

I- Introduction :

La vie a une histoire qui débute il y a environ 3,7 milliards d'années. Depuis l'apparition des premiers organismes unicellulaires, des processus évolutifs ont permis une grande diversité du monde vivant.

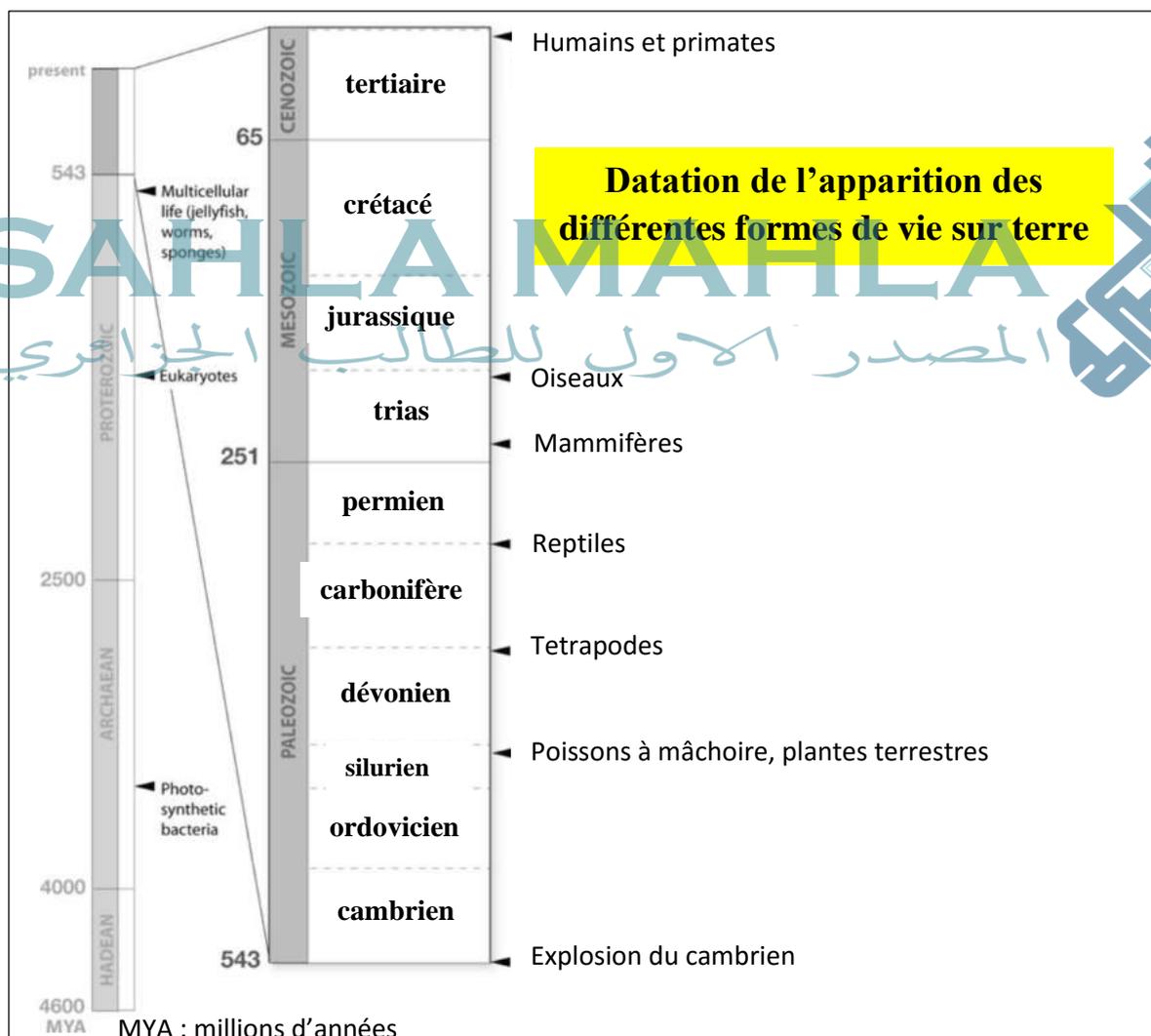
L'évolution se réfère au changement dans le temps à mesure que les espèces se modifient et divergent pour produire des descendants multiples. L'évolution et la sélection naturelle sont souvent confondues, mais l'évolution est l'évènement historique du changement, et la sélection naturelle est un mécanisme - dans la plupart des cas le plus important - qui peut le provoquer.

II- Preuves de l'évolution :

Les preuves évidentes de la réalité de l'évolution reposent sur un grand nombre d'observations que l'on peut résumer en six thèmes.

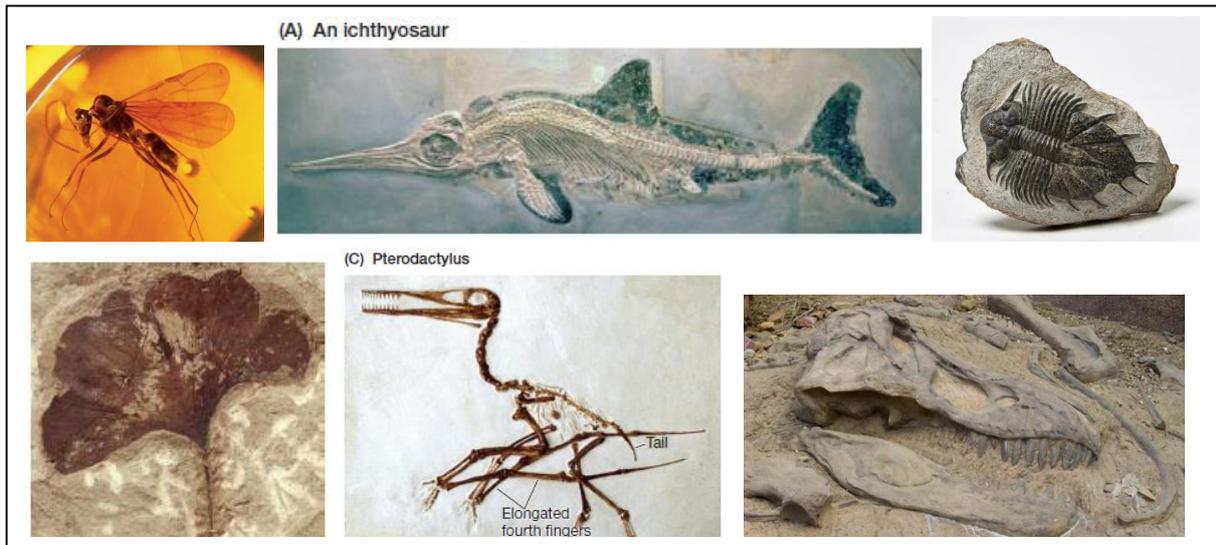
1- Paléontologie (les fossiles) :

L'évidence la plus directe de l'évolution provient des fossiles incrustés dans les couches géologiques, dans lesquelles les changements dynamiques de la vie au fil du temps sont enregistrés, y compris de nombreuses formes de transitions entre grands taxons.



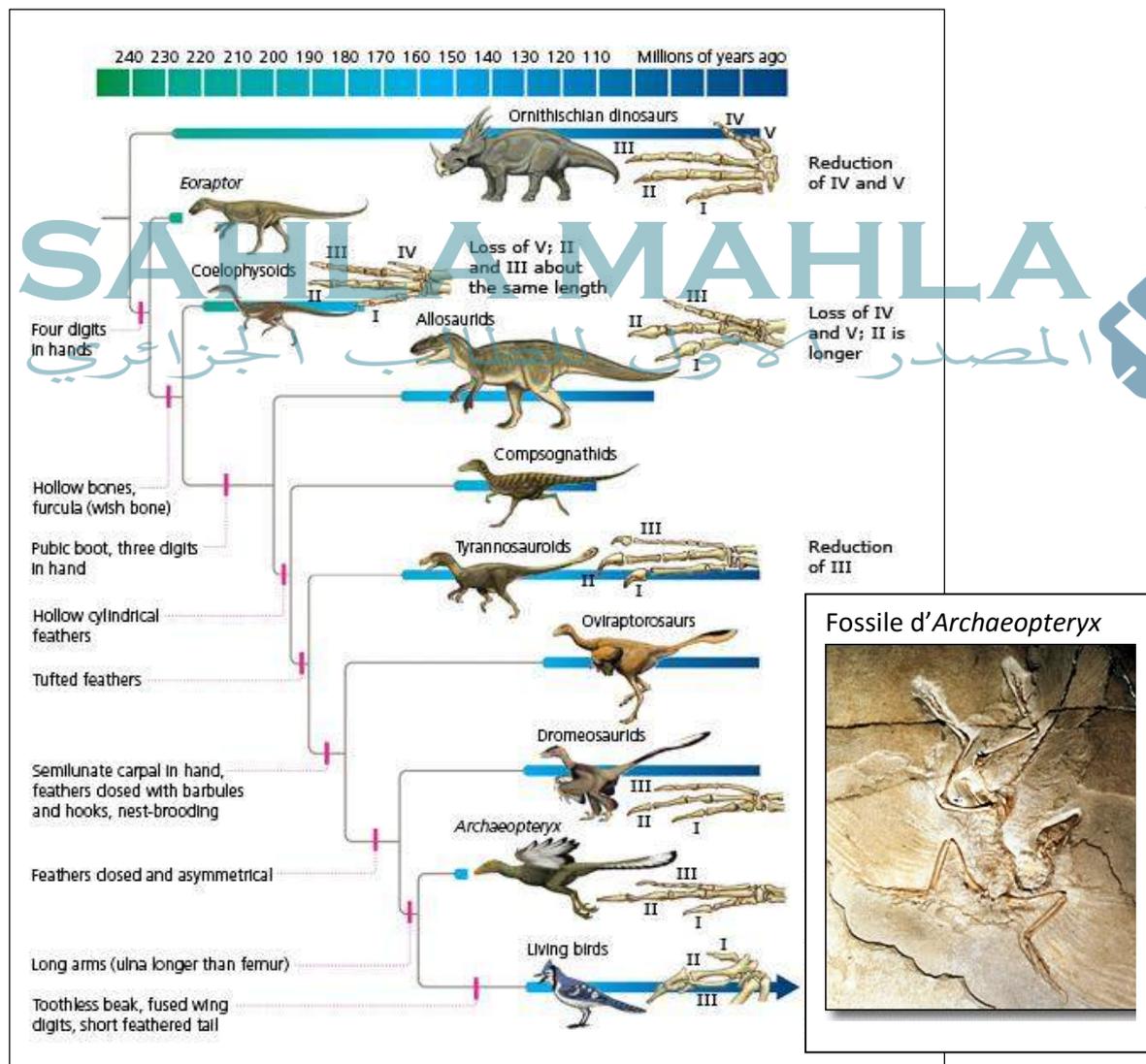
Introduction

Insecte fossilisé dans l'ambre, végétaux et animaux fossilisés dans la roche :



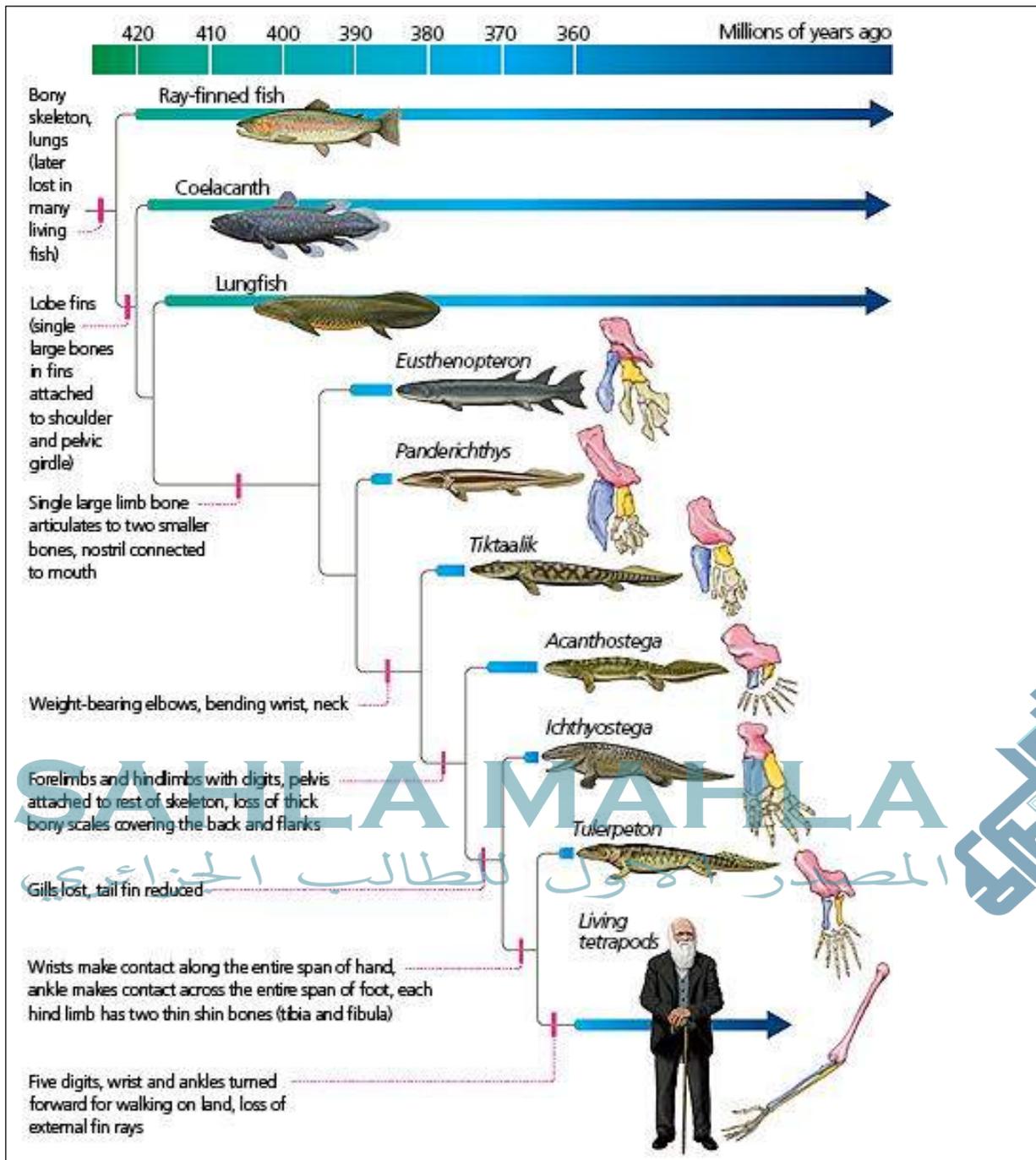
Exemple de formes de transition entre taxons:

- *Archeopteryx* : transition entre les reptiles (particulièrement les dinosaures) et les oiseaux



Introduction

- Tiktaalik : Forme de transition entre les poissons et les tetrapodes (animaux terrestres)



Fossile de *Tiktaalik roseae* trouvé dans la couche géologique correspondant au dévonien, datant de -375 millions d'années. C'est la forme de transition entre le poisson et le tétrapode.

2- Anatomie comparée :

Les comparaisons entre les organismes, vivants ou disparus, montrent que le lien entre les différentes formes à des moments différents est généalogique. Ces similitudes, que l'on trouve dans la morphologie, le développement et même le génome, sont des homologies, c'est-à-dire des similitudes héritées d'un ancêtre commun

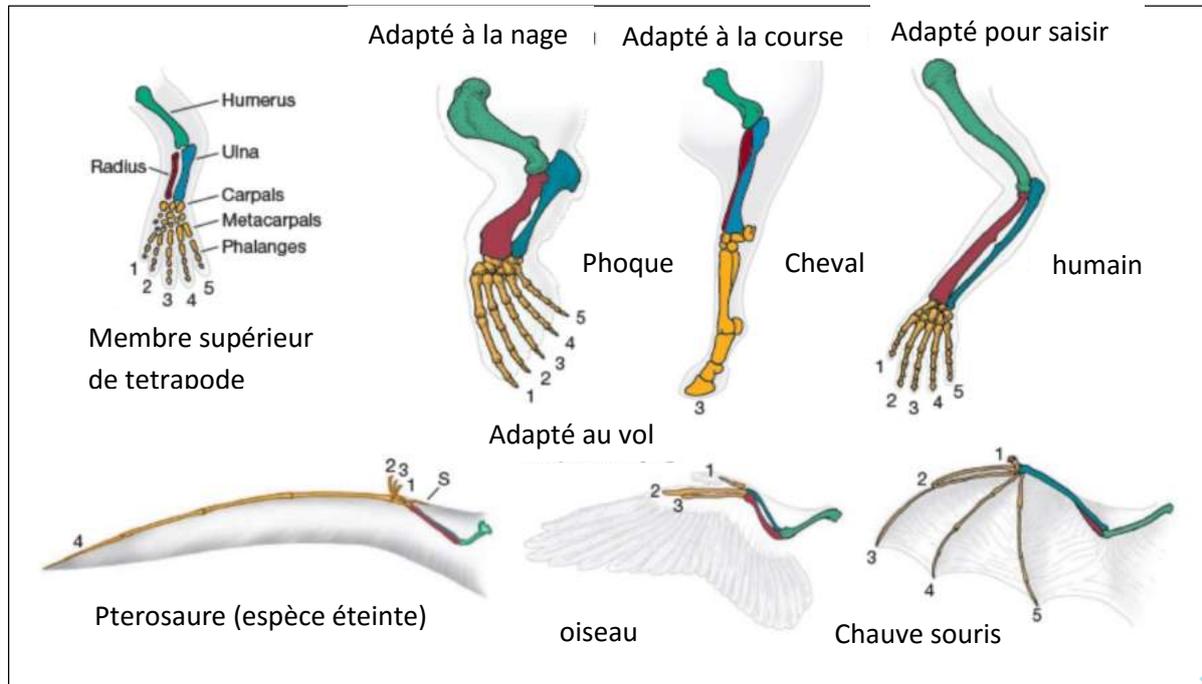


Figure : Squelettes des membres antérieurs de certains vertébrés tétrapodes

Par rapport à « l'architecture » du membre supérieur du tétrapode à ses débuts, les os ont été perdus ou fusionnés (par exemple, cheval, oiseau) ou modifiés en taille et en forme relatives. Les modifications pour la natation ont évolué chez le phoque, pour courir chez le cheval (horse), pour saisir chez l'homme, et pour voler chez l'oiseau, la chauve-souris et le ptérosaure. Tous les os représentés sont homologues parmi ces organismes, à l'exception de l'os sésamoïde (S) du ptérosaure; cet os a une origine développementale différente du reste du squelette du membre.

3- Embryologie comparée :

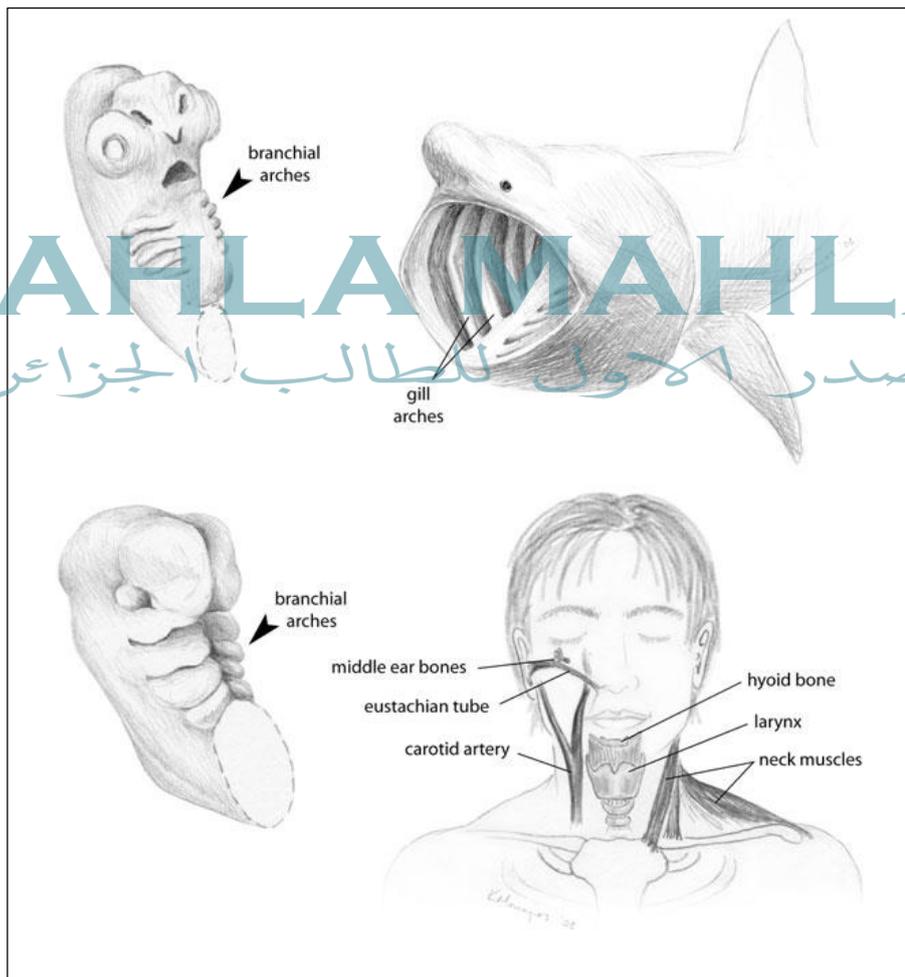
Durant leur développement à partir d'embryons, de nombreuses espèces passent par différentes formes qui changent d'aspect: des organes et d'autres traits apparaissent, puis changent radicalement ou même disparaissent complètement avant la naissance. Les embryons présentent une architecture commune, puis les espèces finissent par se différencier. Parmi les similitudes observables chez les embryons, tous les vertébrés possèdent des arcs branchiaux au cours de leur stade embryonnaire.

Introduction



Embryons d'humains, d'alligators et de souris. Il est difficile de les différencier de par leur ressemblance anatomique.

La figure du dessous montre les Arcs branchiaux d'un embryon de requin (en haut à gauche) et d'un embryon humain (en bas à gauche). Chez les requins et les poissons, les arcs se développent directement dans les structures branchiales adultes, tandis que chez l'homme (et d'autres mammifères) ils se développent en structures diverses dans la tête et le haut du corps adulte.

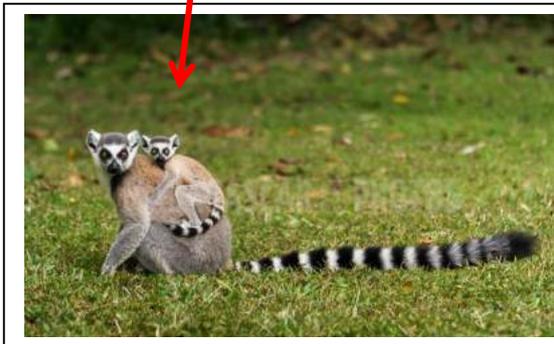


Introduction

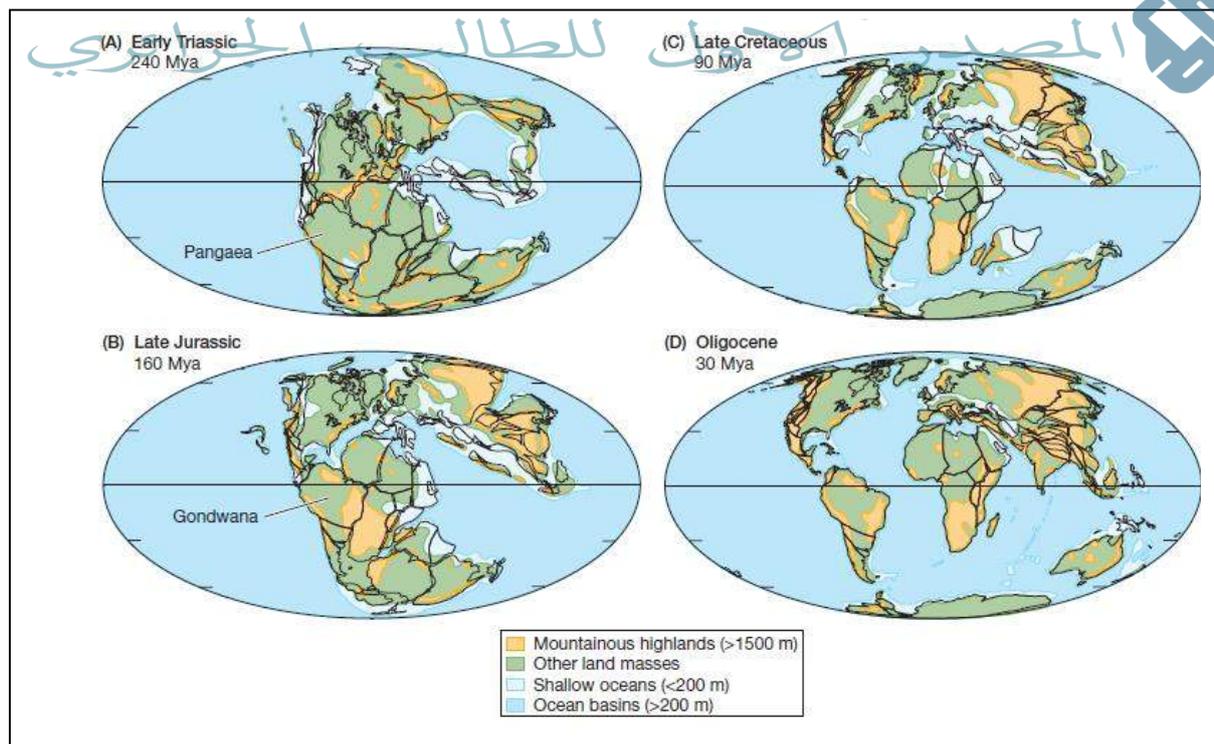
Chez l'homme les tissus du premier arc formeront la mâchoire inférieure et supérieure, 2 os de l'oreille (le marteau et l'enclume) et l'ensemble des vaisseaux et des muscles qui desservent l'oreille. Le deuxième arc formera le troisième osselet de l'oreille (l'étrier) et la plupart des muscles contrôlant l'expression du visage. Le 3^{ème} arc formera les os, les muscles et les nerfs situés plus profondément dans la gorge et qui nous servent à avaler. Le 4^{ème} arc fabrique des éléments du larynx et les muscles et les vaisseaux qui lui assurent son fonctionnement.

4- Biogéographie :

Beaucoup d'espèces proches sont distribuées en divers endroits de la terre, sur différents continents. Par contre dans les Iles, on trouve des espèces endémiques, c'est-à-dire uniques, tels en Australie où on trouve des kangourous ou à Madagascar seul endroit au monde où se trouvent des lémuriers.



La biogéographie s'explique par la dérive des continents



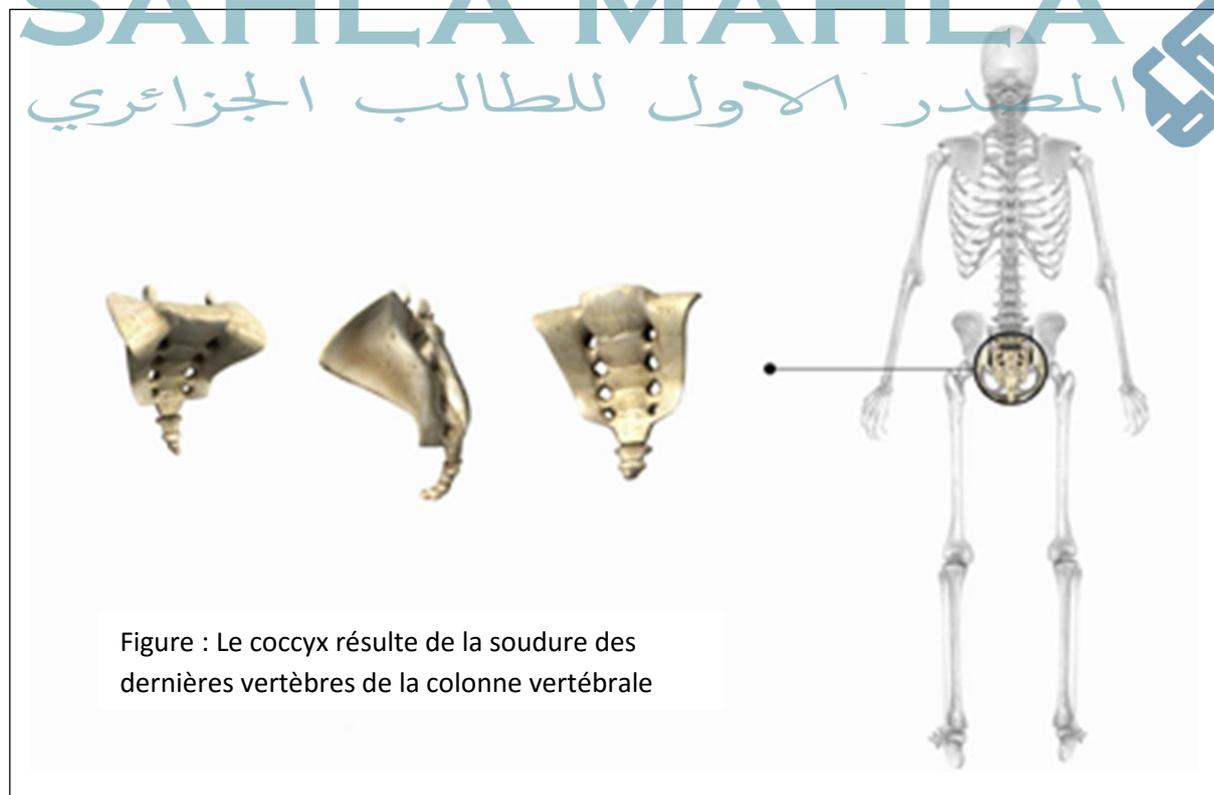
5- Biologie moléculaire et génétique :

Depuis l'avènement de la génétique et le développement des techniques de biologie moléculaire, les séquences d'ADN, de protéines, d'ADN mitochondrial et l'ADN de chloroplastes sont comparées et servent à construire des phylogénies. Elles fournissent des résultats et des analyses robustes qui renseignent plus précisément sur les relations entre espèces.

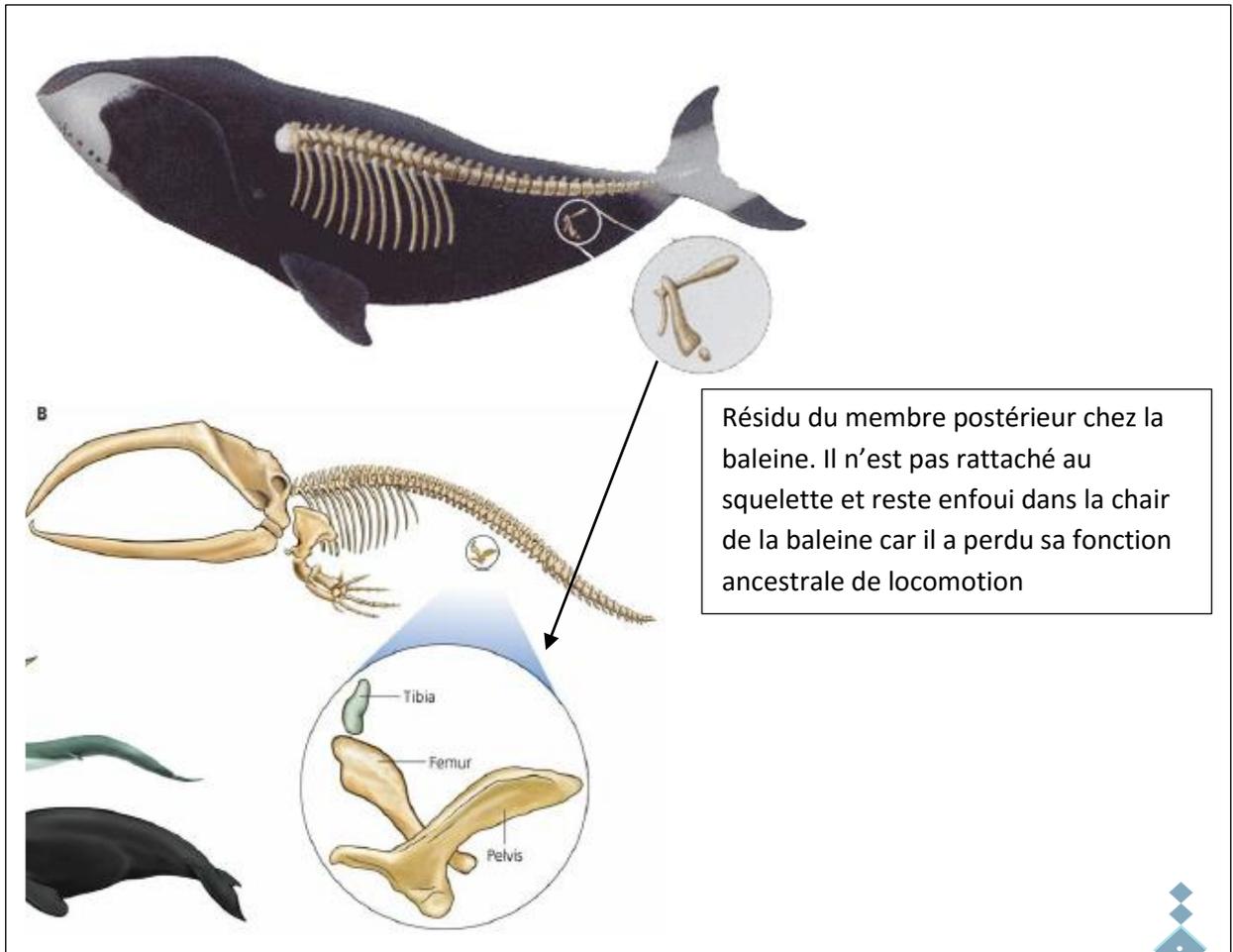
Toutes les espèces actuelles ou passées ont en commun une molécule informative, l'ADN (ARN pour certains virus) dont l'expression décrit toutes les réactions du métabolisme et définit la vie. Cette molécule est capable d'une grande plasticité favorisant l'adaptation et la flexibilité en réponse à des causes externes à l'espèce. Les espèces peuvent ainsi évoluer dans un environnement en constant changement.

