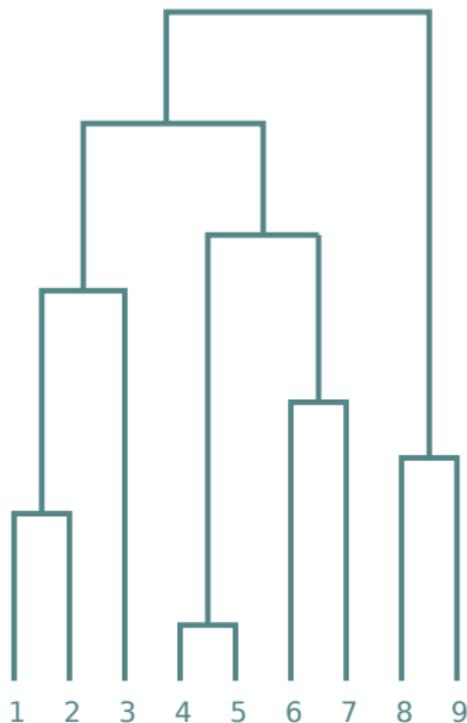
# Bestiaire de modèles de spéciation

## Modèles individus-centrés

Dynamique de population :

- ▶ taille fixée.
- ▶ générations discrètes.
- ▶ choix uniforme du parent dans la génération précédente.

Superposition d'un modèle de spéciation.



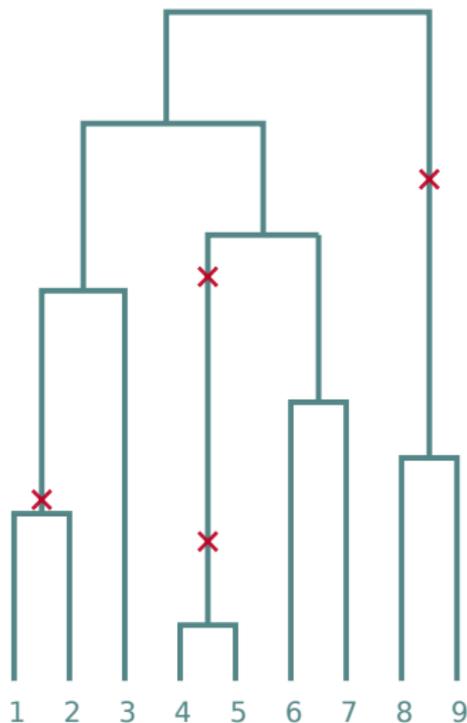
# Bestiaire de modèles de spéciation

## Modèles individus-centrés

Dynamique de population :

- ▶ taille fixée.
- ▶ générations discrètes.
- ▶ choix uniforme du parent dans la génération précédente.

Superposition d'un modèle de spéciation.



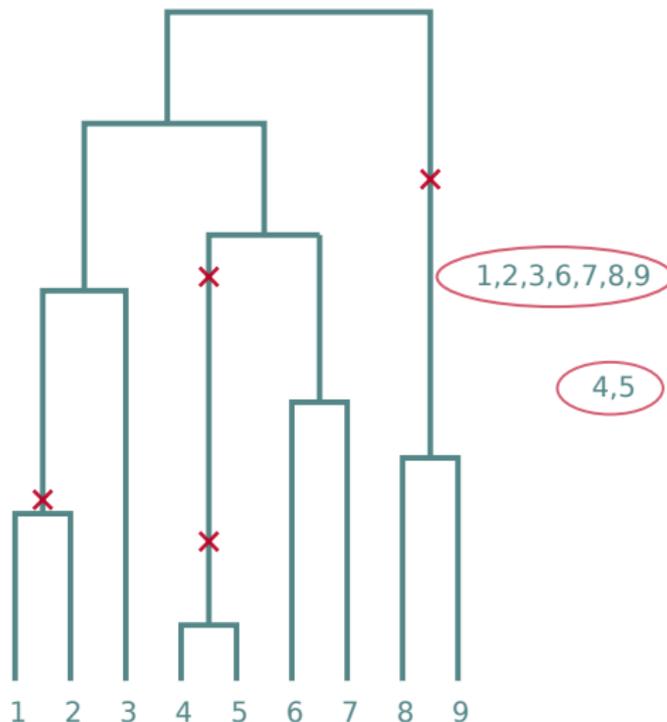
# Bestiaire de modèles de spéciation

## Modèles individus-centrés

Dynamique de population :

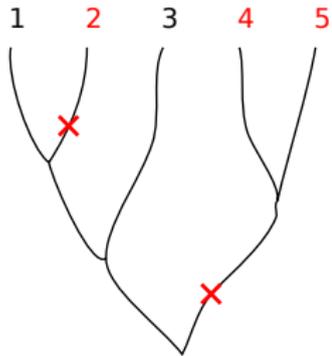
- ▶ taille fixée.
- ▶ générations discrètes.
- ▶ choix uniforme du parent dans la génération précédente.

Superposition d'un modèle de spéciation.

