

UE Biologie Evolutive

2023 – 2024



Cours: Marc Girondot, Myriam Harry

TD/TP: Marc Girondot, Héloïse Bastide



Organisation de l'UE Biologie Évolutive de L3

Cette UE est constituée de 10 cours de 3 h chacun assurés par les professeurs Myriam Harry et Marc Girondot.

08/01/2024	9h-12h	Cours 1	Marc Girondot Evolution-Phylogénie-Systematique
09/01/2024	9h-12h	Cours 2	Marc Girondot Espèce - Homologie Marc Girondot Evolution moléculaire - Méthode de reconstruction
15/01/2024	9h-12h	Cours 3	
22/01/2024	9h-12h	Cours 4	Marc Girondot Dinosaures
29/01/2024	9h-12h	Cours 5	Marc Girondot Autopode
05/02/2024	9h-12h	Cours 6	Myriam Harry Coévolution
12/02/2024	9h-12h	Cours 7	Myriam Harry Sociétés animales
04/03/2024	9h-12h	Cours 8	Myriam Harry Premiers Métazoaires
18/03/2024	9h-12h	Cours 9	Myriam Harry Evo hum 1
25/03/2024	9h-12h	Cours 10	Myriam Harry Evo hum 2

Sont associés à ces 10 cours, 6 séances de TD de 2 heures et une séance de TP de 3 de bioinformatique assurés par Marc Girondot et Héloïse Bastide

Les TD, dont les sujets sont ci-après, sont :

- Lire et interpréter une phylogénie
- Analyse du film : Évolution de la taille chez l'humain



- Co-évolution
- Analyse du film : Les mystères de l'évolution



- Évolution moléculaire de l'émail dentaire
- Évolution humaine

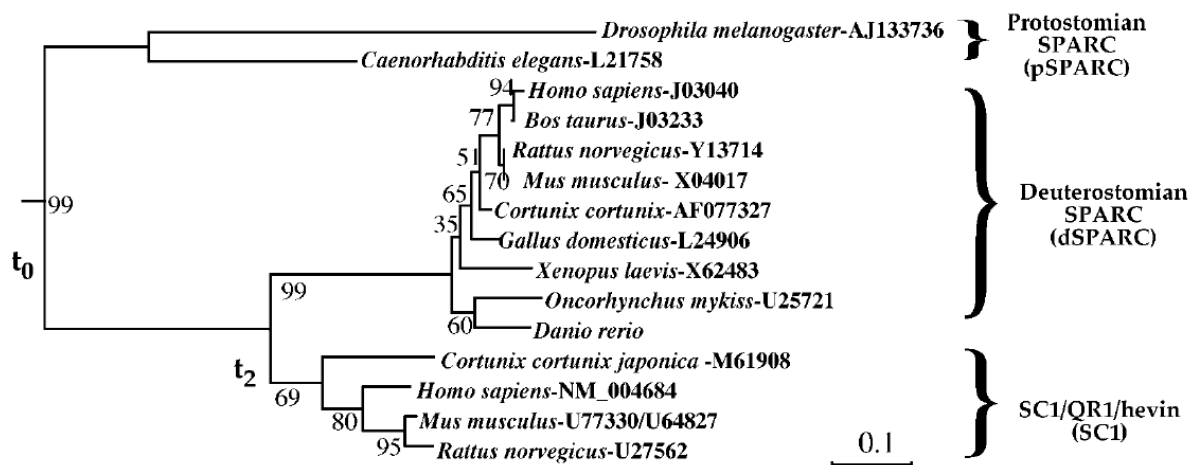
Nous vous demandons de préparer les TD et pour les films, de le regarder (1 heure) et de noter les différentes hypothèses testées et de noter leur niveau de réfutation des hypothèses.

TP de bioinformatique en Biologie Évolutive

Problématique:

Retracer l'évolution de SPARC et SL1

D'après l'article Delgado, S., Casane, D., Bonnaud, L., Laurin, M., Sire, J.-Y., Girondot, M., 2001. Molecular evidence for precambrian origin of amelogenin, the major protein of vertebrate enamel. *Molecular Biology and Evolution* 18, 2146-2153.



- Le but du TP est de vous mettre dans une situation où vous chercherez vous-même les outils à utiliser pour répondre à une question : on ne vous prend pas par la main; on est là seulement pour débloquer des situations qui vous empêcheraient d'avancer.

La procédure à suivre est la suivante :

- Trouver des séquences de SPARC et SPARC-Like 1 permettant de répondre à la problématique
- Aligner les séquences pour que les différentes positions dans la matrice des séquences soient homologues
- Produire une phylogénie
- Enraciner la phylogénie
- Discuter l'évolution des gènes SPARC et SPARC-Like 1

- Faire un compte-rendu d'une page (recto) donnant les principales informations sur les méthodes et le ou les résultats

Matériels et méthodes

- Genbank, GenBank is the NIH genetic sequence database
 - Interrogation par mots-clés ou par recherche de similitudes
- Blast, algorithme de recherche extrêmement rapide permettant de comparer une séquence avec l'ensemble des séquences disponibles dans Genbank
 - Blastn: séquence nucléotidique contre séquences nucléotidiques
 - Bastp: séquence protéique contre séquences protéiques
 - Bastx: séquence protéique contre séquences nucléotiques
 - Megablast: recherche de séquences très proches
- Alignement
 - Clustal-W: This method works by analyzing the sequences as a whole, then utilizing the UPGMA/Neighbor-joining method to generate a distance matrix. A guide tree is then calculated from the scores of the sequences in the matrix, then subsequently used to build the multiple sequence alignment by progressively aligning the sequences in order of similarity.
 - T-Coffee (Tree-based Consistency Objective Function for Alignment Evaluation) is a multiple sequence alignment software using a progressive approach. It generates a library of pairwise alignments to guide the multiple sequence alignment.
 - The MUSCLE algorithm proceeds in three stages: the draft progressive, improved progressive, and refinement stages. In the draft progressive stage, the algorithm produces a draft multiple alignment, emphasising speed over accuracy. In the improved progressive stage, the Kimura distance is used to reestimate the binary tree to create the draft alignment, in turn producing a more accurate multiple alignment. The final refinement stage refines the improved alignment made in step two.

- Phylogénie
 - ME: Minimum evolution is a distance method employed in phylogenetics modeling. It searches for the phylogeny that has the shortest total sum of branch lengths.
 - BioNJ
 - ML: The maximum likelihood method uses standard statistical techniques for inferring probability distributions to assign probabilities to particular possible phylogenetic trees. The method requires a substitution model to assess the probability of particular mutations; roughly, a tree that requires more mutations at interior nodes to explain the observed phylogeny will be assessed as having a lower probability.
 - PhyML
- Outils en ligne
 - Vous trouverez ces outils en ligne à l'adresse:
<http://www.phylogeny.fr>
 - Vous trouverez des informations intéressantes pour comprendre les algorithmes sur internet et en particulier en partant de Wikipedia:
https://en.wikipedia.org/wiki/Computational_phylogenetics

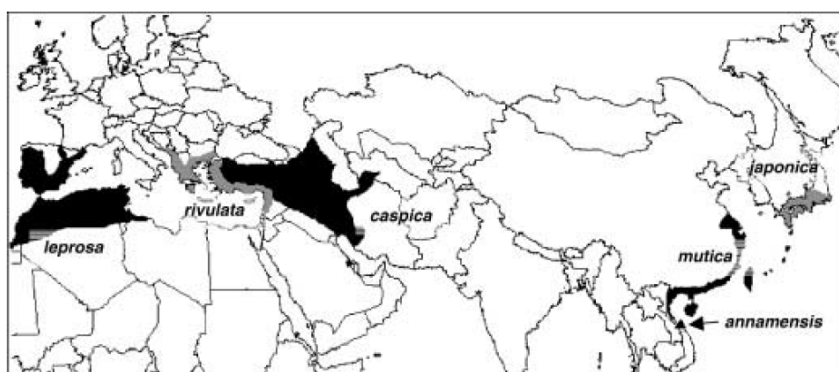
Les modalités de contrôle des connaissances sont un examen en salle de 2 heures avec des sujets portant sur les différents cours, TD et TP.

Biologie Evolutive TD 1 - Lire et interpréter une phylogénie

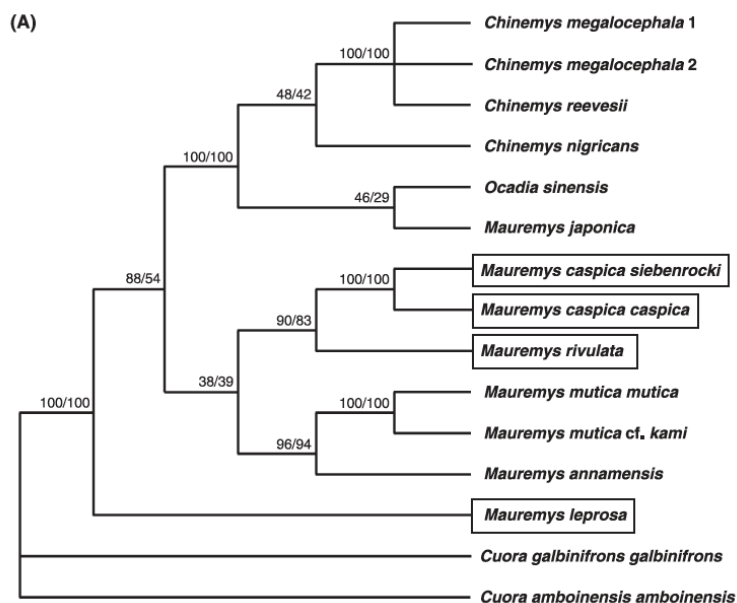
« The freshwater turtle genus *Mauremys* belongs to the Geoemydidae, a family consisting of approximately 60 species in 25 genera. Geoemydid turtles are mainly distributed in eastern Asia. Only three *Mauremys* species occur in the western Palearctic (»

Mauremys has a patchy distribution, including parts of the western Palearctic region and, separated by a huge disjunction, parts of the Oriental and eastern Palearctic regions »

In Barth, D., Bernhard, D., Fritzsich, G., Fritz, U., 2004. The freshwater turtle genus *Mauremys* (Testudines, *Geoemydidae*)-A textbook example of an east-west disjunction or a taxonomic misconception? *Zoologica Scripta* 33, 213-221.



Distribution of all currently accepted *Mauremys* species based on Iverson (1992), Fritz & Wischuf (1997), Keller & Busack (2001) and Wischuf & Busack (2001).



Phylogenetic trees of the genera *Mauremys*, *Chinemys* and *Ocadia* using the two *Cuora* species as outgroups. Names of western Palearctic taxa in boxes, of East Asiatic taxa without boxes.

A. MP tree. The first numbers give the bootstrap values out of 1000 trees. The same topology was inferred with the maximum likelihood (ML) method (HKY85 + G) in PAUP; the second are bootstrap values for 100 bootstrap resamplings.

Nodes supported by values below 50% are given as multifurcations.

Sur la base de ces deux figures de cette publication, que pouvez-vous dire du statut taxonomique du genre *Mauremys* ?

Il est important de savoir lire les légendes : MP, ML, HKY85 + G, PAUP, bootstraps, multifurcations.

L'aire de répartition des autres espèces de cette phylogénie est montrée ci-dessous. Discutez des relations évolutives entre ces différentes espèces d'un point de vue biogéographique.

