



REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE
SCIENTIFIQUE



Université des Sciences et de la Technologie d'Oran Mohamed BOUDIAF

Faculté des Sciences de la Nature et de la Vie

Département de Génétique Moléculaire Appliquée

Polycopié de cours

Génétique Evolutive

Cours destiné aux étudiants de la troisième année Licence Génétique

Dr Khaib Dit Naib Ouahiba

Maitre De Conférences B

2022-2023

Semestre : 6

Unité d'enseignement Fondamentale 1 (UEF 3.2.1) : Génétique Evolutive, Quantitative et des

Populations

Matière 2 : Génétique évolutive

Crédits : 2

Coefficient : 1

La connaissance des forces évolutifs ; la compréhension de l'origine et de la dynamique de la biodiversité ; les interactions entre les gènes, les individus, les populations et l'environnement ; l'adaptation des espèces à l'environnement.

I. Introduction

Les fondements de l'évolution et son essence génétique

II. Notion de l'espèce et les différents modes de spéciation

III. Les forces évolutives

Mutations et migrations

Sélection et adaptation

Dérive génétique

IV. diversité génétique

Notion de polymorphisme

Déterminisme épigénétique

Déterminisme génétique

Du génotype aux phénotypes

V. Evolution des séquences nucléotidiques

- Duplication de gènes et évolution des familles de gène

VI. Phylogénie moléculaire

Mode d'évaluation :

Contrôle continu et Examen semestriel

Références bibliographiques:

Préface

La génétique évolutive occupe aujourd'hui une place centrale dans la compréhension du vivant. Elle relie les processus moléculaires qui modèlent l'ADN à l'échelle individuelle aux grands changements biologiques observés dans les populations, les espèces et l'histoire de la vie. Ce polycopié est conçu pour accompagner les étudiants de Licence dans l'exploration de cette discipline passionnante.

L'objectif de cet ouvrage est double : transmettre les principes fondamentaux de l'évolution du point de vue génétique, et offrir des outils pédagogiques pratiques (résumés, glossaire, QCM) pour faciliter la compréhension et l'appropriation des notions complexes.

Plutôt que de simplement présenter des faits, ce cours invite à la réflexion critique, à l'analyse des mécanismes évolutifs, et à la mise en lien des différents niveaux d'organisation du vivant, du gène à l'espèce.

En espérant que ce support vous accompagne utilement dans votre formation et vous ouvre la voie vers une compréhension plus fine des dynamiques évolutives du vivant.

Dr. Khaib Dit Naïb Ouahiba

Maître de Conférences B

Université des Sciences et de la Technologie d'Oran Mohamed Boudiaf

Résumé du polycopié

Ce polycopié de génétique évolutive propose un panorama structuré des mécanismes biologiques et génétiques responsables de la diversité et de l'évolution des espèces.

Il s'articule autour de six chapitres principaux :

1-Introduction – Présentation des grandes théories de l'évolution, de Lamarck à la théorie synthétique.

2-Notion d'espèce et spéciation – Exploration des différents concepts d'espèce et des mécanismes de formation de nouvelles espèces.

3- Les forces évolutives – Étude des mutations, de la sélection naturelle, de la dérive génétique et du flux génique.

4- Diversité génétique – Analyse du polymorphisme, des déterminismes génétique et épigénétique, et de la relation génotype-phénotype.

5-Évolution moléculaire – Compréhension des changements au niveau des séquences nucléotidiques et de la naissance de familles de gènes.

6- Phylogénie moléculaire – Méthodes d'inférence des relations évolutives à partir des séquences génétiques.

Chaque chapitre est accompagné de résumés, glossaires et QCM pour renforcer l'apprentissage.

Ce support vise à fournir une base solide pour les étudiants de troisième année de Licence en Génétique, tout en leur ouvrant des perspectives vers les approches évolutives les plus récentes.