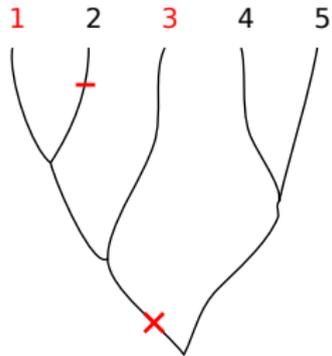


# D'où vient la paraphylie ?

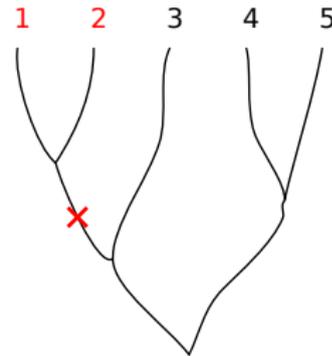
a) Convergence



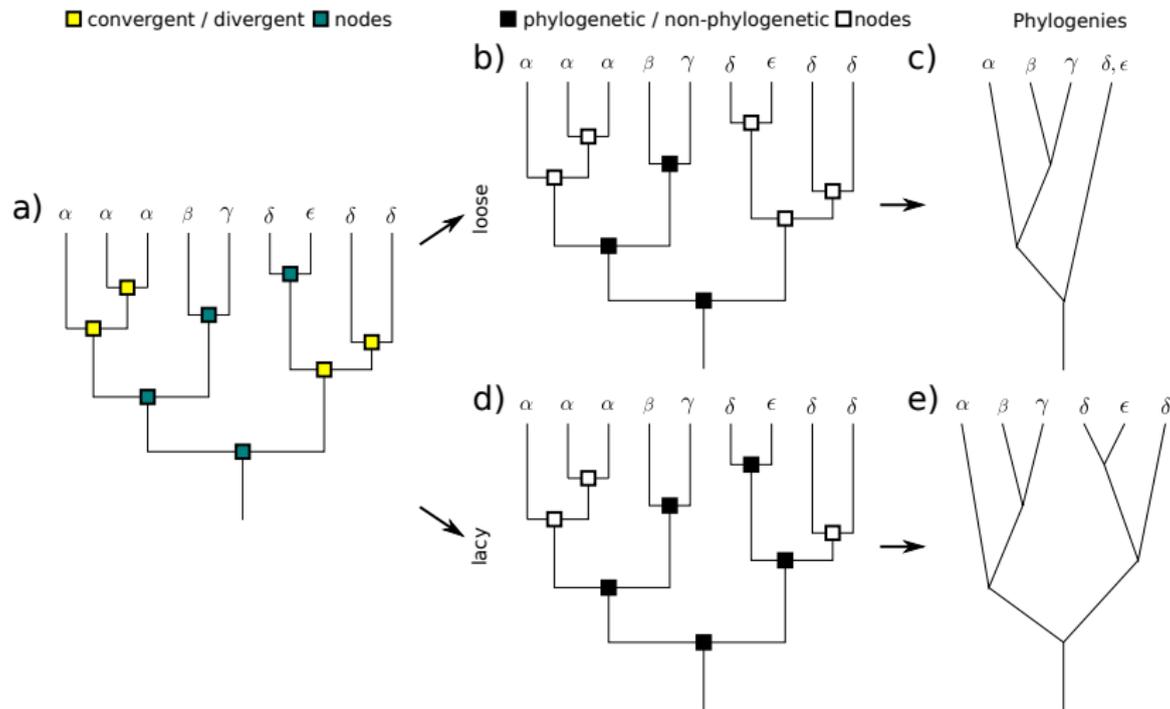
b) Reversal



c) Plesiomorphy



## De la généalogie à la phylogénie



# Processus tri-types

## Loi des phylogénies reconstruites

La loi des phylogénies sous notre modèle est la même que celle :

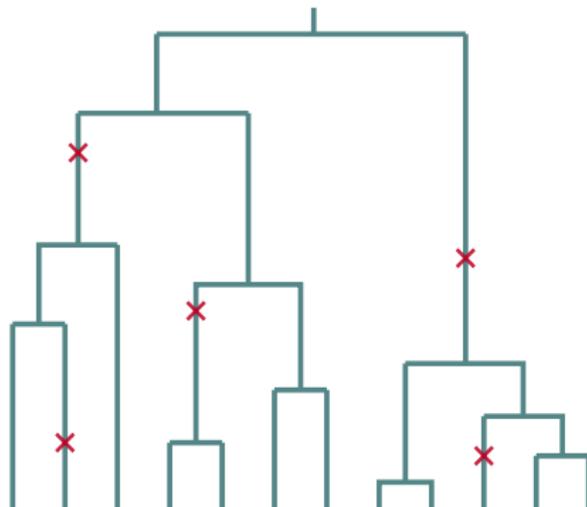
- ▶ d'un processus de branchement à trois types,

type 0 la famille clonale survit jusqu'au présent.

type 1 la famille clonale ne survit pas jusqu'au présent.

type gelé deux familles clonales survivent jusqu'au présent.

- ▶ inhomogène en temps, avec des taux de transitions connus.



# Processus tri-types

## Loi des phylogénies reconstruites

La loi des phylogénies sous notre modèle est la même que celle :

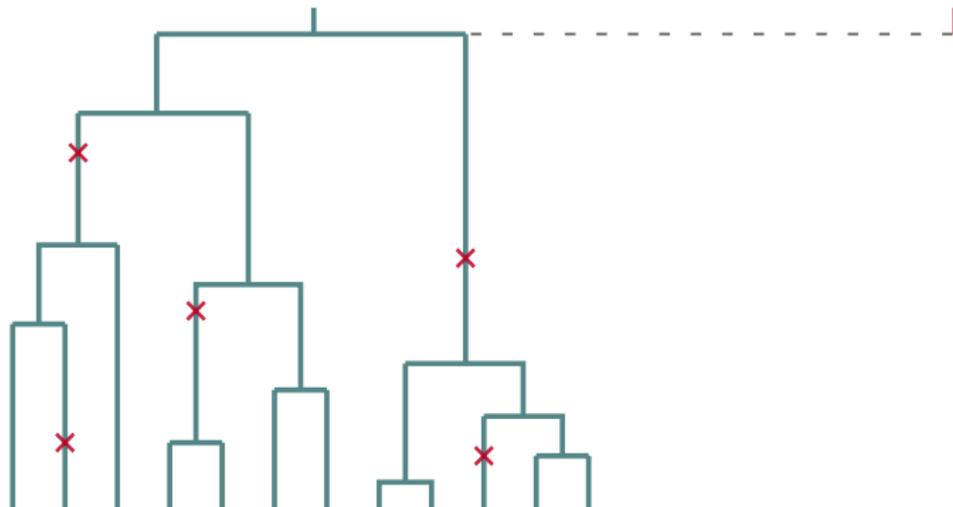
- ▶ d'un processus de branchement à trois types,

type 0 la famille clonale survit jusqu'au présent.

type 1 la famille clonale ne survit pas jusqu'au présent.

type gelé deux familles clonales survivent jusqu'au présent.

- ▶ inhomogène en temps, avec des taux de transitions connus.

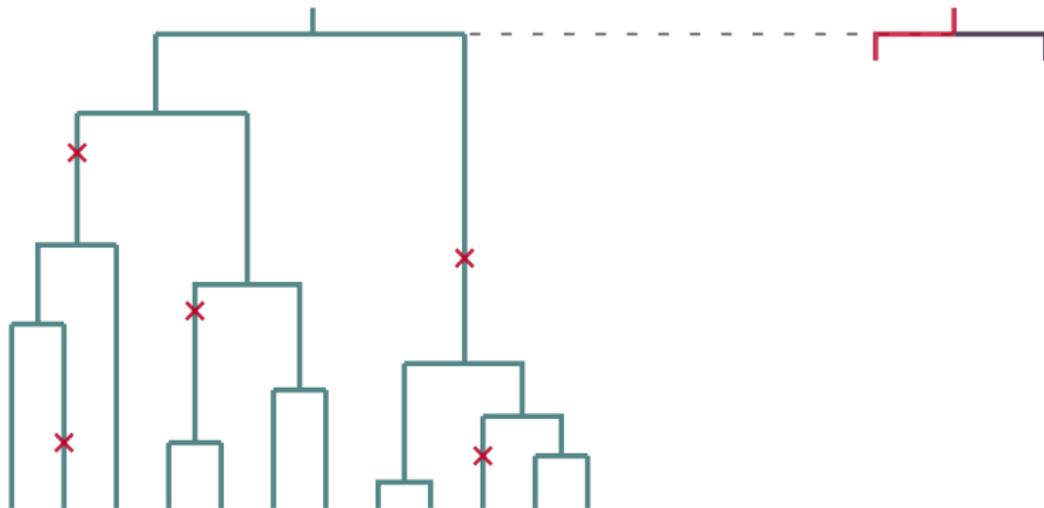


# Processus tri-types

## Loi des phylogénies reconstruites

La loi des phylogénies sous notre modèle est la même que celle :

- ▶ d'un processus de branchement à trois types,
  - type 0 la famille clonale survit jusqu'au présent.
  - type 1 la famille clonale ne survit pas jusqu'au présent.
  - type gelé deux familles clonales survivent jusqu'au présent.
- ▶ inhomogène en temps, avec des taux de transitions connus.