Ocadia sinensis



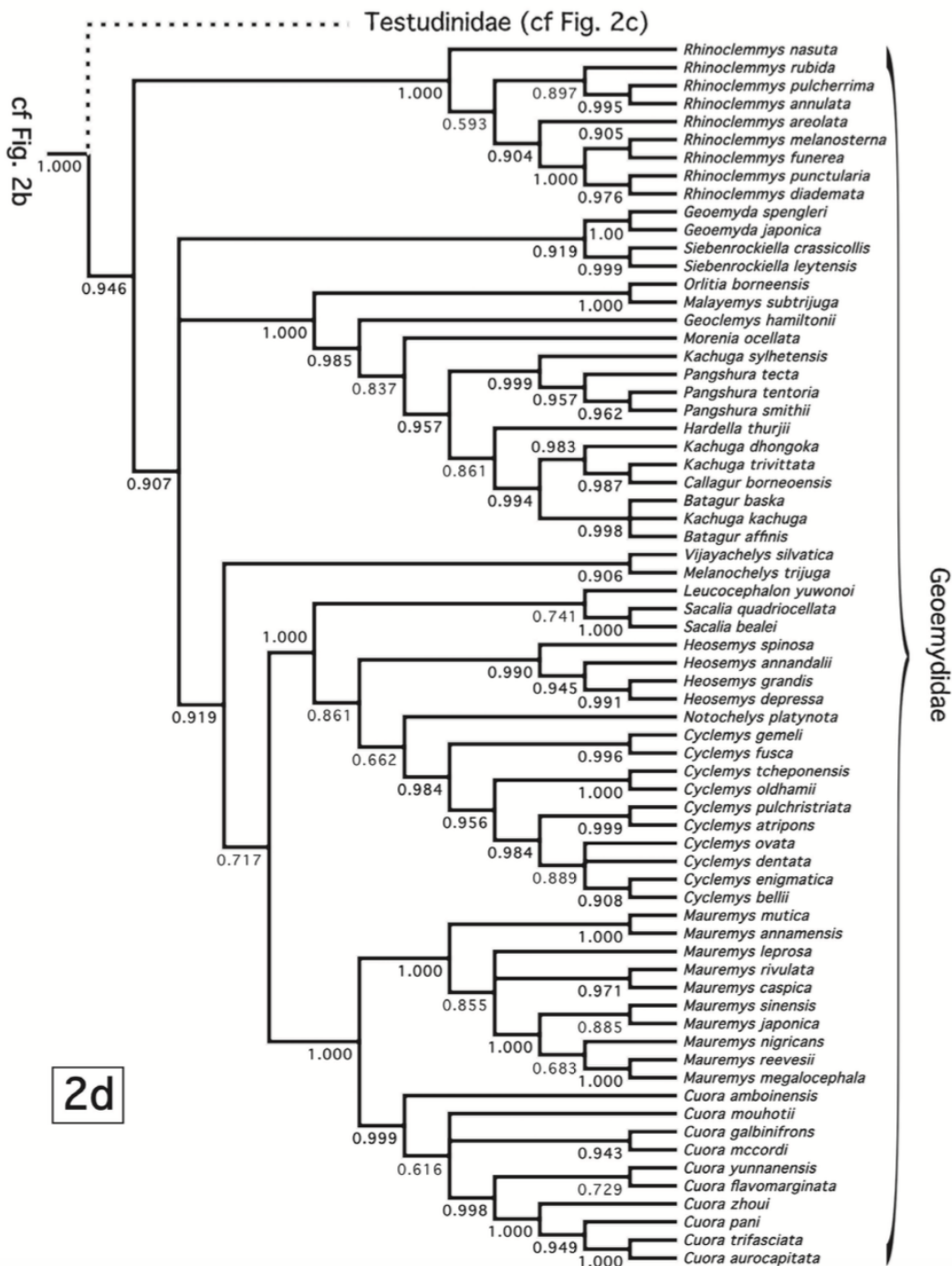
Chinemys nigricans



Chinemys reevesii et *Chinemys megaloccephala*



Une phylogénie complète des espèces de tortues a été publiée :
 Guillon, J.-M., Guéry, L., Hulin, V., Girondot, M., 2012. A large phylogeny of
 turtles (Testudines) using molecular data. Contributions to Zoology 81, 147-158.

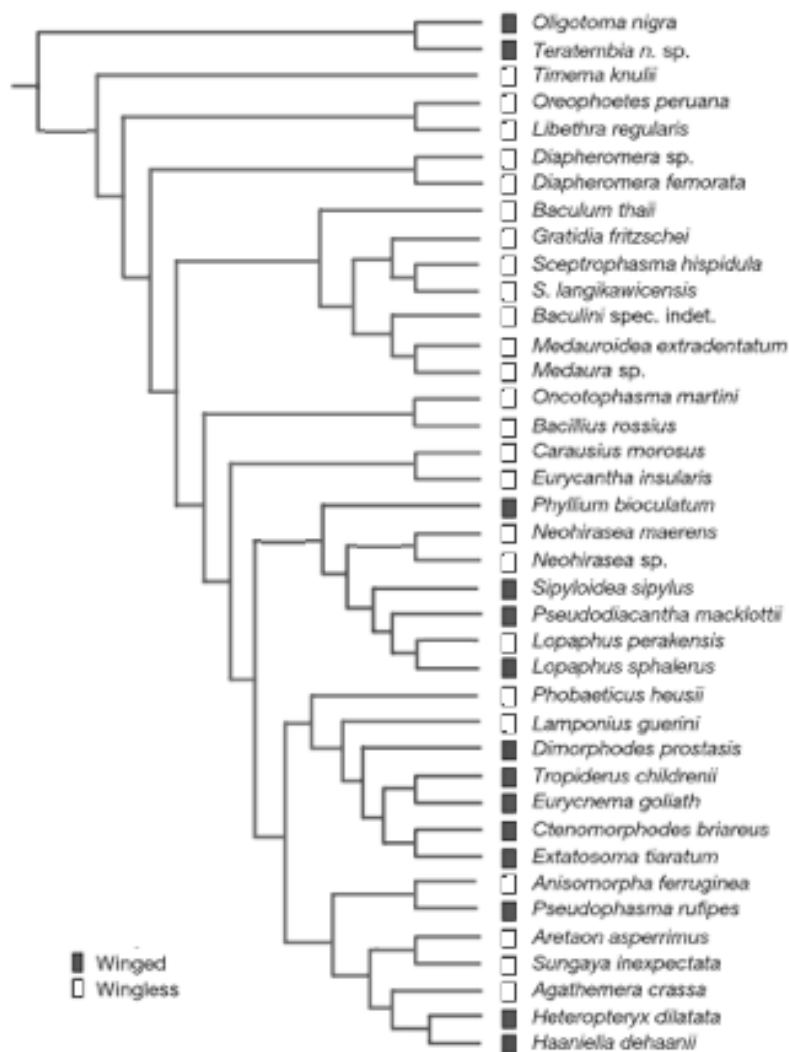


Legend: Maximum likelihood phylogenetic tree obtained using PhyML and the complete DNA dataset with GTR + I + G model: freqA = 0.36109; freqC = 0.29917; freqG = 0.11484; freqT = 0.22491; A <-> C = 0.91133; A <-> G = 5.21616; A <-> T = 0.94284; C <-> G = 0.50873; C <-> T = 10.42092; G <-> T = 1.00000 ; Proportion of invariant sites = 0.233; Gamma shape parameter = 0.512. Confidence values are indicated under each branch. Branches with confidence value lower than 0.5 have been collapsed. Branch supports are in bold when exceeding 0.9.

Ces données sont-elles compatibles avec celles de Barth, D. *et al.* 2004 ?
 Qu'apportent-elles sur la compréhension de la biogéographie du groupe ?

Caractères et phylogénies

Les phasmes sont des insectes mimétiques avec des branches ou des feuilles. La perte des ailes est observée chez plusieurs espèces du groupe :



Whiting MF, Bradler S, Maxwell T (2003) Loss and recovery of wings in stick insects. *Nature* 421: 264-267



Figure 3 Character mapping of wing types on phasmid phylogeny. Parsimony optimization (ACCTRAN) of winged (blue) and wingless (red) states for male phasmids on the optimization alignment topology. This reconstruction requires seven steps with four wing gains and three losses; DELTRAN optimization requires five wing gains and two losses. Maximum likelihood reconstruction produces similar results (see Supplementary Information).

Quels sont le caractère et les états de caractère étudiés ?
 Reconstruire l'évolution du caractère par parcimonie.

Expliquer les notions de ACCTRAN et DELTRAN.

Comment les définiriez-vous dans les clades suivants :

- Le plus proche ancêtre commun de *Lopaphus perakensis* et *Timema knulii*. et tous ses descendants
- Le plus proche ancêtre commun de *Lopaphus perakensis* et *Teratembia n. sp.* et tous ses descendants
- Les groupes plus proches de *Lopaphus perakensis* que de *Timema knulii*
- Les groupes plus proches de *Phyllium bioculatum* que de *Tropidenus childreni*

La loi de Dollo, du nom du paléontologue belge, Louis Dollo, stipule que :
« Mr. Dollo is of the opinion:

1. That evolution occurred by abrupt leaps.
2. That an organism cannot return, even partially, to a previous state already realized in its ancestral series. » in Dollo, L., 1893. Les lois de l'évolution. Bull. Soc. Belge Geol. Pal. Hydr. VII, coll. B, 164-166.



Louis Antoine Marie Joseph Dollo, né à Lille le 7 décembre 1857 et mort à Bruxelles le 19 avril 1931

Que pouvez-vous dire de l'application de cette loi biogénétique sur le cas présenté ? Vous pouvez vous aider de la phylogénie suivante.

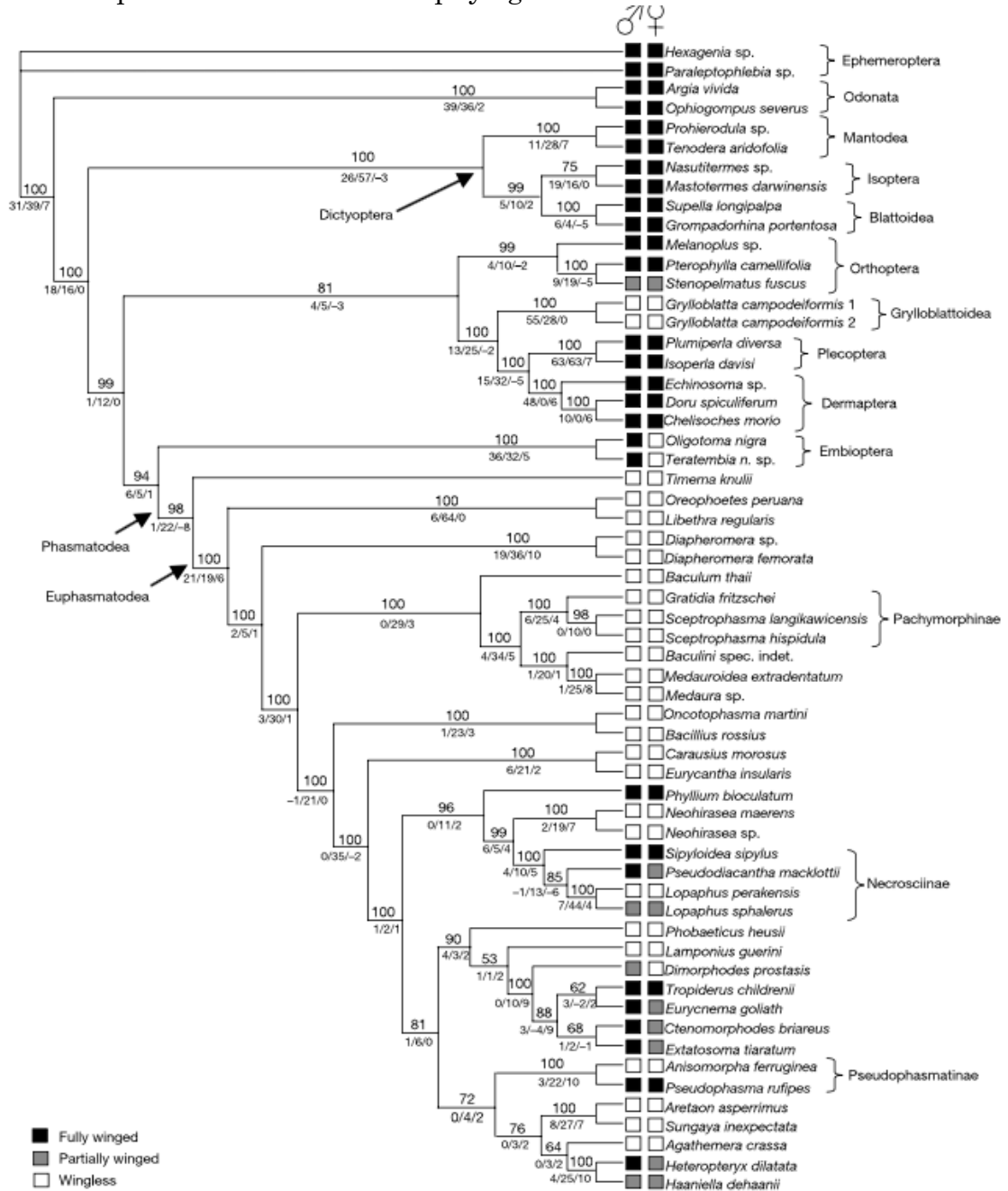


Figure 2 Phylogeny of Phasmatodea on the basis of molecular data. Shown is the single optimization alignment tree based on 18S rDNA, 28S rDNA and histone 3. Nonparametric bootstrap supports are given above each node and partitioned Bremer supports are below each node in the order 18S/28S/H3. This topology is congruent with the maximum

likelihood and bayesian topology (see Supplementary Information). Boxes at the end of nodes represent wing character states for males and females, respectively. n. sp., new species; spec. indet., undetermined species.