6- Structure vestigiales :

Le corps comporte des vestiges de structures anatomiques atrophiées, qui correspondent à des reliquats (des restes) d'organes devenus non fonctionnels. Le corps humain possède de nombreux vestiges de l'ascendance des primates. Par exemple il y a une queue résiduelle: le coccyx, ou l'extrémité triangulaire de la colonne vertébrale, qui est constituée de plusieurs vertèbres fusionnées suspendues sous le bassin. C'est ce qui reste de la longue queue (voir figure ci-dessous). Il a toujours une fonction (certains muscles utiles s'y attachent). Fait révélateur, certains humains ont un muscle de la queue rudimentaire (le «coccygis extenseur»), identique à celui qui déplace la queue des singes et autres mammifères. Il s'attache toujours au coccyx, mais comme les os ne peuvent pas bouger, le muscle est inutile.



Introduction



La baleine ne possède pas de membre postérieur mais des os minuscules enfoncés dans leur corps, qui sont les résidus de la ceinture pelvienne et des pattes postérieures de leur ancêtre terrestre qui marchaient et qui ont disparu.

Les ailes de l'autruche sont un trait résiduel: une caractéristique d'une espèce qui était une adaptation chez ses ancêtres, mais qui a perdu complètement son utilité ou, comme dans l'autruche, a été sélectionnée pour de nouvelles utilisations. Pourtant, les ailes ne sont pas inutiles - elles ont développé de nouvelles fonctions. Elles aident l'oiseau à maintenir son équilibre, à s'accoupler et à menacer ses ennemis.



III- Différentes thèses sur l'évolution du monde vivant :

Le concept d'évolution a longtemps été ignoré puis nié avant de devenir une hypothèse possible et très vraisemblable.

L'avancée des connaissances grâce au développement des sciences et de la technologie a permis une meilleure compréhension du monde. Les thèses qui suivent se sont succéder dans le temps et nous renseignent sur l'évolution de la pensée scientifique :

- **Fixisme**: Les espèces ne changent pas, elles sont fixes tout au long du temps (exemple: le système de classification binomial de Linne (1758).

- **Transformisme** : Les espèces se transforment. Buffon en est l'initiateur (1744). Lamarck (1809) propose que la transformation des espèces entraîne une complexification des formes sous la pression de l'environnement.

- **Catastrophisme** (Cuvier, 1769-1832) : Les catastrophes anéantissent les espèces (fossiles) et sont remplacées par de nouvelles.

- **Darwinisme** (Darwin, 1809-1882) : La sélection naturelle est le mécanisme responsable de l'évolution.

- Mandel (1865) : découvre les lois de l'hérédité et provoque une révolution scientifique qui aura des conséquences capitales par la suite, la **GENETIQUE**.

- **Mutationnisme** (Hugo De Vries, 1901-1903) : Rôle des transformations héréditaires dans le changement des espèces.

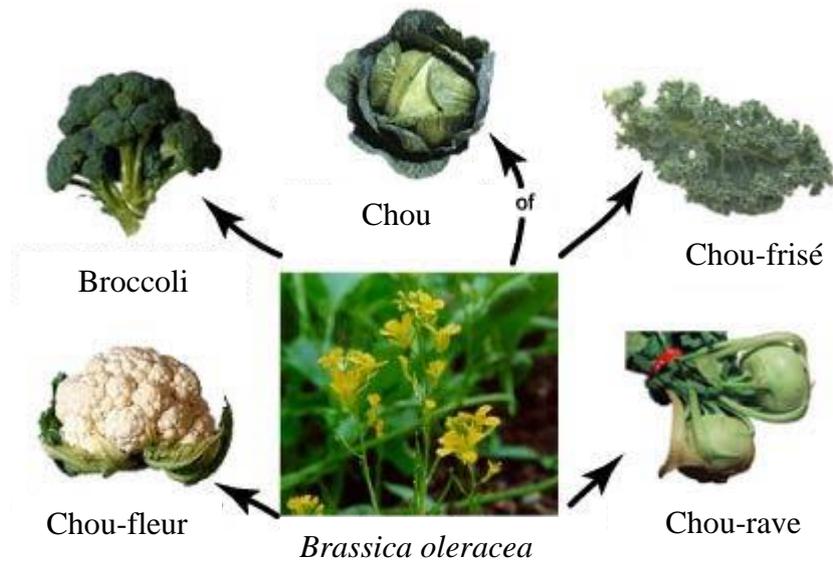
- **Néo-Darwinisme** : Associe la théorie de l'évolution et la théorie Mendélienne. Elle conduit à l'élaboration de la théorie synthétique de l'évolution dans les années 1920. Ernet Meyr propose et considère que la population a une position centrale. L'unité héréditaire est le gène, l'unité sélectionnée est l'individu et l'unité qui évolue est la population. La structure génétique d'une population peut subir une évolution, des modifications sous l'action de différents facteurs et mécanismes. Le mode de reproduction et les forces évolutives (mutations, sélection naturelle, dérive génétique et migration) sont à l'origine des changements qui touchent la population.

IV : Modèles de processus évolutifs :

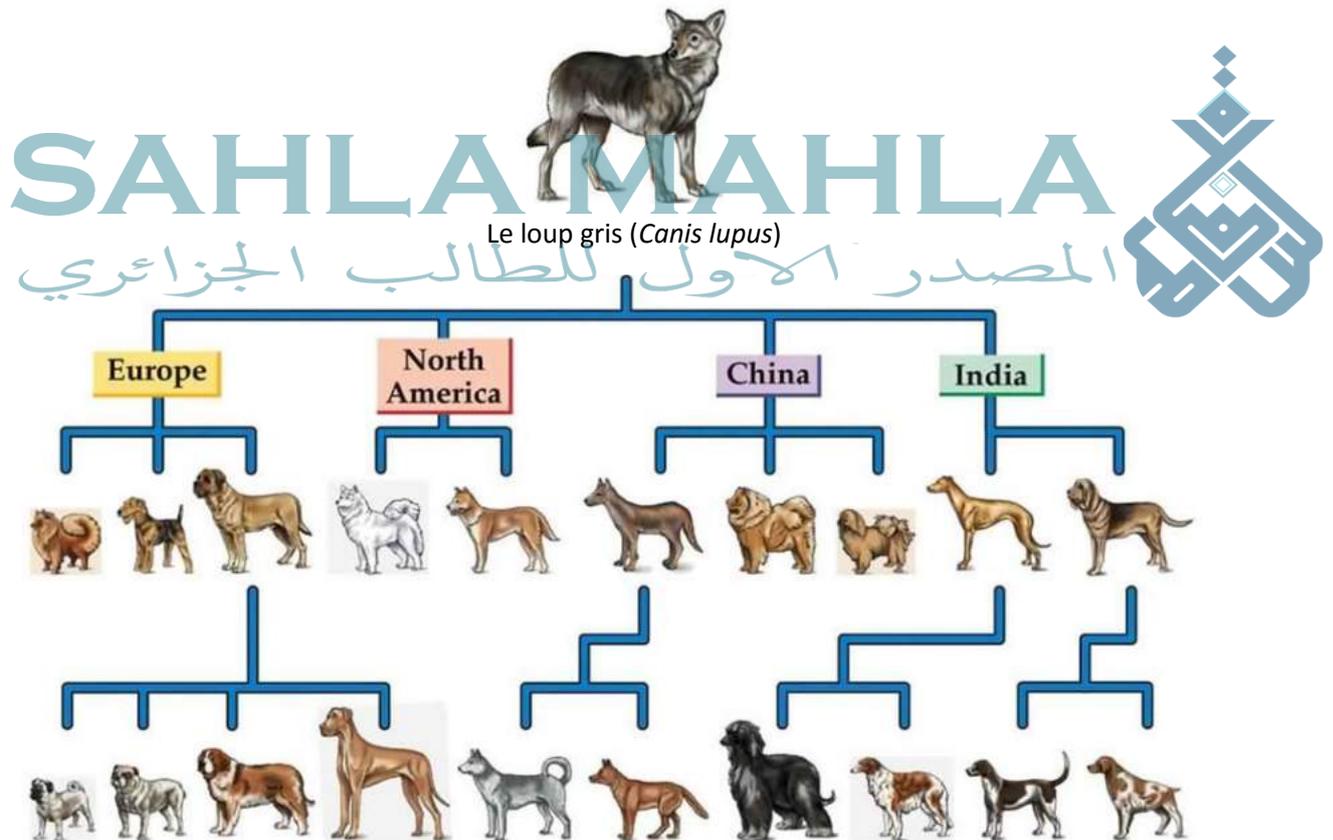
L'évolution des systèmes biologiques peut-être observé à différents niveaux, moléculaires ou au niveau des espèces et populations. Les exemples peuvent-être trouvés dans les populations issues de la sélection artificielle ou naturelle.

a- Les choux actuels (chou-fleur, choux de Bruxelles, chou-rave, etc..) sont le résultat de manipulations de reproduction sélectives, opérées par des agriculteurs pour obtenir des légumes à teneur économique élevée. Ils proviennent d'une espèce sauvage, *Brassica oleracea*, espèce quelconque qui a été amélioré au cours des siècles.

Introduction

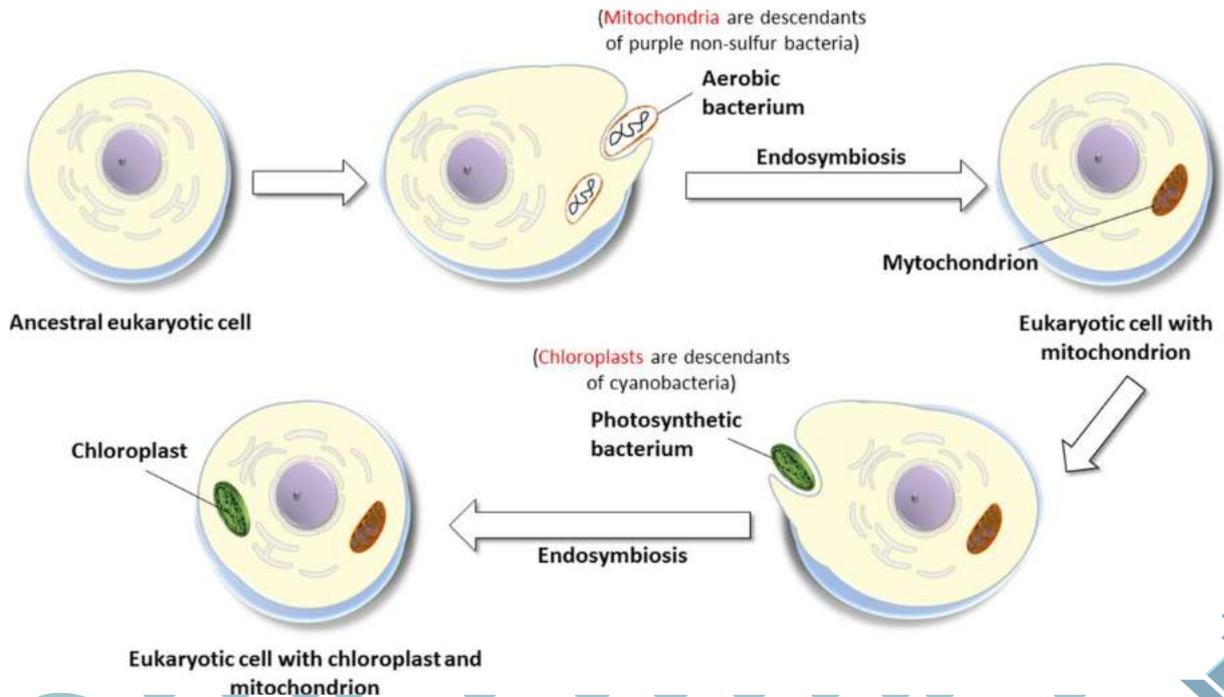


b- Les 200 races de chiens connus, *Canis familiaris*, descendent tous de modifications obtenues sur le loup *Canis lupus*. La domestication du loup sauvage a provoqué des changements dans la morphologie et le comportement des animaux.



Introduction

c- L'ADN des mitochondries et des chloroplastes est indépendant de l'ADN génomique. Il est d'origine bactérienne. Mitochondries et chloroplastes sont le résultat d'une endosymbiose qui a fait évoluer des bactéries en organites.



Origine bactérienne des mitochondries et de chloroplastes (Reiter, 2017)

المصدر الاول للطلاب الجزائري



A- Concept d'espèces

I- Notion d'espèce

Au IV^{ème} siècle avant J.C, Aristote réalisait déjà une classification des animaux. Le mot espèce a connu de nombreuses définitions, répondant à des exigences et des circonstances différentes. Les progrès apportés par les sciences et l'avancée des technologies, va favoriser cette diversité des définitions.

L'espèce reste donc un concept qui recouvre la **CLASSIFICATION** et l'**EVOLUTION** des taxons.

La classification a débuté au XIV^{ème} siècle et a concerné les espèces botaniques en premier lieu, afin de décrire les nouveaux spécimens rapportés par les expéditions qui découvraient de nouveaux mondes (nouveaux continents, nouvelles îles...). Il y avait également la nécessité de décrire de nouvelles plantes utilisables en pharmacologie.

II- Les premières classifications et définitions de l'espèce :

1- Concept de l'espèce typologique :

Linné est le fondateur de la classification (1707-1778). Elle a été conçue à partir du seul critère morphologique. L'identification repose sur une diagnose qui décrit la plante. Il introduit la classification binomale (**Genre espèce en latin** ; exemple: l'olivier *Olea europea*). Les espèces sont définies par rapport à un individu de référence ou « type ».

a- Critère d'interfécondité :

Buffon (1707-1788) introduit le critère d'interfécondité/interstérilité pour décrire l'espèce. Une espèce est caractérisé par des individus capables de produire entre eux des descendants fertiles. Des individus qui produisent des descendants stériles appartiennent donc à des espèces différentes.

Ara bleu et jaune, *Ara ararauna* (à gauche), Ara rouge, *Ara macao* (à droite) et leur hybride stérile *Catalina macaw* (au centre)



b- Critère de descendance :

Cuvier (1769-1839) fonde l'anatomie comparée. Il définit les rapports entre les être vivants. En analysant leur organisation interne. Il définit l'espèce comme une collection de tous les corps organisés nés les uns des autres ou de parents communs et de ceux qui leur ressemblent autant qu'ils se ressemblent entre eux. Il utilise les critères de descendance et de morphologie.

c- Critère évolutif :

Les espèces se succédant les unes des autres selon une continuité évolutive, l'espèce ne peut être décrite selon des critères fixes et figés. Le transformisme rapporté par Lamarck (1744-1829), Darwin (1809-1882) et d'autres, remettent en cause les définitions fixistes précédentes.

Exemple :

Analyse de séquence sur 375 loci des éléphants actuels et de 2 espèces récemment éteintes (mammouth laineux et mammouth d'Amérique) :

Divergences similaires entre les 2 "sous-espèces" africaines et entre éléphants d'Asie et le mammouth laineux

↓

**Deux espèces d'éléphants africains distinctes : *Loxodonta africana*
*Loxodonta cyclotis***

<i>Loxodonta africana africana</i> (éléphant de savane africain)	<i>L. africana cyclotis</i> (éléphant de forêt africain)	<i>Elephas maximus</i> (éléphant d'Asie)	<i>Mammuthus primigenius</i> (Mammouth laineux)
			

Source : Guédon, UMR 1128 INRA-Université de Lorraine, Biologie évolutive, 2018

SAHLA MAHLA

2- Définition biologique de l'espèce :

Proposée par Theodozius Dobzhansky (1900-1975) et Ernst Mayr (1904-2005). C'est le critère d'interfécondité qui est considéré et retenu. Une espèce est un ensemble de gènes qui se transmet entre les générations. **C'est un pool génétique fermé.**

« Les espèces sont **des groupes de populations naturelles réellement** ou **potentiellement** capables de se **croiser, séparés d'un point de vue reproductif des autres groupes.** »

III- Analyse des critères de définition de l'espèce :

a- Critère morphologique :

Il est possible de regrouper dans une même espèce, des individus qui se ressemblent. Au contraire, des individus différents seront séparés dans des espèces distinctes. Cependant 2 erreurs peuvent exister :

a 1- Une morphologie identique entre deux taxons peut s'expliquer par une adaptation au milieu ou un mécanisme de spéciation qui n'implique pas une différenciation morphologique. Ce phénomène s'appelle une **convergence**. Ces espèces isolées d'un point de vue reproductif mais semblables morphologiquement sont dites « **Jumelles** ».

Exemple :

- *Drosophila pseudoobscura* et *Drosophila persimilis* sont semblables morphologiquement mais ne se reproduisent pas entre elles.



Drosophila persimilis



Drosophila pseudoobscura

- Le nématode du genre *Anisakis* était considéré formé d'une seule espèce, *Anisakis simplex*. Avec le développement des techniques moléculaires, il a été démontré qu'il constituait un complexe d'espèces formé par au moins 10 espèces différentes.



a 2- La deuxième erreur est due au fait que la variabilité intra et inter-populationnelle est ignorée.

المصدر الاول للطالب الجزائري



Exemple :

- Chez certains papillons les mâles et les femelles présentent un dimorphisme sexuel.
- Le Lémur aux yeux turquoises, *Eulemur flavifrons*, est originaire de Madagascar le mâle est noir tandis que la femelle est de couleur orangée de plus en plus claire vers le ventre.



b- Critère de descendance :

La descendance avec modifications doit être considérée afin de refléter l'évolution des espèces.

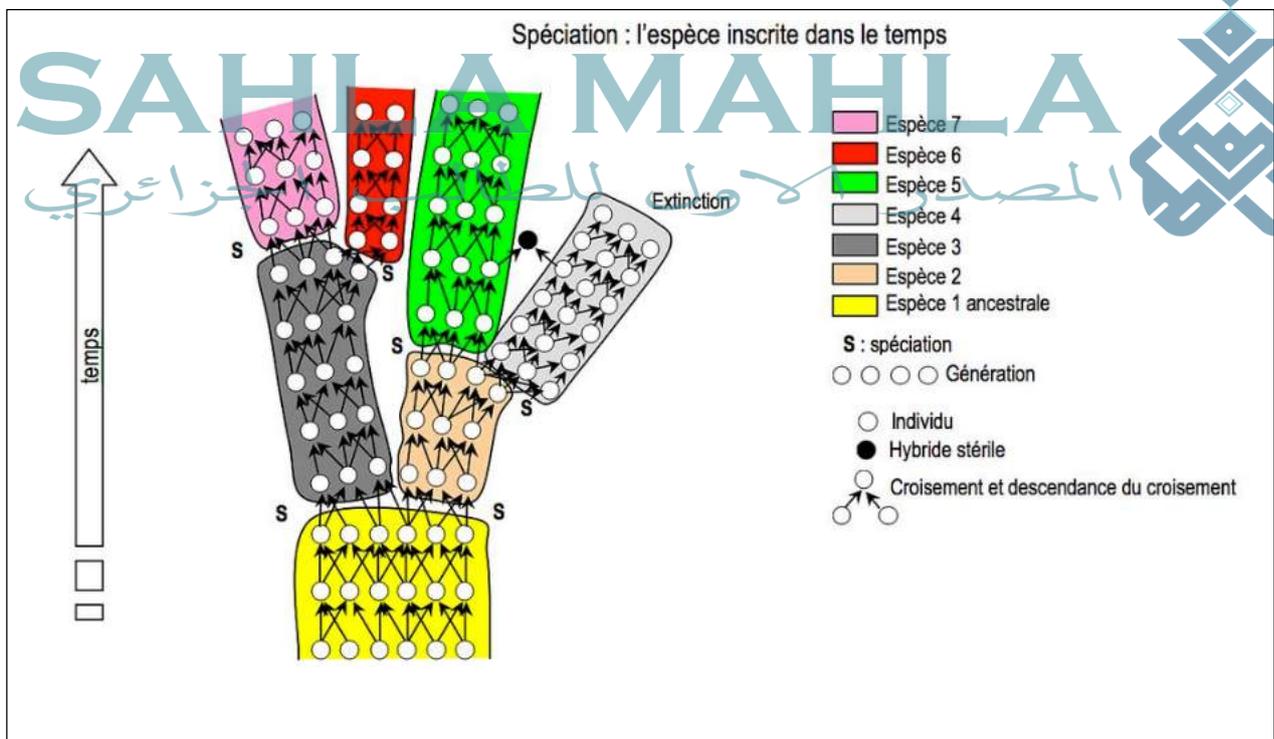
c- Critère d'interfécondité :

Il ne peut s'appliquer qu'à des espèces sympatriques (partagent la même aire géographique). Il ne peut s'appliquer pour des fossiles (raisons évidentes). Il y a également une limite chez certains végétaux, chez qui des espèces génétiquement proches produisent des hybrides fertiles. La même limite est observée chez les animaux à reproduction asexuée (bactéries) ou à reproduction uniparentale (obtention de clones).

IV- Autres définitions :

1- Concept phylogénétique de l'espèce :

- Développé par l'entomologiste Willy Henning (1913- 1976). Il est le fondateur de la **cladistique** et considère que seuls les clades sont à retenir pour la classification, c'est-à-dire les groupes constitués **d'un ancêtre et de tous ses descendants**. Les clades sont des groupes monophylétiques. Groupe d'individus interconnecté par des relations **tokogénétiques** (relation de filiation). Il est donc applicable aussi bien aux espèces sexuées qu'asexuées.



(Thomas *et al.* (2010). Biologie évolutive)

2- Concept d'espèces par reconnaissance :

Paterson dans les années 1980 introduit le concept de reconnaissance des partenaires sexuels et organisme biparentaux. Il s'appuie sur les facteurs positifs permettant aux individus d'une même espèce de se reconnaître et de se reproduire plutôt que sur les mécanismes d'isolement conduisant à l'arrêt des échanges géniques (exclusion des organismes utilisant la voie asexuée).

3- Concept d'espèces cohésives : (Templeton, 1989)

Groupe d'individus capables de supporter les mêmes conditions environnementales grâce à leur patrimoine génétique. Ce concept définit l'espèce du point de vue de **la cohésion génétique** et **phénotypique**. Les populations doivent avoir les mêmes aptitudes à tolérer les variables écologiques. Ce concept est le plus difficile à appliquer car il exige de déterminer les caractères de cohésion et les paramètres écologiques qui définissent la niche écologique de l'espèce.

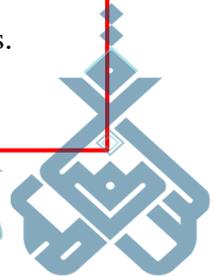
Résumé :

Actuellement 4 disciplines de la biologie essaient de définir le concept d'espèce :

- a- La **morphométrie** qui s'attache à la relation de ressemblance.
- b- La **génétique** qui s'intéresse à l'interfécondité.
- c- La **cladistique** qui se focalise sur la descendance et l'évolution des caractères.
- d- La **biologie moléculaire** qui se préoccupe de l'homologie des séquences

SAHLA MAHLA

المصدر الأول للطلاب الجزائري



B- Spéciation

I- Définition :

La spéciation décrit les mécanismes par lesquels se forment de nouvelles espèces. Sont considérées pour les modalités de spéciation :

- Les mécanismes impliqués.
- Le statut géographique des populations concernées.

II- Description des modalités de spéciation :

A- Spéciation allopatrique :

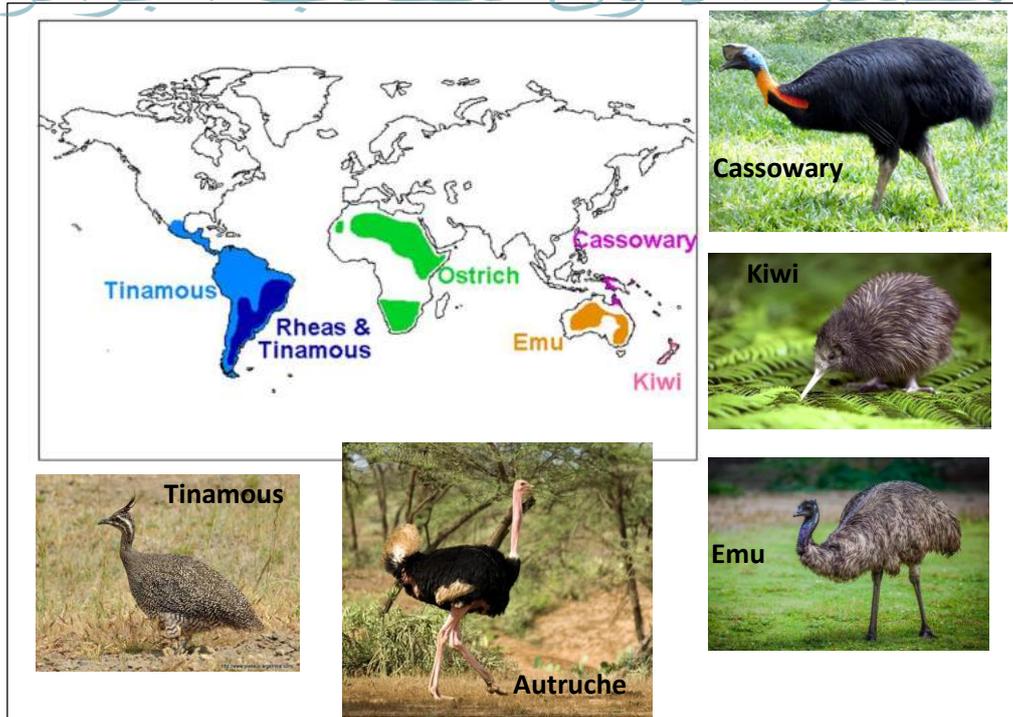
Les populations occupent des aires de distribution différentes.

1- Un facteur environnemental va séparer une population en deux sous-ensembles d'effectifs important. La cause de cette ségrégation peut-être naturelle ou anthropique, provoquée par l'homme (apparition d'un glacier, une rivière, une route, un évènement géologique tel un séisme....).

2- Cette barrière géographique va empêcher tout échange génétique entre les deux populations séparées. Chacune va se différencier sous l'effet de pressions sélectives différentes où de la dérive génétique.

Exemple :

Distribution géographique des oiseaux ratites (groupe d'oiseaux qui comprend autruches (ostrich), émeu, casoars, nandous, kiwis, rheas et tinamous, tous aptères i.e. inaptes au vol même s'ils ont conservé leurs ailes). Ils descendent d'un ancêtre commun, mais ont évolué différemment sur leurs continents respectifs, séparés par des océans.



3- La barrière géographique disparaît, ce qui permet de réunir les deux populations de nouveau en contact.

Deux cas sont possibles :

- a- Les divergences génétiques entre les deux populations sont si importantes qu'elles ne se croisent plus. Il y a donc **SPECIATION**. Les populations constituent deux espèces différentes.
- b- Les deux populations s'hybrident car la divergence était insuffisante pour empêcher leur croisement. Il y a donc homogénéisation des populations par brassage génétique. Il n'y a donc qu'une espèce. La SPECIATION n'a pas eu lieu.

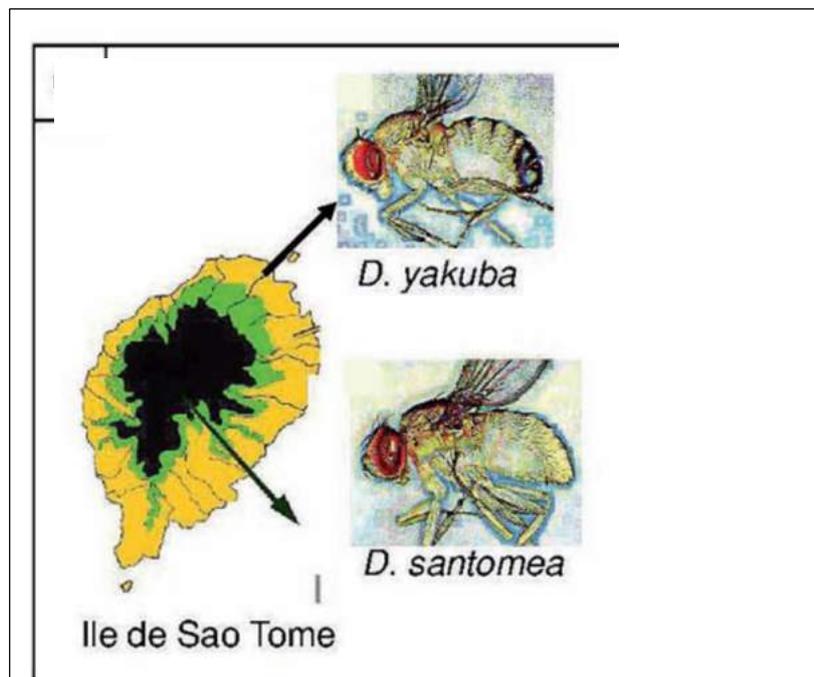
Conclusion : « L'isolement géographique est remplacé par l'isolement reproductif ».

- Spéciation péripatrique :

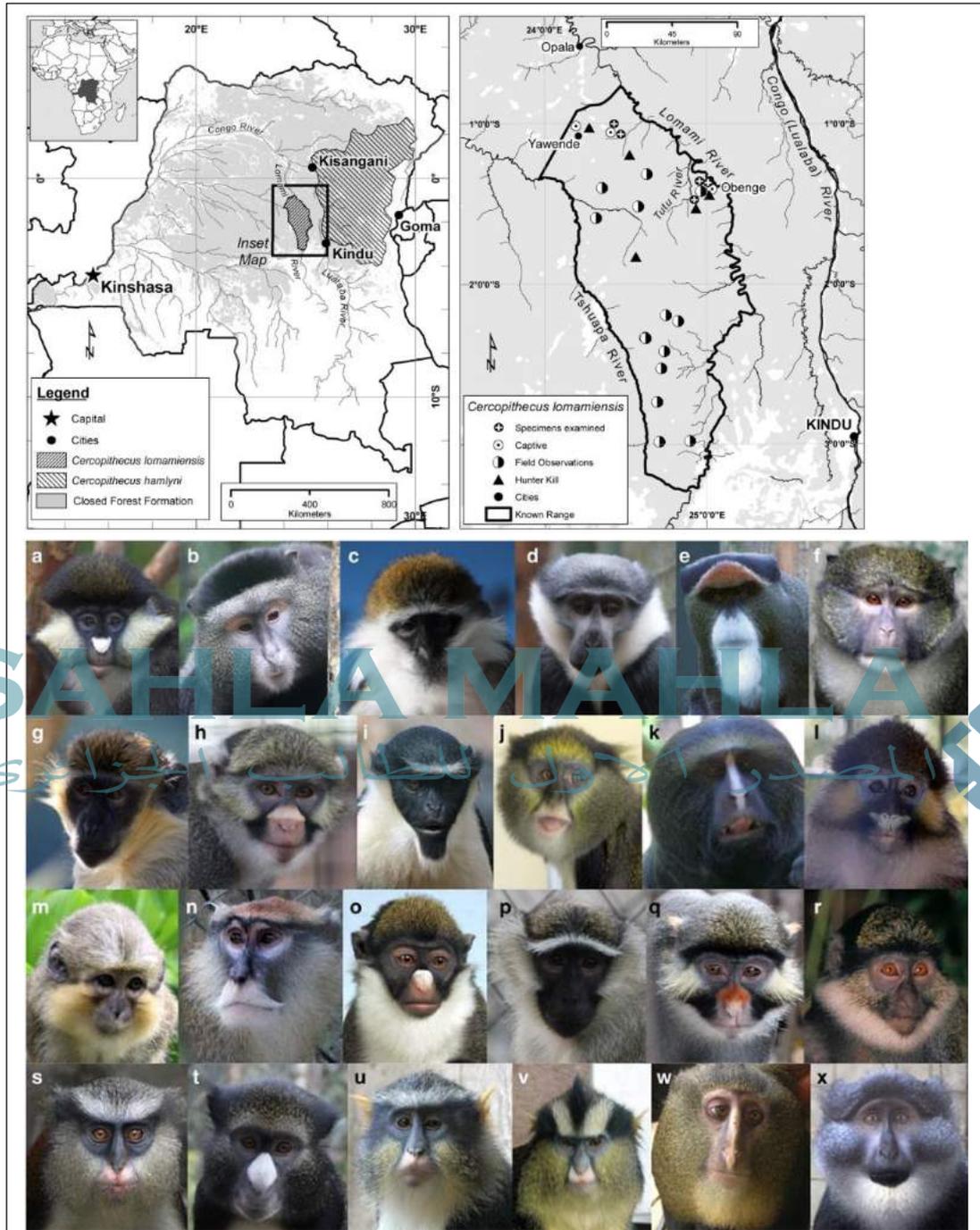
C'est une variante de la spéciation allopatrique, **par effet de fondation**. La population d'origine se scinde en deux ensembles de tailles différentes. Une petite population est isolée de l'autre par une barrière géographique ou par la migration de la petite population. La dérive génétique accentue la différenciation génétique d'avec la population initiale. Il va se former progressivement une nouvelle espèce.