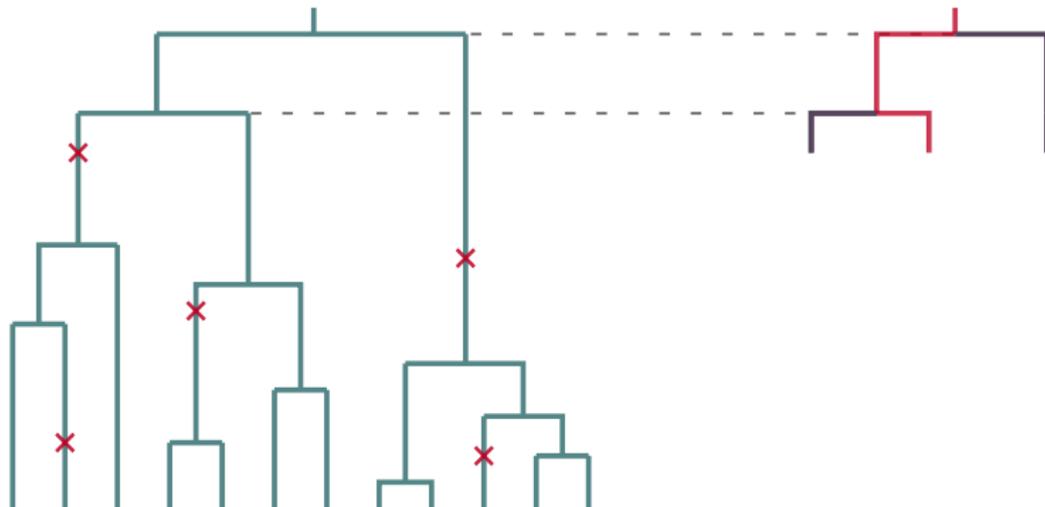


# Processus tri-types

## Loi des phylogénies reconstruites

La loi des phylogénies sous notre modèle est la même que celle :

- ▶ d'un processus de branchement à trois types,
  - type 0 la famille clonale survit jusqu'au présent.
  - type 1 la famille clonale ne survit pas jusqu'au présent.
  - type gelé deux familles clonales survivent jusqu'au présent.
- ▶ inhomogène en temps, avec des taux de transitions connus.



# Processus tri-types

## Loi des phylogénies reconstruites

La loi des phylogénies sous notre modèle est la même que celle :

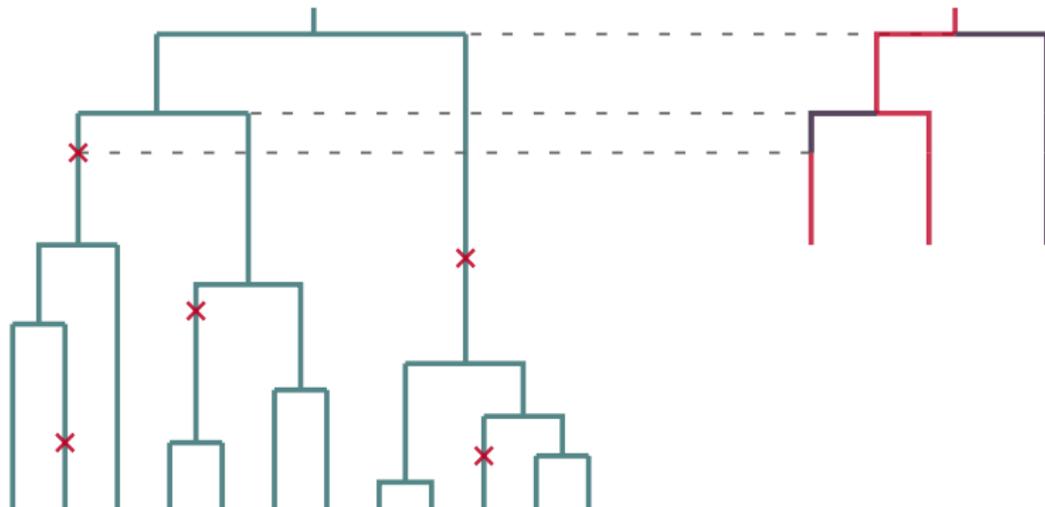
- ▶ d'un processus de branchement à trois types,

type 0 la famille clonale survit jusqu'au présent.

type 1 la famille clonale ne survit pas jusqu'au présent.

type gelé deux familles clonales survivent jusqu'au présent.

- ▶ inhomogène en temps, avec des taux de transitions connus.

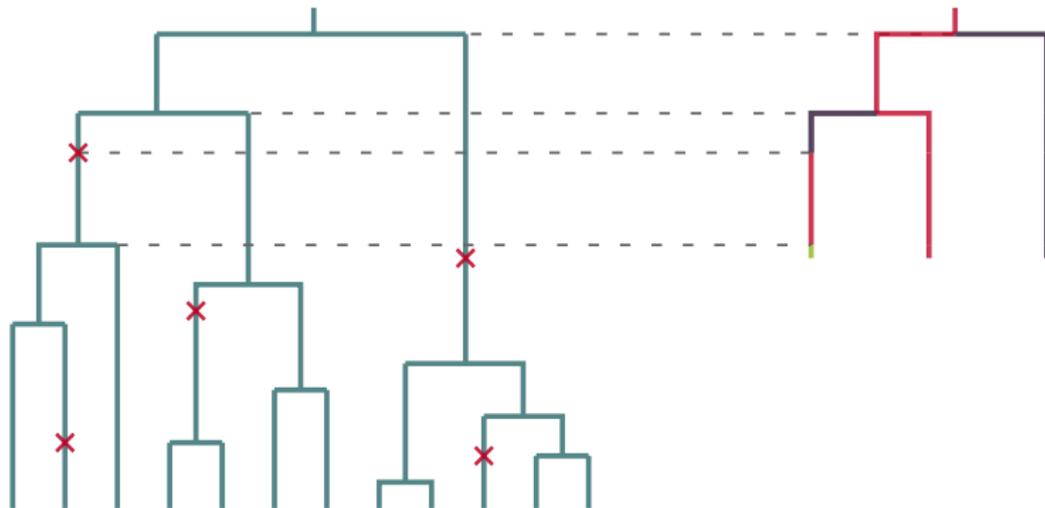


# Processus tri-types

## Loi des phylogénies reconstruites

La loi des phylogénies sous notre modèle est la même que celle :

- ▶ d'un processus de branchement à trois types,
  - type 0 la famille clonale survit jusqu'au présent.
  - type 1 la famille clonale ne survit pas jusqu'au présent.
  - type gelé deux familles clonales survivent jusqu'au présent.
- ▶ inhomogène en temps, avec des taux de transitions connus.