TD 3 : La Coévolution

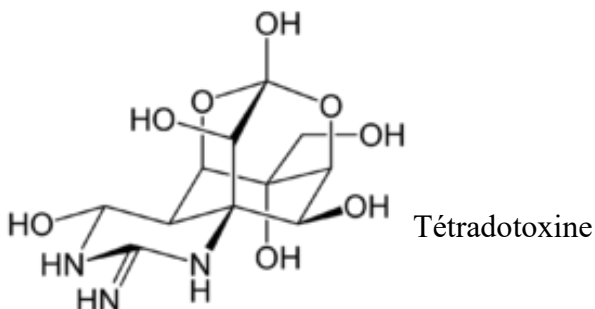
- 1) Quelles sont les 6 types d'interactions possibles entre deux espèces ? Donner un exemple pour chacun d'eux.
- 2) En biologie on appelle « coévolution » le processus par lequel deux espèces (ou plus) qui interagissent évoluent ensemble, chacune changeant en réponse aux changements dans l'autre ou les autres espèces. Parmi les 6 types d'interactions que vous venez de lister, lesquels donneront le plus souvent naissance à un phénomène de coévolution ?

La coévolution antagoniste chez le triton rugueux et la couleuvre rayée

Le triton rugueux *Taricha granulosa* est une espèce de triton vivant dans l'ouest de l'Amérique du Nord qui, comme d'assez nombreuses espèces d'amphibiens, produit dans sa peau de la tétrodotoxine (TTX). Cette molécule, la même que celle qui est produite par les poissons de la famille des tétraodons, est une neurotoxine extrêmement puissante qui en cas d'ingestion bloque les canaux sodium dans les nerfs et les muscles et peut causer la mort par paralysie respiratoire. La dose mortelle pour un homme est environ 1/30^{ème} de la dose contenue dans une salamandre. Contrairement à ce qui se passe pour de nombreux animaux toxiques, les tritons rugueux synthétisent la TTX, elle ne provient pas de leur alimentation.



- 3) Le triton rugueux est gris-brun sur le dessus, et jaune ou orange vif sur le dessous. Quel est l'avantage de ce type de coloration ?
- 4) Quel est l'avantage pour le triton rugueux de produire de la tétrodotoxine ? Est-il surprenant qu'il en produise autant ?



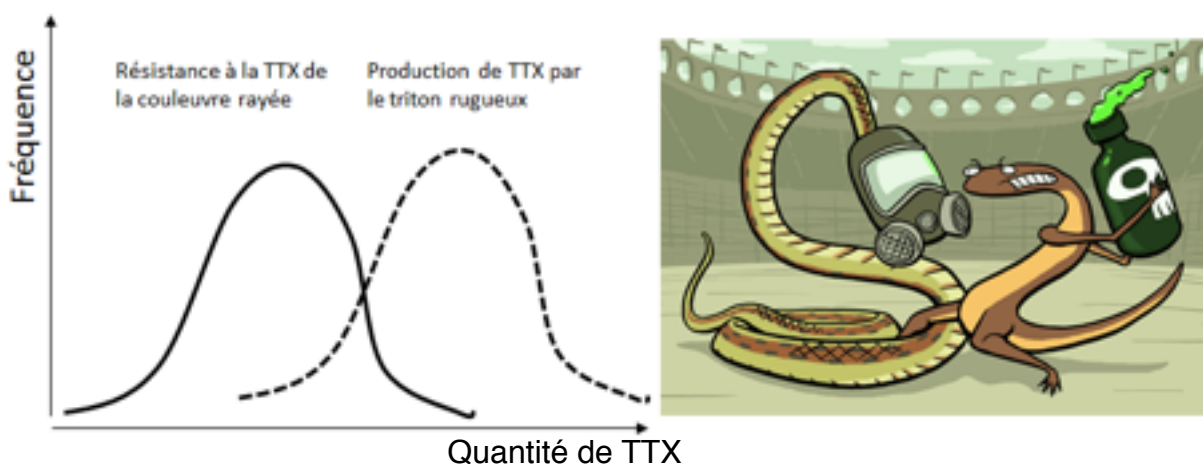
L'aposématisme est la stratégie adaptative qui permet à certains organismes (généralement des animaux, parfois des plantes) d'émettre un signal d'avertissement clairement perceptible, qui peut être visuel (le plus souvent une couleur), sonore ou olfactif. Le terme *aposématisme* vient du grec ancien : ἀπό / *apó*, « repousser », et σήμα / *sēma*, « signal ».

En raison de la grande quantité de TTX qu'ils contiennent, la consommation de triton rugueux est mortelle pour la grande majorité des prédateurs, une exception étant la couleuvre rayée (*Thamnophis sirtalis*), un serpent qui présente une résistance, variable suivant les populations, à la TTX. Lorsqu'une couleuvre rayée attrape un triton rugueux, si celui-ci produit beaucoup de TTX, alors la couleuvre le relâche et il peut s'enfuir.

Williams, B.L., Hanifin, C.T., Brodie, E.D., Brodie lii, E.D., 2010. Tetrodotoxin affects survival probability of rough-skinned newts (*Taricha granulosa*) faced with TTX-resistant garter snake predators (*Thamnophis sirtalis*). *Chemoecology* 20, 285-290.



- 5) Comment la capacité des couleuvres rayées à consommer les tritons rugueux peut-elle expliquer la très grande quantité de TTX produite par les tritons rugueux ? Supposez que, initialement, la distribution de la production de TTX chez le triton et de la résistance chez la couleuvre rayée soit représentée sur la figure ci-dessous. D'après vous comment évolueront ces distributions au cours du temps ? Faites une représentation graphique. Comment peut-on appeler ce type de processus ? Va-t-il continuer indéfiniment ? Si le processus s'arrête, quelles sont les issues possibles ?



Script permettant de visualiser l'évolution de la résistance et celle de la toxicité. A lancer dans RStudio. Modèle individu-centré très simplifié.

```
# Fonction qui permet d'afficher deux histogrammes superposés avec des légendes
hist_both <- function(tritons.synthese, couleuvres.resistance, maxx=100, maxy=100, iteration="") {
  hist(couleuvres.resistance, xlim=c(0, maxx), ylim=c(0, maxy), breaks = seq(from=0, to=100, by=2),
       col=rgb(red = 0.5, green = 0.5, blue = 0.5, alpha = 0.5),
       xlab="Quantité de TTX", ylab="Nombre", main=iteration)
  par(new=TRUE)
  hist(tritons.synthese, xlim=c(0, maxx), ylim=c(0, maxy), breaks = seq(from=0, to=100, by=2),
       col=rgb(red = 0., green = 0., blue = 0.6, alpha = 0.5), xlab="", ylab="", axes = FALSE, main="")
  legend(x="topleft", legend=c("Synthèse TTX chez le triton", "Résistance TTX chez la couleuvre"),
        col=c(rgb(red = 0., green = 0., blue = 0.6, alpha = 0.5),
              rgb(red = 0.5, green = 0.5, blue = 0.5, alpha = 0.5)), pch=15, bty="n")
}

taille.population.couleuvre <- 100
taille.population.triton <- 100
Nombre.iterations <- 100

# Population initiale tirée dans une loi normale de moyenne 20 pour la résistance et 40 pour la synthèse
couleuvres.resistance <- rnorm(taille.population.couleuvre, mean=20, sd=3)
tritons.synthese <- rnorm(taille.population.triton, mean=40, sd=3)

# Je fais tourner le modèle sur Nombre.iterations itérations
for (iteration in 1:Nombre.iterations) {

  hist_both(tritons.synthese, couleuvres.resistance, maxy=40, iteration=as.character(iteration))

  # Rencontre des couleuvres et des tritons
  rencontres <- floor(min(taille.population.couleuvre, taille.population.triton)/2)
  couleuvres <- sample(1:taille.population.couleuvre, rencontres, replace = FALSE)
  tritons <- sample(1:taille.population.triton, rencontres, replace = FALSE)

  # Je regarde qui meurt, couleuvre ou triton
  mort.couleuvre <- ifelse(couleuvres.resistance[couleuvres] < tritons.synthese[tritons], TRUE, FALSE)
  mort.triton <- !mort.couleuvre

  # Je les fais mourir
  if (any(mort.couleuvre)) couleuvres.resistance[mort.couleuvre] <- NA
  if (any(mort.triton)) tritons.synthese[tritons[mort.triton]] <- NA

  couleuvres.resistance <- na.omit(couleuvres.resistance)
  tritons.synthese <- na.omit(tritons.synthese)

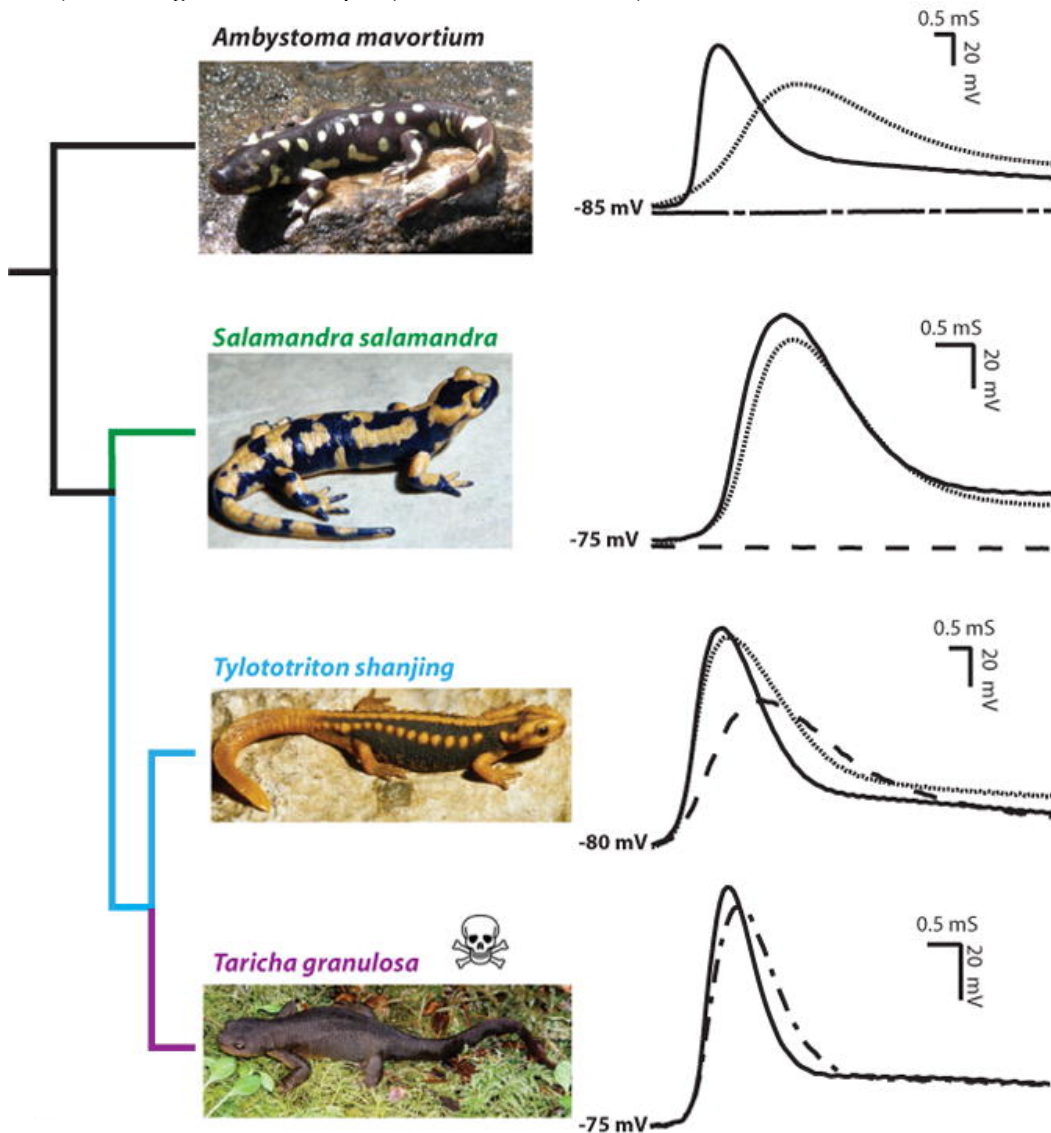
  # S'il y a eu des morts chez les couleuvres, je fais de la reproduction
  if (length(couleuvres.resistance) < taille.population.couleuvre) {
    # Reproduction non-sexuée
    nouvelles.couleuvres <- sample(couleuvres.resistance, taille.population.couleuvre-length(couleuvres.resistance), replace
= TRUE)
    # Mutation
    nouvelles.couleuvres <- rnorm(taille.population.couleuvre-length(couleuvres.resistance), nouvelles.couleuvres, sd=3)
    couleuvres.resistance <- c(couleuvres.resistance, nouvelles.couleuvres)
    # J'élimine les valeurs aberrantes
    couleuvres.resistance <- ifelse(couleuvres.resistance < 0, 0, couleuvres.resistance)
    couleuvres.resistance <- ifelse(couleuvres.resistance > 100, 100, couleuvres.resistance)
  }

  # S'il y a eu des morts chez les tritons, je fais de la reproduction
  if (length(tritons.synthese) < taille.population.triton) {
    # Reproduction non-sexuée
    nouveaux.tritons <- sample(tritons.synthese, taille.population.triton-length(tritons.synthese), replace = TRUE)
    # Mutation
    nouveaux.tritons <- rnorm(taille.population.triton-length(tritons.synthese), nouveaux.tritons, sd=3)
    tritons.synthese <- c(tritons.synthese, nouveaux.tritons)
    # J'élimine les valeurs aberrantes
    tritons.synthese <- ifelse(tritons.synthese < 0, 0, tritons.synthese)
    tritons.synthese <- ifelse(tritons.synthese > 100, 100, tritons.synthese)
  }

  Sys.sleep(0.5)
}
```