Exemple :

- 1- Spéciation péripatrique de la mouche *Drosophila santomea*. Découverte en Avril 1998 sur l'île de Sao-Tome dans le golfe de Guinée (Afrique) (Lachaise et Hany, 2000). Elle est endémique des forêts de brouillard qui ne persistent qu'à partir de 1200 m jusqu'au sommet de l'île (2024 m). Elle est absente des altitudes basses où se trouve une espèce sœur dont elle a divergé, *Drosophila yakuba*.



2- Dans le bassin du Zaire (aujourd'hui RDC), la succession d'évènements climatiques arides et humides ont favorisé la spéciation des Cercopithèques. Ces derniers, petits singes de la sous-famille des Cercopithecinae a connu une radiation évolutive. Il y a actuellement 27 espèces présentant une spécialisation dans l'utilisation des ressources. Le nombre de chromosomes varie de 48 à 72 selon les espèces.



- Spéciation parapatrique :

La spéciation parapatrique est le processus de formation des espèces en présence de flux génique entre populations divergentes. La chaîne d'évènements, ainsi que les mécanismes et les

facteurs généralement impliqués dans le scénario de spéciation parapatrique, sont semblables à ceux de la spéciation allopatrique, sauf que l'isolement spatial des sous-populations divergentes est partiel plutôt que complet.

D'un point de vue théorique, la spéciation parapatrique représente le Scénario de spéciation le plus général et inclut à la fois la spéciation allopatrique et sympatrique comme cas particuliers (qui correspondent à un flux génétique nul et à un flux génétique très important, respectivement).

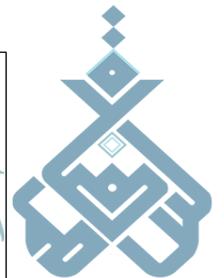
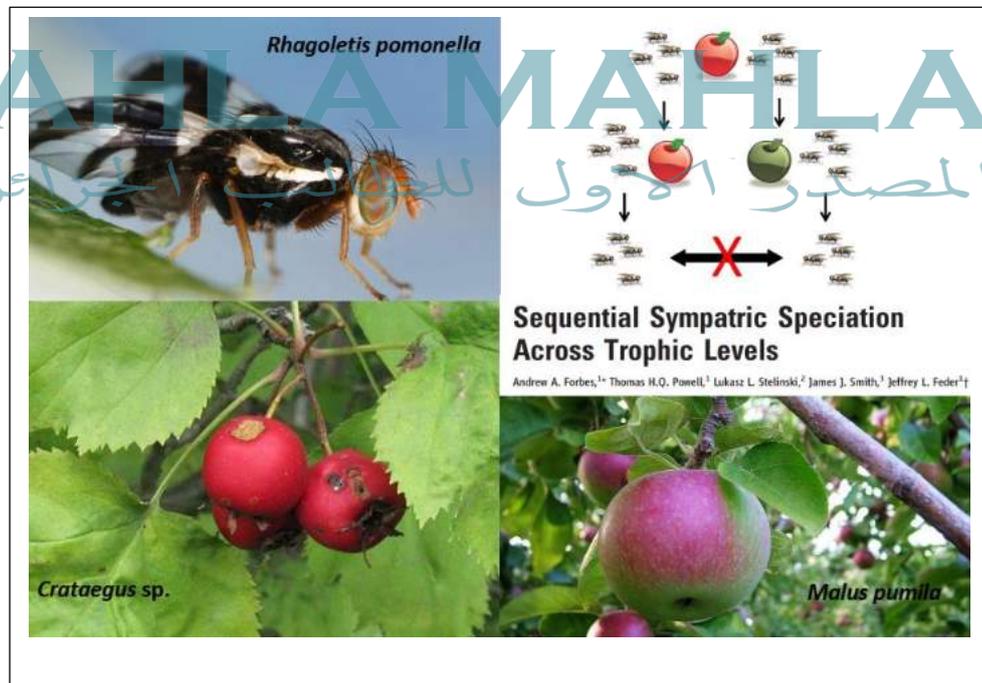
B- Spéciation sympatrique : (Spéciation sans isolement géographique)

Formation de deux espèces au sein de la même population dans la même aire géographique.

Exemple :

Spéciation par changement d'hôtes : elle se fait sur deux ou un ensemble d'hôtes partageant la même aire de distribution.

Les mouches Téphitidés du genre *Rhagoletis*. Des individus se reproduisent sur l'aubépine puis ont changé d'hôtes. Ils se sont déplacés sur les pommiers dont la culture s'était développée. Les adultes issus des larves nées sur les deux hôtes ont tendance à rester sur leur hôte et ne se reproduisent plus qu'entre eux. Une divergence génétique s'installe et les différenciations génétiques s'accumulent. Il s'installe un isolement reproductif. La différenciation porte sur les gènes impliqués dans la reconnaissance des partenaires sexuels.



Remarque :

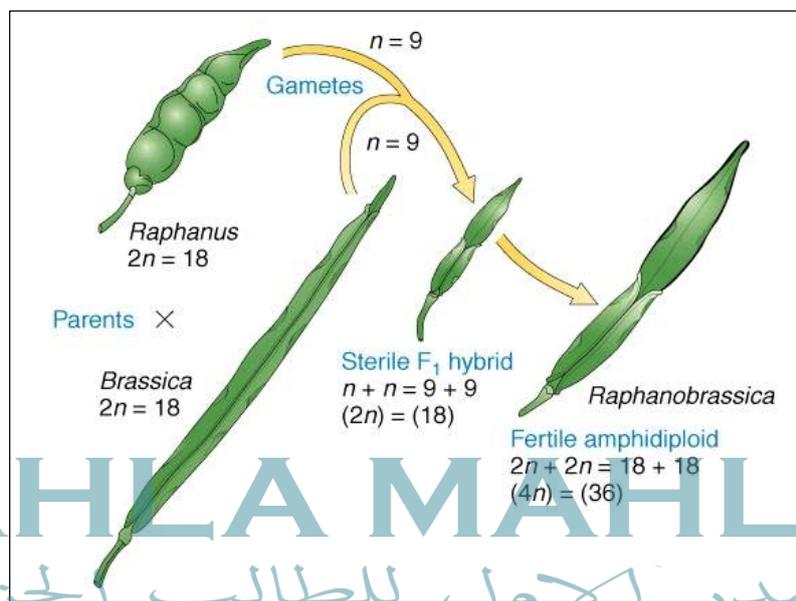
La spéciation STASIPATRIQUE décrit les remaniements portant sur les translocations ou les fusions chromosomiques qui sont à l'origine de l'isolement reproductif sympatrique.

Exemple :

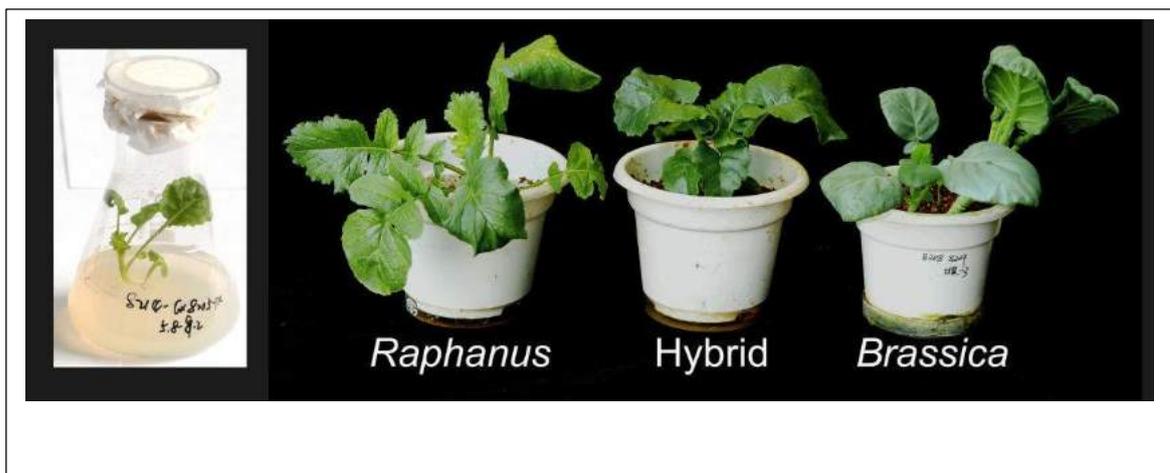
La ploidisation par doublement du lot de chromosomes peut conduire à la différenciation de nouvelles espèces, notamment chez les plantes.

Karpechenkob réalise en 1928 une hybridation entre le chou *Brassica oleacea* et le radis *Raphanus sativus*, plante à $2n=18$ chromosomes.

Une descendance hybride viable est obtenue, mais stérile. Cependant, quelques graines furent un jour produites présentant 36 chromosomes. Une fécondation entre gamètes non réduits s'était réalisée permettant un appariement au cours de la méiose. Une nouvelle espèce fertile était créée *Raphanobrassica*.



SAHLA MAHLA
المصدر الأول للطالب الجزائري



C- Les mécanismes de l'isolement reproductif :

Deux types :

- 1- Mécanismes d'isolement précopulatoire ou prézygotique.
- 2- Mécanisme d'isolement postcopulatoire ou postzygotique.

1- Mécanismes d'isolement précopulatoire ou prézygotique.

a- Les partenaires possibles ne se rencontrent pas pour deux raisons :

- Isolement écologique : région géographique où les populations occupent des habitats différents.
- Isolement saisonnier ou temporel : période de reproduction a lieu à des époques différentes.

b- Les partenaires sexuels se rencontrent mais ne s'accouplent pas du fait d'un isolement sexuel ou éthologique.

c- L'accouplement a lieu mais il n'y a pas de transfert de gamètes du fait d'un isolement mécanique.

d- Les gamètes males sont transférés mais il n'y a pas de fécondation du fait d'un isolement gamétique.

2- Mécanismes d'isolement postzygotique :

On observe :

a- une léthalité hybride, la fécondation a lieu mais les zygotes ne survivent pas.

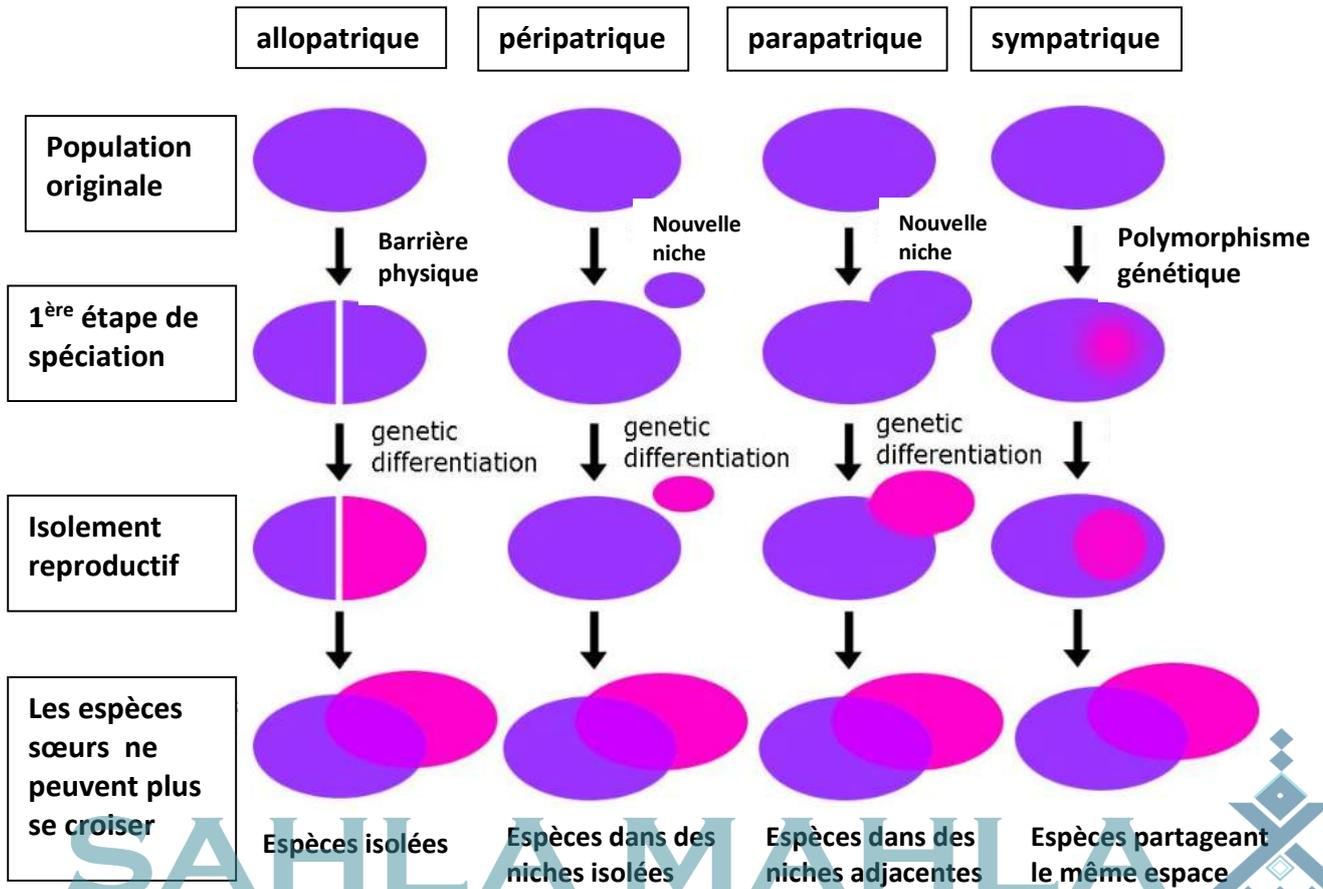
b- une stérilité des hybrides, les hybrides se développent mais ne produisent pas de gamètes fonctionnels chez l'un des deux sexes ou chez les deux.

c- une dépression hybride, la viabilité ou la fertilité en F2 ou des croisements en retour (back-cross) est réduite.

d- une incompatibilité cytoplasmique, résultat d'une infection due à des endosymbiontes cellulaires du genre *Wolbachia* (bactérie intracellulaire présente chez des femelles de nombreux invertébrés et responsables d'un certain nombre de phénotypes, tels la féminisation des mâles, l'incompatibilité cytoplasmique, etc...).



Schéma général des modes de spéciation



المصدر الاول للطلاب الجزائري



Génétique évolutive

MUTATIONS

INTRODUCTION:

La structure de l'ADN et le séquençage ont révélé que l'apparition de nouveautés génétiques comme de nouveaux gènes se produit généralement à partir de modifications d'éléments préexistants (gènes entiers, portions de gènes, exons, introns, promoteurs, segments d'ADN intergéniques...).

Parmi les mécanismes pouvant générer des innovations génétiques nous pouvons citer : la duplication de gènes, la transposition par insertion d'éléments mobiles, le brassage d'exons, l'épissage alternatif, la fusion/fission et le transfert horizontal entre espèces différentes ou proches (en particulier chez les procaryotes) (**Figure 1**).

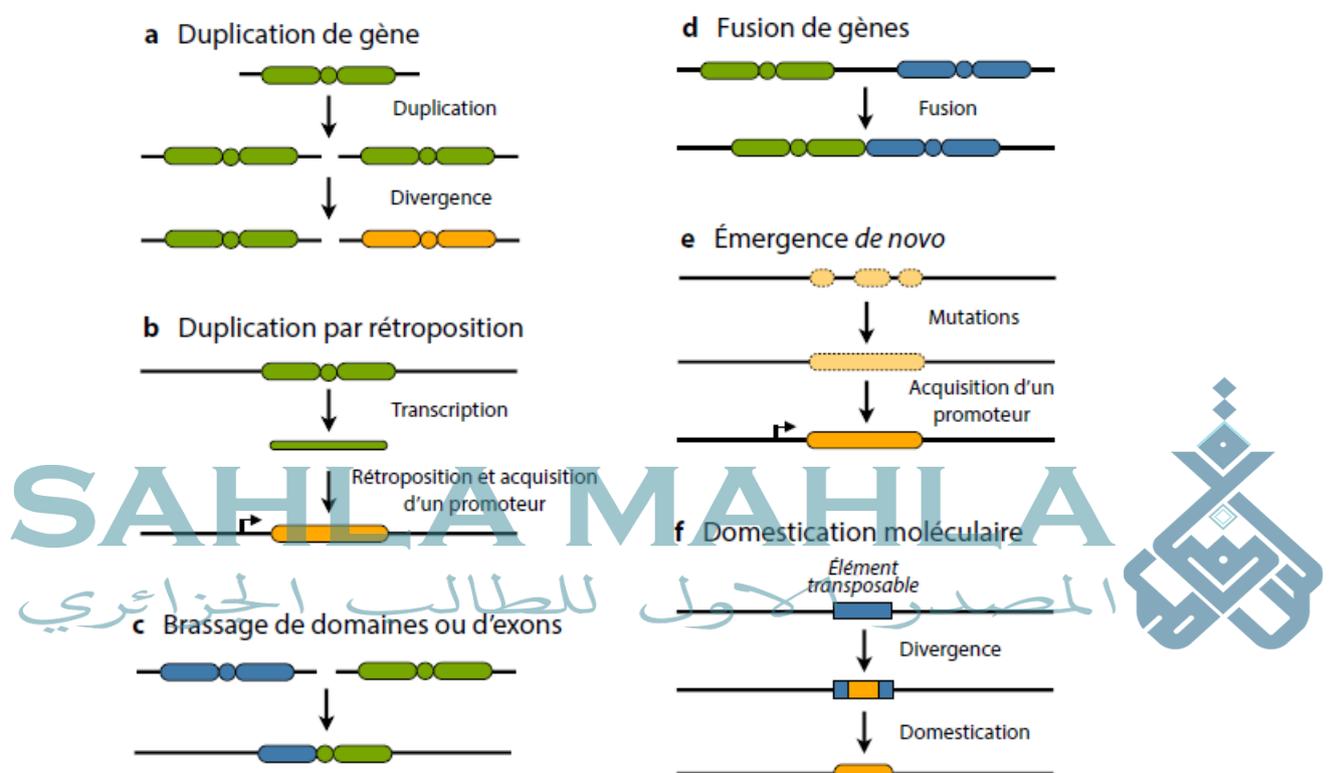


Figure 1. Mécanismes génétiques principaux qui conduisent à l'émergence de nouveaux gènes. Les exons sont représentés par des cylindres et les introns par des cercles. La duplication de gène avec un intermédiaire ADN (depuis la duplication d'un fragment de gène à la duplication de génome) est le mécanisme le plus connu pour l'émergence de nouveaux gènes (a). La duplication avec un intermédiaire ARN nécessite la présence des enzymes de rétroposition dans le cytoplasme et l'acquisition d'un promoteur (b). Le brassage de domaines (ou d'exons) de différents gènes (c) ainsi que la fusion de gènes (d) peut produire des nouveaux gènes chimères. A partir d'une séquence non-codante, des mutations peuvent éliminer les interruptions d'un cadre de lecture potentiel, et suite à l'acquisition d'un promoteur, conduire à l'émergence d'un gène *de novo* (e). La domestication moléculaire (ici d'un élément transposable) peut conduire à de nouveaux gènes (f).

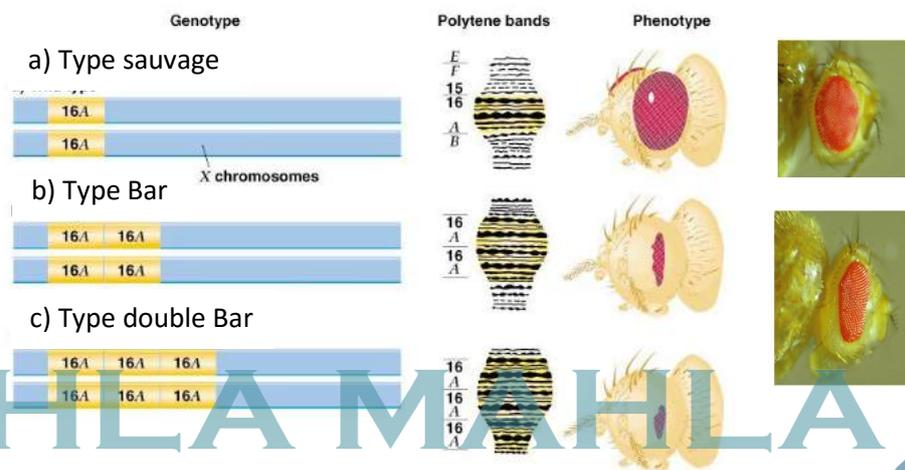
Génétique évolutive

MUTATIONS

I- DUPLICATION :

Chez l'homme, 15 % des gènes sont présents en deux exemplaires au moins. Chez la levure, 26 %. Autant de copies qui peuvent jouer un rôle essentiel dans l'adaptation des organismes.

En 1918, C. Bridges est le premier à mentionner, lors d'un congrès, la possibilité de l'existence de gènes dupliqués. Il évoque déjà l'hypothèse d'une implication de ces gènes dans le processus évolutif. D'après lui, ces gènes identiques peuvent évoluer de manière distincte et diversifier ainsi leur effet. Par la suite, il apportera une des premières observations de duplication, celle du gène *Bar* chez un mutant de *Drosophila melanogaster*, présentant une réduction importante de la taille de l'œil. L'étude cytologique des chromosomes polytènes a mis en évidence un dédoublement de bande sur le chromosome X correspondant à la duplication en tandem du gène *Bar*. Le rôle potentiel des duplications dans le processus évolutif a alors été mis à jour et de nombreuses hypothèses sur le devenir des éléments dupliqués ont été proposées.



En 1959, Markert a montré l'existence de formes multiples d'une même enzyme, les Isozymes, dont l'étude a révélé, entre autres, l'existence de familles de gènes. L'analyse des protéines a précédé, ici, l'identification des gènes qui les codent du fait des techniques disponibles à l'époque. À partir d'exemples de séquences hautement répétées comme les gènes codant les ARN ribosomiques, les histones ou encore l'hémoglobine de mammifères, Susumu Ohno a réellement mis en avant l'importance du rôle des duplications à travers l'évolution dans sa monographie "Evolution by gene duplication" (1970). Il reformule l'hypothèse selon laquelle la duplication de gènes permet l'émergence de nouvelles fonctions. Il a ainsi proposé que la duplication génique était le principal mécanisme d'évolution: une des deux copies d'un gène, en absence de pression de sélection, va pouvoir accumuler des mutations aléatoirement et acquérir ainsi une nouvelle fonction.

Les possibilités par lesquelles des séquences d'ADN au sein d'un génome peuvent se dupliquer sont multiples. L'apparition de gènes paralogues peut se faire par la duplication d'un seul gène, d'un segment composé de plusieurs gènes contigus, d'un chromosome entier ou encore d'un génome entier.

1- Duplication génique

Génétique évolutive

MUTATIONS

Les événements de duplication concernant un seul gène sont qualifiés de duplication génique. Ces éléments dupliqués peuvent être situés l'un à côté de l'autre, ce que l'on appelle une **duplication en tandem par crossing-over inégal** ou bien ils peuvent être dispersés au sein du génome.

1.1. Les duplications par recombinaison inégale :

La recombinaison homologe survient pendant la méiose, lors de l'appariement des chromosomes homologues (**Figure 2**).

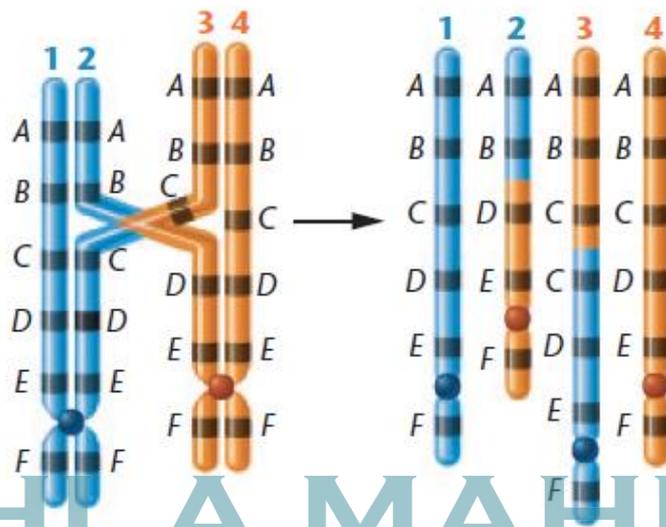


Figure 2 : Représentation schématique du principal mécanisme à l'origine des duplications en tandem: la recombinaison inégale. La tétrade de gauche est mal appariée lors de la prophase I. Un seul croisement entre les chromatides 2 et 3 entraîne les régions chromosomiques déficientes (chromosome 2) et dupliquées (chromosome 3) illustrées à droite. Les deux chromosomes non impliqués dans le croisement restent normaux dans la séquence et le contenu des gènes (chromosomes 1 et 4).

Sous l'influence des nombreuses séquences répétées présentes dans les grands génomes comme ceux des animaux, il peut arriver que cet appariement se trouve **légèrement décalé**. En conséquence, si une recombinaison se produit, l'un des chromosomes vient s'enrichir d'un fragment provenant de l'autre chromosome, emportant avec lui un ou plusieurs gènes déjà présents sur le premier. Si la recombinaison n'implique qu'un seul gène, celui-ci et sa copie se retrouvent dans le prolongement l'un de l'autre, d'où leur disposition **en tandem**. Ainsi, la mise en évidence de gènes apparentés présentant une disposition en tandem est un argument très fort pour penser qu'ils se sont formés par ce mécanisme. L'inverse n'est toutefois pas toujours vrai, car des gènes initialement engendrés par duplication en tandem peuvent être par la suite séparés à la faveur de remaniements chromosomiques, comme par exemple des translocations. Au cours de l'évolution, il arrive parfois que des duplications en tandem se répètent à plusieurs reprises en un même locus. Elles conduisent alors à la formation de complexes multigéniques. Les

Génétique évolutive

MUTATIONS

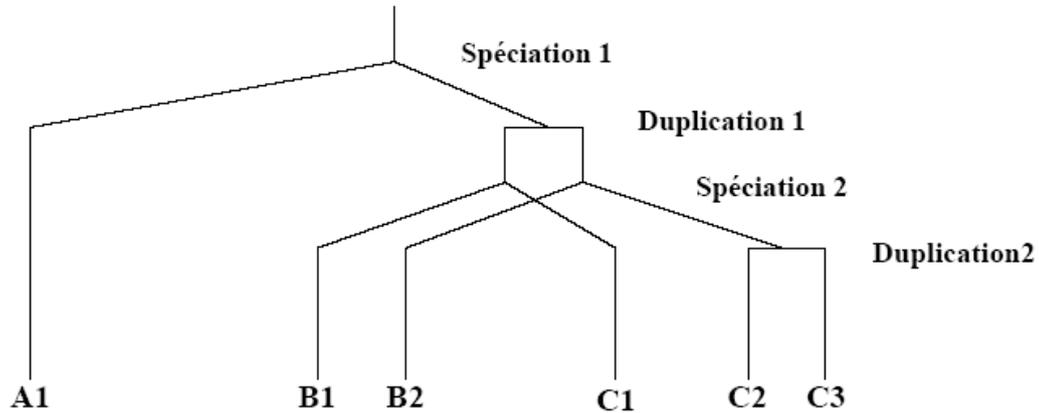
complexes de gènes Hox, qui existent chez l'ensemble des bilatériens, sont clairement le résultat d'un tel processus.

1.2. Terminologie des gènes dupliqués :

Deux gènes sont dit homologues lorsqu'ils possèdent une origine commune.

Deux gènes sont dits **orthologues** lorsqu'ils ont divergé lors d'un événement de spéciation.

Deux gènes sont dits **paralogues** lorsqu'ils résultent d'une duplication au sein d'une même espèce.



Gènes	Orthologues	Paralogues
A1	B1, B2, C1, C2, C3	aucun
B1	C1	B2, C2, C3
B2	C2, C3	B1, C1
C1	B1	B2, C2, C3
C2	B2	B1, C1, C3
C3	B2	B1, C1

Le tableau précédent montre les rapports reliant les différents gènes représentés dans l'arbre phylogénétique plus haut. Ainsi le gène A1 est orthologue avec tous les gènes issus des différentes duplications (B1, B2, C1, C2, C3) et n'a pas de gènes paralogues.

1.3. Le devenir des gènes dupliqués :

Le devenir des gènes engendrés par duplication est une question complexe. Selon le modèle proposé par Susumu Ohno (1970), sur les deux copies, initialement identiques et donc répétitives sur le plan fonctionnel, une seule est suffisante pour continuer à assurer les fonctions du gène ancestral. La structure de cette copie est donc en général fortement conservée. En

Génétique évolutive**MUTATIONS**

revanche, l'autre copie, échappant de ce fait aux pressions de sélection, peut évoluer beaucoup plus librement et donc accumuler de nombreuses mutations. Ohno (1970) postule que cette évolution peut suivre deux destinées différentes (**Figure 3**) :

- la **non-fonctionnalisation**, lorsque les mutations viennent à supprimer les capacités d'expression du gène
- la **néo-fonctionnalisation**, lorsque les mutations permettent l'émergence de fonctions nouvelles, susceptibles par la suite d'être sélectionnées

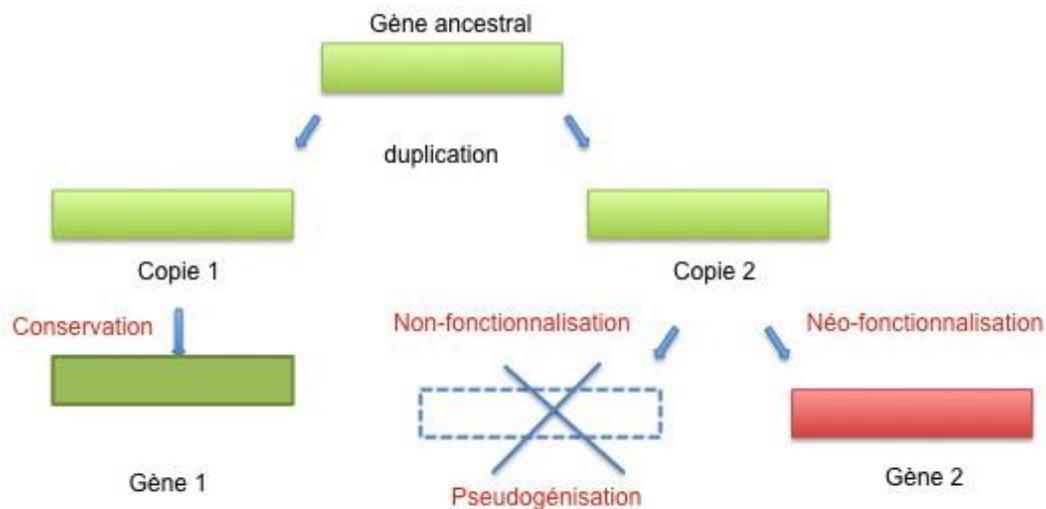


Figure 3. Devenir des copies d'un gène après duplication : non-fonctionnalisation et néo-fonctionnalisation

La première de ces trois voies est de loin la plus fréquente de sorte que, dans la majorité des cas, les copies surnuméraires non fonctionnelles du gène ancestral évoluent sous la forme de pseudogènes et finissent par être perdues. Ainsi, dans le cas des gènes Hox de la souris, seuls trois des treize gènes de chaque complexe sont en quatre exemplaires. Dans tous les autres cas, de une à deux copies sont systématiquement manquantes.

Exemple :

Cas du gène EDN (Eosinophil-Derived Neurotoxin) et ECP (Eosinophil Cationic Protein). Dans la lignée des Platyrrhiniens (singes du nouveau monde), le gène EDN qui fait partie de la famille des RNase, est resté sous forme de copie unique et a conservé la structure et la fonction du gène d'origine (rôle antiviral). Par contre, dans la lignée évolutive des Catarrhiniens (qui conduit à la lignée des Hominidés et des singes de l'ancien monde), ce gène a subi une duplication. L'une des copies (EDN) a gardé la fonction d'origine tandis que l'autre a acquis une nouvelle fonction, en effet elle code pour une protéine anti bactérienne. Cette activité n'est pas liée à l'activité RNase de l'EDN mais elle est la conséquence de mutations qui ont permis le remplacement de plusieurs acides aminés (**Figure 4**).