Mots clés : Evolution. Spéciation. Forces Evolutives. Diversité Génétique. Déterminismes Génétique. Epigénétique, Évolution Moléculaire. Phylogénie Moléculaire

Liste des figures

Figure 1 : séries phylétiques étudiée par Lamarck.....	4
Figure 2 : Anatomie comparée des membres supérieurs	6
Figure 3 : illustration des processus anagenèse et cladogénèse.....	8
Figure 4 : exemple d'une classification phénétique.....	15
Figure 5 : exemple de dimorphisme sexuel chez le canard.....	15
Figure 6 : exemple de dimorphisme sexuel chez une espèce pélagique de poissons osseu...16	
Figure 7 : Ecophénotypes chez <i>Sida fallax</i>	16
Figure 8 : Sturnelle de l'ouest à gauche et Sturnelle des prés à droite.....	17
Figure 9 : Représentation de l'isolement reproductif assuré par des barrières biologiques qui s'opposent aux mélanges entre espèces différentes.....	18
Figure 10 : différents types de barrières biologiques.....	19
Figure 11 : Barrières reproductibles comportementales.....	20
Figure 12 : Barrière géographiques.....	21
Figure 13 : Barrière chronologique.....	21
Figure 14 : Barrière anatomique isolement gamétique.....	21
Figure 15 : exemples de barrières postzygotiques.....	22
Figure 16 : exemples de classification cladiste.....	24
Figure 17 : Evolution d'une espèce dans le temps selon le processus Anagenèse.....	26
Figure 18 : Evolution d'une espèce dans le temps selon par Cladogénèse.....	26
Figure 19 : Spéciation allopatrique.....	28
Figure 20 : Exemple de spéciation par hybridation interspécifique.....	29
Figure 21 : Comparaison des caryotypes de la poule et la dinde.....	30
Figure 22 : spéciation des espèces de spartine.....	32
Figure 23 : les différents modes de spéciation.....	33
Figure 24 : La duplication d'un gène essentiel génère deux loci paralogues indépendants...35	
Figure 25 : exemple de cospéciation.....	36
Figure 26 : Modèle de dérive génétique (p est la fréquence initiale de l'allèle, N le nombre d'individus de la population).....	39
Figure 27 : principe de l'effet fondateur.....	41
Figure 28 : profil électrophorétique des marqueurs AFLP des vignes marronne de la Réunion et du Vitnam.....	42
Figure 29 : principe du goulot d'étranglement.....	43

Figure 30 : diminution de la diversité génétique chez l'éléphant de mer.....	44
Figure 31 : exemple de substitutions dans le gène β de l'hémoglobine.....	46
Figure 32 : Remaniements chromosomique.....	46
Figure 33 : principe de la conversion génique.....	48
Figure 34 : conséquences de la conversion génique.....	48
Figure 35 : illustration de la non disjonction méiotique.....	51
Figure 36 : Histoire évolutive du Blé.....	52
Figure 37 : Illustration des différentes formes de sélection au niveau phénotypique.....	53
Figure 38 : Les mécanismes épigénétiques influant sur l'état de la chromatine.....	56
Figure 39 : exemple d'étude du déterminisme génétique de la résistance à l'oïdium.....	57
Figure 40 : résultat de l'étude du déterminisme génétique de la résistance à l'oïdium.....	58
Figure 41 : Illustration du génotype au phénotype.....	59
Figure 42 : Annotation des gènes codants.....	61
Figure 43 : Exemple de duplication de gènes qui donne lieu à deux lignées de gènes paralogues.....	63
Figure 44 : Arbre phylogénétique	67

Liste des tableaux

Tableau 1 : Comparaison des caryotypes de la poule, la dinde, le canard et la caille.....31

Tableau 2 : fréquences des allèles du complexe HLA et des groupes sanguins chez les Huttérites, les Amish et la population européenne et américaine.....42

Avant-propos

La théorie de l'évolution aussi ancienne qu'elle soit n'est toujours pas confirmée et consolidée car nous ignorons toujours ses mécanismes. Dans ce polycopié, seront exposées les différentes facettes anciennes et nouvelles de cette théorie ainsi que les critiques d'une façon scientifique loin de toutes croyances, car il s'agit d'une œuvre scientifique.