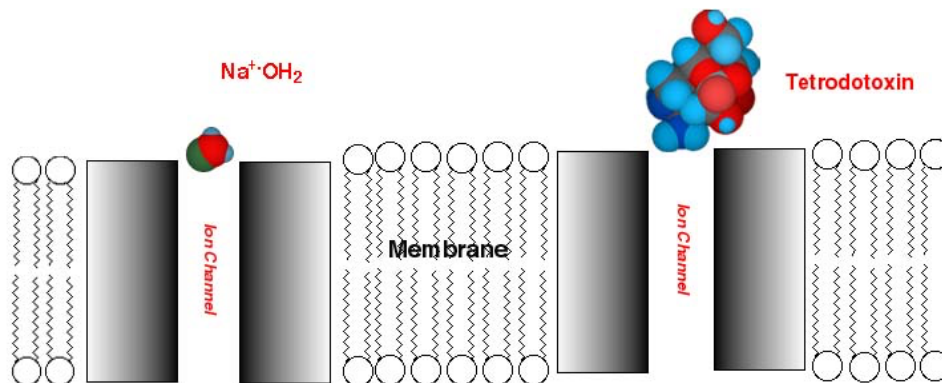
Pour pouvoir utiliser la TTX comme défense contre les prédateurs, il faut bien sûr ne pas être sensible soi-même à la TTX. La figure ci-dessous vous présente des mesures de résistance à la TTX qui ont été réalisées chez le triton rugueux et 3 autres espèces. Seul le triton rugueux produit de la TTX.

Figure 1. Hanifin et al. 2014. Evolutionary history of a complex adaptation: Tetrodotoxin resistance in salamanders. Intracellular action potentials (APs) recorded in skeletal muscle fibers from the indicated species exposed to a range of TTX concentrations. Control APs (no TTX) are shown in solid black lines. *Ambystoma mavortium*: 0.025 μM TTX (staggered dashed line), 0.010 μM TTX (dotted line). *Salamandra Salamandra*, *Tylotriton shanjing*: 0.050 μM TTX (dashed line), 0.010 μM TTX (dotted line). *Taricha granulosa*: 300 μM (dashed and dotted line).



6) Analysez ce document. Quel mécanisme évolutif pourrait avoir été impliqué dans l'apparition de l'utilisation de la TTX chez le triton rugueux ?

- 7) La TTX cause son effet toxique en bloquant les canaux sodium (cf figure ci-dessous), dans les nerfs et les muscles. La fixation sur un canal sodique « classique » est très étroite. Quel type de mutation pourrait conduire à une résistance à la TTX ?



- 8) Le document ci-dessous vous présente la séquence protéique du canal sodium chez plusieurs espèces de tritons et salamandres. Des expériences de mutagenèse ainsi que des observations chez différentes espèces ont permis de montrer que les substitutions observées aux positions 1116, 1424, 1431 et 1432 conduisent toutes à une résistance, plus ou moins importante, à la TTX. Ces données confirment-elles les hypothèses que vous avez formulées précédemment ?

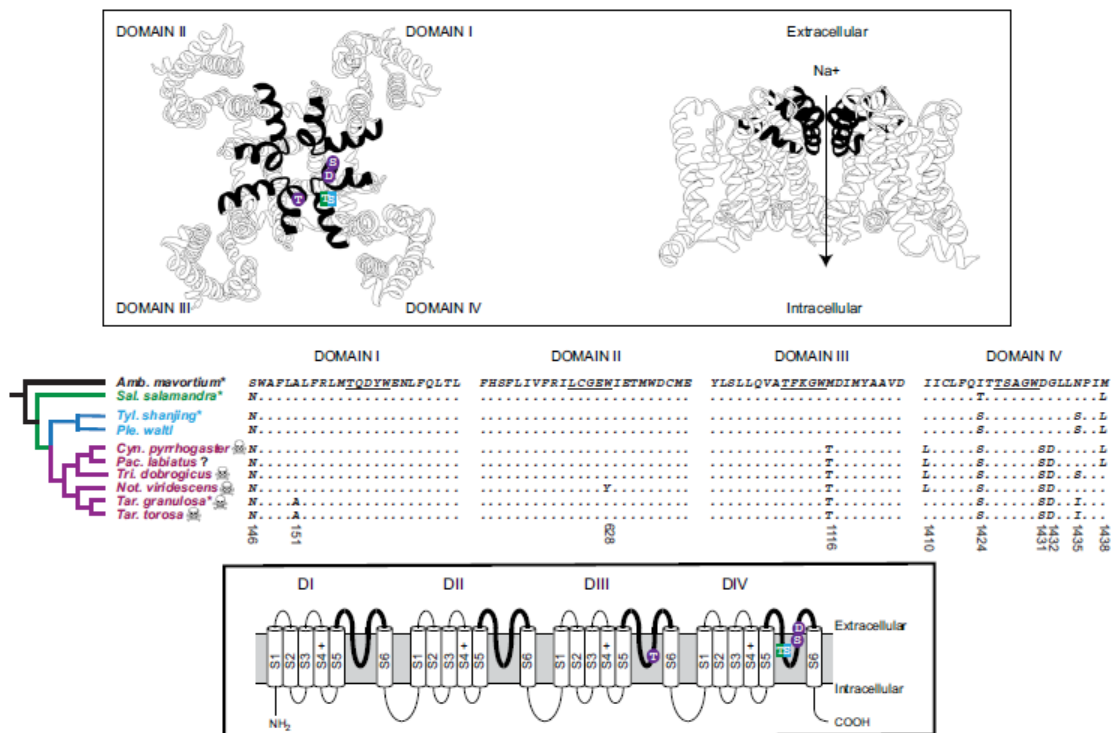
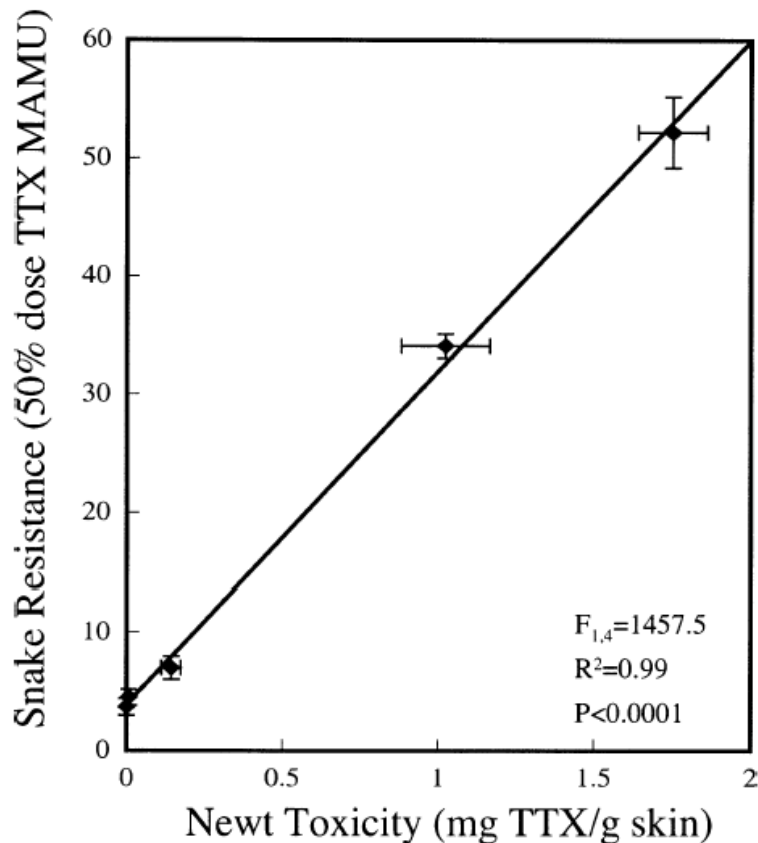


Figure 2. Substitutions in the salamander voltage-gated sodium channel, Na_v 1.4. The phylogenetic distribution of amino acid substitutions in the outer pore of salamander Na_v 1.4 (middle) is shown with a model of the TTX-binding site/outer pore (in black, above) and the transmembrane organization of Na_v 1.4 (below).

Après consommation d'un triton rugueux, la couleuvre rayée sera seulement immobilisée ou ralentie plus ou moins longtemps. On peut donc mesurer la résistance des couleuvres rayées à la TTX en leur faisant consommer une certaine dose de TTX et en mesurant leur réaction. Une étude a mesuré à la fois la vitesse de déplacement des serpents de plusieurs localités et leur résistance à la TTX. On observe que les deux variables sont corrélées négativement.

9) Comment pourrait-on expliquer cette corrélation négative ?