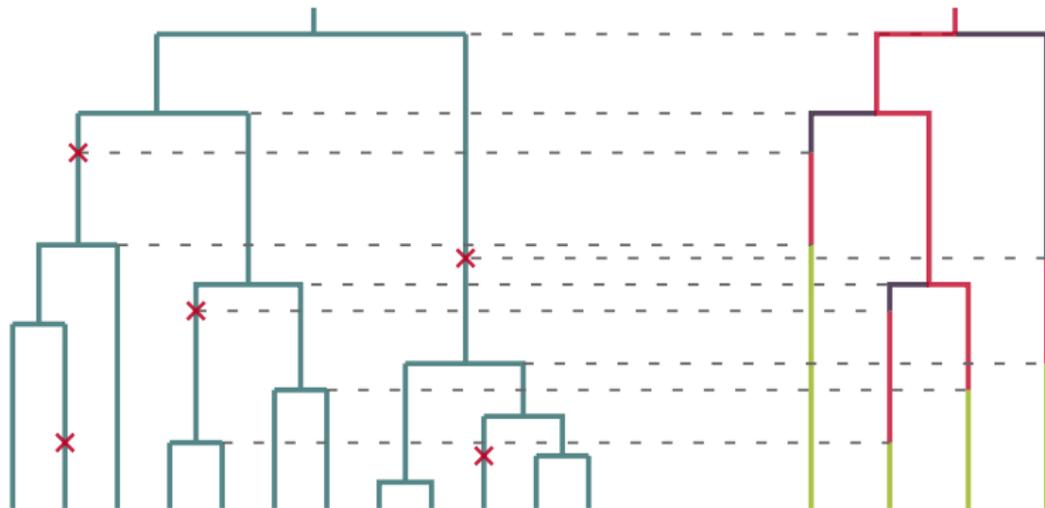


# Processus tri-types

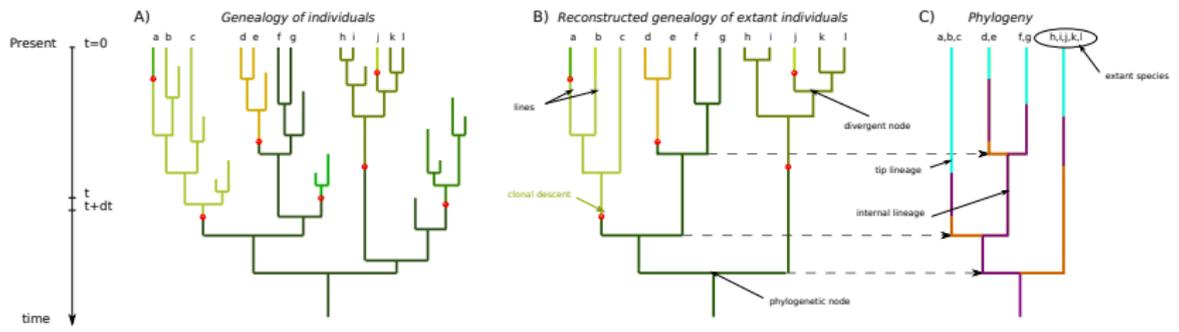
## Loi des phylogénies reconstruites

La loi des phylogénies sous notre modèle est la même que celle :

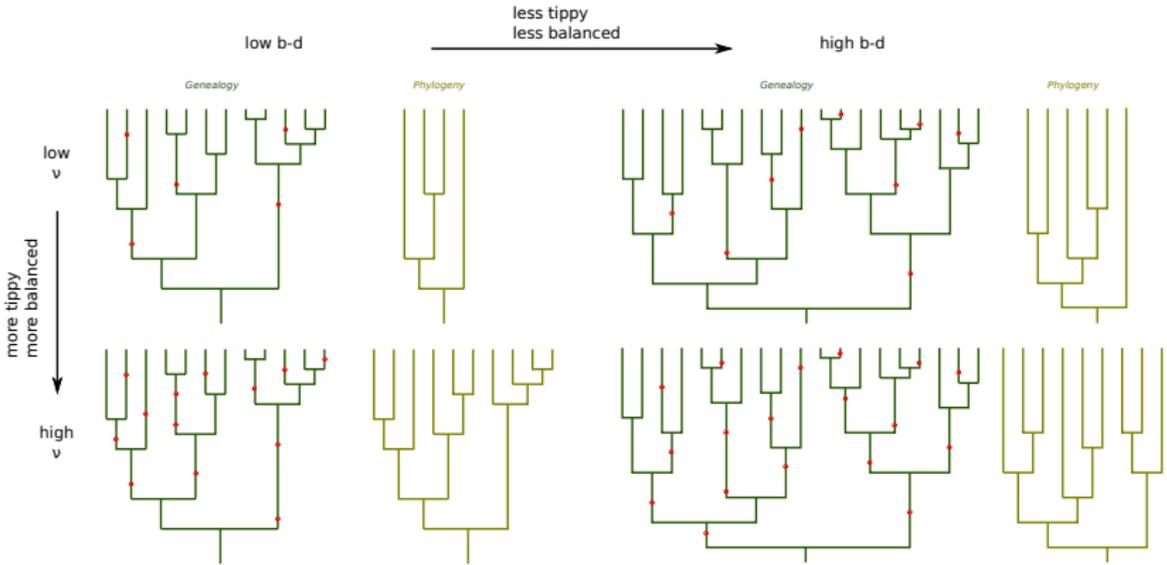
- ▶ d'un processus de branchement à trois types,
  - type 0 la famille clonale survit jusqu'au présent.
  - type 1 la famille clonale ne survit pas jusqu'au présent.
  - type gelé deux familles clonales survivent jusqu'au présent.
- ▶ inhomogène en temps, avec des taux de transitions connus.