Génétique évolutive

MUTATIONS

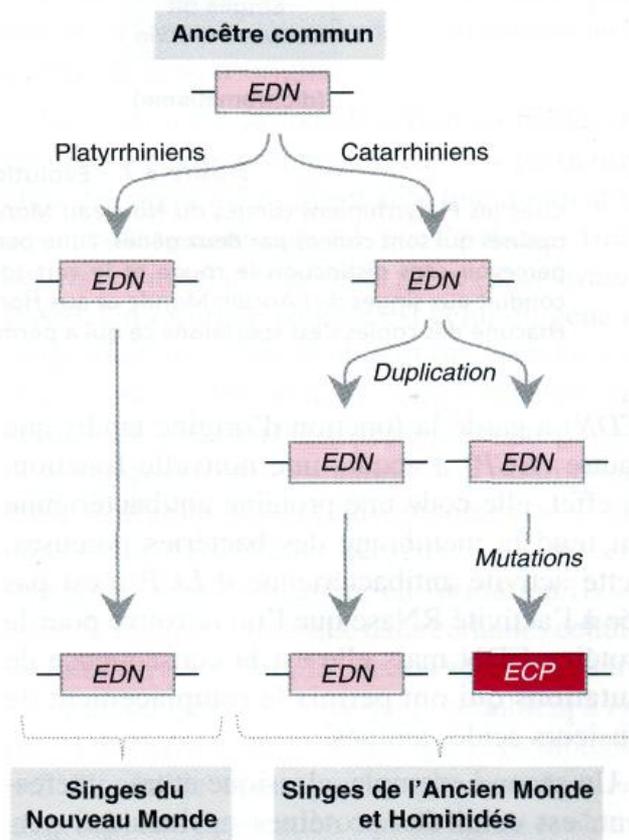


Figure 4 : Origine du gène *ECP* par néo-fonction.

Dans la lignée des Catarrhiniens, qui conduit aux singes de l'Ancien Monde et aux Hominidés, le gène *EDN* (*Eosinophil-Derived Neurotoxin*) a subi une duplication. L'une des copies a conservé sa fonction d'origine tandis que l'autre, appelée *ECP* (*Eosinophil Cationic Protein*), a subi des mutations (remplacement de plusieurs acides aminés par des arginines) lui permettant d'acquérir des propriétés antibactériennes.

2- Famille de gènes et de protéines:

La duplication de gènes va permettre la formation de familles multigéniques qui comprend tous les éléments possédant un lien de parenté. L'un des exemples incontournables est constitué par la superfamille des globines chez les eucaryotes supérieurs. Ces protéines à hème sont adaptées à la fixation réversible de l'oxygène. Chez l'homme, la super famille des globines est constituée, dans l'état actuel des connaissances, de quatre familles:

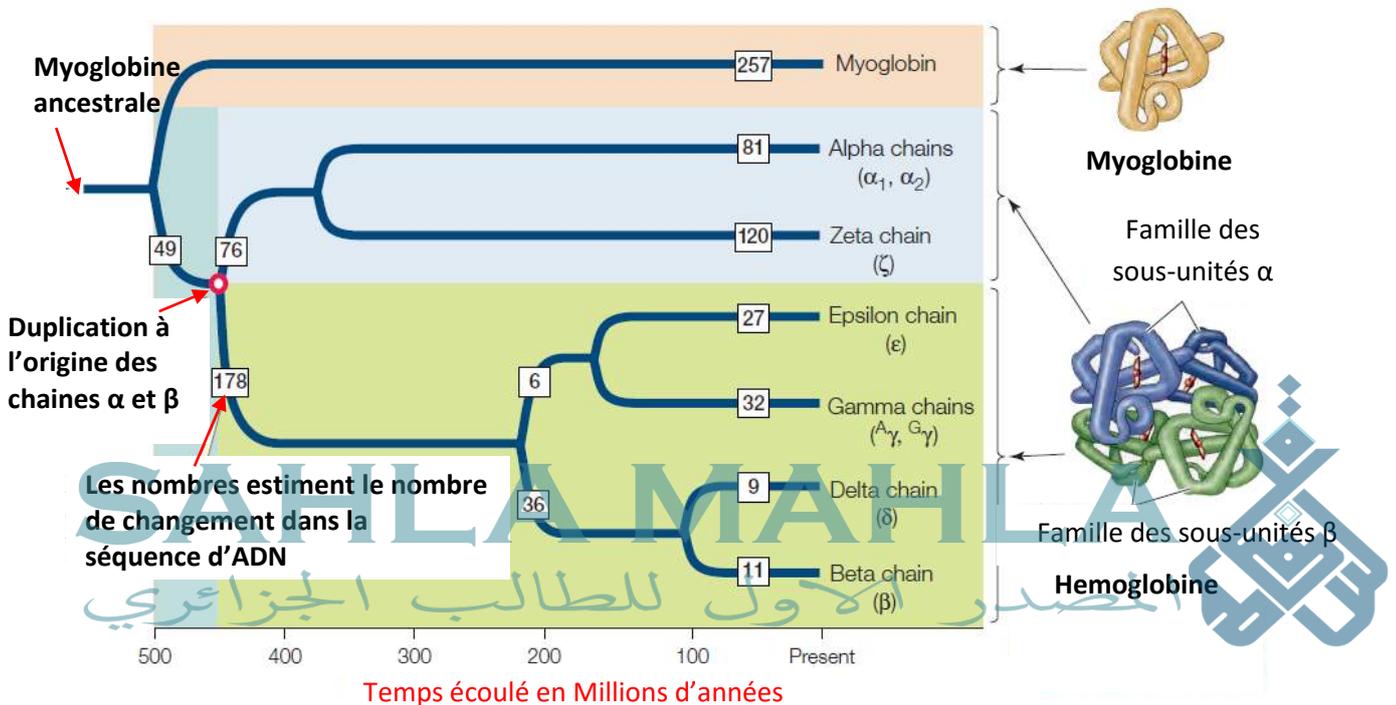
- 1 - la famille de la neuroglobine dont l'unique membre est localisée sur le chromosome 14
- 2 - la famille de la myoglobine dont l'unique membre est localisé sur le chromosome 22 ;
- 3 - la famille multigénique de **α -globine** localisée sur le chromosome 16 ;

Génétique évolutive

MUTATIONS

4 - la famille de la **β -globine** localisée sur le chromosome 11.

Les quatre familles produisent trois types de protéines ayant une activité spécifique: La myoglobine, la neuroglobine et l'hémoglobine. Elles ont divergé à partir d'un ancêtre commun il y a environ 450 millions d'années et sont devenues spécialisées sous plusieurs aspects. Ainsi, chaque protéine a une spécificité tissulaire. La myoglobine est devenue la protéine de stockage de l'oxygène dans les muscles, la neuroglobine celle des cellules nerveuses. L'hémoglobine, quant à elle, permet le transport de l'oxygène dans le sang. Par ailleurs, alors que la myoglobine et la neuroglobine ont gardé une structure monomérique, l'hémoglobine est devenue un hétérotétramère composé de deux sous-unités α et de deux sous-unités β .



Chez l'homme, la famille α comprend 3 gènes fonctionnels et 2 pseudogènes alors que la famille β comprend 5 gènes et un pseudogène (Figure 5).

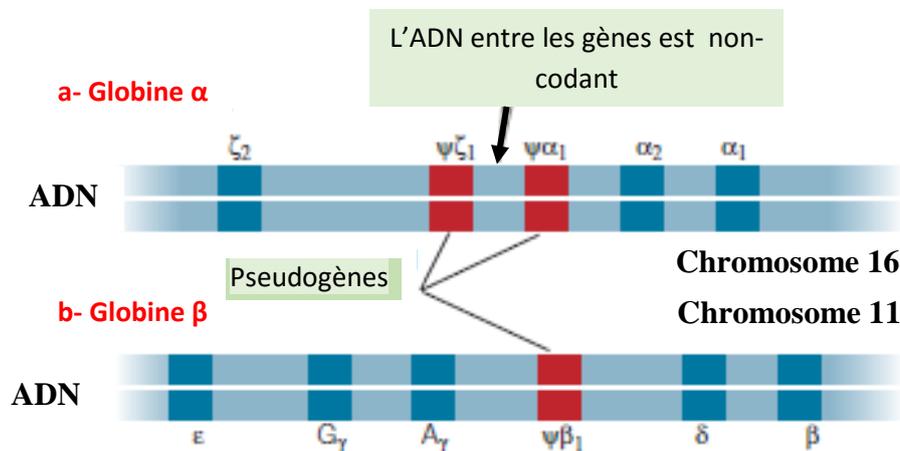


Figure 5. Gènes de l'hémoglobine chez l'homme. Les pseudogènes non fonctionnels sont indiqués par la lettre grecque psi (ψ).

Génétique évolutive**MUTATIONS**

Ces différents gènes ont divergé à la fois au niveau de la régulation et de la fonction. En effet, différents hétérotétramères d'hémoglobines apparaissent à différents stades du développement: $\zeta 2\varepsilon 2$ et $\alpha 2\varepsilon 2$ dans l'embryon, $\alpha 2\gamma 2$ dans le fœtus, puis $\alpha 2\beta 2$ et $\alpha 2\delta 2$ chez l'adulte. Des différences d'affinité pour l'oxygène sont à l'origine de la présence de ces différents hétérotétramères. Par exemple, l'hémoglobine foetale ($\alpha 2\gamma 2$) possède une affinité à l'oxygène plus importante que celle de l'adulte. Ceci permet au fœtus de recevoir assez d'oxygène à partir du sang maternel. La base moléculaire de cette affinité provient de la présence d'une sérine en position 143 de la globine γ alors qu'il s'agit d'une histidine dans la chaîne β .

La création d'une nouvelle fonction à partir d'un gène dupliqué peut nécessiter un nombre important de substitutions d'acides aminés comme ou une seule substitution comme pour les globines γ et β .

Conclusion :

Les duplications géniques représentent, avec les mutations, la source majeure de l'enrichissement du contenu informationnel des génomes au cours de l'évolution. La contribution la plus importante de la duplication de gènes à l'évolution est de fournir un nouveau matériel génétique pour différents mécanismes d'évolution, tels la mutation, la dérive génétique et la sélection naturelle, dont le résultat serait le développement de fonctions spécialisées ou de nouveaux gènes avec de nouvelles fonctions. Identifier chez les espèces actuelles les traces des duplications passées et reconstituer ainsi l'histoire de leur génome est l'un des enjeux majeurs des recherches actuelles.

II- Impact des éléments transposables:

Découverts dans les années 1940 par Barbara Mc Clintock dans le maïs et qui obtiendra le prix Nobel en 1983 pour ses travaux.

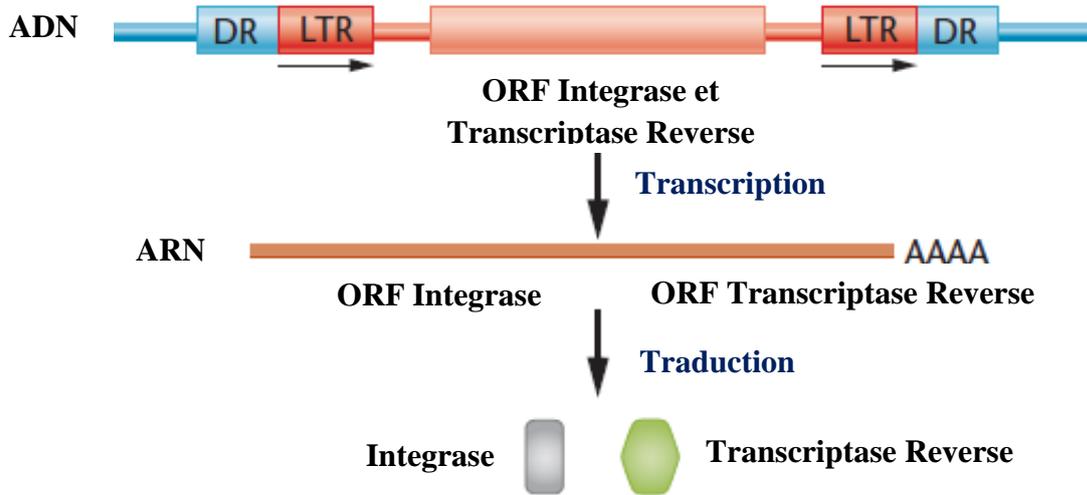
Il s'agit d'une séquence d'ADN capable de se déplacer de façon autonome et de se multiplier dans le génome. On distingue deux types de transposons :

- 1- Les éléments **transposables de classe I (Rétrotransposons, Figure 6)**, qui transposent via un intermédiaire ARN. Leur ARN est transcrit de façon inverse en un ADNc (complémentaire) qui sera intégré dans le génome. Ils sont subdivisés en deux groupes, Rétrotransposons à LTR (Longue séquence Terminale Répétée ou Long Terminal Repeat) et les Rétrotransposons sans LTR. Ils se différencient en fonction de
 - leur capacité à se transposer.
 - la présence de longues répétitions terminales (LTR).
 - leur capacité à former des particules virales.

Génétique évolutive

MUTATIONS

a) Retrotransposons à LTR



b) Retrotransposons sans LTR

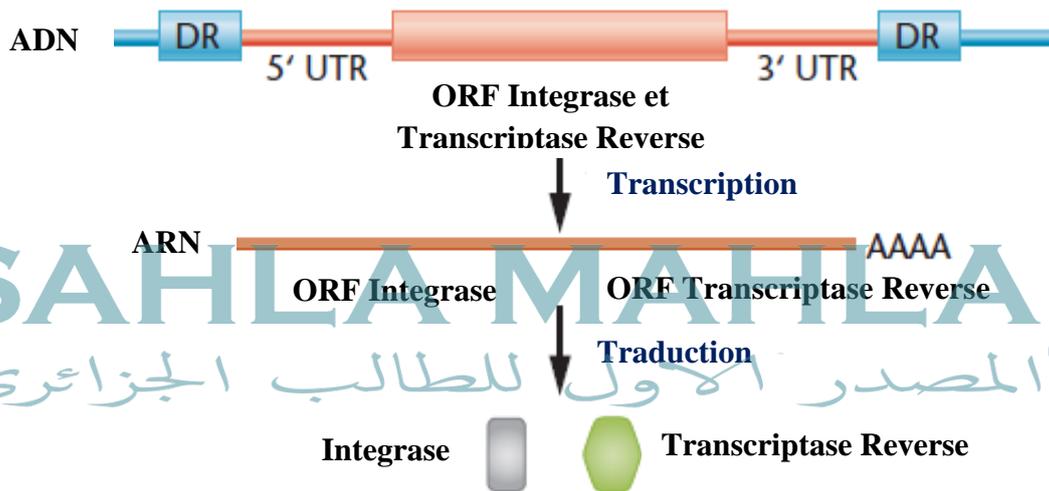


Figure 6. Structures de transposon de classe I

(a) Les rétrotransposons LTR, représentés en rouge, contiennent des cadres de lecture ouverts (ORF) qui codent pour les enzymes intégrase et transcriptase inverse (RT). Les promoteurs de transcription et les sites de polyadénylation sont situés dans les (LTR) 5' et 3'. La partie inférieure du diagramme montre la transcription du rétrotransposon LTR et la traduction en intégrase et transcriptase reverse. (b) Les rétrotransposons non LTR codent également l'intégrase et la transcriptase reverse, mais sans LTR. Les promoteurs de transcription et les sites de polyadénylation sont situés en 5' et 3' des régions non traduites (UTR).

- 2- Les éléments **transposables de classe II** (Figure 7), transposent via un intermédiaire ADN. Ils peuvent être flanqués par des séquences d'insertions ou des séquences inversées répétées (ITR inverted terminal repeats). Leur longueur varie de 2500 à 7000 pb. On les retrouve souvent en tant que familles de séquences répétées à travers le génome, car il existe de légères différences entre les séquences traduisant l'ancienneté de leur production et de leur divergence. En plus de la Transposase, ils contiennent des

Génétique évolutive

MUTATIONS

gènes autres que ceux liés à la transposition. Les plus simples sont les séquences d'insertions (IS) : éléments transposables d'une longueur de 700 à 2500 pb, leurs extrémités sont constituées de séquences courtes répétées inversées. Ils contiennent 2 cadres de lecture ouverts (Ins A et Ins B) codant pour une ou deux formes de la Transposase. Ainsi, les séquences d'insertions ne contiennent que les éléments nécessaires à la transposition.

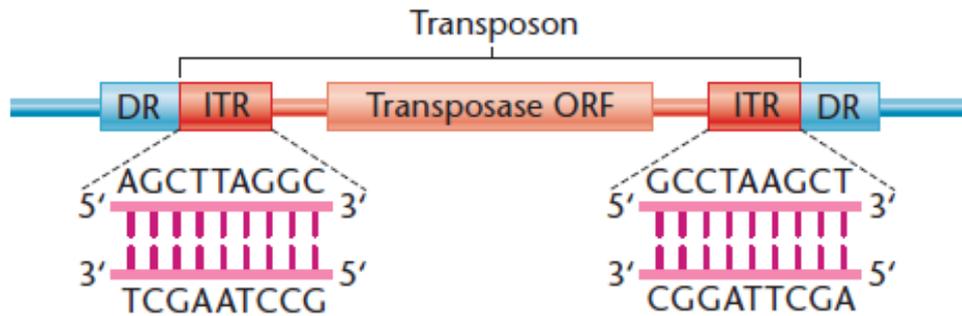


Figure 7. Caractéristiques structurales des transposons de classe II.

Les transposons d'ADN, représentés en rouge, contiennent un cadre de lecture ouvert (ORF) qui code pour l'enzyme Transposase. Certains transposons d'ADN contiennent également des ORF codant pour d'autres protéines en plus de la Transposase. Les répétitions terminales inversées (ITR), sont de courtes séquences d'ADN qui sont inversées les unes par rapport aux autres. Des répétitions directes (DR, montrées en bleu) encadrent le transposon d'ADN dans l'ADN chromosomique.

Certains transposons de la bactérie contiennent des gènes de résistance aux antibiotiques (Figure 8).

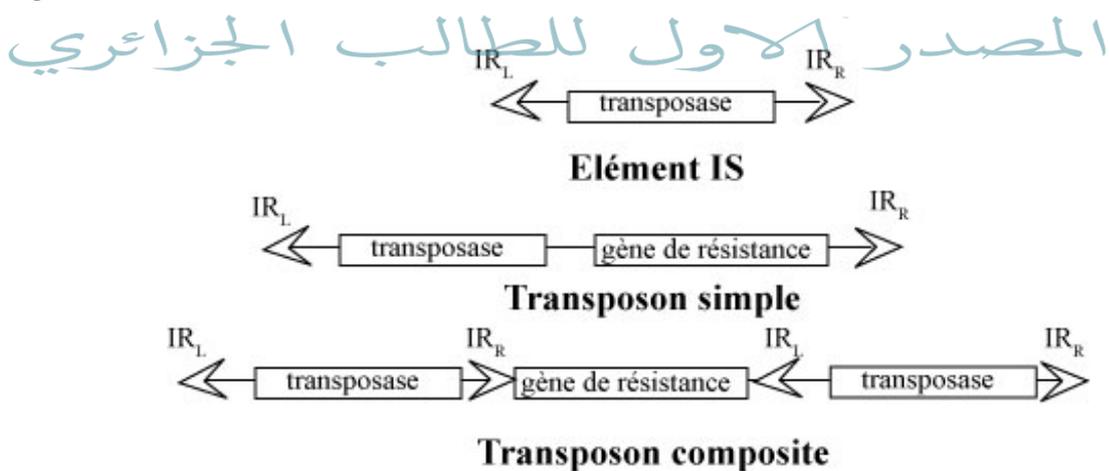


Figure 8. Quelques transposons de classe II.

Génétique évolutive**MUTATIONS****1- Evolution des génomes**

Les transposons (en particulier les rétro transposons qui se répandent dans les génomes par copié-collé) sont responsables au moins en partie de l'augmentation de la taille des génomes (10% pour les eucaryotes supérieurs).

Chez les bactéries, ils permettent la transmission de gènes de résistance.

L'expression de certains gènes peut être altérée suite à l'intégration d'un transposon dans un gène ou à proximité (effet délétère de la transposition). Un gène peut ainsi être transformé en pseudogène, par changement du cadre de lecture ou altération de la séquence d'acides aminés. La restauration de la fonctionnalité est possible à condition qu'il y ait par la suite une excision parfaite de l'élément transposable, ce qui est rare. Le plus souvent le transposon s'excise de manière incomplète, conduisant à l'insertion ou à la délétion de nucléotides. Il peut parfois y avoir augmentation de l'expression de certains gènes en aval de l'insertion d'éléments transposables, s'ils contiennent des séquences régulatrices ou promotrices.

Exemple :

- Chez la drosophile *Drosophila melanogaster*, l'insertion d'un transposon *hobo* dans l'intron 3-4 du gène *vestigial* donne un phénotype alaire (forme de l'aile) mutant ne permettant pas de voler. Une délétion de cet élément transposable restitue un phénotype alaire similaire au phénotype sauvage.

- L'insertion d'un transposon dans la séquence codante d'un gène inactive ce gène. Ceci a été observé chez le tabac au niveau de l'apoenzyme de la Nitrate Reductase (NR) codée par le gène *nia2*. L'insertion du Rétrotransposon à LTR « Tnt1 » dans une région codante de ce gène entraîne un arrêt de la transcription.

- Les transposons sont plus représentés dans les parties régulatrices des gènes. Chez les plantes on a trouvé que plus de 120 gènes possédaient dans leur région régulatrice les transposons « *Tourist* » et « *Stowaway* ». Chez l'homme 27% des gènes étudiés renferment des transposons dans les UTRs 5' ou 3' pouvant créer un nouveau promoteur, une région régulatrice ou un nouveau site d'initiation pour la traduction.

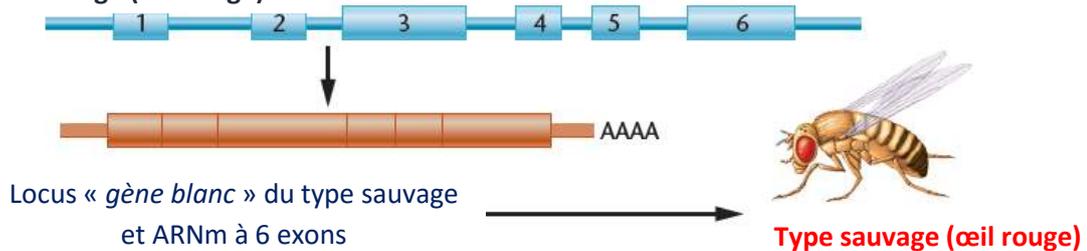
- Le gène CYP19 code l'aromatase P450, enzyme clé dans la synthèse des œstrogènes qui est exprimée dans les gonades, le cerveau et le tissu adipeux de la plupart des mammifères. Chez les primates ce gène est exprimé dans le placenta. L'expression dans les différents tissus est liée à la présence de trois promoteurs alternatifs. Chez les primates un quatrième promoteur a été mis en évidence en amont du gène qui correspond à une LTR de Rétrotransposon. La présence de ce rétro transposon chez les singes de l'ancien et du nouveau monde suggère qu'il a été acquis durant l'évolution des primates, permettant de mettre en place un promoteur spécifique du placenta contrôlant ainsi le niveau d'œstrogène durant la grossesse.

Génétique évolutive

MUTATIONS

- En 1975, une classe de rétrotransposons désignés « *copia* » a été identifiée chez *Drosophila melanogaster*. Les éléments « *copia* » sont présents en 10 à 100 copies dans les génomes des cellules de drosophile. Elles sont transposables à différents emplacements chromosomiques et sont dispersées dans tout le génome. Chaque élément « *copia* » se compose d'environ 5000 à 8000 pb d'ADN, y compris une longue séquence de répétition terminale (LTR) de 267 pb à chaque extrémité. Comme les autres rétrotransposons LTR, la transcription de l'élément « *copia* » commence dans 5'LTR, qui contient un promoteur et un site de début de transcription. La transcription est clivée et polyadénylée (ajout de la queue poly-A) dans 3' LTR, qui contient un site de polyadénylation. Ces caractéristiques permettent au rétrotransposon d'être transcrit par l'ARN polymérase de la cellule. L'insertion de « *copia* » dépend de la présence des séquences LTR et semble se produire préférentiellement à des sites cibles spécifiques dans le génome. Les éléments « *copia* » donnent des effets régulateurs après leur insertion dans le chromosome. Certaines mutations, y compris celles affectant la couleur des yeux et la formation de segments, sont dues à des insertions de « *copia* » dans les gènes. L'une des premières descriptions des effets du « *copia* » est venue de la recherche sur la mutation « *abricot-blanc* » chez la drosophile. Cette mutation est apparue spontanément pour la première fois en 1923 et a changé la couleur des yeux de la drosophile d'un rouge sauvage à une couleur jaune orangé. Des études ultérieures de séquençage d'ADN ont démontré que la mutation était causée par une insertion de « *copia* » dans le deuxième intron du gène codant la « couleur blanche ». À la suite de cette insertion, la plupart des transcrits provenant du promoteur du gène « couleur blanche » se terminent prématurément dans le 3' LTR du rétrotransposon « *copia* ». Ces transcrits prématurément terminés ne codent pas pour le produit du gène blanc fonctionnel, ce qui entraîne une perte de pigment rouge dans l'œil. Parce que certains transcrits de gènes blancs lisent l'élément « *copia* », suffisamment de produits géniques blancs sont produits pour produire un œil de couleur orange clair (Figure 9).

a) Type sauvage (œil rouge)



b) Type mutant (œil de couleur orange)

Insertion de « *copia* » dans le second intron

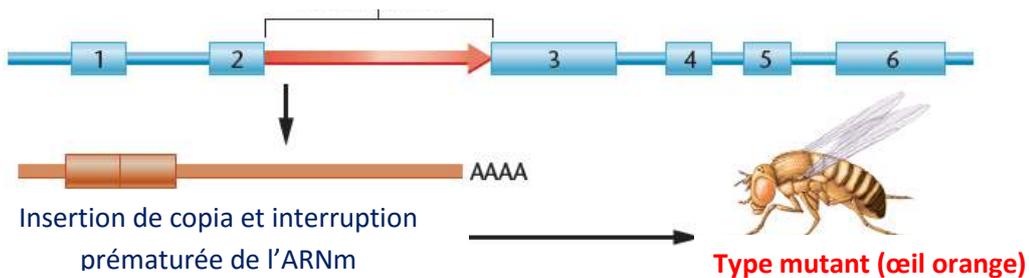


Figure 9. Effets de l'insertion de *copia* dans le gène blanc

Génétique évolutive

MUTATIONS

- Au début des années 1990, le gène *jingwei*, a été découvert et décrit. Des analyses génétiques, de biologie moléculaires et phylogénétiques ont montré qu'il s'agissait d'un nouveau gène fonctionnel apparu il y a environ 2,5 millions d'années chez l'ancêtre commun de deux espèces de drosophiles africaines, *Drosophila yakuba* et *Drosophila teissieri* (**Figure 10**).

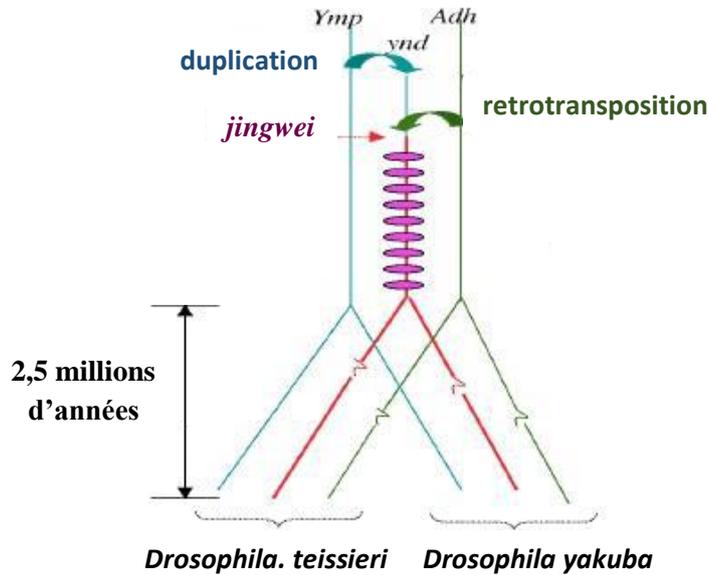


Figure 10. Histoire évolutive du gène *jingwei*

Dans l'espèce ancestrale, il y avait deux gènes à copie unique, le *yellow-emperor* (*ymp*) et l'*Adh*, qui code pour la protéine *alcool déshydrogénase*. Le gène *ymp* a été dupliqué en deux exemplaires: l'un également appelé *ymp* et l'autre appelé *yande* (*ynd*). Alors que *ymp* a conservé ses fonctions d'origine, *yande* a été impliqué dans l'origine du gène *jingwei*. L'ARNm d'*Adh* a été retrotransposé dans le troisième intron de *yande* en tant qu'exon fusionné et recombinaison avec les trois premiers exons de *yande*. Ce qui a eu pour résultat, la formation du nouveau gène *jingwei*, dont la traduction produit une protéine chimérique (mélange de 2 gènes) qui s'exprime dans le testicule des drosophiles mâles (**Figure 11**).

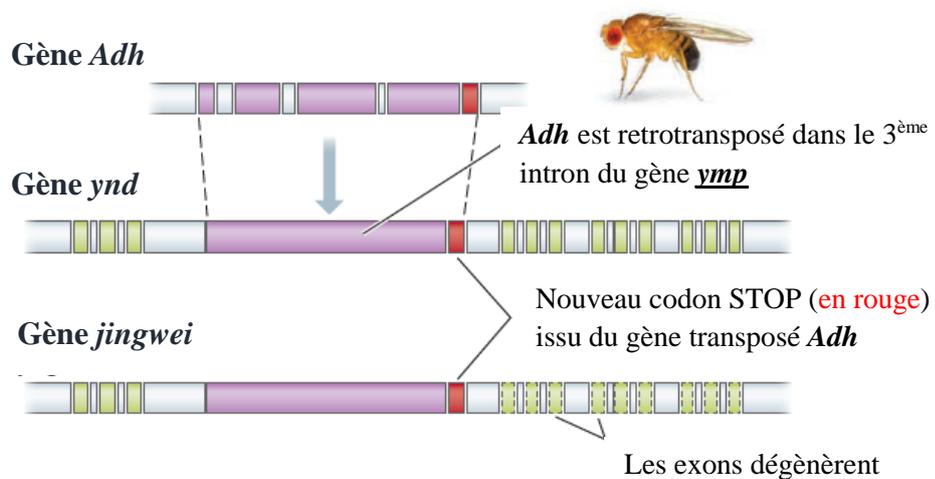


Figure 11. Formation du gène *jingwei*

Génétique évolutive**MUTATIONS**

- Les éléments transposables peuvent contribuer à des réarrangements de portions du génome:

- Inversions
- Translocations
- Duplications

Les transposons peuvent parfois augmenter le taux de mutation dans leur voisinage.

- Chez l'homme 44% du génome est constitué d'éléments transposables (**Tableau 1**).

Tableau 1. Éléments transposables dans le génome humain (Klug *et al.*, 2019)

Element Type	Length	Copies in Genome	% of Genome
LINEs	1–6 kb	850,000	21
SINEs	100–500 bp	1,500,000	13
LTR elements	<5 kb	443,000	8
DNA transposons	80–300 bp	294,000	3
Unclassified	—	3,000	0.1

Conclusion :

Les transposons font partie des séquences moyennement répétées du génome. Au niveau évolutif, ils peuvent générer de la variabilité génétique sur laquelle la sélection naturelle pourra agir. Ils peuvent moduler l'expression des gènes. Ils jouent un rôle important dans la plasticité des génomes.



Introduction :

Chaque population possède des variations génétiques qui sont obtenues par mutation génétique, recombinaison génétique, duplication de gènes, etc... Ces variations génétiques fournissent les éléments nécessaires aux forces évolutives pour créer des espèces mieux adaptées et plus efficaces. Le polymorphisme génétique se reflète à tous les niveaux dans les populations, par exemple aux niveaux phénotypique, chromosomique, protéique et ADN.