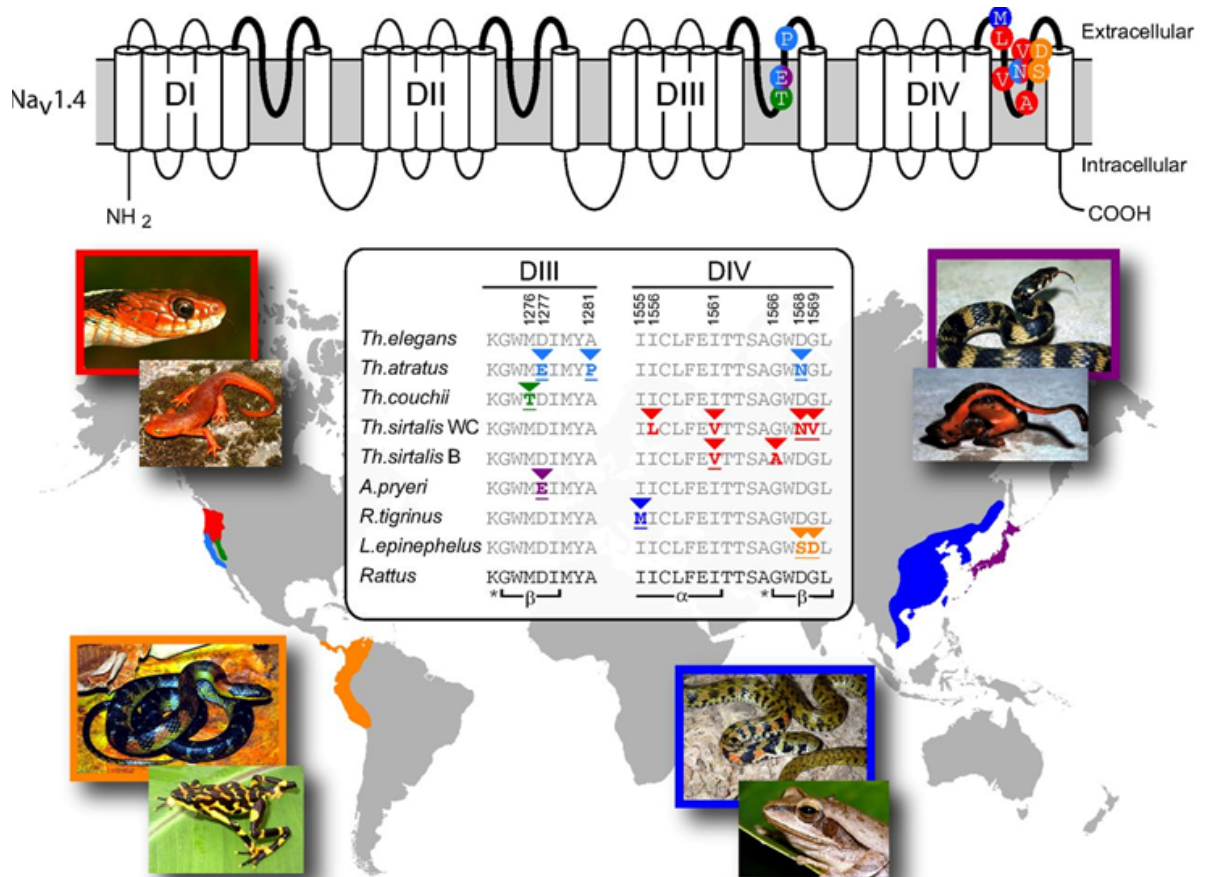
La figure ci-dessous présente la relation entre la toxicité des tritons et la résistance des serpents pour 4 populations de l'ouest des États-Unis.



La résistance des serpents est estimée par la dose de TTX nécessaire pour réduire leur vitesse de déplacement de 50%, exprimée en MAMU (a Mass-Adjusted Mouse Unit [MAMU] is the amount of TTX required to kill a 20-g mouse in 10 min.)

10) Comment pouvez-vous expliquer la relation observée entre la toxicité des tritons et la résistance des serpents ?

11) La TTX est présente dans la peau d'assez nombreuses espèces d'amphibiens. Les seuls prédateurs capables de consommer ces amphibiens toxiques sont certaines espèces de serpent. 6 serpents résistant ont été étudiés. La séquence protéique des canaux sodiques de ces serpents, plus celle de la couleuvre rayée, est représentée ci-dessous. Que pouvez-vous déduire de ce document sur les contraintes existant pour l'évolution de la résistance à la TTX ?



Chris R. Feldman et al. PNAS 2012;109:4556-4561

TD 5 : Evolution de l'émail et des protéines de l'émail

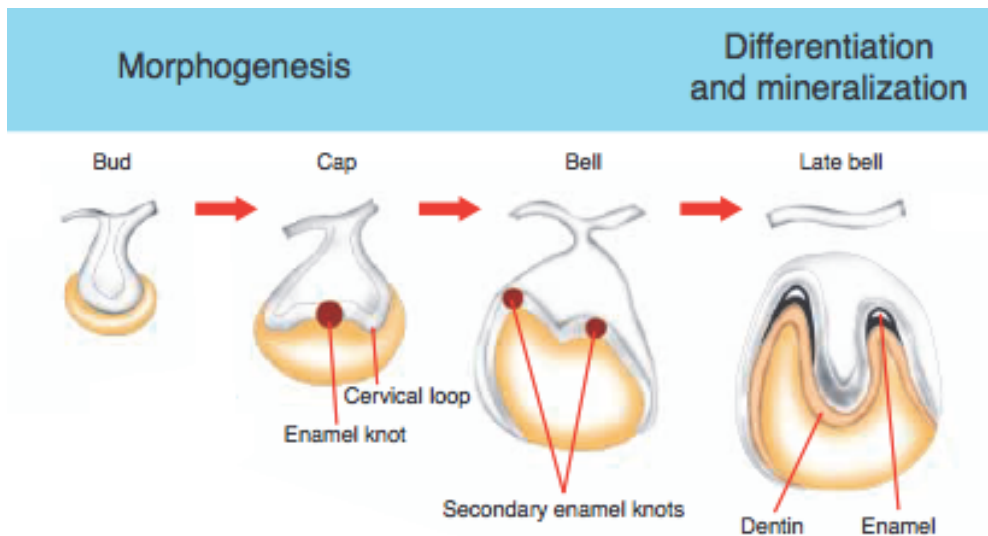
Modifié d'après Gasse, B., 2015. Les phosphoprotéines sécrétées liant le calcium (SCPP) impliquées dans la formation de l'émail dentaire : expression chez le lézard *Anolis carolinensis* et évolution chez les amniotes, Ecole doctorale Complexité du vivant, UMR7138 Evolution Paris-Seine, Equipe Evolution et Développement du Squelette. Université Pierre et Marie Curie, Paris, France, p. 186.

Le recrutement des dents sur les mâchoires des premiers gnathostomes est une innovation majeure dans l'évolution des vertébrés, notamment grâce à leur rôle fondamental dans la prédation. Elles sont composées d'une cavité pulpaire entourée de dentine, recouverte, sauf exceptions, d'un tissu protecteur hyperminéralisé, généralement de l'émailloïde chez les chondrichthyens et les actinoptérygiens, ou de l'émail chez les sarcoptérygiens (Smith, 1989; Wakita, 1993; Donoghue & Sansom, 2002; Donoghue et al., 2006).



Un fossile de placoderme du carbonifère

Le développement dentaire se divise en quatre stades mettant en jeu des interactions épithélio-mésenchymateuses : bourgeon, capuchon, cloche et couronne (Fig. 1) (Thesleff, 2003). C'est lors de cette dernière étape que se déroule la formation de l'émail, ou l'amélogénèse, elle-même généralement répartie en deux phases principales : une phase de sécrétion, puis une phase de maturation (Moradian-Oldak, 2012).



C'est au contact de la première couche de dentine synthétisée par les odontoblastes, que les cellules de l'épithélium dentaire interne se différencient en améloblastes et commencent la synthèse de la matrice de l'émail : c'est la phase de sécrétion. La matrice de l'émail est principalement composée de trois protéines : amélogénine (AMEL), améloblastine (AMBN) et énaméline (ENAM). Ces protéines ont la capacité d'initier la formation de cristaux de phosphate de calcium et d'en contrôler la forme et sont clivées, peu après leur sécrétion, par une métalloprotéinase, la MMP20. A la fin de l'étape de sécrétion, l'émail immature a atteint son épaisseur finale, mais n'est minéralisé qu'à environ 30% (Moradian-Oldak, 2012). Commence alors la phase de maturation de l'émail pendant laquelle 50% des améloblastes disparaissent par apoptose et les 50% restant se raccourcissent et s'élargissent (Smith &

Warshawsky, 1977). Une lame basale contenant d'autres protéines secrétées par les améloblastes : Odontogenic Ameloblast-associated protein (ODAM), Secretory Calcium-binding PhosphoProtein rich in Proline et Glutamine 1 (SCPP-PQ1) et amélotine (AMTN), est synthétisée entre la surface de l'émail et les améloblastes (Ganss & Abbarin, 2014). Une seconde protéase, la kallikréine 4 (KLK4), est synthétisée par les améloblastes et dégrade les fragments protéiques restant dans la matrice amélaire, permettant ainsi la croissance en largeur des cristaux. A la fin de la maturation, l'émail mature est composé de 96% d'hydroxyapatite, de 3,2% d'eau et de seulement 0,8% de matière organique (Smith, 1998).

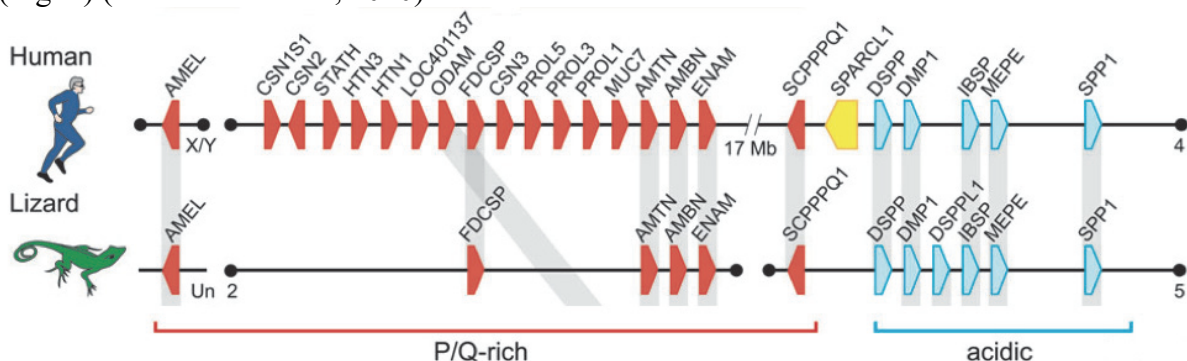
Quel avantage peut présenter ce tissu par rapport à d'autres tissus dans un contexte de biologie évolutive ?

Les protéines de la matrice de l'émail secrétées pendant l'amélogenèse (AMEL, AMBN et ENAM) font partie de protéines impliquées dans la minéralisation de l'os ou de la dent, les phosphoprotéines secrétées liant le calcium (SCPP, pour Secretory calcium-binding PhosphoProtein) (Kawasaki & Weiss, 2003).

Que peut-on dire de l'origine évolutive de cette famille de protéine ?

Chez les mammifères les gènes des SCPP, à l'exception du gène AMEL, résident tous sur un même chromosome, et sont distribués en deux clusters de part et d'autre de SPARC-L1 (Fig. 4). La protéine SPARC-like 1 interagit avec la matrice extracellulaire et est impliquée dans les métastases cancéreuses. SPARC-L1 est une protéine sécrétée présentant une grande similitude structurelle avec SPARC qui est une glycoprotéine présente dans l'os qui fixe le calcium et présente également une affinité pour le collagène. Elle est sécrétée par les ostéoblastes pendant la formation de l'os, initiant la minéralisation et favorisant la formation de cristaux minéraux.