# Processus tri-types

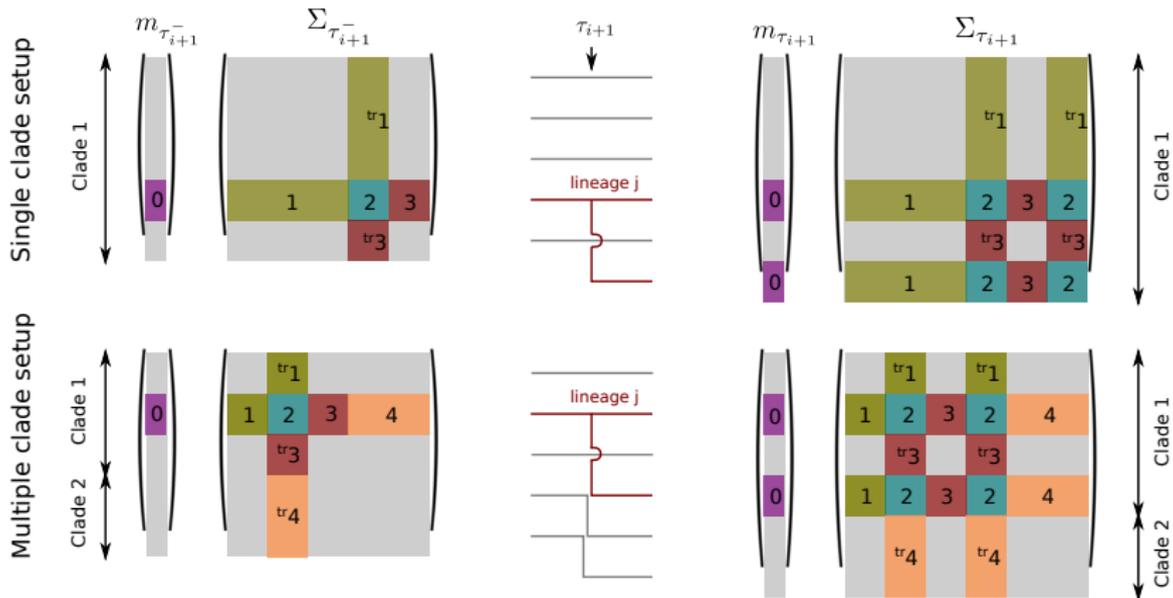


# Influence des paramètres sur la forme

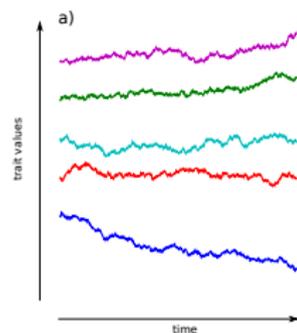


## Matrice de variance

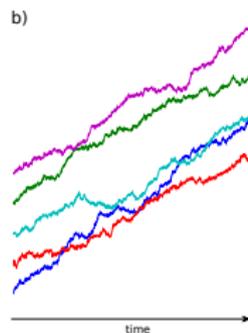
Pre-branching distribution   Ordering of lineages   Post-branching distribution



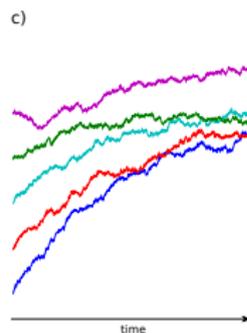
## Cartoon compréhension



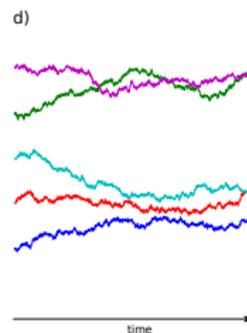
$$a = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \quad A = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix}$$



$$a = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \quad A = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix}$$

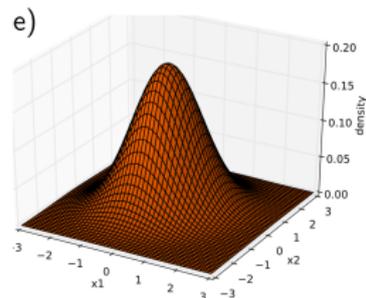
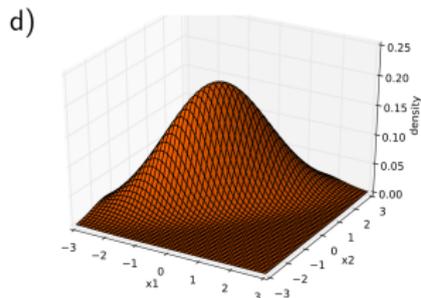
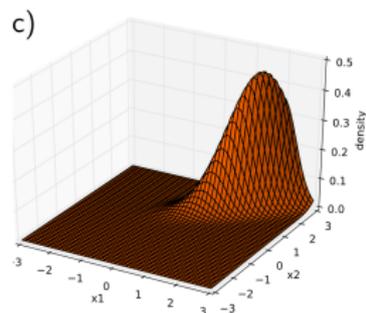
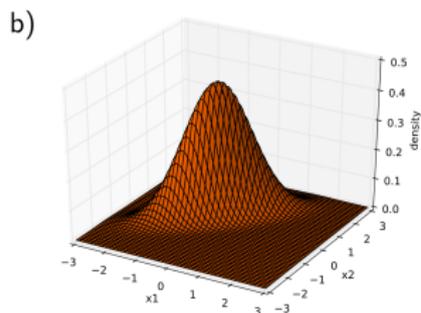
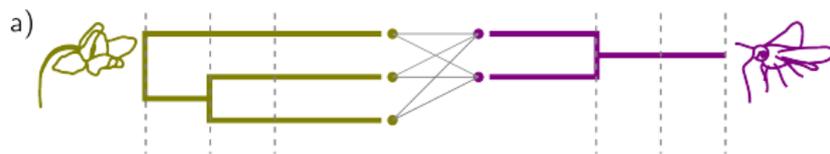


$$a = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \quad A = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 1 \end{pmatrix}$$



$$a = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \quad A = \begin{pmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ -\frac{1}{2} & 1 & -\frac{1}{2} & 0 & 0 \\ -\frac{1}{2} & -\frac{1}{2} & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}$$

# Generalist Matching Mutualism



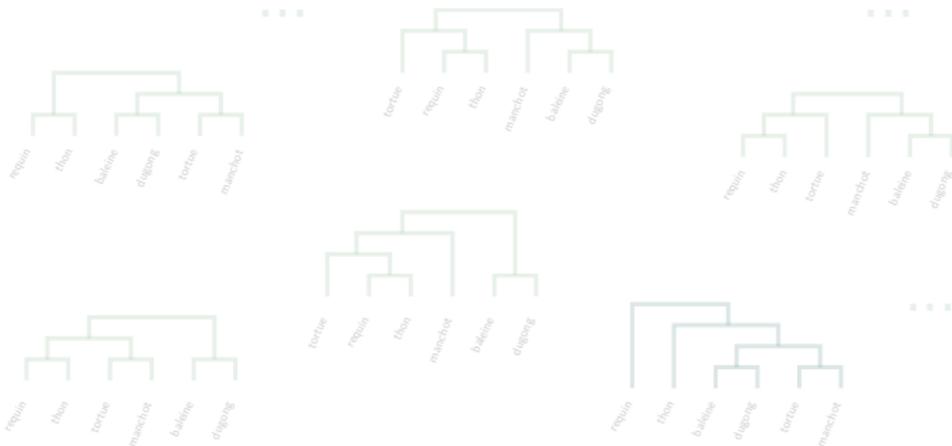
## Exploration de l'espace des arbres

## Principe de reconstruction

DONNEES

requin	A	T	C	G	A	A	G	C	...
thon	-	-	G	-	C	T	-	-	...
baleine	-	A	G	-	-	-	-	-	...
dugong	-	A	G	-	-	-	-	-	...
tortue	-	-	G	-	-	-	A	-	...
manchot	-	-	G	-	-	-	A	-	...