En 1931 le chimiste Arthur L. Fox découvre que la molécule **PhénylThioCarbamide (PTC)**, a un goût très amer pour certaines personnes, mais aucun goût pour d'autres, La variation parmi les individus d'un caractère tel que la capacité de goûter au PTC est un **polymorphisme**. La sensibilité au **PTC** dépend d'un gène à 2 allèles : un allèle dominant, qui confère la sensibilité, et un allèle récessif, associé au phénotype d'insensibilité. En 2003 le gène du récepteur au PTC est cloné. Nommé **TAS2R38** il est localisé sur le chromosome 7, en position 7q34. Il code pour un récepteur membranaire à 7 domaines transmembranaires, couplé aux protéines G. Il fait partie de la famille des récepteurs **TAS2R**, qui compte au moins 25 gènes fonctionnels connus. Ces gènes sont portés par les cellules sensorielles des bourgeons du goût situés sur les papilles linguales (sur la langue). Ils sont impliqués dans la détection des substances amères, de la famille des glucosinolates, en particulier celles contenues dans les Brassicacées (la famille des choux). Trois variants ponctuels de l'ADN (SNP, Single Nucleotide Polymorphism) ont été identifiés dans ce gène :

Position du nucléotide	Changement de nucléotide	Changement de codon	Changement d'acide aminé	Position de l'acide aminé
145	C -> G	CCA -> GCA	Proline -> Alanine	49
785	C -> T	GCT -> GTT	Alanine -> Valine	262
886	G -> A	GTC -> ATC	Valine -> Isoleucine	296

L'allèle conduisant à la combinaison d'acides aminés **Proline-Alanine-Valine (PAV)** est celui qui confère la sensibilité. L'allèle conduisant à la combinaison **Alanine-Valine-Isoleucine (AVI)** produit un récepteur membranaire qui ne fixe plus le PTC.

Un autre exemple de variation génétiques (**polymorphisme**) est illustré par les groupes sanguins A, B, O chez l'homme, qui sont déterminés par 3 allèles (I_A , I_B , I_O). Ces trois allèles coexistent chez la plupart des populations, qui sont donc **polymorphes** pour ce locus.

I- Définitions :

Le polymorphisme exprime une variabilité intra populationnelle (chez les individus d'une même population).

Il existe des polymorphismes nucléotidiques (révélés au niveau des sites nucléotidiques de la séquence).

Des polymorphismes biochimiques (affectant la structure d'une enzyme).

Des polymorphismes chromosomiques (affectant la structure des chromosomes) correspondant à des remaniements non associés à des états délétères (n'entraîne pas la mort de l'individu porteur de

Un locus est polymorphe lorsqu'il présente plus d'un allèle.

II- Variabilité génétique :

Bien qu'il peut sembler qu'une population bien adaptée à son environnement doit être homozygote, car on suppose que l'allèle le plus favorable à chaque locus est présent à une fréquence élevée, la plupart des populations contiennent un degré élevé d'hétérozygotie. La plupart des populations de plantes ou d'animaux présentent de nombreuses similitudes phénotypiques entre les individus, cependant la variation génétique n'est pas nécessairement apparente dans le phénotype (**Figure 1**); par conséquent, sa détection n'est pas simple. Néanmoins, l'ampleur de la variation au sein d'une population peut être révélée par plusieurs méthodes. Pour cela on utilise des marqueurs. Le marqueur génétique, de nature biochimique, chromosomique ou moléculaire, permet de révéler un polymorphisme.



Figure 1. Les papillons de l'espèce *Panaxia dominula* diffèrent par les marques sur leurs ailes

Avant le développement des techniques de biologie moléculaire ou d'analyse moléculaire, il existait un moyen de révéler la variation génétique au sein de populations animales ou végétales.

1- Sélection artificielle :

Une façon de déterminer s'il existe une variation génétique dans une population est d'utiliser la sélection artificielle. S'il y a peu ou pas de variation présente dans les gènes contrôlant un phénotype particulier, la sélection aura peu ou pas d'effet sur le phénotype. Si une variation génétique est présente, le phénotype changera sur quelques générations. Un exemple est le chien domestique. Le large éventail de tailles, formes, couleurs et comportements observés chez différentes races de chiens résulte des effets de la sélection sur la variation génétique présente chez les loups, dont tous les chiens domestiques sont les descendants (**Figure 2**). L'utilisation de l'élevage sélectif a permis de créer des centaines de races de chiens dans l'Angleterre du dix-neuvième siècle, sur une période de moins de 75 ans.



Figure 2. La différence de taille entre un Chihuahua et un grand Danois illustre le degré élevé de variation génétique présent dans le génome du chien.

2- Polymorphisme des protéines:

L'avènement des nouvelles techniques de biologie moléculaire, telles l'électrophorèse sur gel mise au point par Lewontin et Hubby à la fin des années 1960, a permis de mettre en évidence qu'une grande proportion des locus enzymatiques était polymorphe. On appelle **alloenzymes** les enzymes que l'on différencie par leur mobilité électrophorétique, c'est-à-dire par leur vitesse de migration, déterminée par leur charge ionique totale de la molécule. Ces enzymes ont la même fonction et sont codées par un seul gène (locus). Chaque **alloenzyme** représente un variant allélique à ce locus.

Exemple :

- Parmi les différentes espèces de drosophile, il existe une grande variation des niveaux de polymorphismes des allozymes qui dépend de l'espèce, de la région, du nombre et de la nature des loci en question, etc...

Les polymorphismes d'allozymes simples, sont souvent définis par un seul changement d'acides aminés, c'est le cas des enzymes suivantes : **ADH** (alcohol dehydrogenase), **G6PD** (glucose-6-phosphate dehydrogenase), **TPI** (triosephosphate isomerase), **GPDH** (glycerol-3-phosphate dehydrogenase), **6PGD** (6-phosphogluconate dehydrogenase). Par contre **PGM**

Polymorphisme

(phosphoglucomutase) montre un grand nombre de changements d'acides aminés associés aux trois allèles allozymatiques communément identifiés.

- Chez le poisson *Fundulus heteroclitus*, un petit poisson (5 à 10 cm de long) qui vit dans les criques, les baies et les estuaires le long de la côte atlantique de l'Amérique du Nord, de la Floride à Terre-Neuve, une étude a mesuré la fréquence des allèles au locus codant pour l'enzyme lactate déshydrogénase-B (LDH-B), qui est produite dans le foie, le cœur et le muscle squelettique rouge. La LDH-B convertit le lactate en pyruvate et joue donc un rôle central dans la fabrication du glucose et le métabolisme aérobie. Deux allozymes de LDH-B sont observées sur des gels d'électrophorèse; ils diffèrent en deux positions d'acides aminés. Les allèles codant pour les allozymes sont appelés **Ldh-B^a** et **Ldh-B^b**.

Dans les populations du nord, où la température moyenne de l'eau est d'environ 6 °C, le **Ldh-B^b** prédomine. Dans les populations du sud, où la température moyenne de l'eau est d'environ 21 °C, **Ldh-B^a** prédomine.

L'enzyme codée par **Ldh-B^b** a une efficacité catalytique plus élevée que Ldh-B^a à basse température, alors que le produit génique de Ldh-B^a est plus efficace que Ldh-B^b à haute température (**Figure 3**). Un mélange des deux formes a une efficacité intermédiaire à toutes les températures.

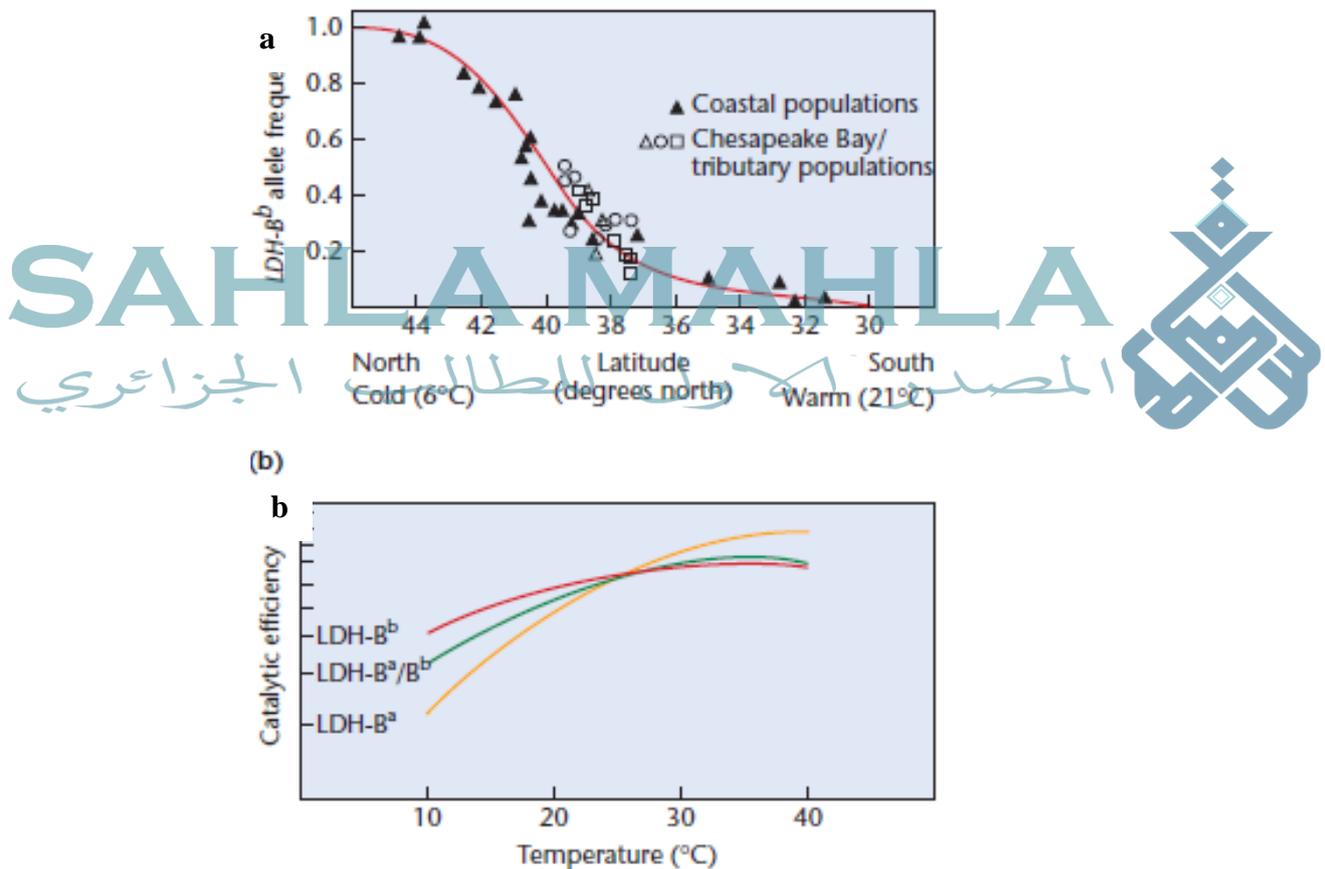


Figure 3. Variation de la structure génétique parmi les populations de poissons

- a) Fréquences de l'allèle B^b dans les populations le long de la côte atlantique de l'Amérique du nord.
- b) Efficacité catalytique des allozymes LDH-B en fonction de la température.

Les différences dans l'efficacité catalytique et le taux de transcription des deux allèles Ldh-B sont

cohérentes avec l'hypothèse que les populations de poissons *Fundulus heteroclitus* sont adaptées aux températures auxquelles elles vivent.

3- Variations nucléotidiques :

- Parmi les gènes humains les plus étudiés, le locus codant pour la protéine CFTR (le régulateur de conductance transmembranaire de la fibrose kystique (Cystic Fibrosis Transmembrane Conductance Regulator)). Les mutations perte de fonction du locus CFTR provoquent la **mucoviscidose**, une maladie qui provoque une production en excès de mucus des les poumons, et une sensibilité élevée aux infections bactériennes. Plus de 1900 mutations différentes du gène CFTR ont été identifiées. Parmi celles-ci, il y a des mutations faux-sens, des délétions d'acides aminés, des mutations non-sens, des décalages de cadre et les anomalies d'épissage. La **figure 4** montre une carte des 27 exons du gène CFTR, avec la plupart des exons identifiés par leur fonction. L'histogramme au-dessus de la carte montre les emplacements de certaines des mutations pathogènes et le nombre de copies de chaque mutation qui ont été identifiées. Une seule mutation, une délétion de 3 paires de bases dans l'exon 10 appelée $\Delta F508$, représente 67 pour cent de tous les allèles de mucoviscidose mutants, mais plusieurs autres mutations étaient également présentes dans au moins 100 des chromosomes étudiés. Dans les populations d'ascendance européenne, entre 1 sur 44 et 1 sur 20 individus est hétérozygote porteur de l'allèle mutant de la maladie.

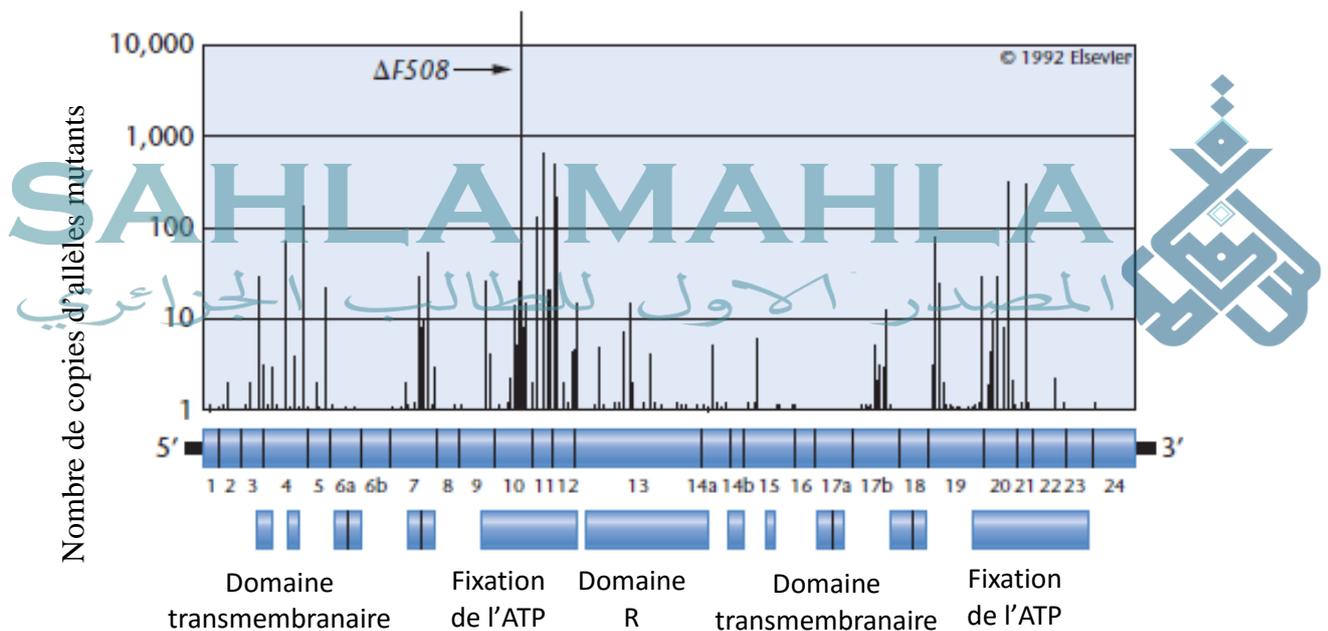


Figure 4. Localisation des mutations responsables de la mucoviscidose dans le gène CFTR. L'histogramme montre le nombre de copies de chaque mutation identifiée. (L'axe vertical est une échelle logarithmique.) La carte génétique sous l'histogramme montre l'emplacement et la taille relative des 27 exons du locus CFTR. Les cases en bas indiquent les fonctions des différents domaines de la protéine CFTR.

- Le locus de l'alcool déshydrogénase (*Adh*) chez *Drosophila melanogaster* (**Figure 5**) a deux allèles, les allèles *Adh-f* et *Adh-s*. Les protéines codées ne diffèrent que par un seul acide aminé (Thr ou Lys au codon 192). Le clonage et le séquençage des gènes *Adh* d'individus représentant cinq populations naturelles de drosophile (**Figure 6**) a montré que les 11 loci clonés de ces cinq populations contenaient

Polymorphisme

un total de 43 variations nucléotidiques dans la séquence *Adh* de 2721 paires de bases. Ces variations sont réparties dans tout le gène: 14 dans les régions codant pour les exons, 18 dans les introns et 11 dans les régions flanquantes et non codantes. Sur les 14 variations dans les régions codantes, une seule conduit à une substitution des acides aminés - celui du codon 192, produisant les deux allèles. Les 13 autres substitutions de nucléotides de la région codante ne conduisent pas à des changements d'acides aminés et sont des variations silencieuses de ce gène.

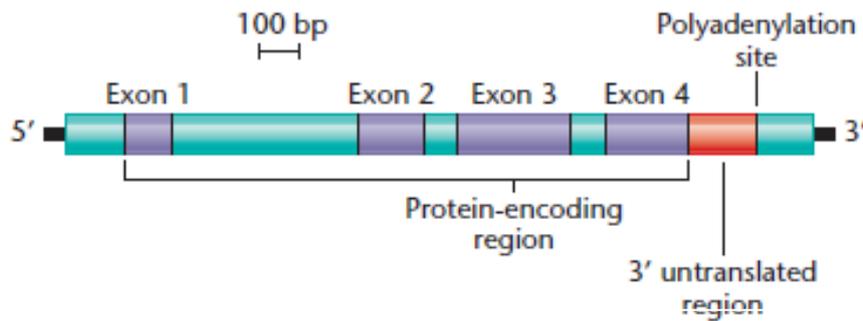


Figure 5. Locus du gène *Adh* de *Drosophila melanogaster*

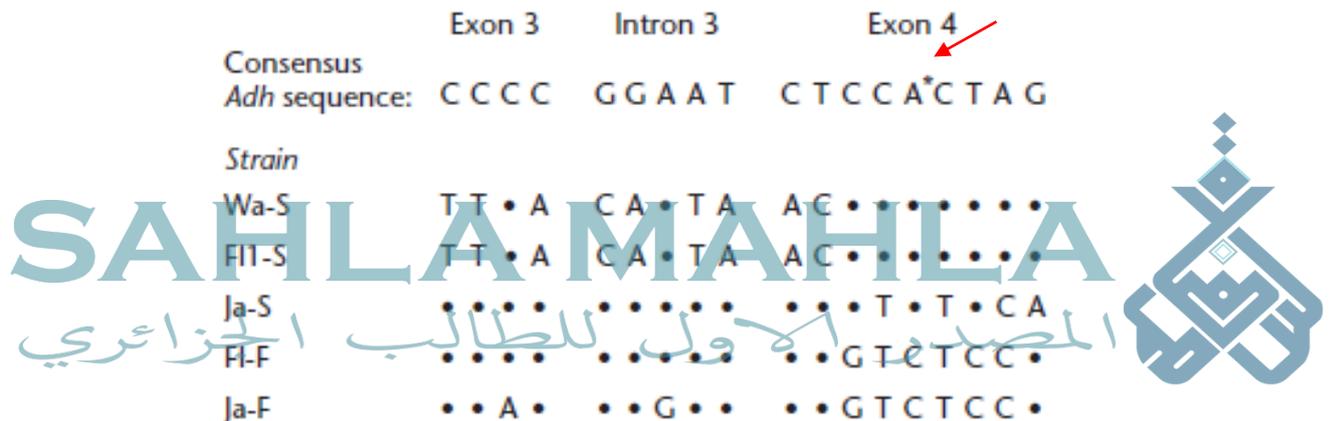


Figure 6. Variations de l'ADN dans certaines parties du gène *Adh* de la drosophile

Les points représentent des nucléotides identiques à la séquence consensus; les lettres représentent les polymorphismes nucléotidiques. Un polymorphisme A / C (A *) dans le codon 192 crée les deux allèles *Adh* (*F* et *S*). Tous les autres polymorphismes sont silencieux ou non codants.

4- Variations chromosomiques :

La mouche *Drosophila pseudoobscura* est trouvée dans de nombreux habitats variés de l'ouest et du sud-ouest des États-Unis. Bien que les mouches dans cette aire de répartition soient morphologiquement similaires, l'équipe de Dobzhansky a découvert que les populations de différents endroits varient dans l'arrangement des gènes sur le chromosome 3 (**Figure 7**). Ils ont trouvé plusieurs inversions différentes dans ce chromosome, qui peuvent être détectées par la formation de boucles dans les chromosomes de larves. Chaque séquence d'inversion est nommée d'après le lieu dans lequel elle a été découverte pour la première fois (par exemple, AR = Arrowhead, Colombie-Britannique; et CH = Chiricahua Mountains, Arizona). Les séquences d'inversion ont été comparées à une séquence standard, désignée arbitrairement ST.

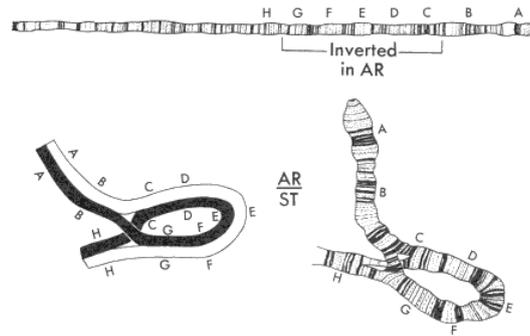


Figure 7. La mouche *D. pseudoobscura* et les inversions du chromosome 3

La **figure 8** compare les fréquences de trois arrangements du chromosome 3 à différentes altitudes dans les montagnes de la Sierra Nevada en Californie. L'arrangement ST est le plus commun à basse altitude; à 2438 mètres, AR est le plus fréquent et ST le moins abondant. Dans les populations étudiées, la fréquence de l'arrangement de CH augmente progressivement avec l'altitude, un phénomène qui est probablement le résultat de la sélection naturelle et qui suit les changements environnementaux graduels se produisant avec l'altitude, comme la baisse de la température de l'air.

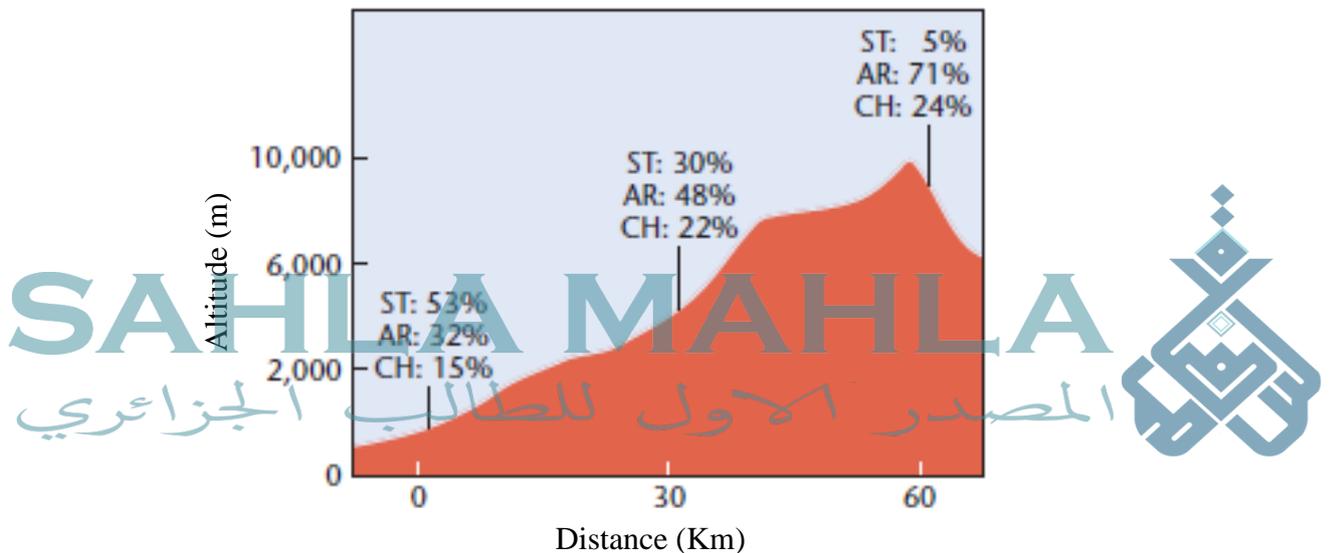


Figure 8. Inversions dans le chromosome 3 de *D. pseudoobscura* à différentes altitudes dans la chaîne de la Sierra Nevada.

L'équipe de Dobzhansky a également constaté que si les populations sont collectées sur un seul site tout au long de l'année, les fréquences d'inversion changent également de manière cyclique au fil du temps. Autrement dit, des variations cycliques des arrangements chromosomiques se produisent au fil des saisons, comme le montre la **figure 9**. Cette variation a été observée de façon constante sur une période de plusieurs années. Au printemps, la fréquence de ST diminue toujours, et celle de CH augmente.

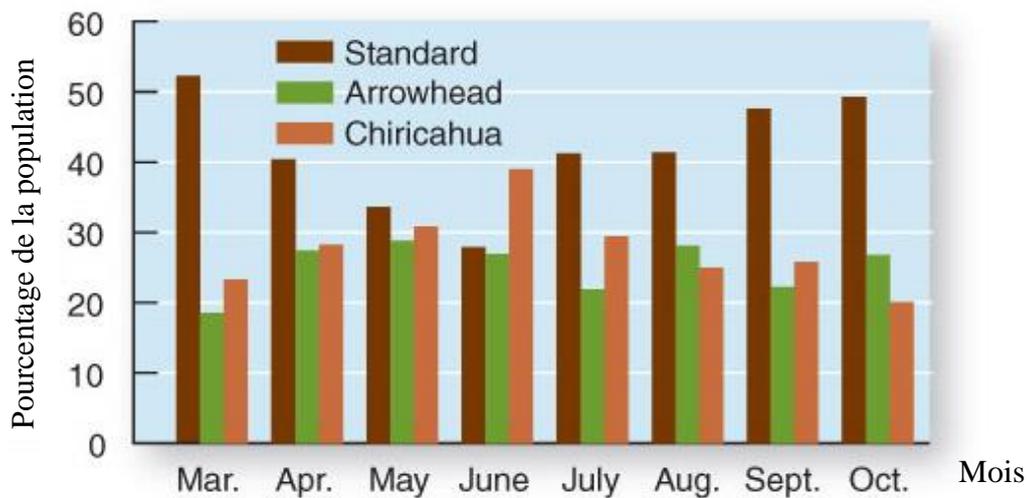


Figure 9. Variations des fréquences des réarrangements ST, AR et CH chez *D. pseudoobscura*

5- Origine de la variabilité génétique dans les populations:

La variation est quasiment présente dans toutes les populations. Elle constitue un réservoir de possibilités offertes aux organismes pour élaborer des stratégies adaptatives, en réponse à des modifications de l'environnement. La **sélection naturelle** était considérée comme la force évolutive qui permettait l'adaptation et la formation d'une espèce meilleure. Cependant, elle agit en favorisant les mutations favorables, ce qui entraîne la perte du polymorphisme.

La **théorie neutre** de l'évolution moléculaire, proposée par **Motoo Kimura** en 1968, explique les niveaux élevés de variation génétique en affirmant que les mutations conduisant à des substitutions d'acides aminés sont généralement délétères, avec une très petite fraction favorable. Certaines mutations sont *neutres* ou fonctionnellement équivalentes à l'allèle remplacé.

Les polymorphismes favorables ou nuisibles sont respectivement préservés ou éliminés de la population par **sélection naturelle**.

Cependant, la fréquence des allèles neutres dans une population sera déterminée par les **taux de mutation** et une **dérive génétique** aléatoire. Certaines mutations neutres dériveront jusqu'à être fixés dans la population (c'est-à-dire maintenues); d'autres mutations neutres seront perdues. À tout moment, la population peut contenir plusieurs allèles neutres dans différents loci.

Il existe plusieurs exemples dans lesquels les variations enzymatiques ou protéiques sont maintenues par adaptation à certaines conditions environnementales. L'avantage bien connu des hétérozygotes de l'anémie falciforme lorsqu'ils sont infectés par des parasites paludéens en est un exemple

Exemple :

L'anémie falciforme est une maladie génétique due à une mutation ponctuelle sur la chaîne β de l'hémoglobine. L'allèle muté est noté S. La solubilité de l'hémoglobine S est diminuée et elle cristallise dans les globules rouges qui prennent une forme de faucille (**Figure 10**). Les homozygotes S/S, développent une anémie très forte qui en général aboutit à la mort des individus. Cette maladie génétique existe en très faible fréquence dans la plupart des populations humaines. Elle est récessive.

Les hétérozygotes A/S ont une espérance de vie semblable à celle des homozygotes normaux A/A (A est l'allèle normal).

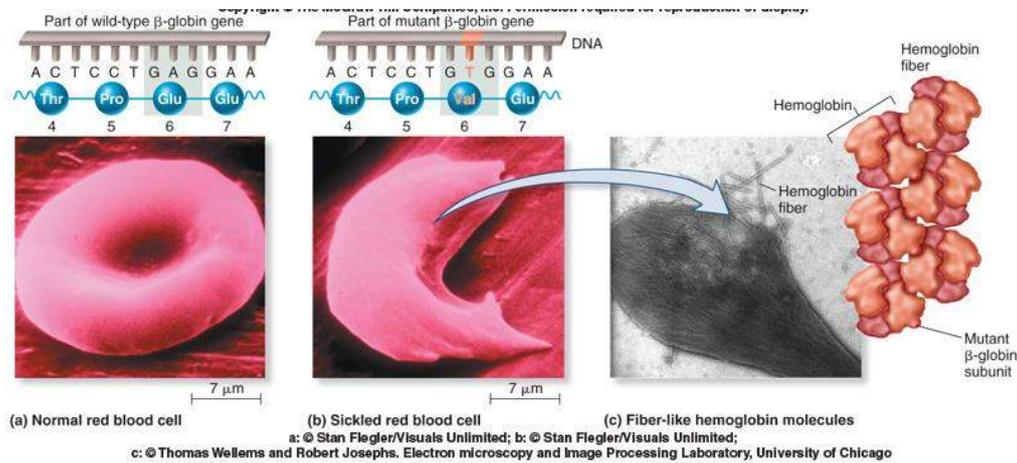


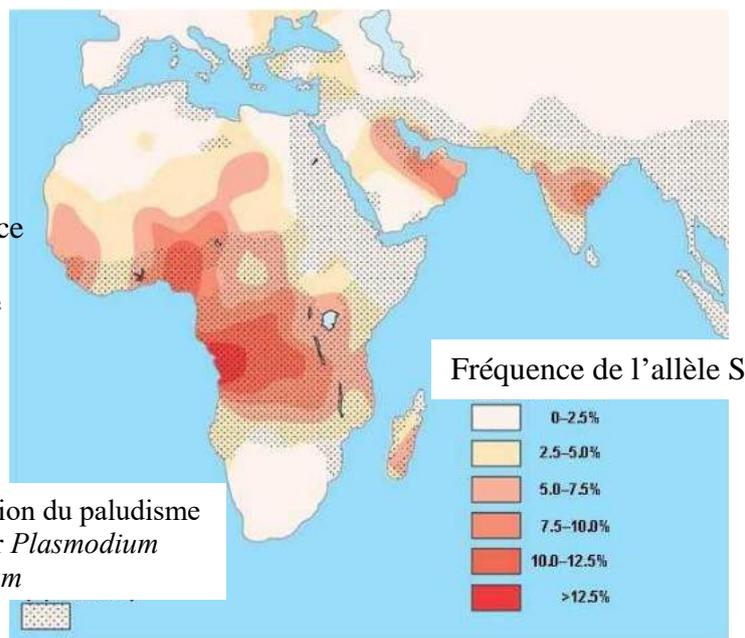
Figure 10. Anémie falciforme provoquée par une mutation ponctuelle

a) Globule rouge normal. b) Globule rouge en forme de faucille. c) Fibre d'hémoglobine

La prévalence de la maladie est cependant élevée en Afrique tropicale ou dans certaines régions d'Asie. Cette prévalence est géographiquement corrélée au paludisme (**Figure 11**). En effet, les individus hétérozygotes A/S sont résistants au parasite *Plasmodium falciparum*, responsable de la maladie.

C'est un cas de superdominance : l'hétérozygote A/S (non anémié et résistant au paludisme) est favorisé par rapport à l'homozygote A/A (non anémié mais sensible au paludisme) ou à l'homozygote S/S (insensible au paludisme mais gravement anémié). Cette situation maintient donc l'allèle S en fréquence non négligeable alors qu'il est pratiquement inexistant dans les régions où il n'y a pas de paludisme. La superdominance est donc un moyen de stabiliser le polymorphisme

Figure 11. Distribution du paludisme et fréquence de l'allèle S de l'anémie falciforme



Distribution du paludisme causé par *Plasmodium falciparum*

Conclusion :

A partir des années 60 et le développement de techniques moléculaires et de méthodes innovantes et performantes, il a été possible de mesurer la variation moléculaire entre organismes à trois niveaux : la variation protéique observée par électrophorèse, la variation des séquences des protéines en acides aminés et celle des séquences d'ADN. Il en est ressorti deux observations majeures :

- 1- Les populations actuelles présentent un polymorphisme important
- 2- Les taux de substitution au sein d'une espèce sont élevés : l'apparition et la fixation de nouveaux variants sont fréquentes.

SAHLA MAHLA

المصدر الأول للطلاب الجزائري



Introduction

La phylogénie a pour but de reconstruire les relations de parenté entre les organismes ou les gènes.