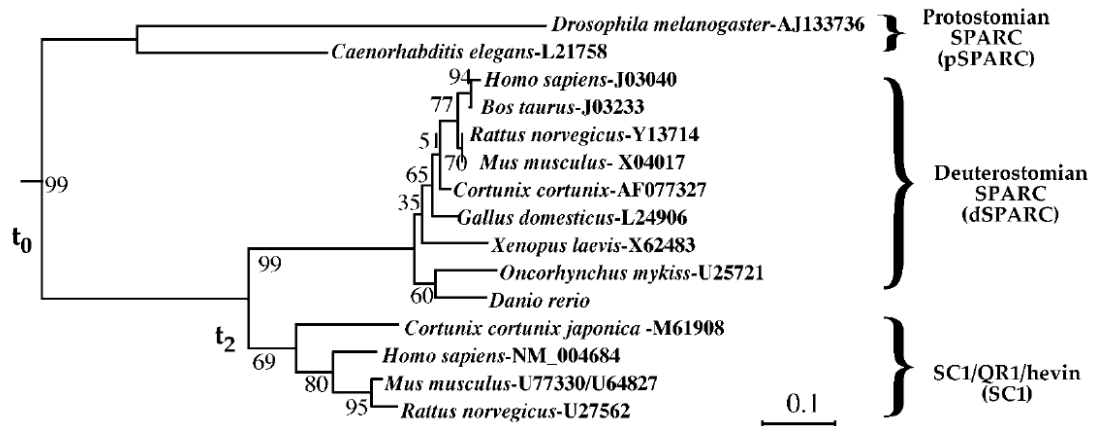
Ces deux clusters regroupent, d'une part, les gènes codant des protéines riches en acides aminés acides ("acid-rich SCPP cluster"), impliquées dans la minéralisation de l'os et de la dentine, et, d'autre part, des gènes codant des protéines riches en proline et glutamine ("P/Q-rich SCPP cluster"), comprenant les protéines impliquées dans la minéralisation de l'émail, les caséines du lait et des protéines salivaires (Kawasaki & Weiss, 2003, 2006). Chez les sauropsides (reptiles et oiseaux), les deux clusters résident sur des chromosomes différents (Fig. 4) (Al-Hashimi et al., 2010).



Que nous apporte cette information ?

Bien que les protéines appartenant à cette famille ne présentent pas de similitude de séquences évidentes, elles possèdent un peptide signal qui permet la sécrétion de la protéine. De plus, elles possèdent un ou plusieurs sites de phosphorylation SXE (Ser-Xaa-Glu où Xaa représente n'importe quel acide aminé), l'un étant codé par la région 3' de l'exon 3. Ces protéines lient le calcium via leurs résidus phosphorylés.

Ces caractéristiques permettent-elle de définir une homologie ?



Arbre génique de SPARC et de gènes apparentés à SPARC produit à partir des séquences d'acides aminés. Les distances de Gonnet ont été utilisées avec la méthode de construction de l'arbre par neighbor-joining. Les bootstraps sont indiqués aux nœuds correspondants en pourcentage.

La matrice de Gonnet est construite en comparant de très nombreuses protéines homologues et en dénombrant le nombre de fois où un acide aminé A est remplacé par un acide aminé B. Plus cet évènement est fréquent, plus on considère que les fonctions de A et B dans une protéine sont interchangeable.

Gonnet, G. H., Cohen, M. A. & Benner, S. A. 1992. Exhaustive matching of the entire protein sequence database. *Science*, 256, 1443-1445.

Quelle information nous apporte cette phylogénie ? Quels évènements évolutifs se sont déroulés aux dates t₀, t₁ et t₂ ? Selon vous, quelle méthode a été utilisée pour enraciner l'arbre ? Les conclusions dépendent-elles de cet enracinement ?

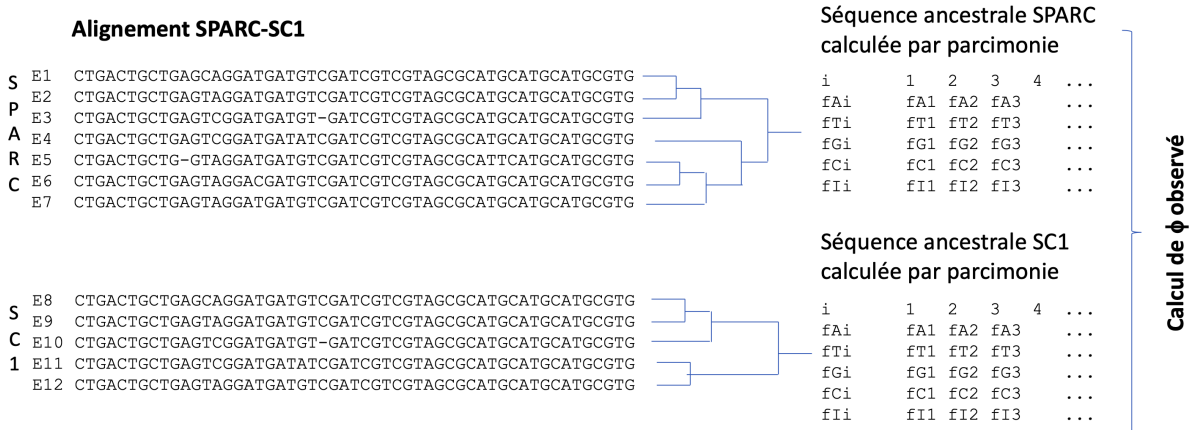
Il était important de s'assurer que les séquences dSPARC, SC1 et amélogénine partageaient un ancêtre commun, c'est-à-dire que les ressemblances observées reflétaient une réelle homologie et non une convergence évolutive. Pour se faire, la ressemblance de l'exon 2 de dSPARC, SC1 et amélogénine a été analysée.

Les proportions d'identité des différentes paires de séquences ne peuvent pas être directement comparées car ces valeurs ne sont pas indépendantes. En effet, nous devons prendre en compte que toutes les amélogénines, d'une part, et dSPARC et SC1, d'autre part, ont une histoire phylogénétique commune.

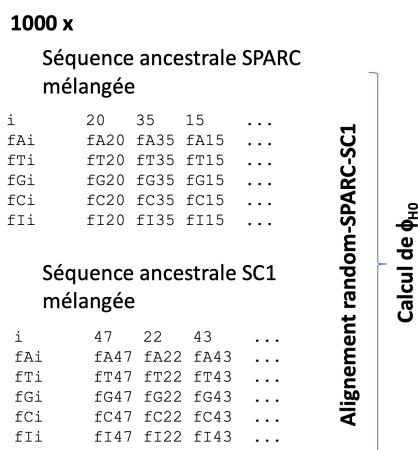
La séquence ancestrale la plus parcimonieuse pour le second exon de chaque gène a été calculée séparément à l'aide de MacClade 3.07 (Maddison and Maddison 1992). En raison des ambiguïtés de la séquence ancestrale, la position i dans une séquence j est entièrement décrite par les cinq fréquences suivantes f_{Aij}, f_{Gij}, f_{Cij}, f_{Tij}, et f_{Iij}, qui sont, respectivement, les probabilités que soit un A, un G, un C, un T, ou I un indel est présent à cette position pour

cette séquence. Ensuite, la probabilité moyenne que deux bases identiques soient présentes au hasard dans un alignement de l bases est estimée comme suit

$$\Phi = \frac{\sum_{i=1}^l (f_{Ai1} \cdot f_{Ai2} + f_{Ti1} \cdot f_{Ti2} + f_{Gi1} \cdot f_{Gi2} + f_{Ci1} \cdot f_{Ci2})}{\left(l - \sum_{i=1}^l (f_{Ii1} \cdot f_{Ii2}) \right)}$$



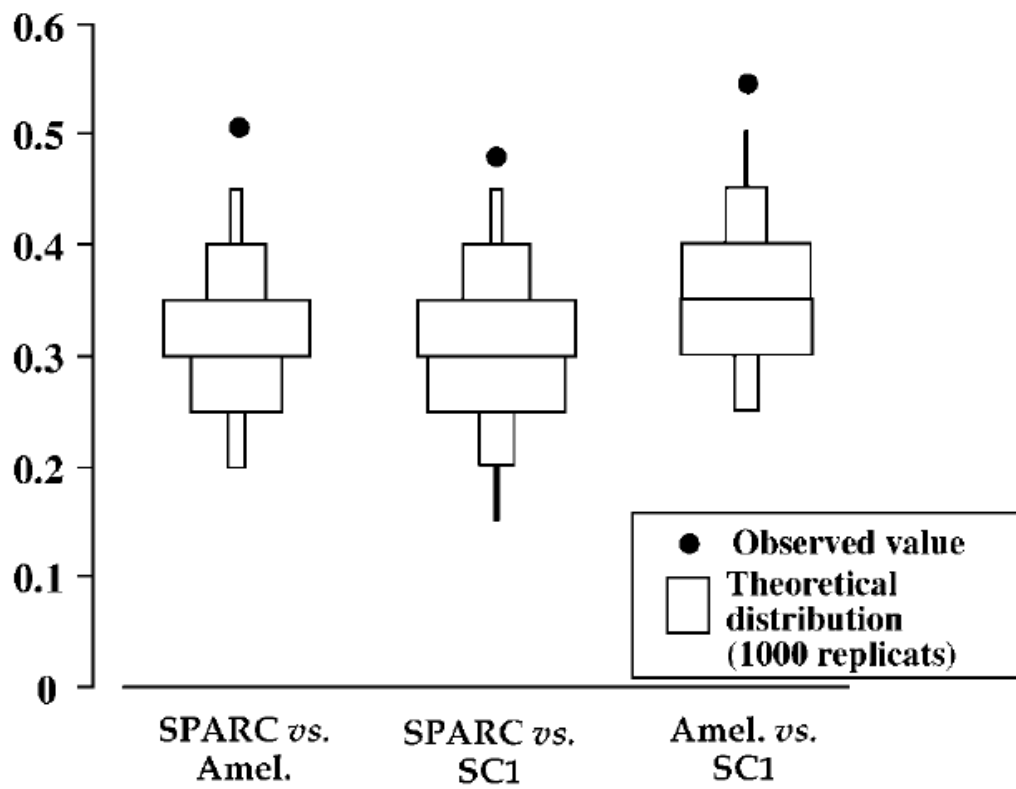
Pour chaque paire, la distribution de Φ sous l'hypothèse nulle (c'est-à-dire que la valeur observée Φ n'est pas significativement différente de ce qui est attendu pour des séquences aléatoires) a été établie en utilisant 1 000 répliqués de deux séquences aléatoires générées par ordinateur partageant les mêmes caractéristiques que les séquences analysées (même longueur, même niveau d'ambiguïté et même composition en bases). Une simple permutation de position a été utilisée pour générer ces séquences aléatoires. Chaque paire de séquences générées par ordinateur a été alignée pour maximiser la valeur estimée. La valeur correspondante a été utilisée pour estimer la distribution de Φ sous H_0 (Φ_{H_0}).



Ce test a été effectué avec les seules bases 1 et 2 afin de ne pas tenir compte des substitutions synonymes qui se produisent principalement à la troisième position de chaque codon. La figure montrée ci-dessous représente la distribution des valeurs de Φ_{H_0} pour 1000 répliqués comparée avec la valeur observée (point noir).

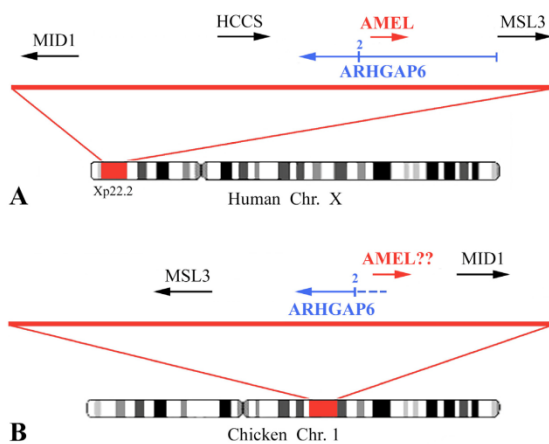
Décrire l'intérêt de chaque étape.