ESTIMATION



## Exploration de l'espace des arbres

## Principe de reconstruction

DONNEES

requin	A	T	C	G	A	A	G	C	...
thon	-	-	G	-	C	T	-	-	...
baleine	-	A	G	-	-	-	-	-	...
dugong	-	A	G	-	-	-	-	-	...
tortue	-	-	G	-	-	-	A	-	...
manchot	-	-	G	-	-	-	A	-	...

ESTIMATION



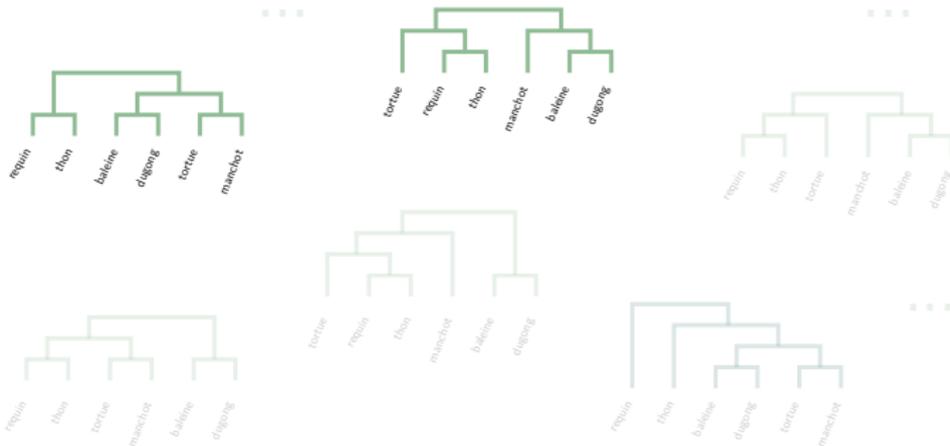
## Exploration de l'espace des arbres

## Principe de reconstruction

DONNEES

requin	A	T	C	G	A	A	G	C	...
thon	-	-	G	-	C	T	-	-	...
baleine	-	A	G	-	-	-	-	-	...
dugong	-	A	G	-	-	-	-	-	...
tortue	-	-	G	-	-	-	A	-	...
manchot	-	-	G	-	-	-	A	-	...

ESTIMATION



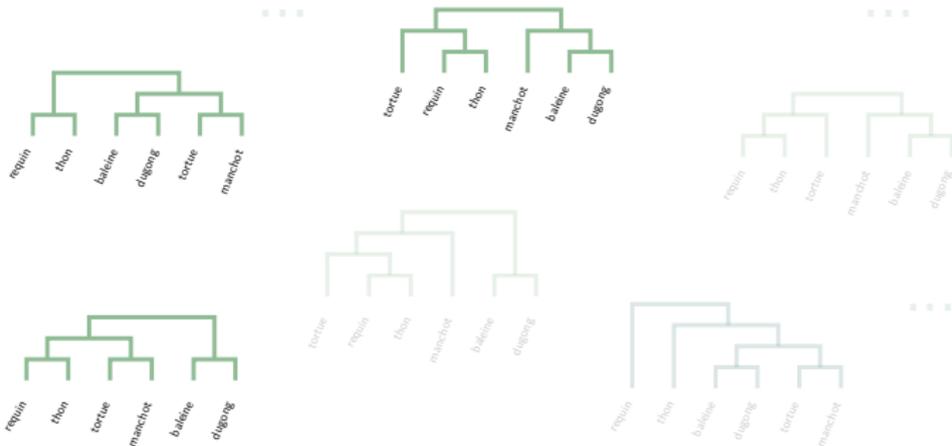
## Exploration de l'espace des arbres

## Principe de reconstruction

DONNEES

requin	A	T	C	G	A	A	G	C	...
thon	-	-	G	-	C	T	-	-	...
baleine	-	A	G	-	-	-	-	-	...
dugong	-	A	G	-	-	-	-	-	...
tortue	-	-	G	-	-	-	A	-	...
manchot	-	-	G	-	-	-	A	-	...

ESTIMATION



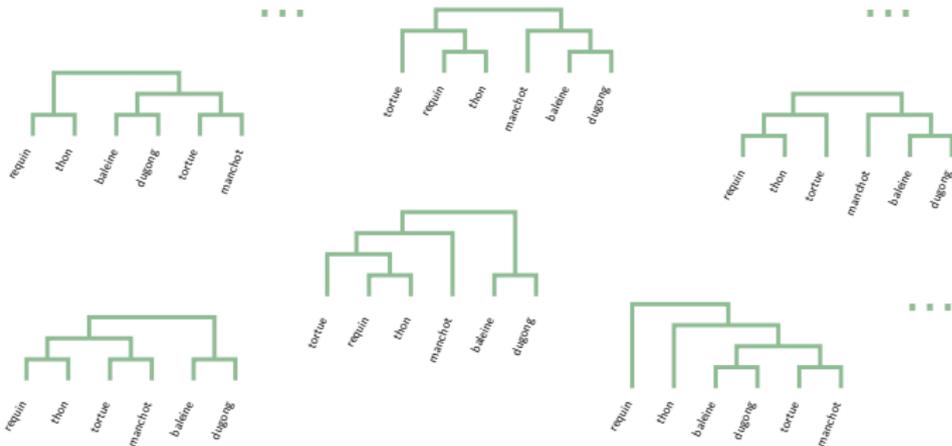
## Exploration de l'espace des arbres

## Principe de reconstruction

DONNEES

requin	A	T	C	G	A	A	G	C	...
thon	-	-	G	-	C	T	-	-	...
baleine	-	A	G	-	-	-	-	-	...
dugong	-	A	G	-	-	-	-	-	...
tortue	-	-	G	-	-	-	A	-	...
manchot	-	-	G	-	-	-	A	-	...