Différentes méthodes permettent de construire des arbres phylogénétiques.

Un arbre phylogénétique des espèces est basée sur la comparaison de l'état de caractères entre espèces, mais les caractères peuvent être de différentes nature : morphologiques, biochimiques, comportementales ou moléculaires (nucléotides ou acides aminés).

Les caractères peuvent correspondre à une **homologie** (similitude héritée d'un ancêtre commun) ou à une **homoplasie** (caractère non hérité d'un ancêtre commun, mais qui peut résulter d'une convergence ou d'une réversion).

I- Représentations phylogénétiques

Les représentations phylogénétiques sont des objets géométriques ayant pour but de modéliser l'histoire évolutive d'un groupe d'organismes. L'arbre phylogénétique répond à la question: " qui est plus proche parent de qui ? ".

Un arbre phylogénétique (**Figure 1**) est formé de nœuds et de branches. Les branches relient les nœuds qui peuvent être internes ou externes. Les nœuds internes correspondant au dernier ancêtre commun hypothétique à partir duquel dérivent des descendants. Les nœuds terminaux correspondent aux **OTUs** (**O**perational **T**axonomic **U**nits, unités taxonomiques opérationnelles).

L'ensemble des branches et des nœuds définissent la topologie (les propriétés) de l'arbre.

L'arbre peut être représenté de deux manières :

- la longueur des branches peut-être proportionnelle au nombre de modifications évolutives qui se sont déroulés le long de cette branche.
- Les OTUs sont alignés et la position des différents nœuds interne sera proportionnelle au temps qui s'est écoulé

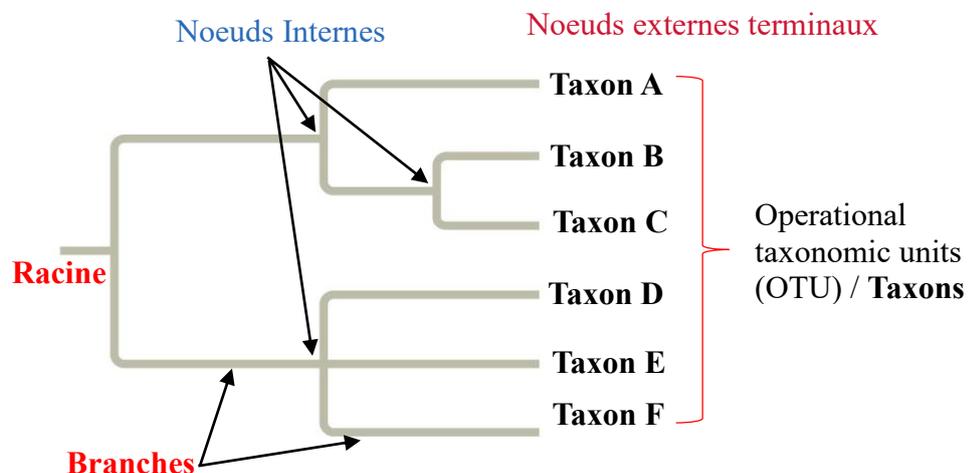


Figure 1. Structure d'un arbre phylogénétique

Un nœud interne et l'ensemble des OTUs qui en dérivent constitue un **groupe monophylétique**. C'est un ensemble de taxons incluant un ancêtre hypothétique et l'intégralité de ses descendants (**Figure 2**).

Un groupe qui inclue un nœud interne et une partie des OTUs qui en dérivent, est un **groupe paraphylétique**. C'est un groupe incluant un ancêtre hypothétique et **une partie** de ses descendants (**Figure 2**).

Un groupe qui rassemble des OTUs et plus d'un nœud interne, est un **groupe polyphylétique**. C'est un groupement de taxons non apparentés présentant des ressemblances (**Figure 2**).

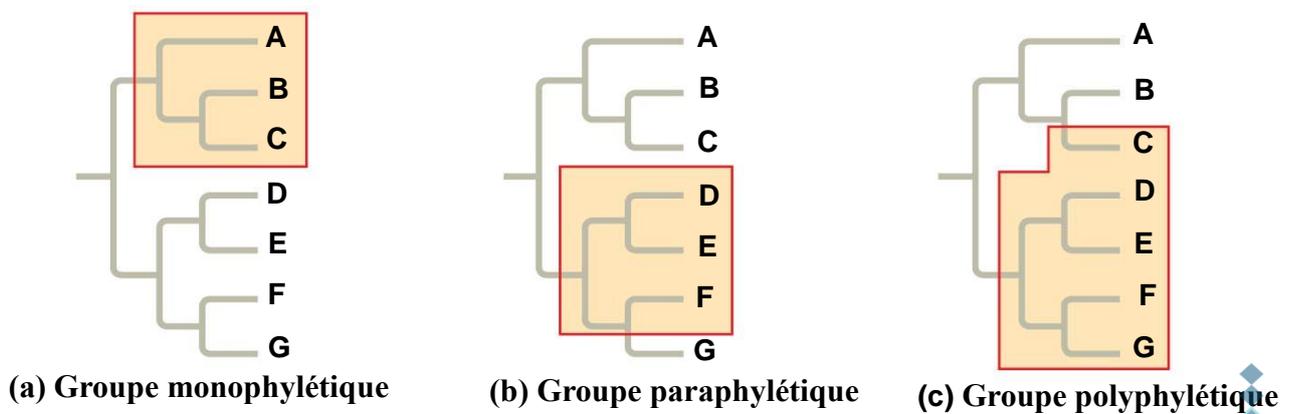


Figure 2. Notions de monophylie, paraphylie et polyphylie

Un arbre peut être enraciné ou pas. La racine indique la direction du temps évolutif vers les OTUs, elle représente l'ancêtre commun de toutes les OTUs. Dans un arbre sans racine, seules les relations entre les OTUs sont représentées, le temps évolutif n'est pas indiqué (**Figure 3**).

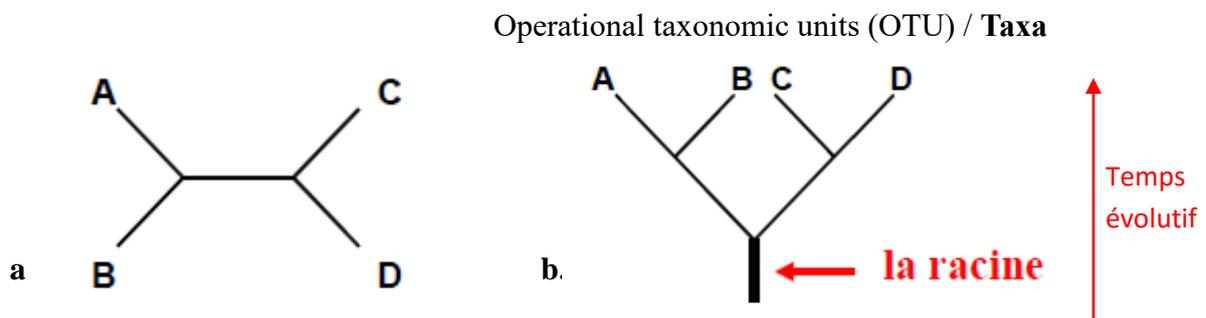


Figure 3. Représentation des arbres
a. Arbres non enraciné. b. Arbre enraciné

1- Notion de similitude et de ressemblance :

La similitude entre deux organismes représente leur degré de ressemblance, mais ne fait aucune hypothèse quant à l'origine de cette ressemblance (ce qui la distingue de la notion d'**homologie**). Ressemblance et apparentement ne sont pas équivalents. Deux organismes étroitement apparentés peuvent être très dissemblables (différents morphologiquement), et inversement, deux taxons phylogénétiquement éloignés peuvent être très semblables. Ce cas de figure est fréquent dans la nature, comme illustré par la morphologie très semblable des taupes européennes (famille des Talpidés) et des taupes dorées africaines (famille des Chrysochloridés), alors que ces dernières sont phylogénétiquement plus apparentées aux éléphants (ordre des Proboscidiens) et font partie avec eux du groupe des Afrotheriens (Delsuc *et al.* 2003) (**Figure 4**).

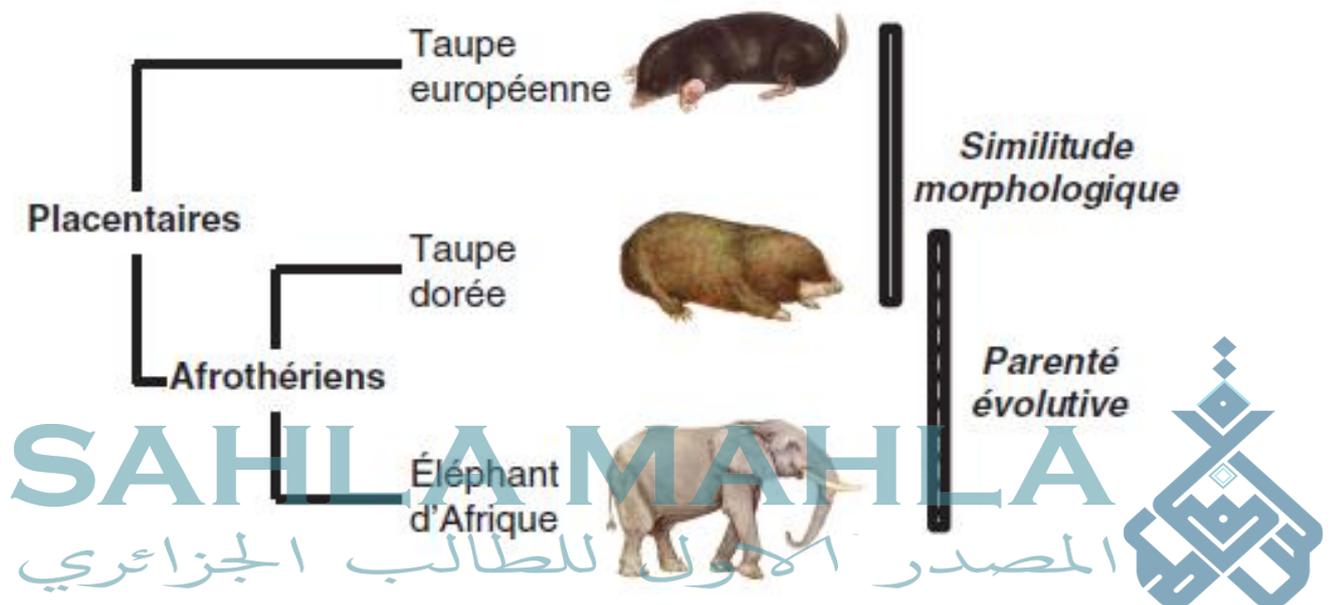


Figure 4. La phylogénie moléculaire a démontré que, du point de vue évolutif, les taupes dorées sont plus apparentées aux éléphants qu'aux taupes européennes, contrairement à ce que leur similitude morphologique pourrait laisser supposer (D'après Delsuc *et al.*, 2003).

2- Propriétés des caractères évolutifs :

Toutes les constructions phylogénétiques, quelles que soient leurs natures (morphologiques, biochimiques, comportementales ou moléculaires) sont basées sur l'utilisation de caractères. Un caractère est une particularité identifiable d'un objet biologique (cet objet pourrait être une espèce ou un gène). L'état du caractère correspond quant à lui à une des possibilités qui est prise par ce caractère. Par exemple, "**couleur de l'œil**" est un **caractère**, tandis que "bleu" est un **état** du caractère "**couleur de l'œil**". La présence et l'absence d'un caractère correspondent également à des possibilités. Ainsi la **présence** ou l'**absence** d'yeux correspond à un des états du caractère "**yeux**".

Dans le cas d'une analyse de séquences nucléotidiques, le caractère correspond au "nucléotide" occupant une position dans la séquence, l'état correspondra à l'identité du nucléotide, c'est-à-

dire un des 4 nucléotides, “A, T, C ou G”. L’absence ou la présence du nucléotide à une position dans la séquence correspond aussi à un état du caractère nucléotide dans la séquence étudiée (**Figure 5**).

human	5'	AACAGACACC	ATG	GTG	CAT	CTG	ACT	CCT	GAG	GAG	AAG	TCT	GCC	GTT	3'
chimpanzee	5'	AACAGACACC	ATG	GTG	CAC	CTG	ACT	CCT	GAG	GAG	AAG	TCT	GCC	GTT	3'
mouse	5'	AACAGACATC	ATG	GTG	CAC	CTG	ACT	GAT	GCT	GAG	AAG	TCT	GCT	GTC	3'
dog	5'	AACAGACACC	ATG	GTG	CAT	CTG	ACT	GCT	GAA	GAG	AAG	AGT	CTT	GTC	3'
codon			↑	1	2	3	4	5	6	7	8	9	10	11	12

Figure 5. Alignement des séquences nucléotidiques de régions conservées de globines provenant de 4 espèces différentes: 6 positions ont des changements synonymes (nucléotides en bleu) et 6 positions avec des changements non synonymes (nucléotides rouges)

Pour pouvoir être utiles, les caractères étudiés doivent être comparables entre les différents objets dont on recherche les liens de parenté. Ces objets doivent donc avoir une ressemblance suffisante pour pouvoir les comparer et poser l’hypothèse qu’il existe un certain nombre de caractères de ces objets qui sont homologues les uns aux autres c’est-à-dire dérivés d’un ancêtre commun.

2.1. Notion d’homologie et d’homoplasie

Le partage d’un état par plusieurs objets, ne signifie pas forcément qu’il a été hérité d’un ancêtre commun (ce qui représente une **homologie**). Parfois des événements indépendants peuvent aboutir à un même état, par hasard, partagé par des objets divers (ce que l’on appelle une **homoplasie**). L’**homoplasie** peut être le résultat d’une **convergence évolutive** ou d’une **réversion évolutive** (l’état de caractère retourne à son état ancestral). (**Figure 6**)

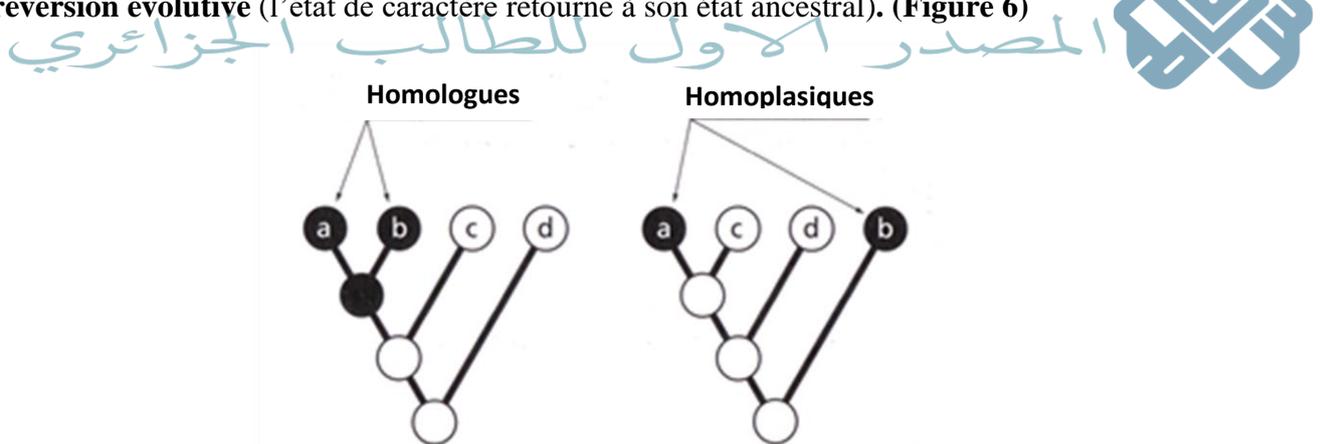


Figure 6. Représentation graphique de l’homologie et de l’homoplasie

Exemples :

L’évolution des ailes reflète l’adaptation des espèces au vol, à partir de membres homologues ou analogues. Dans ce dernier cas, les ailes des chauves-souris et des oiseaux sont le résultat d’une homoplasie, c’est-à-dire une évolution convergente chez deux espèces très éloignées phylogénétiquement (respectivement un mammifère et un oiseau), mais adaptées au vol (**Figure 7**).

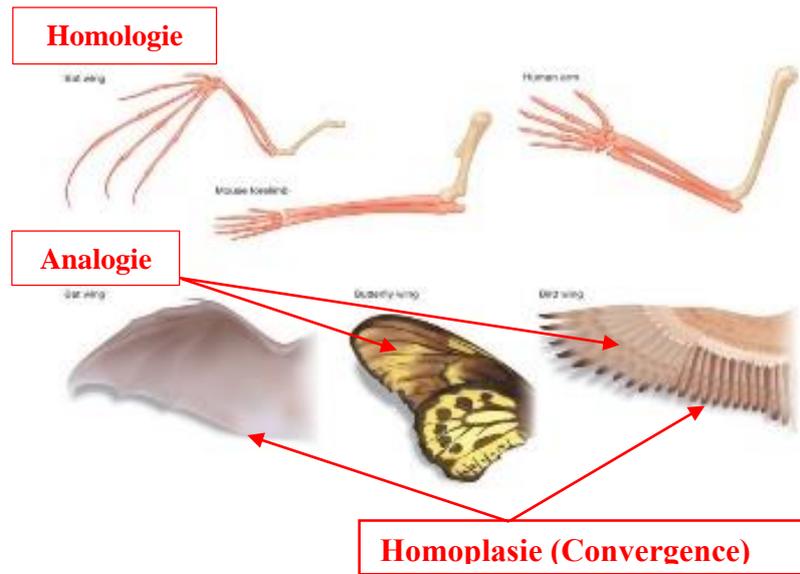


Figure 7. Homologie et homoplasie

2.2. Notions de gènes orthologues et paralogues

Les gènes orthologues sont des gènes qui ont été séparés par un évènement de spéciation. Les gènes paralogues sont issus d'une séparation liée à une duplication. Pour reconstruire l'histoire des organismes, **seuls les gènes orthologues doivent être utilisés**. Par contre, c'est l'ensemble des copies (ensemble des paralogues) qui devra être considéré si le but est de reconstruire l'histoire évolutive de la famille de gènes (Figure 7).

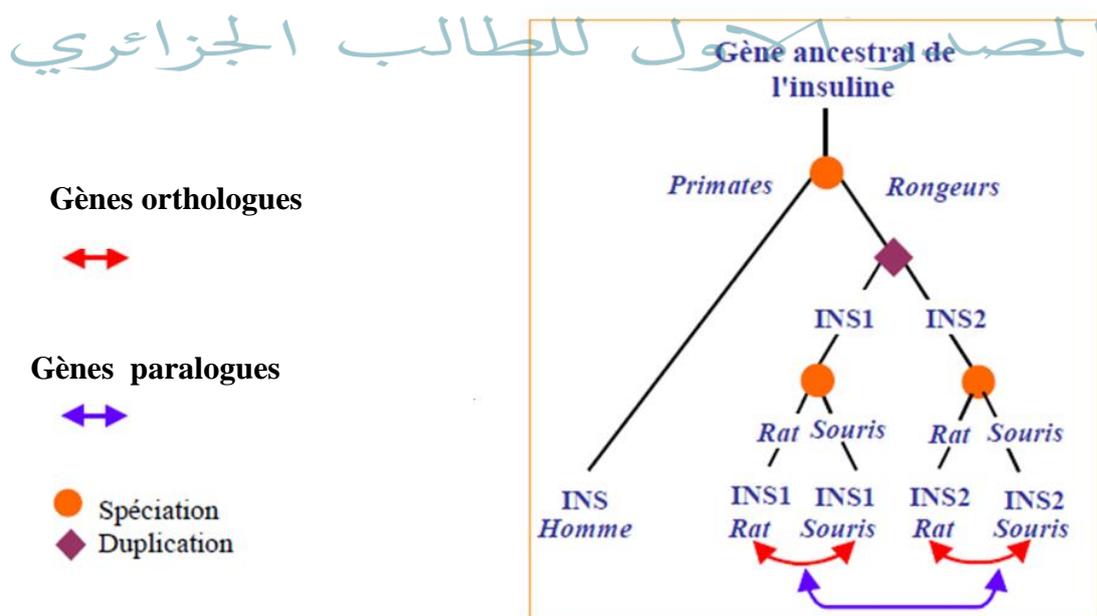


Figure 7. Représentation graphique des gènes orthologues et paralogues : cas de l'évolution des gènes de l'insuline

3. Propriétés des caractères utilisés pour les constructions phylogénétiques

Les **caractères morpho-anatomiques** ont été à la base des premières reconstructions phylogénétiques. Issus de l'anatomie comparée et de l'embryologie, les caractères externes (morphologie) et internes (anatomie) représentent une source d'information relativement accessible. Le squelette interne des vertébrés ou la forme des pièces copulatrices des insectes (genitalia) sont des exemples classiques de caractères morpho-anatomiques utilisables en reconstruction phylogénétique.

La biochimie et la biologie moléculaire constituées principalement par l'analyse des protéines et des acides nucléiques, constituent une autre source d'information. Les phylogénies moléculaires se sont généralisées grâce au développement des techniques de biologie moléculaire, notamment le séquençage.

La force des caractères moléculaires consiste :

a- en leur *abondance* : ces caractères sont disponibles en nombre gigantesque (les génomes de certains virus, bactéries, mammifères et plantes peuvent respectivement comporter des milliers, millions, milliards et dizaines de milliards de paires de bases).

b- en leur *objectivité* : la détermination de leur état est indépendante de l'expérimentateur (s'il s'agit d'analyser des séquences nucléotidiques, il n'y a que 4 états possibles pour les nucléotides, A, T, C ou G ; s'il s'agit d'analyser des protéines il y a 20 états possibles, chacun correspondant à l'identité de l'acide aminé de la séquence polypeptidique, Glu, ou Phe, ou Met, ou Asp, ou His, etc...).

c- leur *ubiquité* (présent partout): l'expérience acquise pour un groupe taxonomique est facilement transposable à un taxon complètement différent, et un même gène ou une même protéine peuvent permettre de comparer des organismes à l'apparence morpho-anatomique complètement différente

d- une partie de la variation moléculaire peut être considérée comme neutre d'un point de vue évolutif, c'est-à-dire qu'elle va s'accumuler sans conséquence biologique (donc avec des risques réduits de convergence). Ceci est typiquement le cas des substitutions en région non codante ou des substitutions synonymes (c'est-à-dire n'affectant pas la séquence protéique) en région codante. (voir théorie neutre de l'évolution dans le cours polymorphisme)

L'analyse phylogénétique repose sur l'analyse de caractères dont nous pouvons *à priori* penser qu'ils sont suffisamment identiques, c'est-à-dire qu'ils sont comparables. Ces caractères constituent des hypothèses d'homologie.

II. Alignement des séquences moléculaires

De même qu'il faut vérifier que les séquences étudiées sont orthologues, il faut aussi s'assurer que les nucléotides (ou acides aminés) ayant la même position dans les différentes séquences sont issus d'un même nucléotide (ou acide aminé) ancestral. Cette étape d'alignement des séquences permet par la suite de comparer des nucléotides qui sont effectivement comparables. Cette étape d'alignement multiple est très délicate, et elle se fait en général à l'aide de logiciels (par exemple *Clustal*). Les séquences alignées constituent les données fondamentales de la reconstruction phylogénétique. L'alignement est donc une étape décisive qui conditionne la qualité de l'inférence phylogénétique (**Figure 8**).

Séquences nucléotidiques avant alignement

Human	ACAT	TATGG	ACAGG	TAAAG	TAAAA	AACAT	TATT
Chimpanzee	ACAT	TATGG	ACAGG	TAAAG	TAAAA	AACAT	TATT
Macaque	ATA	TACATT	ACGG	ACAGG	TAAAG	TAAAA	AACAT

Séquences nucléotidiques après alignement

Human	ACA	TTAT	TGGAC	CAGGT	AAGT	TAAAA	AACAT	TATT
Chimpanzee	ACA	TTAT	TGGAC	CAGGT	AAGT	TAAAA	AACAT	TATT
Macaque	ATA	TACA	TTAC	CGGAC	CAGGT	AAGT	TAAAA	AACAT

Figure 8. L'alignement de trois séquences d'humains, de chimpanzés et de macaques permet de déterminer quelles positions de la séquence doivent être comparées les unes aux autres. Les séquences alignées et non alignées sont identiques, mais les séquences alignées incluent des espaces supplémentaires qui améliorent l'alignement. Dans les séquences non alignées, 15 spots sont identiques; dans les séquences alignées, 22 spots sont identiques. Les spots sont colorés.

III. Méthodes de construction phylogénétiques :

Différentes méthodes aboutissent à la réalisation d'arbres phylogénétiques. Trois principales catégories sont possibles.

Les méthodes de Maximum de Parcimonie (méthode *cladistique*), les méthodes de distance (méthodes *phénétiqes*), et les méthodes probabilistes (Maximum de *vraisemblance*).

1- Méthodes cladistiques (Maximum de Parsimony)

Basée sur l'étude de l'état des caractères. Ces derniers peuvent correspondre à l'état ancestral ou à un état dérivé. L'état du caractère étudié, peut être soit ancestral et être partagé par un ou plusieurs OTUs. Il peut être dérivé et être partagé ou pas par des OTUs.

L'état ancestral d'un caractère est une **plésiomorphie**. Un état dérivé à partir de l'état ancestral est une **apomorphie**. Un état ancestral d'un caractère partagé par plusieurs taxons est une **symplesiomorphie**. Un état dérivé d'un caractère partagé par plusieurs taxons est une **synapomorphie**. Un état dérivé d'un caractère mais présent chez un unique taxon est une **autapomorphie**. La **Figure 9** résume ces différents états.

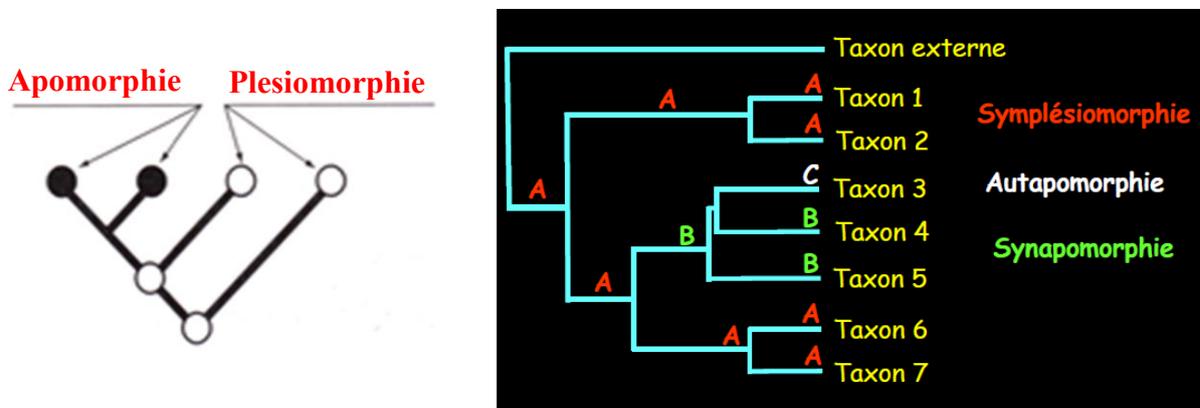


Figure 9. Etat des caractères et leur évolution à travers les taxons au cours du temps

« Lorsque la nature ancestrale ou dérivée du caractère est déterminée », c'est la **polarité** du caractère qui est déterminée.

Les bases de la cladistique (ou classification phylogénétique) sont principalement le fait des travaux de l'entomologiste allemand Hennig (1966). La cladistique recherche les caractères dérivés partagés par différents taxons. Elle est basée sur la notion d'homologie.

Pour retracer l'arbre phylogénétique (**ou cladogramme**), on cherche à **minimiser** le nombre de changements évolutifs. Elle utilise la méthode du **Maximum de parcimonie**. Toutes les données ne sont pas utilisées pour reconstruire l'arbre. On ne tient compte que **des sites (caractères) informatifs**. Un caractère ou un site est invariable si tous les taxons présentent le même état.

Un caractère est variable si il présente au moins deux états, et il est informatif ou pas. Un caractère ou un site est informatif s'il présente plus de deux états et si l'un de ses états est partagé par au moins deux taxons étudiés.

Lorsque le nombre de taxons à étudier est petit, il est possible de rechercher tous les arbres possibles. Pour chaque arbre, est recherché le nombre d'évènements qui permettent de passer d'une séquence à une autre. Ce nombre détermine la **longueur de l'arbre**. Parmi tous les arbres possibles, il n'est retenu que l'arbre qui présente la longueur **la plus petite** (le plus **parcimonieux**).

Par exemple, considérons trois phylogénies possibles et une matrice de données de deux caractères. Les données sont reportées sur chaque arbre de la manière la plus parcimonieuse possible, mais l'un des arbres est clairement plus parcimonieux que les autres. L'arbre 1 ne nécessite que deux changements de caractères, tandis que les arbres 2 et 3 nécessitent trois changements. Parmi ces trois arbres, l'arbre 1 sera sélectionné comme le plus susceptible d'être précis, car il émet l'hypothèse de la trajectoire évolutive la plus simple (**Figure 10**).

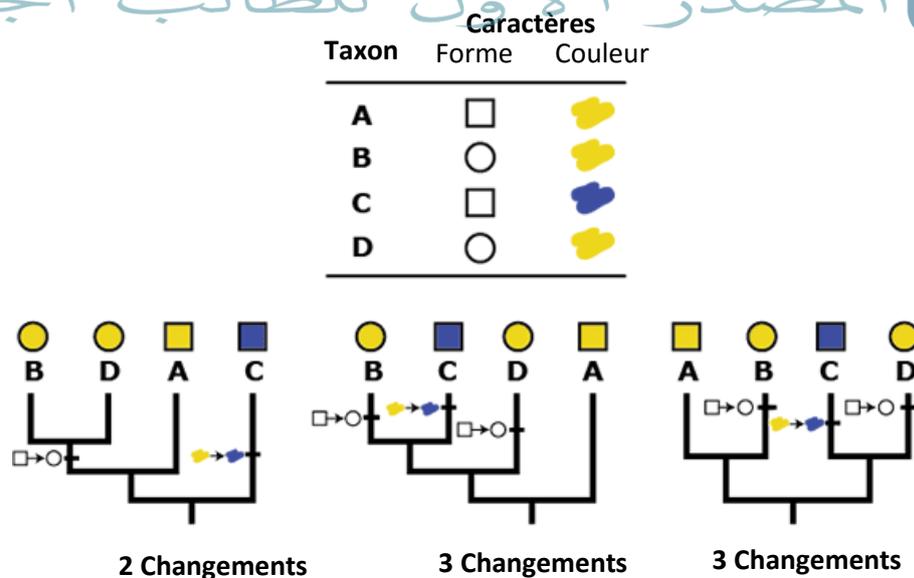


Figure 10. L'arbre qui demande le moins de pas évolutifs (2 changements) est choisi L'arbre le plus à gauche est préféré car il nécessite le moins de changements évolutifs

2- Méthodes Phénétiques (de distance)

Méthodes les plus simples et les plus rapides, mais moins fiables et moins performantes que les autres méthodes. Basée sur l'utilisation d'une matrice des distances (pourcentage de différences de séquence).

Cette méthode ne considère que les ressemblances entre taxons, on ne recherche que **l'état ancestral qui a été hérité d'un ancêtre commun.**

Dans l'arbre phylogénétique (ou **phénogramme**) La longueur des branches est proportionnelle à la distance.

Il existe deux principales méthodes :

- **La méthode UPGMA** (Unweighted Pair Group Method with Arithmetic mean).

Méthode la plus simple. Elle utilise un algorithme de regroupement séquentiel (en série) dans lequel les relations topologiques sont inférées par ordre de similarité décroissante.

- **La méthode de Neighbor-Joining (NJ)** : elle consiste à regrouper les séquences les plus similaires entre elles. Elle donne toujours un seul arbre phylogénétique pour une matrice de distances donnée (**Figure 11**).