Probability of identical bases in alignment



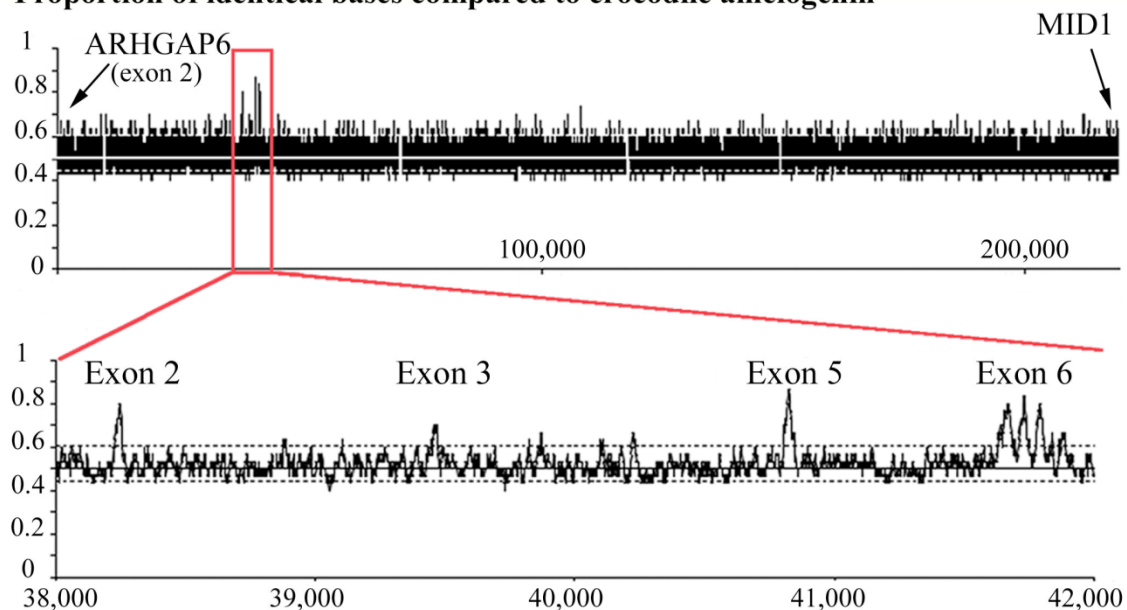
Que peut-on conclure ?

Le gène de l'amélogénine réside, sur le brin antisens, dans un intron du gène ARHGAP6 (ARHGAP6 Rho GTPase activating protein 6) chez tous les sarcoptérygiens (Kawasaki & Amemiya, 2014). Il code pour une protéine activatrice des GTPases de type Rho en les convertissant à un état inactif lié au GDP. Pourrait réguler les interactions des molécules de signalisation avec le cytosquelette d'actine. Favorise l'allongement continu des processus cytoplasmiques pendant la motilité cellulaire et la rétraction simultanée du corps cellulaire, modifiant la morphologie de la cellule (en fait on ne sait pas trop à qui il sert !). Chez les mammifères placentaires, il est localisé sur les chromosomes sexuels, tandis que chez les marsupiaux, les monotrèmes et les non-mammaliens il est autosomal (Sire et al., 2005).



Le gène de l'amélogénine n'avait pas été détecté chez le poulet malgré le séquençage complet du génome. Une recherche par une technique de recherche de similitude sur des fenêtres de petite taille a permis de détecter chez le poulet une zone présentant des similitudes avec la séquence d'amélogénine chez le crocodile.

Proportion of identical bases compared to crocodile amelogenin



Quel est l'intérêt de rechercher le gène de l'amélogénine chez le poulet ?

```

Crocodile  ATGGAGGGCTGG--ATG--TTGATCACTTGCCTACTAGGTGCAACATTTGCTATACCA
Ψ-Chicken ATGGAGGACAGACTATTATTGACTGCTTGCCTCCTAGGAGCACTGTTTGTATGCCA
***** * *   * *   * *   * *   * *   * *   * *   * *   * *
Ψ-Chicken  MEDRLFIDCLPPRSTVCYA

Crocodile  MEGWMLITCLLGATFAIP
"Chicken"  MEDRILTACLLGALFAMP
* *   * *   * *   * *

```

Chicken Ψ-AMEL exon 2 analysis. Top: Alignment of chicken and crocodile AMEL exon 2 sequences. Four nucleotides are inserted (red) in chicken Ψ-AMEL exon 2 (signal peptide), leading to a shift in the reading frame. Middle: Putative deduced amino acid sequence from chicken Ψ-AMEL exon 2. Bottom: The four inserted codons were removed from the Ψ-AMEL sequence, which was translated and aligned to the crocodile sequence; both amino acid sequences are highly similar.

Que conclure sur la fonction de la protéine Amélogénine ?

TD 6 : Evolution humaine

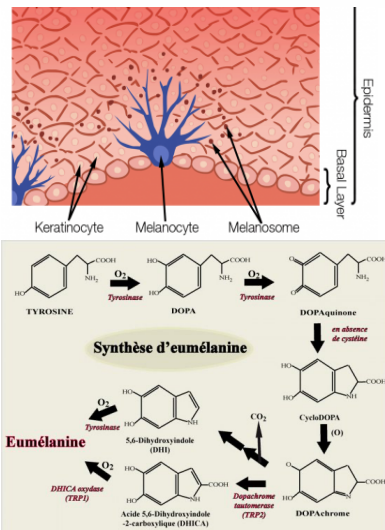
A) Évolution de la pigmentation chez l'homme



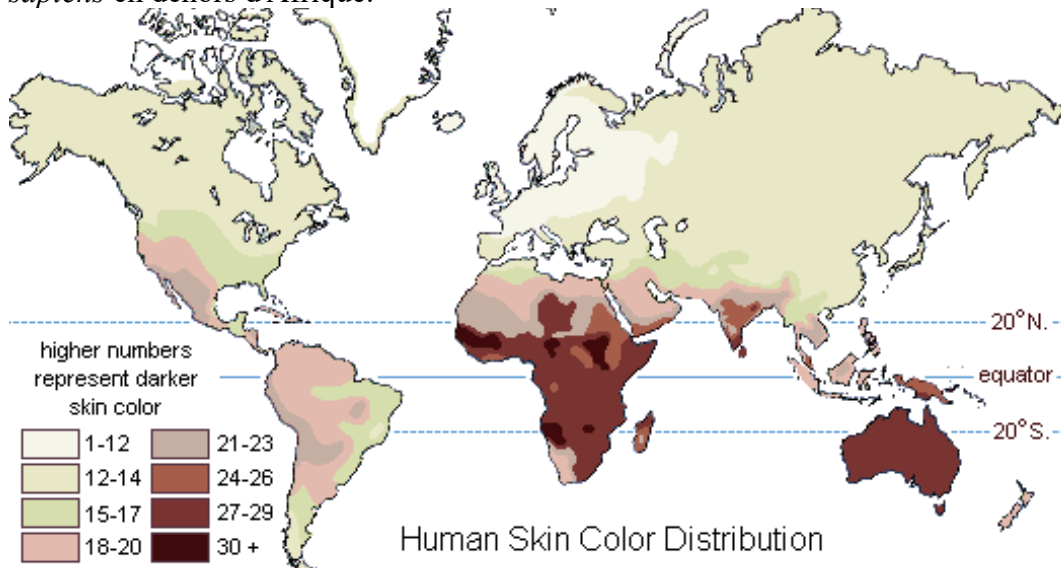
(b) Humans show a wide range of skin tones, from almost white to very dark brown

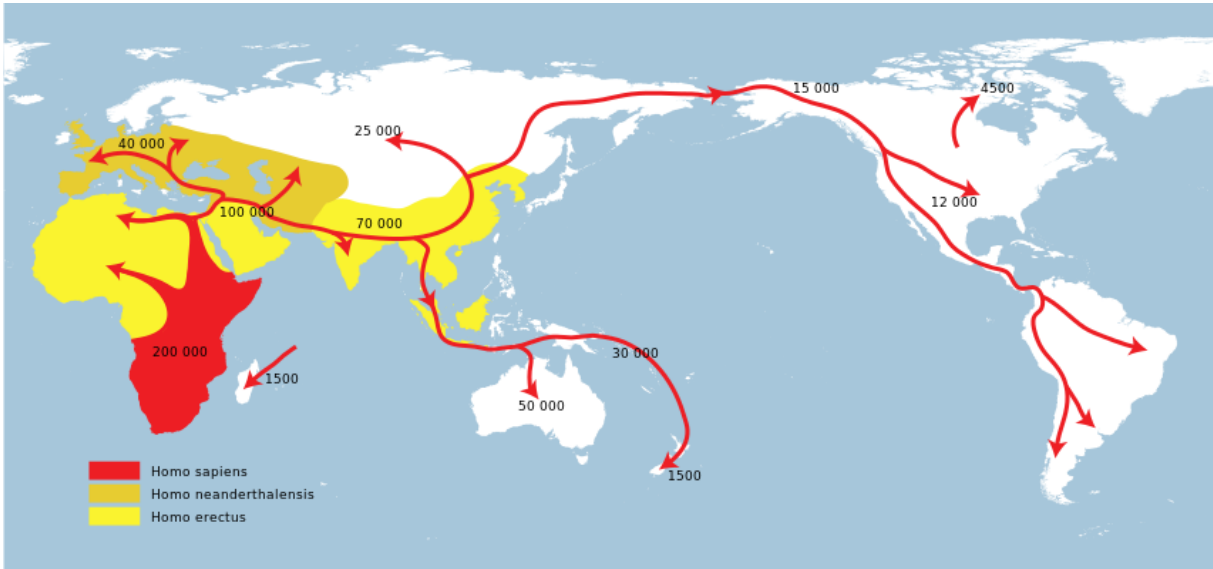
© 2011 Pearson Education, Inc.

Les mélanocytes sont les cellules qui pigmentent la peau, les poils ou les plumes des vertébrés. Ils produisent des mélanosomes, des granules pigmentés riches en mélanines qui chez l'homme sont évacués en direction des kératinocytes. Les êtres humains présentent une variation importante de l'intensité de la pigmentation de la peau, due à des différences de nombre, de taille et de densité en pigment des mélanosomes présents dans la peau. Le nombre de mélanocytes varie, lui, assez peu. Chez l'homme, la variation de pigmentation de la peau est sous le contrôle d'une dizaine de gènes environ.

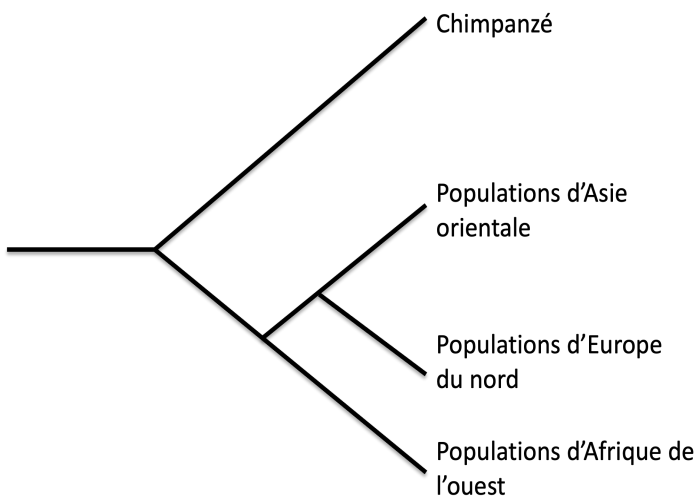


Les cartes ci-dessous présentent la distribution de la coloration de la peau observée historiquement dans notre espèce et une datation approximative des migrations d'*Homo sapiens* en dehors d'Afrique.





1) En vous aidant des deux cartes précédentes, sur la figure ci-dessous, reconstruisez par parcimonie la pigmentation ancestrale et sur quelle(s) branche(s) de l'arbre ont eu lieu des modifications de la pigmentation.



2) Quelles pressions de sélection peuvent avoir causé les modifications de pigmentation observées dans l'histoire évolutive de notre espèce ?

3) Pourquoi certaines populations vivant dans les latitudes élevées ont-elles la peau claire ? Proposez deux hypothèses reposant chacune sur une force évolutive différente. Comment pourriez-vous tester quelle hypothèse est correcte ?

La vitamine D intervient dans l'absorption du calcium et du phosphore par les intestins, ainsi que dans leur réabsorption par les reins. Une carence modérée en vitamine D cause faiblesse, douleurs musculaires et fatigue. Une carence sévère peut causer le rachitisme chez les enfants et une décalcification osseuse chez les adultes. Environ 90% de la vitamine D est synthétisée dans la peau et les reins à partir d'un précurseur proche du cholestérol sous l'effet des UV. Les 10% restant proviennent de la nourriture (poissons gras, jaune d'œuf).