ESTIMATION



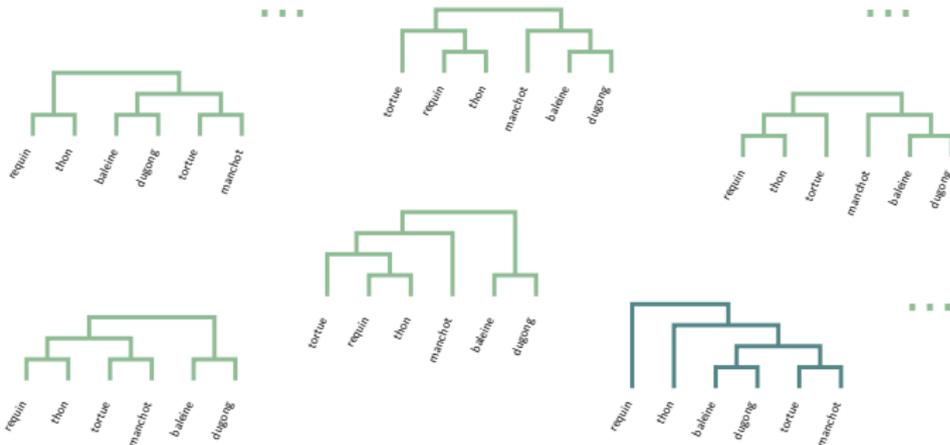
## Exploration de l'espace des arbres

## Principe de reconstruction

DONNEES

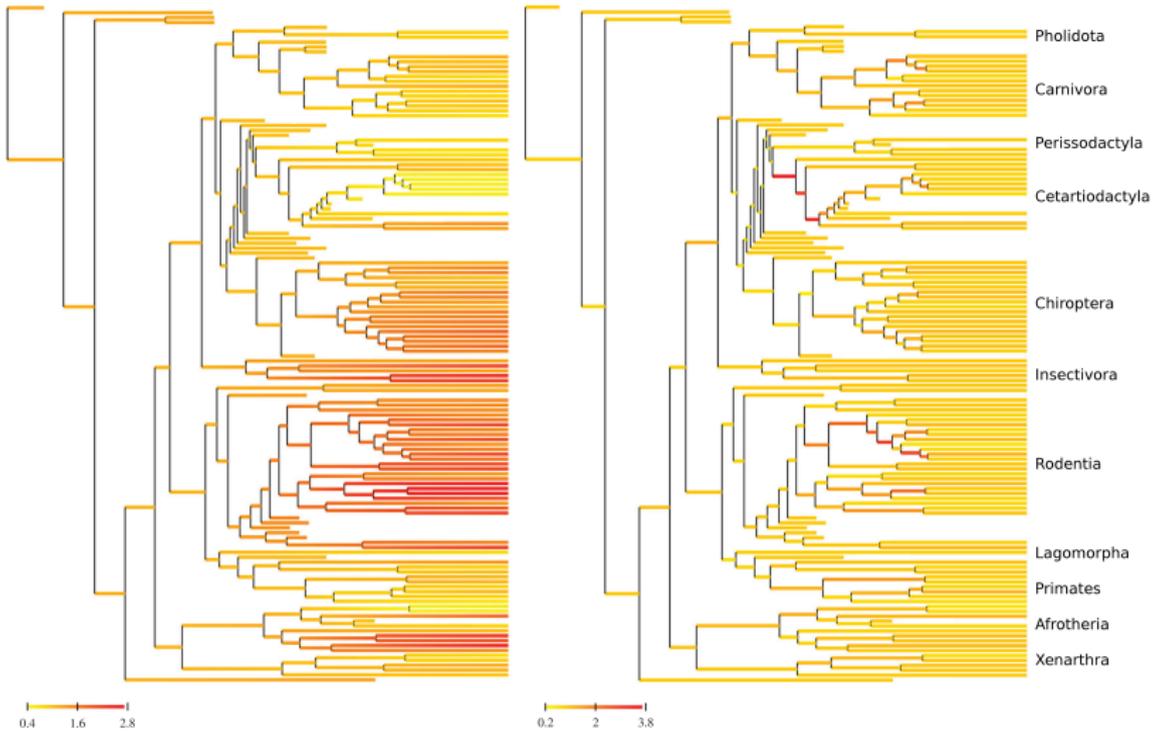
requin	A	T	C	G	A	A	G	C	...
thon	-	-	G	-	C	T	-	-	...
baleine	-	A	G	-	-	-	-	-	...
dugong	-	A	G	-	-	-	-	-	...
tortue	-	-	G	-	-	-	A	-	...
manchot	-	-	G	-	-	-	A	-	...

ESTIMATION



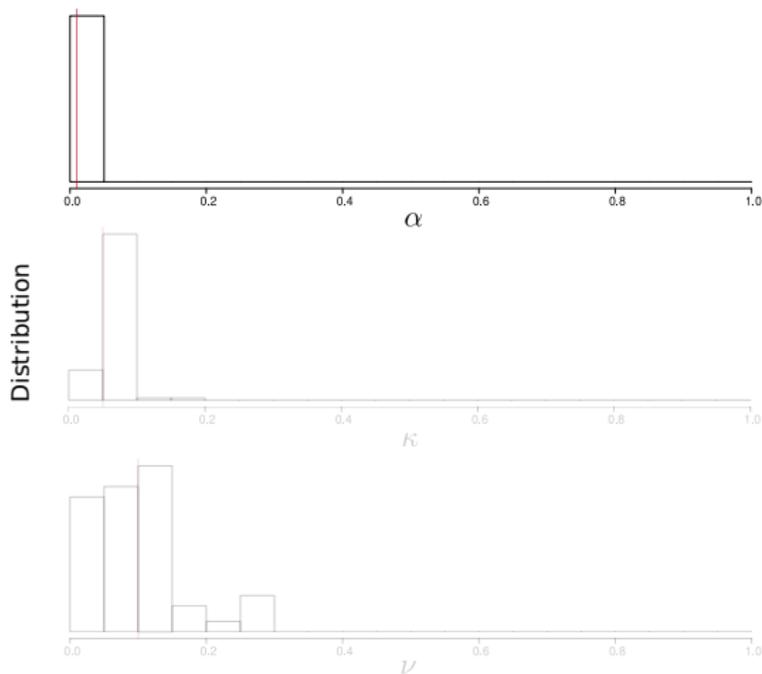
# Horloge moléculaire relâchée

## Résultats biologiques



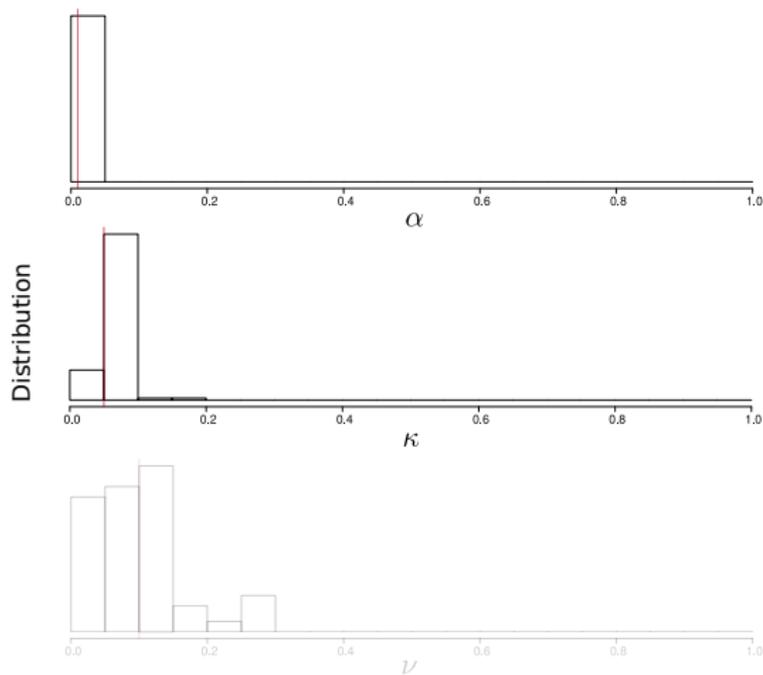
## Avec un prior sur les "paramètres"