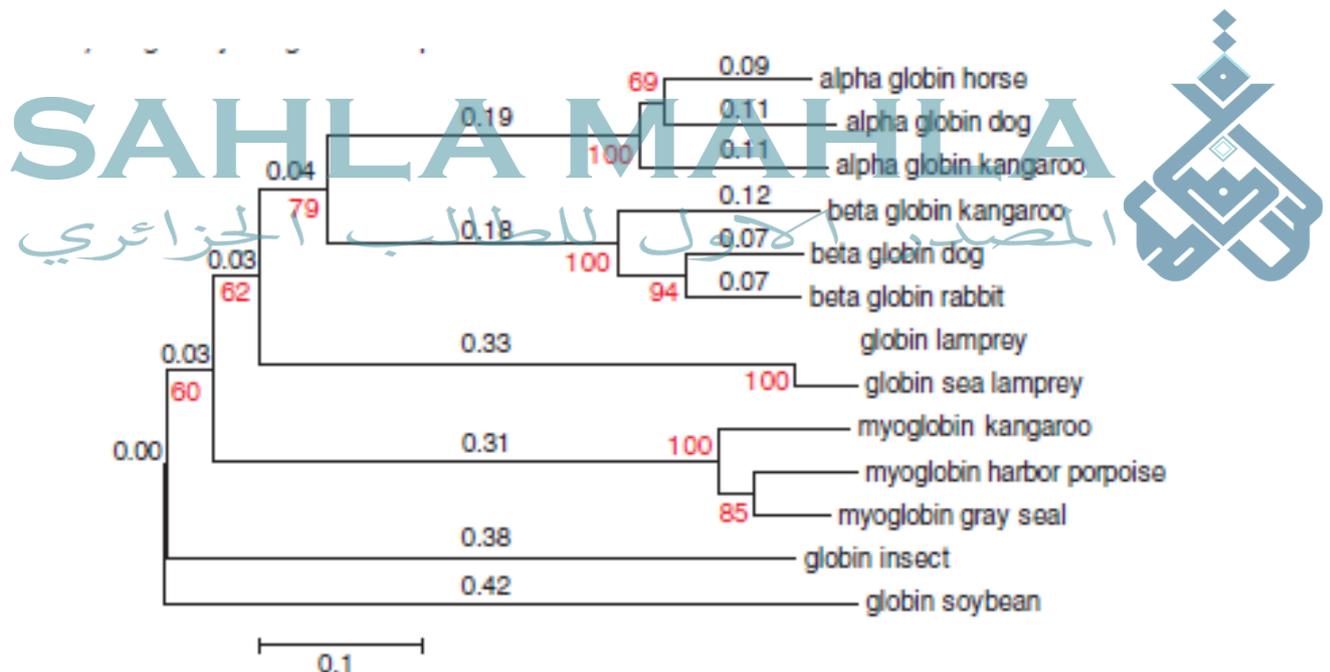


Figure 11. Arbre phylogénétique de 13 protéines de globine créé par la méthode de NJ. Les longueurs des branches expriment les distances évolutives entre séquences (divergences en acides aminés entre les séquences) pour déduire chaque arbre. Le bootstrap (en rouge) indique la robustesse des arbres

3- Méthodes probabilistes

Elle utilise des **méthodes d'évolution**, telle la probabilité de changement d'un nucléotide vers un autre, différente pour les transitions et les transversions. Parmi les arbres possibles, l'arbre retenu est celui qui présente la plus grande **vraisemblance**. La méthode la plus utilisée est la méthode du maximum de vraisemblance (**Maximum of Likelihood**).

Très performante et très fiable, plus efficace que les méthodes cladistiques et de distances, mais demande beaucoup de temps à traiter.

Les méthodes probabilistes sont basées sur des modèles d'évolution. Ces dernières années, les méthodes décrivant les modèles d'évolution ont été améliorés. En effet, pendant longtemps il a été postulé que tous les sites dans une séquence évoluent à la même vitesse, or il est maintenant reconnu que ce n'est pas le cas. Ainsi les acides aminés qui constituent des sites actifs des enzymes évoluent plus lentement. Par contre d'autres sites peuvent évoluer plus vite. Les modèles utilisés actuellement prennent en compte ces différents modes d'évolution.

4. Méthodes Bayésiennes

Les méthodes bayésiennes s'appuient sur l'interprétation dite *subjectiviste* des probabilités. Elles sont basées sur des méthodes statistiques où l'on recherche la plausibilité que l'arbre phylogénétique construit à partir des données moléculaires en possession du chercheur, soit le plus probable. Dans son principe, l'inférence bayésienne ne cherche pas forcément à choisir un arbre. Elle cherche plutôt à décrire la distribution de probabilités *a posteriori* sur l'ensemble des arbres possibles. Elles sont de plus en plus souvent utilisées, car elles prennent en compte l'incertitude liée à l'estimation de la probabilité *a posteriori* de la topologie des arbres reconstruits.

5. Evaluation de la robustesse des arbres phylogénétiques

Quelle que soit la méthode utilisée, il faut tester la validité du résultat obtenu. Une approche très répandue pour mesurer la solidité d'une phylogénie consiste à mesurer celle de chacun de ses nœuds par la technique dite de bootstrap. Il s'agit d'un ré échantillonnage qui repose sur le principe que les sites évoluent de façon indépendante. Ce ré échantillonnage est répété 100 à 10000 fois. Les valeurs du bootstrap sont exprimées en pourcentage. Plus la valeur est élevée, plus l'arbre est dit **robuste**. Il a ainsi été montré qu'à une valeur de bootstrap d'au moins 80 % correspondait une probabilité très élevée que le clade correspondant soit valide (**Figure 12**).

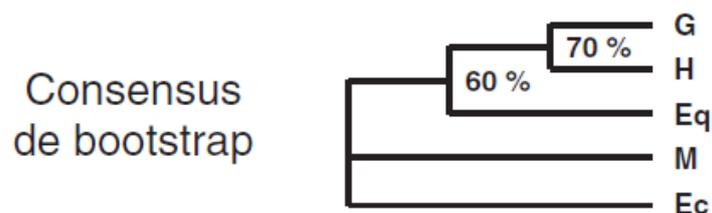


Figure 12. Le bootstrap figure en pourcentage au niveau de chaque nœud

IV. Exemples d'application des phylogénies :

1. Application à la systématique moléculaire d'un rongeur

Un rongeur observé pour la première fois dans les marchés du Laos en Asie et décrit scientifiquement qu'en 2005 comme le représentant d'un nouveau genre et d'une nouvelle espèce de rongeurs, nommée *Laonastes aenigmamus*. Une étude phylogénétique a également permis de révéler que ce genre présentait de nombreuses synapomorphies signant une proche parenté entre *Laonastes* et le genre fossile *Diatomys* (Dawson *et al.* 2006). Les auteurs ont donc proposé d'inclure *Laonastes* au sein de la famille des Diatomyidae dont le dernier représentant s'est éteint il y a environ 11 millions d'années. L'analyse phylogénétique par maximum de vraisemblance de la combinaison de deux gènes mitochondriaux et quatre gènes nucléaires a conduit à rapprocher sans ambiguïté *Laonastes aenigmamus* (Diatomyidae) de la famille des Ctenodactylides à laquelle appartiennent seulement cinq espèces de goundis (petit rongeur d'Afrique) réparties en Afrique. Les méthodes de datation moléculaire (Huchon *et al.* 2007) et les données paléontologiques (Dawson *et al.* 2006) soulignent l'ancienneté de la divergence entre *Laonastes* et les Ctenodactylides qui est estimée autour de 40 à 44 millions d'années. Ceci renforce encore un peu plus le caractère distinctif de cette espèce dont la conservation est aujourd'hui une priorité.

2. Evolution des familles multigéniques

Les gènes qui existent sous de multiples copies résultent d'évènements de duplication et sont qualifiés de paralogues. Ils constituent des familles dites multigéniques. Comme il est possible de reconstruire des phylogénies de taxons il est également possible d'inférer des phylogénies de gènes. Le meilleur exemple est le cas de la famille multigénique de la globine β . Dans le génome humain, il existe 6 gènes de globines localisés sur le chromosome 11 : β et δ s'expriment chez l'adulte et participent à la constitution de l'hémoglobine, ϵ ne s'exprime qu'à l'état embryonnaire, $\gamma 1$ et $\gamma 2$ ne s'expriment qu'à l'état fœtal, tandis que $\psi\eta$ n'est qu'une version non-fonctionnelle, c'est-à-dire un pseudo-gène.

3. L'estimation des états de caractères ancestraux

Les crabes des neiges (genre *Chionoecetes*) et les crabes royaux (genre *Lithodes*) (**Figure 13**), appartiennent à deux groupes phylogénétiquement très distincts. En effet, parmi les Crustacés Décapodes, les crabes des neiges appartiennent à la famille des Majidés au sein de l'infra-ordre des Brachyours ou « crabes-vrais », alors que les crabes royaux sont classés dans la famille des Lithodidés au sein de l'infra-ordre des Anomoures qui comprend aussi les bernard-l'ermite et les galathées (Scholtz et Richter 1995). Par rapport aux « crabes-vrais » dont des fossiles sont connus depuis 380 millions d'années (Ma), les crabes royaux sont en fait d'origine évolutive relativement récente (environ 20 Ma), puisque les premières études moléculaires basées sur l'ARN ribosomique 16S mitochondrial les placent au sein des bernard-l'ermite (Cunningham *et al.* 1992). Les crabes royaux constituent donc un exemple extrême d'évolution morphologique vers la forme crabe à partir d'ancêtres dont le plan d'organisation était de type bernard-l'ermite.

La phylogénie de 26 espèces de Décapodes a été reconstruite en combinant l'information issue des séquences de deux gènes mitochondriaux et deux gènes nucléaires avec celle contenue dans l'ordre des gènes du génome mitochondrial (Morrison *et al.* 2002). La phylogénie obtenue par cette approche montre l'occurrence de cinq évènements indépendants d'évolution de la forme crabe. L'estimation des états de caractères ancestraux sur cette phylogénie par maximum de vraisemblance, permet d'inférer de façon non ambiguë que ces cinq évènements correspondent bien à des phénomènes indépendants de carnisation (formation de la carapace du crabe) à partir d'ancêtres de type non-crabe, tels que les morphotypes crevettes et homards (**Figure 13**). Cet exemple est particulièrement illustratif du pouvoir des données moléculaires pour comprendre l'évolution de la morphologie en révélant des cas extrêmes de convergence.

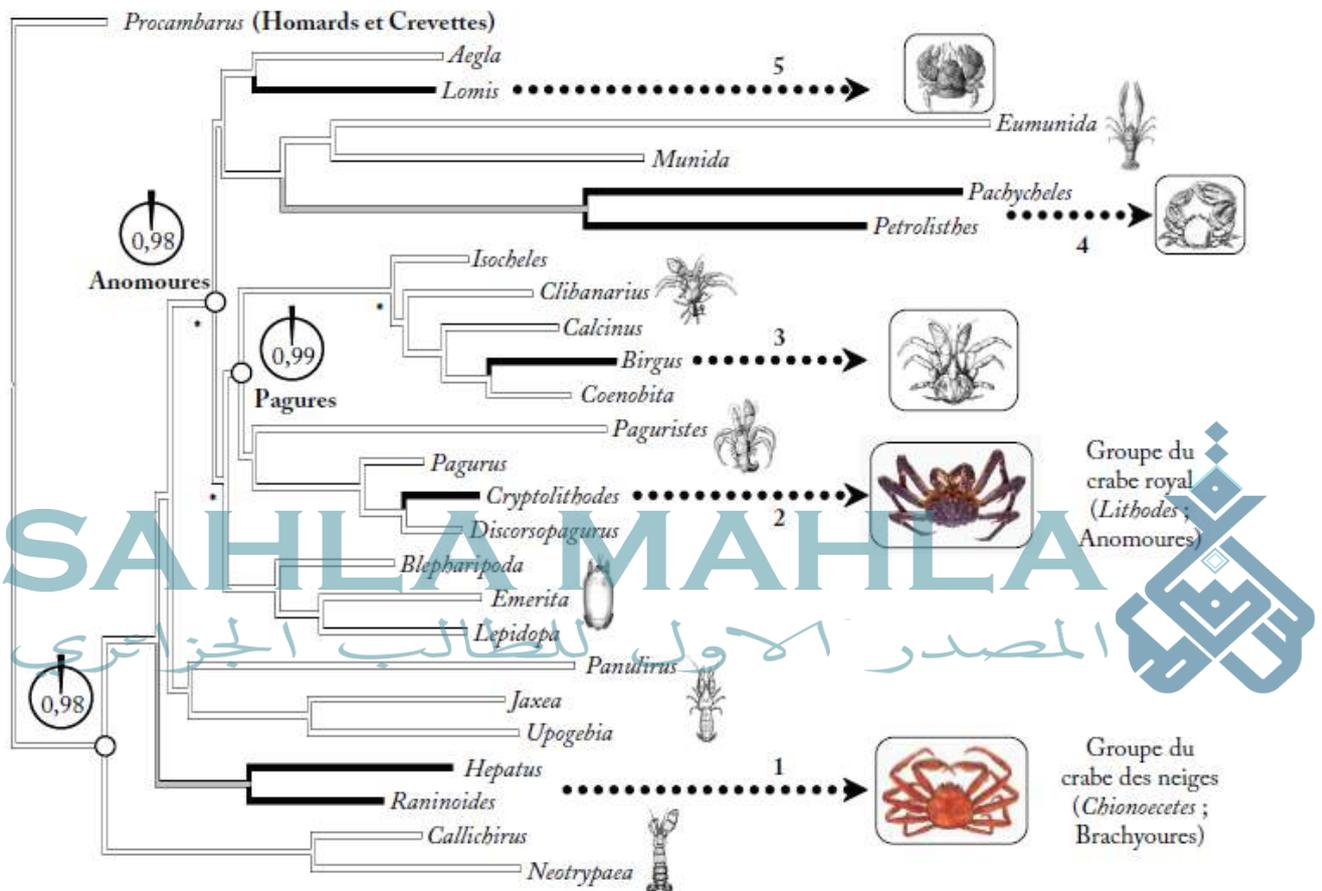


Figure 13. L'apparition du plan d'organisation de type crabe s'est produite au moins 5 fois indépendamment au cours de l'histoire évolutive des Crustacés Décapodes. La phylogénie représentée a été inférée par maximum de vraisemblance à partir de la concaténation de 4 gènes, dont deux mitochondriaux (*cox2* et ARNr 16S) et deux nucléaires (ARNr 18S et ARNr 28S). Certains nœuds ont été résolus grâce à l'existence de réarrangements originaux de l'ordre des gènes mitochondriaux (Etoiles). La coloration des branches de l'arbre représente la reconstruction des états de caractère ancestraux par une méthode de maximum de vraisemblance (noir = morphotype crabe ; blanc = morphotype crevette / homard). Les diagrammes circulaires aux nœuds indiquent les probabilités respectives des deux états. Ces reconstructions montrent de manière non-ambiguë que la forme homard / crevette représente le morphotype ancestral de toutes les lignées dont sont issues les différentes formes de crabes. (d'après Morrison *et al.* (2002) in Thomas *et al.*, 2010).

4. La phylogéographie

La phylogeographie est l'étude des processus qui contrôlent la distribution géographique des populations d'une espèce ou d'espèces proches par la reconstruction de leur phylogénie. Elle permet de comprendre les phénomènes historiques, autant au niveau génétique que démographique, qui ont conduit à la distribution et à la structuration actuelle des populations. L'analyse phylogeographique a ainsi été utilisée pour étudier l'origine et la propagation de l'épidémie de grippe aviaire (Janies *et al.* 2007). L'origine de cette épidémie est due à l'émergence en Asie du Sud-Est d'une forme particulièrement pathogène de la souche H5N1 (possédant les antigènes Hémagglutinine de type 5 et Neuraminidase de type 1) du virus de la grippe (*influenza A*) que l'on retrouve chez de nombreuses populations d'oiseaux de la région. Bien que d'origine strictement aviaire, la souche H5N1 s'est avérée transmissible à l'homme notamment par contact direct avec les volailles, mais a également infecté de nouveaux hôtes mammaliens et même des arthropodes. Même si les cas de transmission directe entre individus semblent être très rares chez l'homme, le développement potentiel par mutation d'une souche capable de transmission directe d'humain à humain nécessite une surveillance accrue de la panzootie en cours. Ce que révèle d'ailleurs la pandémie due au SARS-Covid 19.

Une collaboration internationale a ainsi été mise en place afin de séquencer le génome d'un très grand nombre d'échantillons de virus de la grippe, y compris les souches de la grippe aviaire (<http://msc.tigr.org/influenza/>). C'est à partir de ces données que la phylogénie de 351 génomes complets échantillonnés sur différents hôtes et couvrant une large distribution géographique a été reconstruite (Janies *et al.* 2007). Au début de l'épidémie, de 1996 à 2002, plusieurs lignées de H5N1 ont été retrouvées chez des oiseaux anseriformes (comme les canards) et galliformes (comme les poulets) à Hong-Kong et dans le Sud-Est de la Chine. En 2003 et 2004, une lignée chinoise de H5N1 arrive en Corée et au Japon où elle se diversifie en infectant plusieurs nouveaux hôtes tels que des corvidés et même des arthropodes. A la même époque, trois lignées distinctes se différencient en Asie : une est retrouvée exclusivement au Vietnam, une deuxième exclusivement en Indonésie et une dernière conquiert plusieurs pays tels que la Thaïlande, la Malaisie, le Vietnam, le Laos et le Cambodge. Chacune de ces trois lignées va également se diversifier localement essentiellement au travers des anseriformes et galliformes (Chen *et al.* 2006). Si la lignée exclusivement vietnamienne n'infecte que des hôtes aviaires, ce n'est pas le cas pour les deux autres lignées qui infectent de nombreux mammifères, y compris l'homme en Asie du Sud-Est et en Indonésie. En 2004, une lignée de H5N1 qui a débuté sa radiation en Asie se retrouve pour la première fois en Europe (**Figure 14**). Cette contamination s'est avérée être le fait d'aigles contaminés par la souche thaïlandaise de H5N1 et importés illégalement en Belgique à partir de Thaïlande (Van Borm *et al.* 2005). Cependant, les aigles ayant été saisis avant de pouvoir disséminer le virus, ce sont d'autres lignées indépendantes qui vont être à l'origine des épizooties européenne et africaine observées en 2005 et 2006, avec des relais en Russie et Mongolie (**Figure 14**). De façon intéressante, ces lignées semblent correspondre à un génotype particulier de H5N1 (figuré en rouge), caractérisé par la présence d'une Lysine en position 627 (Lys-627) dans la protéine de la polymérase basique 2 (PB2), et connue pour conférer au virus une capacité accrue à se répliquer chez des hôtes mammaliens (Janies *et al.* 2007).

Cet exemple démontre à quel point l'analyse phylogénétique de données épidémiologiques dans un contexte géographique peut s'avérer informative pour comprendre l'origine et la mise en place d'une épidémie virale. Il illustre également le pouvoir prédictif de ce type d'analyse pour prévenir les éventuelles sources d'épidémies futures : sans la phylogénie des virus *influenza A*, il n'aurait pas été aussi simple d'identifier l'origine asiatique et la migration transcontinentale de la souche H5N1 à l'origine du foyer de contamination en Europe occidentale.

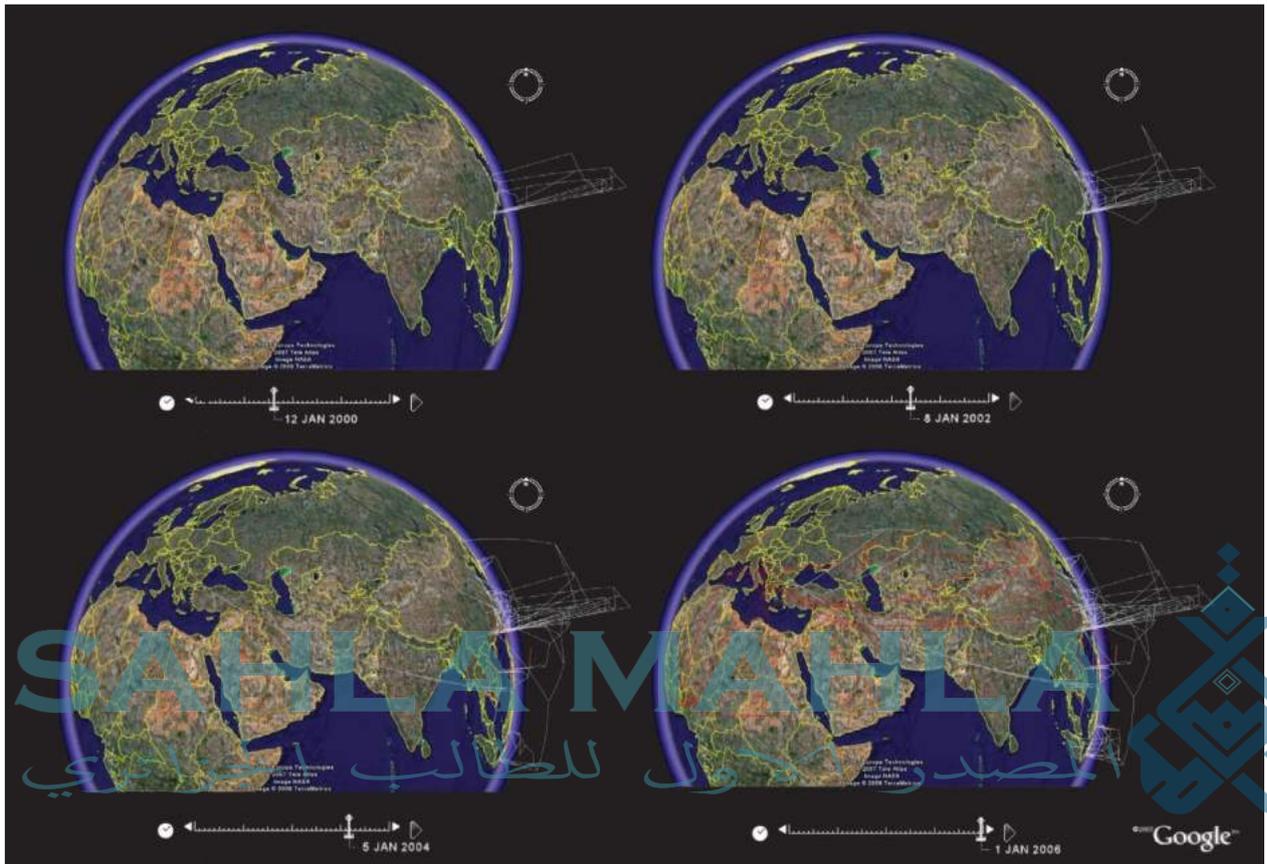


Figure 14. Approche phylogéographique de la dispersion de la grippe aviaire. Série temporelle de captures d'écran réalisées avec Google Earth (<http://earth.google.com>) montrant la propagation du virus *influenza A* de la grippe aviaire (génotype Lys-627 dans la protéine polymérase basique 2 pour la souche H5N1) sous la forme d'arbres phylogeographiques inférés par parcimonie maximale à partir de 351 génomes complets du virus (Janies *et al.* 2007). (d'après Thomas *et al.*, 2010)

Les forces évolutives

Il existe 4 "**forces évolutives**" qui agissent en interactions et font évoluer les fréquences alléliques et génotypiques en populations naturelles :

- la **mutation**
- la **dérive génétique** (variations des fréquences alléliques dues aux effets d'échantillonnages)
- la **migration** (ou flux de gènes)
- la **sélection naturelle**

1- La mutation :

Elles sont à l'origine de l'innovation génétique et sont produites par différents mécanismes (voir le chapitre « Mutation »). Elles génèrent du polymorphisme si de nouveaux allèles sont formés.

Rappel :

Les mutations peuvent-être classés en 4 catégories :

- **Mutations nucléotidiques** affectant un nombre limité de nucléotides (vu dans le chapitre mutation)
- **Mutations d'insertions** dues à la transposition d'ADN étranger (vu dans le chapitre mutation)
- **Remaniements chromosomiques** affectant une proportion de chromosomes.
- **Mutations caryotypiques** affectant le nombre de chromosomes.

a- Remaniements chromosomiques:

Le nombre de chromosomes et leur structure (caryotype) caractérisent chaque espèce. Tout changement de structure des chromosomes par perte, addition ou réarrangement de chromosomes est appelé remaniements chromosomiques. Il existe catégories :

Délétions, duplications, inversions et translocations. Les deux premières ont déjà été abordés, il ne sera donc questions que des deux dernières catégories dans ce chapitre.

a 1- Inversions :

Correspondent à u retournement de 180° d'un segment de chromosome sans perte de matériel chromosomique. Selon que le centromère est impliqué ou non dans l'inversion, on distingue :

- Les inversions paracentriques si le centromère n'est pas inclus dans l'inversion.
- Les inversions péricentriques si le centromère est inclu dans l'inversion (**Figure 1**).

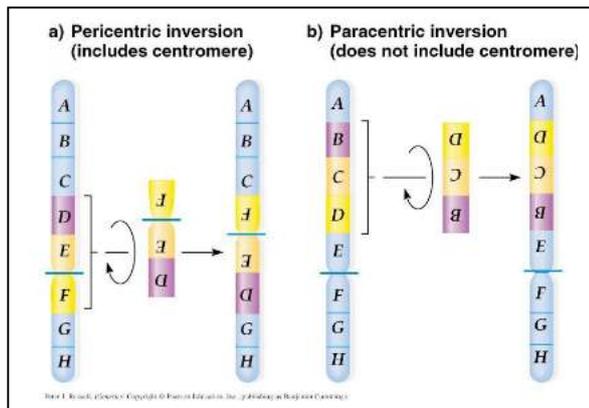


Figure 1: Inversion chromosomique :
 a- Péricentrique
 b- Paracentrique

Exemple:

Les Populations de *Drosophila subobscura* sont **polymorphiques** pour au moins une inversion dans 5 des 6 chromosomes. La fréquence des inversions chez les populations varie en fonction de l'altitude et de la latitude, formant des clines (lente divergence évolutive morphologique ou physiologique au sein d'une même population provoquée par une barrière géographique). (**Figure 2**)

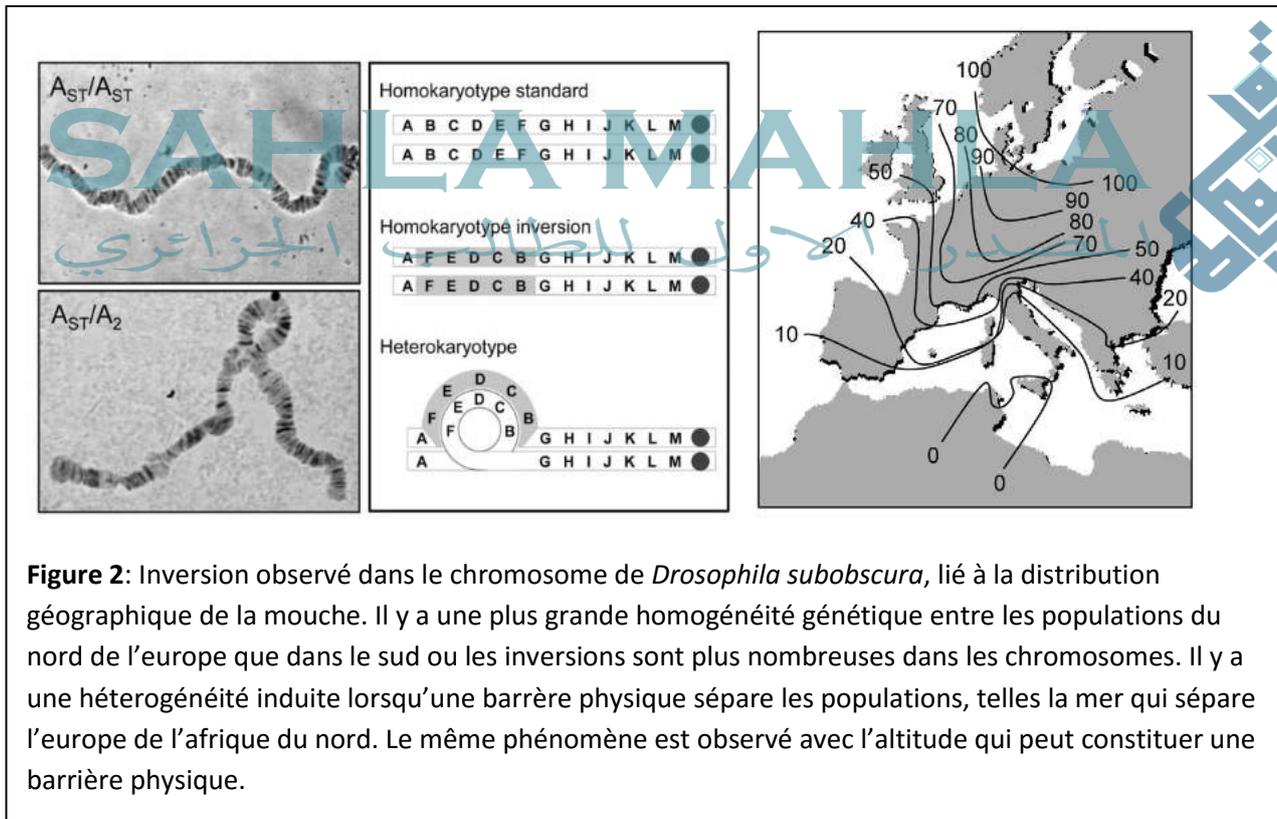


Figure 2: Inversion observé dans le chromosome de *Drosophila subobscura*, lié à la distribution géographique de la mouche. Il y a une plus grande homogénéité génétique entre les populations du nord de l'Europe que dans le sud où les inversions sont plus nombreuses dans les chromosomes. Il y a une hétérogénéité induite lorsqu'une barrière physique sépare les populations, telles la mer qui sépare l'Europe de l'Afrique du Nord. Le même phénomène est observé avec l'altitude qui peut constituer une barrière physique.

a 2- Translocations : (Figure 3)

Transfert d'un segment de chromosome sur une autre région du génome.

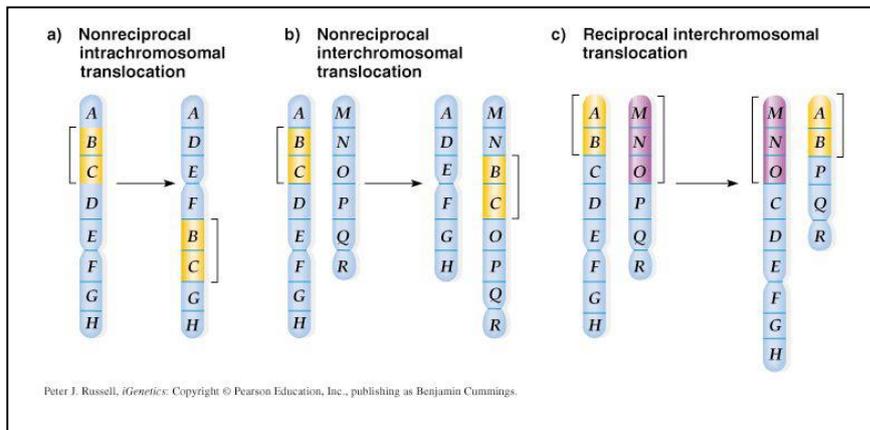
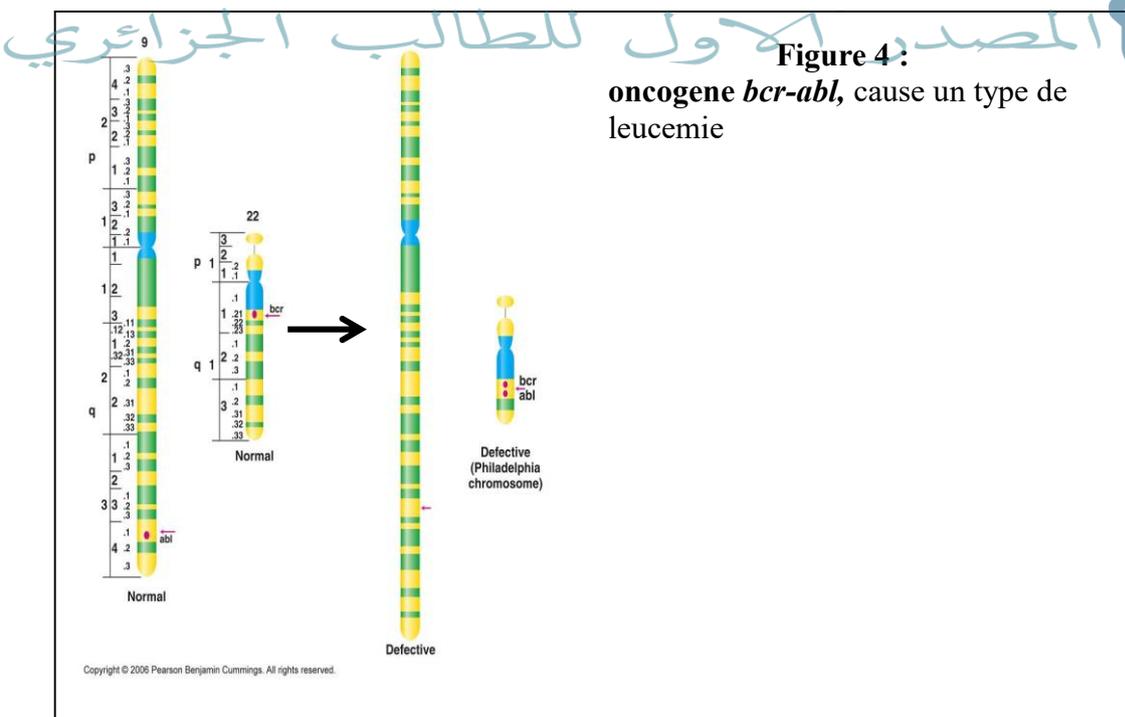


Figure 3 :
Translocations

Il peut y avoir des **translocations simples** : il y a une seule cassure d'un segment qui se soude à l'extrémité d'un autre chromosome. Cas rare.

