

# Le parc jurassique

Équipe Algorithmique Avancée

Semaine 4

## 1 Se préparer à l'expédition

La phylogénèse ou phylogénie (du grec ancien *phylon*, signifiant « race, tribu ») est l'étude des relations de parenté entre êtres vivants. L'un des objectifs de la phylogénie est de reconstituer des arbres phylogénétiques. Les nœuds d'un arbre phylogénétique correspondent à des espèces passées ou présentes (ou à des groupes d'espèces). Une espèce est descendante d'une autre dans l'arbre si la première est effectivement issue de la seconde.

Le dernier ancêtre commun de deux espèces est l'espèce la plus récente dont les deux espèces descendent.

### Question 1

Quel est le dernier ancêtre commun du gorille et du chameau ? Celui du lynx et de la chauve-souris ?

### Question 2

Quel est le dernier ancêtre commun du gorille et du chameau ? Celui du lynx et de la chauve-souris ?

## 2 Au labo !

On a récupéré des fragments d'ADN de plusieurs espèces (dans des fichiers sur Célène). Ces fragments sont des suites de *bases* prises dans l'alphabet "ACTG". Au cours de l'évolution, l'ADN d'une espèce peut subir des mutations. Dans la vraie vie, ces mutations peuvent détruire des parties de l'ADN, ajouter des morceaux qui n'existaient pas, ou encore changer une *base* à un endroit du génome en une autre.

Dans l'univers fictif de ce tp, les seules mutations sont des changements d'une base (pas d'insertions, pas de suppressions). Ainsi, tous les génomes ont la même longueur (vous pouvez vérifier). Nous pouvons donc comparer les brin d'ADN pour estimer combien de mutations les séparent et savoir quelles sont les espèces les plus proches. Pour cela, on suppose que la probabilité de deux mutations sur la même base est nulle, donc le nombre

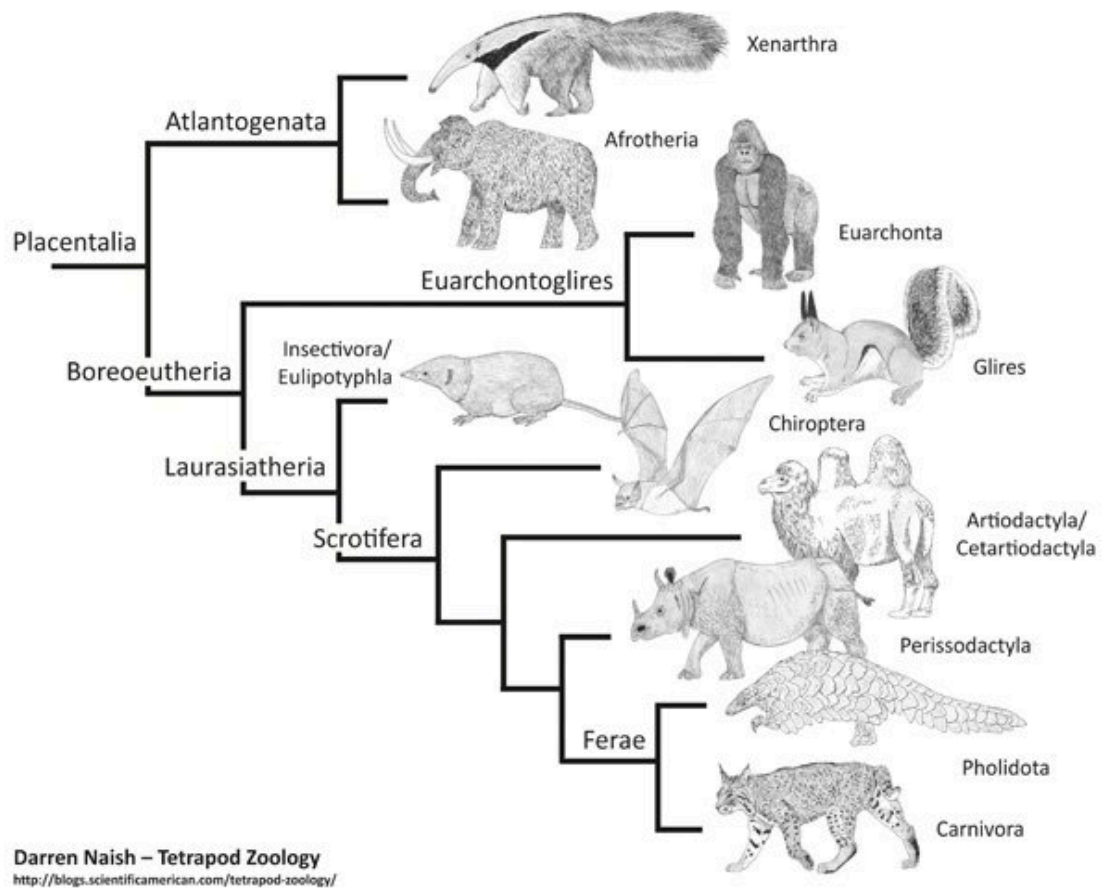


FIGURE 1 – Un exemple d'arbre phylogénétique ; toutes les espèces ne sont pas nommées, mais toutes celles dont vous avez besoin dans le devoir le sont.

de mutations entre deux séquences d'ADN est le nombre de bases qui diffèrent entre les deux.

ACTCAG  
 AGTCAA

FIGURE 2 – Ces deux séquences sont séparées par deux mutations.

### Question 3

Créer une classe **Espec**. Une espèce a un nom, et un adn (extrait d'un des fichiers .adn)

#### Question 4

Combien de mutations séparent l'abeille de l'éponge ? Combien de mutations séparent l'abeille du lapin ? Écrire une fonction `mutations(espece1, espece2)` qui donne le nombre de mutations qui séparent les deux espèces.

#### Question 5

Quelles sont les espèces les plus proches parmi nos échantillons ?

Nous voulons maintenant reconstituer l'arbre phylogénétique comme au-dessus. Nous avons pour cela besoin de représenter les ancêtres communs des espèces dont nous avons l'ADN. Nous n'avons pas l'ADN de ces ancêtres à notre disposition.

#### Question 6

Créer une classe `EspeceHypothetique` qui dérive de `espece`. Son nom est obtenu en concaténant les noms des espèces qui en dérivent (par exemple, «trex-roadrunner»). Son ADN est `None`, et elle a un attribut `especesFilles` qui contient les espèces descendant de cette espèce.

#### Question 7

Créer le `trex-roadrunner`.

Pour évaluer le nombre de mutations qui séparent une `EspeceHypothetique` A d'une autre espèce B, on fait la moyenne du nombre de mutations qui séparent chacune des `especeFilles` de B.

#### Question 8

Calculer la distance entre le «trex-roadrunner» et le lapin.

#### Question 9

Calculer la distance entre le «trex-roadrunner» et le «lapin-garou» (ou humain lapin).

Pour reconstruire l'arbre phylogénétique, on procède comme suit :

- On crée toutes les espèces
- Tant qu'il reste au moins deux espèces qui n'ont pas d'ancêtre commun, on trouve les deux espèces dont la distance est la plus petite, et on les remplace par l'espèce-hypothétique représentant leur dernier ancêtre commun.

**Question 10**

Reconstituer l'évolution de la vie sur terre.

**Question 11**

Écrire une fonction qui reconstitue l'évolution de la vie sur terre pour nous.