Loi de  $\alpha, \kappa, \nu$  sachant les longueurs de branche et l'alignement



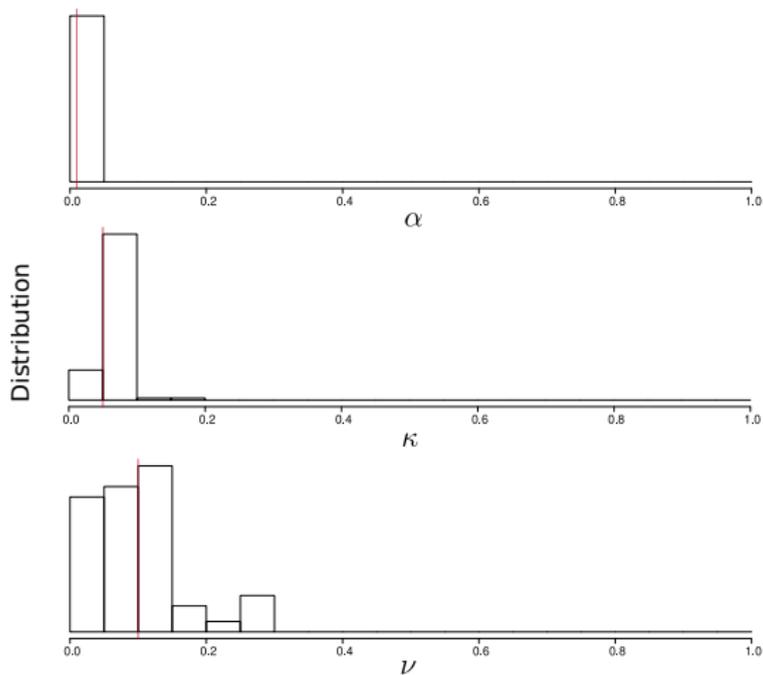
## Avec un prior sur les "paramètres"

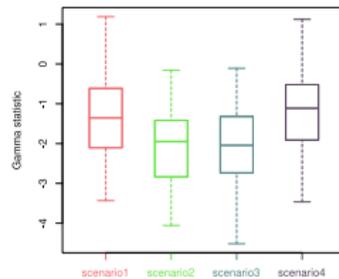
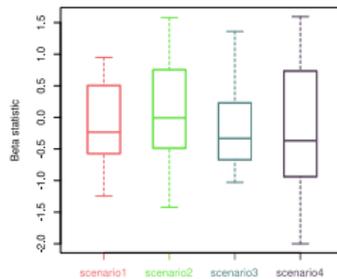
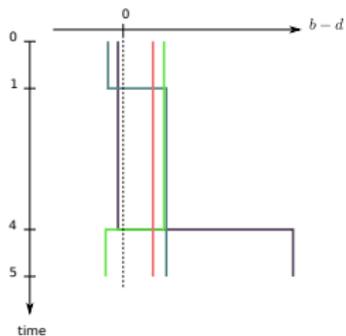
Loi de  $\alpha, \kappa, \nu$  sachant les longueurs de branche et l'alignement



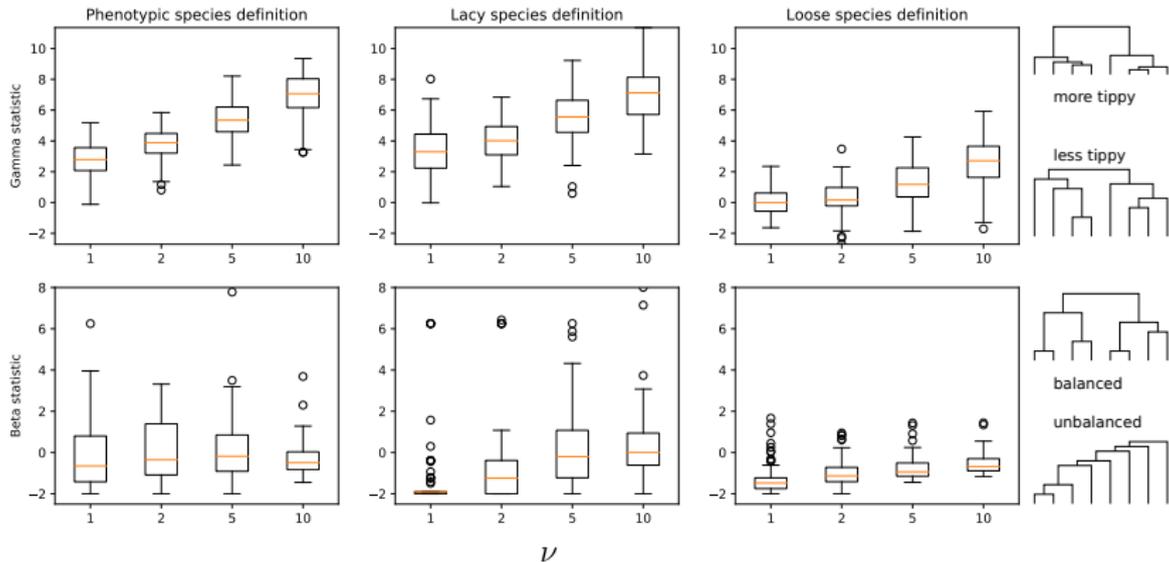
## Avec un prior sur les "paramètres"

Loi de  $\alpha, \kappa, \nu$  sachant les longueurs de branche et l'alignement

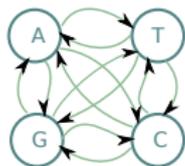




## Conclu



# Liens entre les chapitres



Molecular evolution

$$p(A | T, R)$$



Substitution rate evolution

$$p(R | T)$$



Trait evolution

$$p(C | T)$$



Diversification

$$p(T)$$

$$p(T, R | A, C) \propto p(A | T, R) p(R | T) p(C | T) p(T)$$