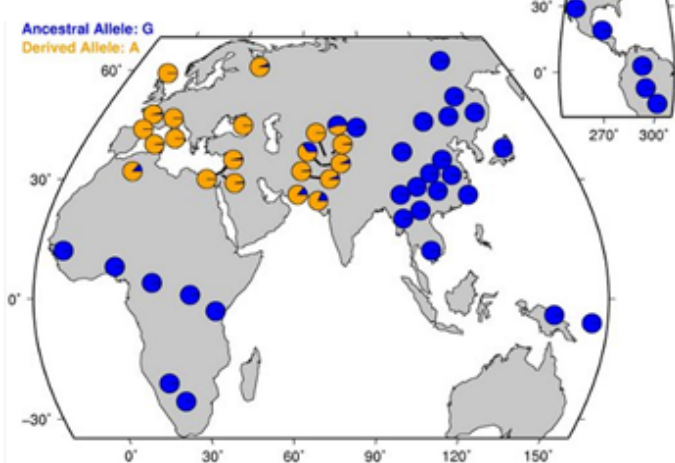


4) En Europe, on a assisté à une multiplication des cas de rachitisme aux XVIII^{ème} et XIX^{ème} siècles, d'après vous pourquoi ?

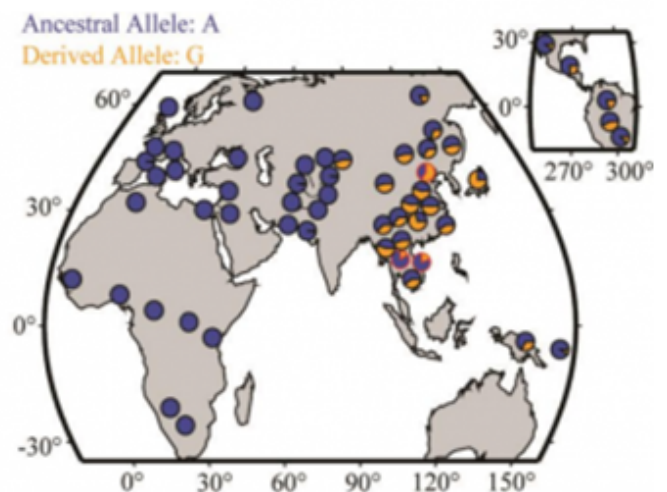
Les cartes ci-dessous vous présentent pour deux des gènes impliqués dans la variation de pigmentation chez l'homme, *SLC24A5* et *OCA2* la fréquence des allèles ancestraux et des allèles dérivés (l'allèle dérivé de ces gènes étant associé à une pigmentation plus claire) parmi des populations africaines, européennes et asiatiques.

6) Ces cartes vont-elles plutôt dans le sens d'une convergence ou d'une homologie pour la pigmentation claire observée à la fois chez certaines populations européennes et est-asiatique ? Etant donnée la carte des migrations de *Homo sapiens* en dehors d'Afrique, cette réponse était-elle attendue ?

Gène *SLC24A5*



Gène *OCA2*



7) Peut-on penser qu'il y a à l'heure actuelle des discordances entre l'environnement dans lesquels vivent certaines populations et la pigmentation de leur peau ? Quelles conséquences cela peut-il avoir ?

B) Homo sapiens et Néandertal

L'homme de Néandertal est un représentant fossile du genre *Homo* qui a vécu en Europe et en Asie occidentale et centrale au Paléolithique moyen, entre environ 250 000 et 28 000 ans avant le présent. La position phylogénétique exacte de l'homme de Néandertal a provoqué de nombreux débats. Certains le regardent comme une sous-espèce au sein de l'espèce *Homo sapiens* et le nomment donc *Homo sapiens neanderthalensis* ; d'autres considèrent qu'il constitue une espèce indépendante et le nomment *Homo neanderthalensis*. Il ne s'agit pas d'un simple problème de classification, il s'agit aussi de savoir si l'Homme de Neandertal représente une lignée parallèle et éteinte (espèce indépendante), ou bien s'il a pu contribuer en partie au patrimoine génétique de l'homme actuel.

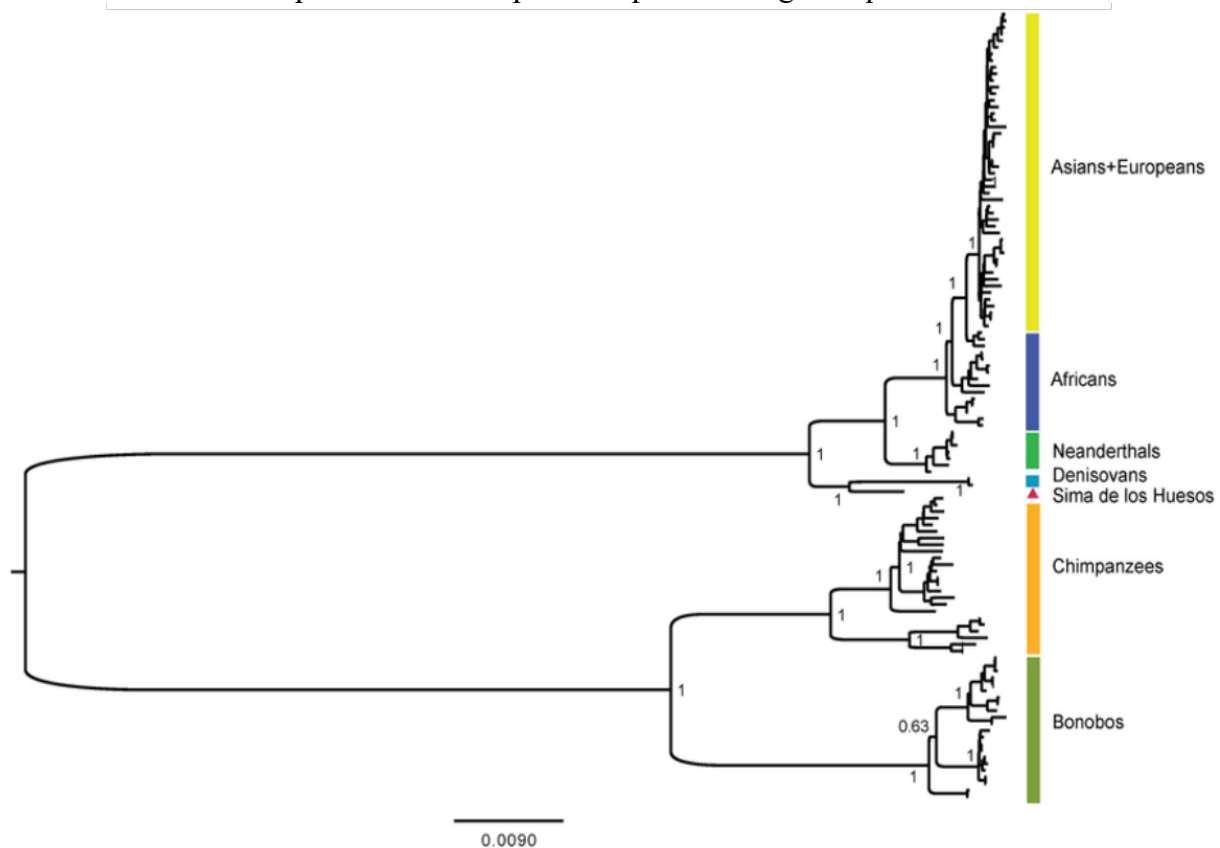


Reconstitution d'une femme de Néandertal à partir de fossiles et d'analyses ADN

8) Quelles données ont pendant longtemps été les seules disponibles pour répondre à cette question de la position phylogénétique de Neandertal ?

Au début des années 2000, les progrès techniques en matière d'analyse d'ADN ancien ont permis d'obtenir des séquences d'ADN mitochondrial d'hommes de Neandertal et de construire des phylogénies.

9) D'après vous, pourquoi seules des données d'ADN mitochondrial étaient-elles disponibles ? Commentez la figure ci-dessous. Quel élément de réponse apporte-t-elle à la question de savoir si Neandertal a pu contribuer en partie au patrimoine génétique de l'homme actuel ?



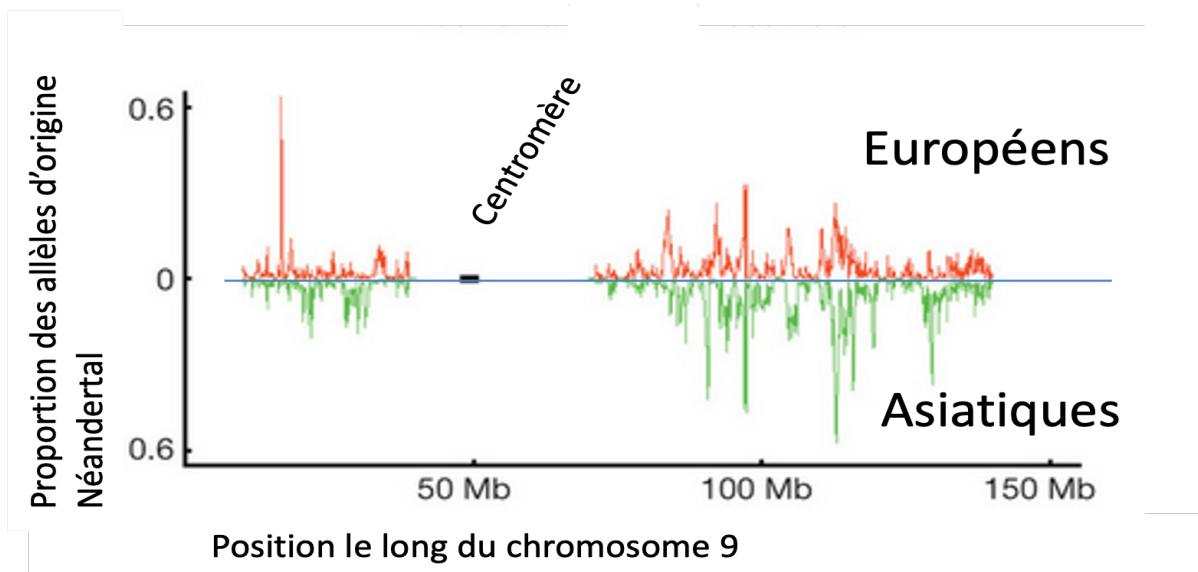
Phylogénie reconstruite à partir de données de séquences mitochondriales.

En 2010, un nouveau progrès des techniques de séquençage de l'ADN ancien a permis d'obtenir des données de séquences nucléaires d'homme de Neandertal. Ces nouvelles données ont permis de montrer que toutes les populations humaines actuelles à l'exception des populations africaines présentent un petit pourcentage (entre 1 et 4%) d'ADN de Neandertal dans leur génome.

10) Quel scénario pouvez-vous proposer pour expliquer à la fois les données mitochondriales et les données nucléaires ?

11) Quelles sont les conséquences possibles de la présence d'allèles provenant de d'homme de Neandertal dans le patrimoine de *H. sapiens* ?

Dans les populations non africaines, en moyenne un à quelques pourcents des individus présentent des allèles provenant de l'homme de Neandertal. Mais cette proportion est très variable le long du génome : elle peut aller de 0 à plus de 60%. La figure ci-dessous vous présente la proportion d'allèles provenant de l'homme de Neandertal en fonction de la position le long du chromosome 9 pour une population européenne et une population asiatique.



Une analyse de ce type effectuée sur l'ensemble du génome a montré qu'une partie importante du génome de notre espèce présente une proportion nulle d'allèles provenant de Neandertal. C'est le cas notamment de beaucoup de régions contenant des gènes et de beaucoup de régions du chromosome X. Certains locus comme ceux impliqués dans la synthèse de la kératine et la coloration de la peau présentent au contraire des fréquences élevées d'allèles provenant de Neandertal.

12) Proposez une explication pour ces observations.

Pour aller plus loin :



<https://www.arte.tv/fr/videos/043200-000-A/le-neandertal-en-nous/>