Il y a des **translocations réciproques** : transfert de segments de chromosomes entre deux chromosomes non homologues. Ce transfert peut-être équilibré s'il n'y a pas d'altération quantitative de l'ADN (aucune perte de la quantité de matériel génétique chez les deux chromosomes). Par contre cet échange entre les 2 chromosomes peut-être déséquilibré si il y a une perte de matériel génétique chez un ou chez les deux chromosomes.

Exemple : chromosome de Philadelphie, responsable de certains cancers (Figure 4)



Il y a également **la fusion** : qui correspond à la formation d'un seul chromosome à partir de deux chromosomes. L'exemple le plus remarquable est celui du chromosome 2 humain qui correspond à la fusion des chromosomes 2A avec le chromosome 2B du chimpanzé (**Figure 5**).

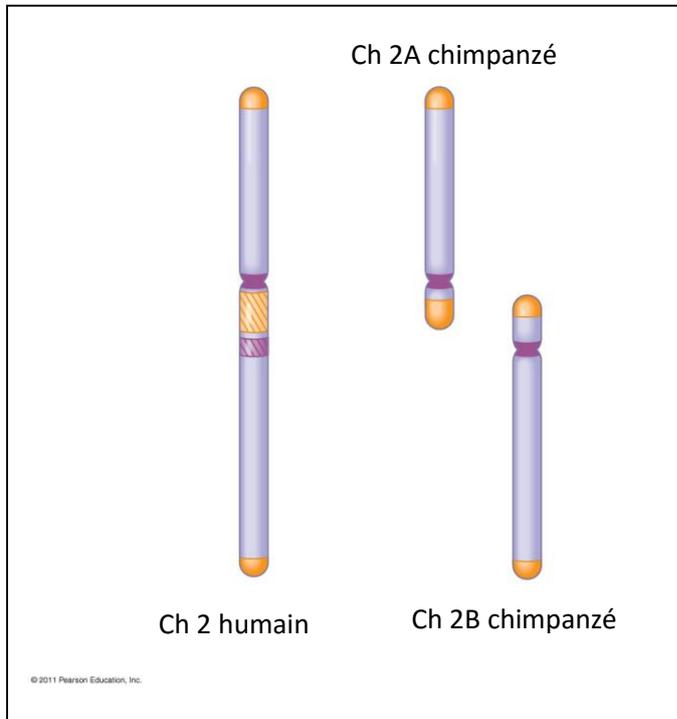


Figure 5 :
Fusion des chromosomes 2A et 2B du chimpanzé correspond au chromosome 2 humain

b- Mutations caryotypiques:

Correspond à la modification du nombre de chromosomes.

b 1- Aneuploidie :

Correspond au changement du nombre de chromosomes, soit un ou plusieurs chromosomes ou tout le lot de chromosomes. Chez l'homme elle peut-être à l'origine de maladies. La trisomie 21 en est un exemple, mais également le syndrome de Klinefelter (XXY, 2 chromosomes sexuels X au lieu d'un seul X).

Chez les espèces animales ou végétales, la nullisomie (perte de deux chromosomes homologues) entraîne la létalité des zygotes formés.

Un nombre impair de chromosomes (aneuploidie) provoque le plus souvent une stérilité chez l'organisme porteur de ce type de mutations car il n'y a pas d'appariements de chromosomes durant la méiose.

b 2- Polyploidie :

Correspond à un nombre supérieur à deux lots de chromosomes.

Trois principaux types de polyploïdisation:

1- *La non réduction gamétique*: due à l'absence de division cellulaire (cytocinèse) ou la non disjonction des chromosomes homologues au cours de la méiose.

2- *La fertilisation multiple* (non viable).

3- *L'endoréplication*: mitoses anormales pendant lesquelles il y a une duplication du matériel héréditaire non suivie de cytocinèse.

De nombreuses espèces cultivées ou domestiquées sont des mutants polyploïdes, hexaploïdes comme le prunier à fruits comestibles (*Prunus domesticus*).

Les espèces cultivées du blé et de la vigne *Vitis vinifera* sont des Allopolyploïdes (Des gamètes d'espèces différentes mais proches peuvent fusionner et donner naissance à un hybride qui pourra évoluer en une nouvelle espèce).

Le blé est l'espèce de céréales la plus consommée dans le monde. Les deux variétés de blé les plus cultivées, le blé dur (pour la fabrication des pâtes et de la semoule) et le blé tendre (pour la fabrication des farines destinées à l'alimentation humaine et animale) résultent d'une longue histoire évolutive mêlant hybridation spontanée et domestication. Le blé tendre moderne est le résultat de deux événements, « *la non réduction gamétique* » et « *l'endoréplication* », qui s'est produite il y a environ 8 000 à 10 000 ans, résultant en un hexaploïde. Les propriétés de formation des grains et d'adaptation à l'environnement du blé proviennent donc de trois espèces ancestrales différentes.

Il y a environ 500 000 ans une hybridation a eu lieu entre un blé diploïde sauvage *Aegilops speltoides* (BB) des Balkans et du moyen-orient et *Triticum monococcum* (AA), blé sauvage de la même région. Cela a abouti à la formation du blé sauvage tétraploïde *T. turgidum*.

Entre 9000 et 12000 ans, à la naissance de l'agriculture une hybridation s'est réalisée entre un blé tétraploïde et un autre blé sauvage diploïde *A. tauschii* (DD) originaire d'Asie donnant un hybride triploïde ABD.

Une endoréplication chez cet hybride a résulté en la formation du blé tendre actuel *T. aestivum* (AA BB DD) (**figure 6**).

D'autres cultures allopolyploïdes comprennent le coton, l'avoine et la canne à sucre.

Génération parentale P

Blé diploïde sauvage **Blé diploïde sauvage**

Triticum monococcum *Aegilops speltoides*

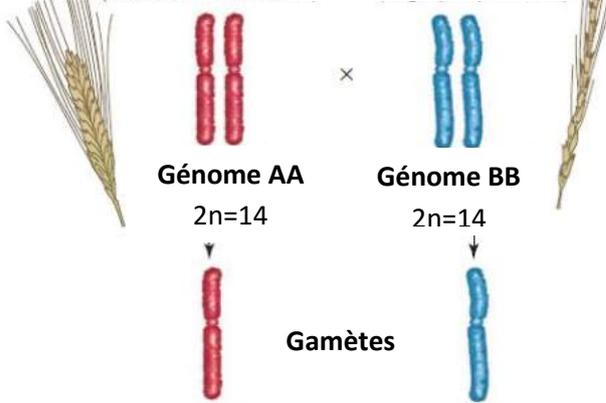
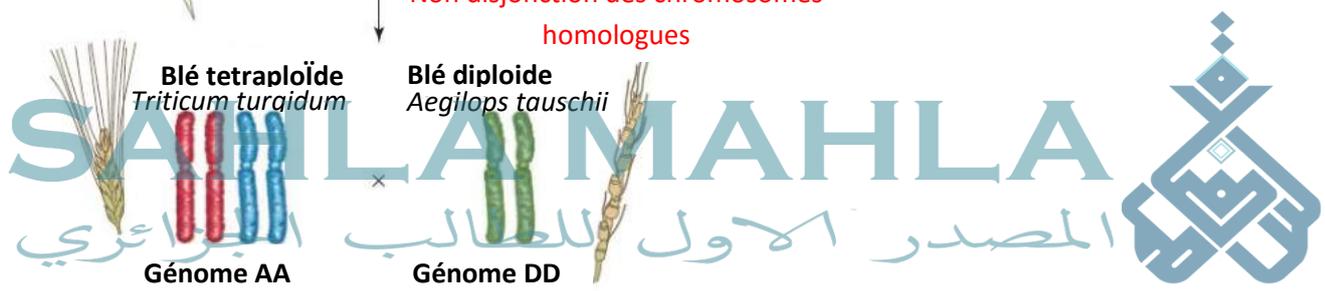
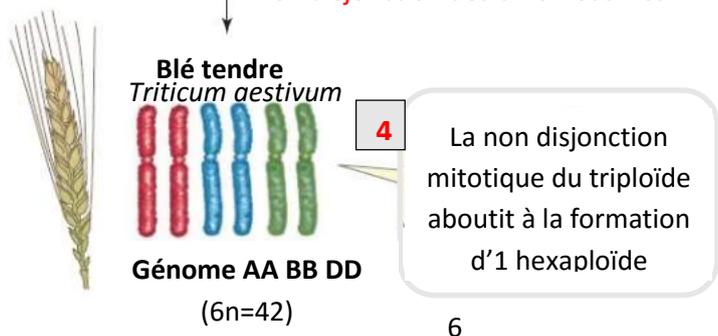
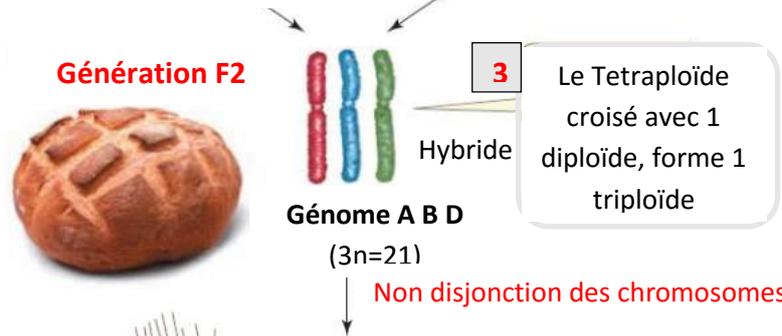
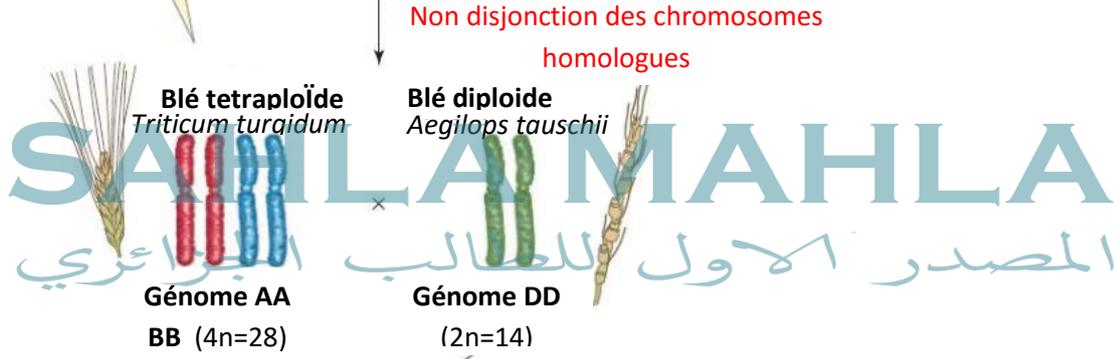
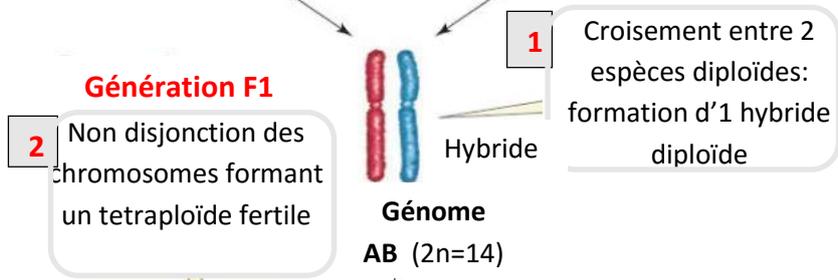


Figure 6 : Polyplôïdie et origine du blé tendre :
3 espèces et 2 événements de non-disjonction mitotique ont conduit au blé moderne.



2- Sélection naturelle :

A partir de ses observations sur les pinsons des Galapagos, après son long voyage autour du monde, Darwin a élaboré le concept de la sélection naturelle ou de « la survie du plus apte ».

Depuis l'avènement de la Génétique, la théorie s'est affirmé et s'est renforcé.

Trois paramètres doivent être requis pour que la sélection opère :

- 1- Existence d'une **reproduction** qui produira une descendance.
- 2- Existence d'une **variabilité** de caractères héritable et héréditaires, portés par les individus de la population
- 3- Existence d'une différence **d'aptitude** entre les organismes due à la variation présente dans le caractère considéré

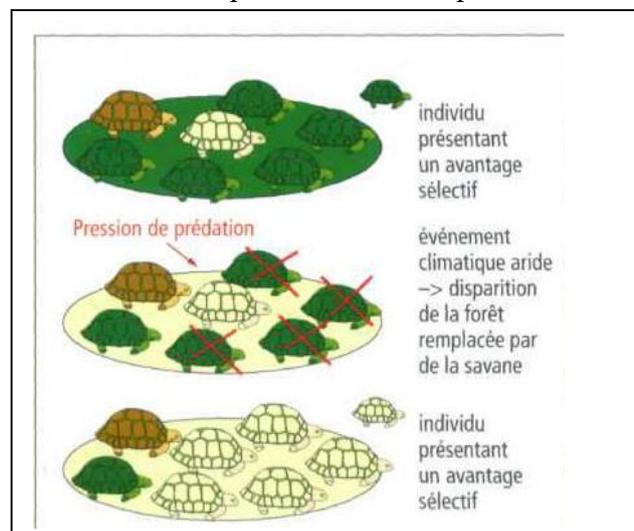
L'**aptitude** correspond au rapport entre le nombre de descendants de l'organisme considéré et le nombre moyen de descendants des membres de la population.

Les individus présentant une **meilleure** aptitude ont un **avantage sélectif** sur le reste de la population.

Exemple : (**Figure 7**)

Dans une population de tortues polymorphes pour la couleur de la carapace, vivant dans un environnement boisé dont les individus sont verts, marrons ou jaunes. Au début les tortues vertes sont prédominantes, car mieux adaptées à la verdure, elles échappent plus aux prédateurs, car leur couleur les confond avec la couleur de la végétation. Puis un changement climatique entraîne une désertification de la région, ce sont les tortues jaunes qui sont favorisées par le nouvel environnement. La savane a remplacé la forêt, les tortues vertes deviennent vulnérables face aux prédateurs. Les tortues jaunes laisseront une progéniture plus importante à chaque génération et occuperont plus l'environnement car favorisées par leur couleur qui leur donne un avantage sélectif.

Figure 7 : La Sélection naturelle



Un exemple naturel est représenté par la phalène du bouleau. En Angleterre au cours de la révolution industrielle du XIX^{ème} siècle les forêts ont subi une importante pollution par dépôt de résidus de charbon sur les troncs d'arbres. Le papillon *Biston betularia*, phalène du bouleau, présente deux formes, l'une aux ailes blanches, dite *typica*, et l'autre aux ailes noires, dite *carbonaria*. En 1950, un biologiste anglais, Kettlewell, montre que la forme noire représente 90% des papillons dans toutes les forêts exposées au vent provenant des usines. Par contre dans les régions non polluées, la forme blanche du papillon est dominante. Le lien entre pollution et fréquence des formes noires a donné le nom de « **mélanisme industriel** » à ce cas d'adaptation. Jusqu'au début du dix-neuvième siècle on ne connaissait que la forme typique aux ailes blanches tachetées de noir. En 1849, le premier spécimen aux ailes et au corps noirs est signalé en Angleterre près de Manchester. Il est nommé *carbonaria* ; on distingue alors la forme blanche ou *typica* et la forme noire mélanique ou *carbonaria*. A la fin du dix-neuvième siècle, à proximité des centres industriels anglais la plupart Phalènes sont noires dans des proportions qui peuvent atteindre 98%. En moins d'un demi-siècle la proportion des phénotypes s'est inversée. Kettlewell a démontré que l'effet de la sélection fait intervenir une pression de prédation et un avantage sélectif lié à un camouflage. Sur les troncs noircis par le charbon, les papillons noirs sont moins visibles aux prédateurs (**Figure 8, a**) alors que les blancs sont plus visibles et donc susceptibles d'être chassés (**Figure 8 b**).



Figure 8. Le papillon *Biston betularia*, phalène du bouleau

La couleur noire des Papillons est héréditaire et l'allèle qui la contrôle est dominant sur l'allèle responsable de la couleur blanche. L'apparition de la couleur noire ne peut être due qu'à une mutation brusquement apparue dans la population de Papillons blancs.

Le mélanisme est lié à la production d'un pigment foncé, la mélanine, par les cellules de l'épiderme.

Le gène contrôlant le mélanisme de *Biston betularia* constitue le locus *carbonaria* de 200 kbases située dans une région déjà connue pour contenir les gènes contrôlant la forme et la couleur des ailes chez les papillons. Cette zone a été le siège de très nombreuses mutations expliquant l'extraordinaire diversité des motifs colorés des ailes de ces insectes.

En 2016, une équipe de chercheurs britanniques (Van'T Hof *et al.*, 2016) a publié une étude dans laquelle ils identifient la mutation à l'origine de l'apparition du phénotype *carbonaria*. À partir des études génétiques classiques ayant permis de localiser le locus *carbonaria*, les chercheurs ont identifié le gène porteur de la mutation. Il s'agit du gène *cortex*. Ce gène présente une organisation morcelée : un premier exon ayant de multiples formes alternatives, séparé par un grand intron des huit exons suivants (**Figure 9**). La mutation identifiée est l'insertion au sein du premier intron d'un élément transposable (ou transposon), long de 21 925 nucléotides. Il a été nommé *carb-TE*.



Figure 9. Structure schématique du gène *cortex*

Les allèles *typica* et *carbonaria* possèdent chacun 9 exons, numérotés de 1 à 9. Les deux allèles diffèrent par l'insertion dans le premier intron de l'allèle *carbonaria* d'une séquence d'ADN mobile (un transposon), nommé *carb-TE*.

Modes de sélection naturelle :

Les traits phénotypiques contrôlés par un seul locus se distinguent souvent par des qualités discrètes (par exemple la couleur : noir contre blanc, ou la forme : lisse contre froissé comme dans le cas des petits pois de l'expérience de Mendel), et sont appelés traits qualitatifs. Cependant, de nombreux traits sont influencés par des allèles à plusieurs loci. De tels traits sont susceptibles de présenter une variation quantitative continue plutôt qu'une variation qualitative discrète, et sont donc appelés traits quantitatifs. La taille du corps, par exemple, est influencée par les gènes dans de nombreux loci, ainsi que par l'environnement (par exemple : la nutrition). Par conséquent, la répartition des tailles des individus dans une population ressemble probablement à une courbe continue en forme de cloche.

Un individu est considéré comme un lot de gènes qui contrôlent l'expression des phénotypes (les caractères). La sélection va donc agir au niveau des phénotypes. Dans une population (ensemble d'individus), en fonction de l'évolution des caractéristiques du phénotype étudié, on peut distinguer trois modes de sélection :

a- La sélection directionnelle : modifie les caractéristiques d'une population en favorisant les individus qui varient dans une direction par rapport à la moyenne de la population.

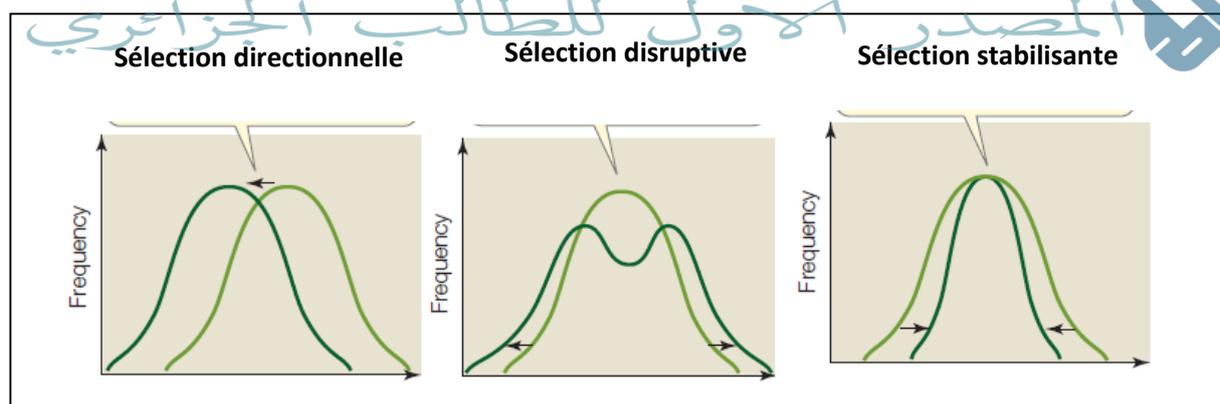
Exemple : avantage sélectif de la forme *carbonaria* de la phalène du bouleau en milieu pollué

b- La sélection disruptive : modifie les caractéristiques d'une population en favorisant des individus qui varient dans les deux sens par rapport à la moyenne de la population.

Exemple : L'escargot, *Cepaea nemoralis* (Linnaeus, 1758) qui vit dans les champs et forêts britanniques a plusieurs formes de couleur distinctes qui lui servent de camouflage pour différents milieux naturels.

c- La sélection stabilisante : préserve les caractéristiques moyennes d'une population en favorisant les individus moyens. La distribution des phénotypes au cours des générations reste inchangée. La sélection stabilisante réduit la variation dans les populations, mais elle ne change pas la moyenne.

Exemple : La sélection stabilisante fonctionne, par exemple, sur le poids des humains à la naissance. Les bébés qui sont plus légers ou plus lourds à la naissance que la moyenne de la population, meurent à des taux plus élevés que les bébés dont le poids est proche de la moyenne.

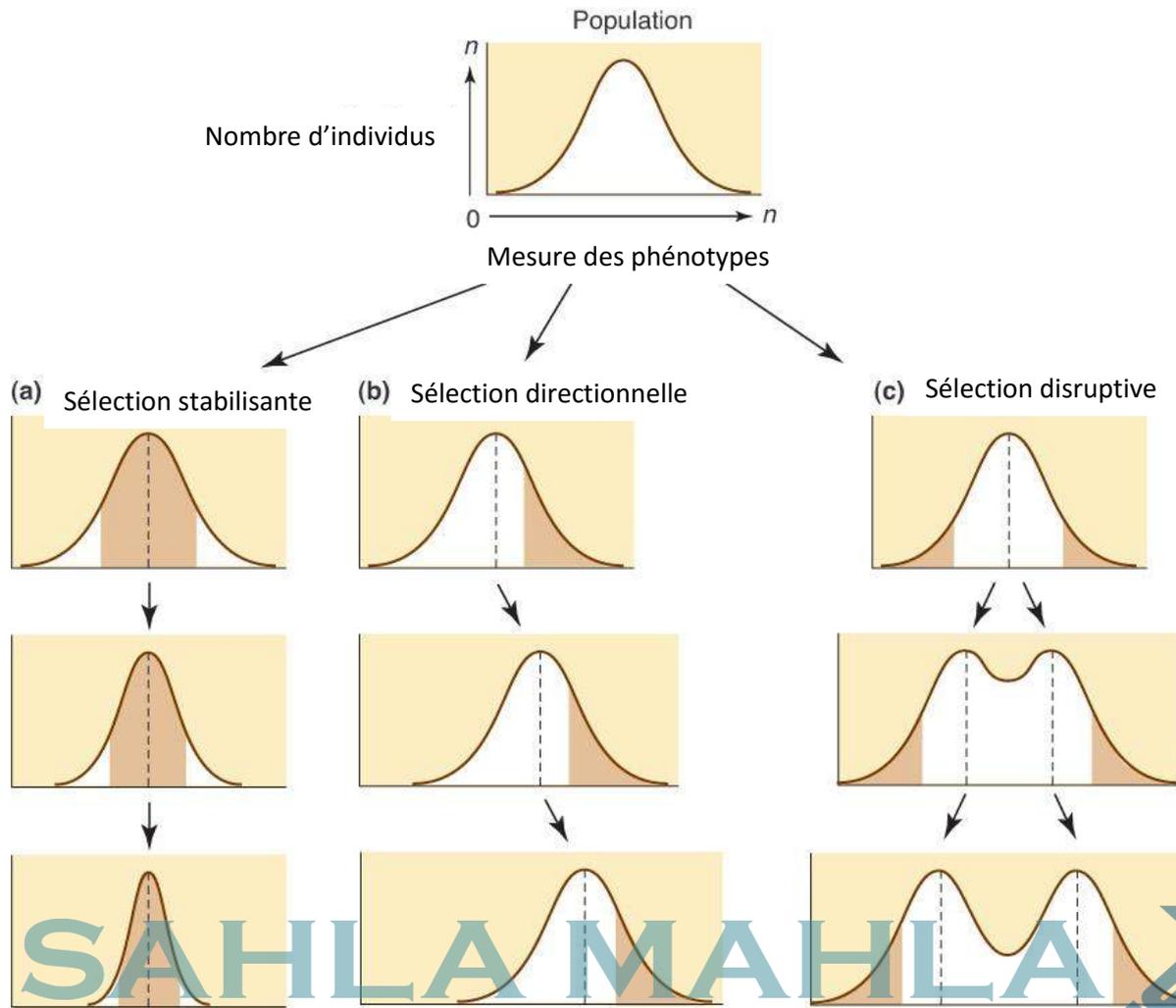


Remarque

La sélection ne crée pas la variabilité. La variation génétique est aléatoire. La sélection opère un tri parmi la variabilité présente

Le schéma qui suit, résume la distribution de la variation des phénotypes dans les populations dans les 3 modes de sélection.

Les forces évolutives



المصدر الاول للطالب الجزائري

3- Dérive génétique :

En 1968, Motoo Kimura propose la théorie neutraliste de l'évolution. Les allèles sont neutres, c'est-à-dire que les mutations se répandent par elles-mêmes dans une population sans conférer un avantage sélectif particulier. Le devenir de l'allèle dépend donc du hasard. La fréquence d'un caractère fluctue au cours du temps à cause d'un tirage aléatoire à chaque génération des gamètes au cours de la reproduction. Pendant ce tirage au hasard, que l'on nomme **dérive génétique**, certains allèles se perdent et d'autres se fixent dans la population. Par conséquent la composition génétique des populations varie. **L'effet de fondation** illustre ce mécanisme de dérive génétique (**Figure 10**). Supposons une population de tortues polychromes (de couleurs différentes). Une femelle féconde est emportée par un radeau ou un arbre flottant vers une île. Cette tortue sera à l'origine de la naissance d'une nouvelle population, mais qui ne portera qu'une partie de la variabilité génétique de la population mère (d'origine). Dans l'exemple présent, la tortue emportée est de couleur foncée, rouge, elle donnera donc naissance à une population rouge. C'est ce qui certainement passé avec les tortues des Galapagos. En effet sur chaque île des Galapagos ne s'est développé qu'une seule espèce de tortue, chacune étant

adaptée à son île. Il est à noter, que des études ont montré que toutes les tortues des îles Galapagos descendent de tortues terrestres vivant sur le continent sud-américain.

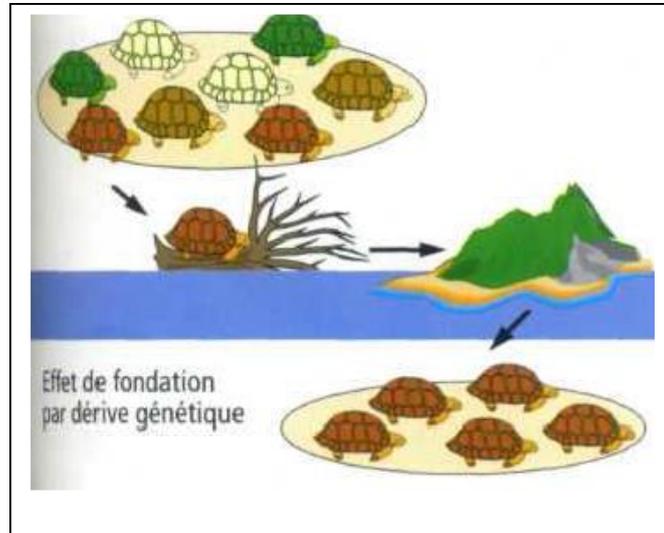


Figure 10 : Dérive génétique

La composition génétique de la population fille sera différente de la population mère suite à l'effet de fondation par une seule mère fécondée

Un tel phénomène a été observé sur une île au large de la Croatie l'île de **Pod Mrcaru**, en Méditerranée, où ont été implanté cinq couples de lézards *Podarcis sicula* en 1971 ; En 2008, une équipe belge a été sur cette île et a trouvé une population nombreuse de lézard qui s'était développé, alors qu'à l'origine il n'en existait pas sur l'île (**Figure 11**).

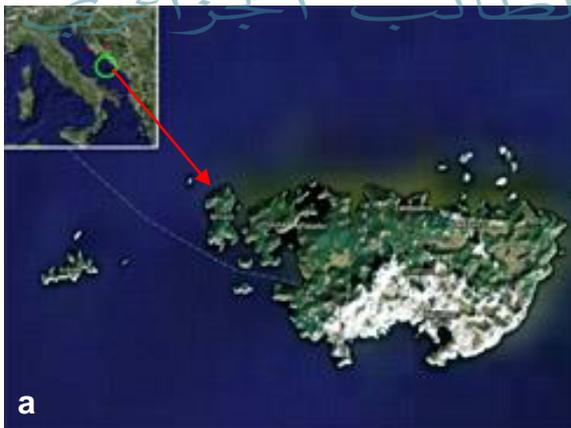


Figure 11. a) île de Pod Mrcaru.

b) Lézard *Podarcis sicula* introduit dans l'île en 1971

L'effet du goulot d'étranglement :

Effondrement des effectifs d'une population qui conduit à un tri au hasard des allèles pour la fondation d'une nouvelle Population. Le goulot d'étranglement en réduisant le polymorphisme réduit la variabilité génétique. Les effets démographiques s'appliquent à **l'ensemble du génome** alors que les effets sélectifs sont localisés à une région génomique.

C'est le cas des Guépards du Sahara *Acinonyx jubatus hecki* et d'Afrique en général, dont la population est en grand danger d'extinction, mais qui se reconstitue péniblement (**Figure 12**).



Figure 12. Le guépard du Sahara *Acinonyx jubatus hecki*

SAHLA MAHLA

المصدر الأول للطالب الجزائري

**Remarque :**

L'effet du tri au hasard des allèles lors de la reproduction sera d'autant plus grand que la population sera petite. L'effet de fondation est un cas extrême de dérive génétique

4- Migration :

Tout mouvement d'individus ou de gamètes d'une population à une autre est une migration. Il y a passage et échange de gènes, c'est un **flux génique**.

Si les populations possèdent des fréquences alléliques différentes, ces flux géniques vont atténuer ces différences.

Exemple :

Dans le cas des tortues polychromes, l'échange d'individus entre populations diminue la divergence génétique et homogénéise les populations (**Figure 13**).

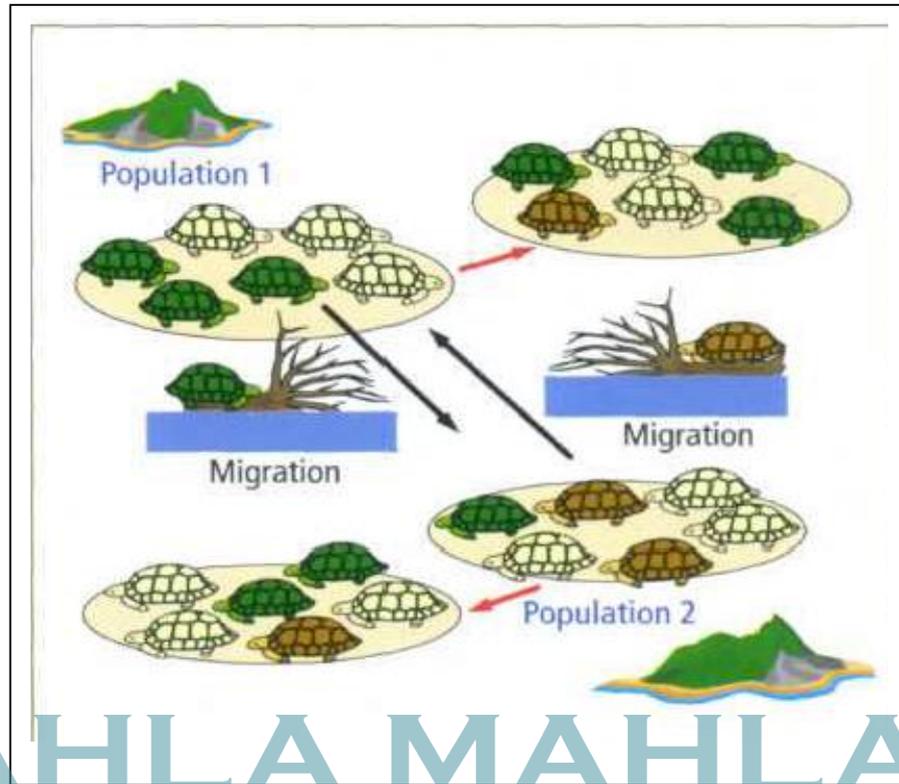


Figure 13 : La migration atténue la différenciation génétique

SAHLA MAHLA
المصدر الأول للطالب الجزائري



Exemple :

En Amérique latine, les populations autochtones n'ont que le groupe sanguin O. C'est l'arrivée de migrants européens ou autres qui ont introduit les autres groupes sanguins