Quand on parle de l'évolution on peut la considérer sur plusieurs niveaux ; création de la vie sur terre, évolution des espèces mais aussi évolution des génomes qui peuvent avoir lieu sans forcément donner naissance à une nouvelle espèce, mais c'est une évolution à petite échelle, comme par exemple l'apparition d'une nouvelle mutation qui donnera naissance à un allotype ou une protéine non fonctionnelle, ou encore les épimutations. Ces changements moléculaires et phénotypiques sont héréditaires mais est-ce qu'ils vont conduire un jour à l'apparition d'une nouvelle fonction ? Ou encore une nouvelle espèce ? où, ces changements participent seulement à la diversité intra-populationnelle.

Table des matières

Liste des figures	I
Liste des tableaux	II
Liste des abréviations	III
Avant-propos	IV
I. Introduction	3
I.1. Définition de l'évolution	3
I.2. Fondements de l'évolution et son essence génétique	3
I.2.1. Théories ignorant la génétique et la géochronologie	3
I.2.1.1. Créationnisme	3
I.2.1.2. Evolutionnisme	4
I.2.2. Théories intégrant la génétique et la durée des temps géologiques	7
I.3. Critiques la théorie de l'évolution	10
a- Mutation	10
b- Sélection	11
À retenir	11
Glossaire	12
QCM	12
II. Notion de l'espèce et les différents modes de spéciation	13
II.1. introduction	13
II.2. Critères de classification d'une espèce	14
II.2.1. Concept morphologique	14
II.2.2. Concept biologique	17
II.2.2.1. Barrières reproductives	17
II.2.2.2. Exemples de barrières reproductibles	19
II.3. Concept écologique	23
II.4. Concept cladiste (phylogénétique)	23
II.5. Notion de population	25
II.6. Spéciation	25
II.6.1. Mécanismes de spéciation	25
II.6.1. Modes de spéciation	27
II.6.2. Isolement reproductif : quand l'épigénétique s'en mêle	33
II.6.3. Cospéciation	36
À retenir	37

Glossaire.....	37
QCM	37
III. Forces évolutives.....	39
III.1. Dérive Génétique.....	39
Effet fondateur	40
Goulot d'étranglement.....	43
III.2. Mutation	45
III.3. Sélection naturelle	52
À retenir	53
Glossaire.....	53
QCM	54
IV. Diversité génétique.....	55
IV. 1. Notion de polymorphisme	55
IV.2. Déterminisme épigénétique	55
IV.3. Déterminisme génétique.....	57
À retenir	59
Glossaire.....	59
QCM	60
V. Evolution des séquences nucléotidiques	61
V.1. Annotation des gènes : premier pas vers la fonction des génomes.....	61
V.2. Familles de gènes	62
À retenir	64
Glossaire.....	64
QCM	64
VI. Phylogénie Moléculaire.....	66
VI.1. Reconstruction des relations de parenté de séquences homologues.....	66
VI.2. Arbre phylogénétique	66
VI.3. Construction d'un arbre phylogénétique.....	67
À retenir	69
Glossaire.....	69
QCM	69
VII. Références :	71

I. Introduction

La théorie est apparue, au XIX^e siècle, le concept d'évolution recouvrait un ensemble de théories dont les plus importantes sont les théories transformistes, dont les principales sont le lamarckisme (transmission des caractères adaptatifs acquis pour répondre à des instructions de l'environnement), et le darwinisme (variation, sélection des individus adaptés aux conditions de l'environnement). À l'époque, il s'agissait d'hypothèses capables d'expliquer des observations, mais, en l'état des connaissances sur les mécanismes de l'hérédité, ces hypothèses ne pouvaient reposer sur aucun support matériel tangible.

Partant de l'observation et de la comparaison à différents niveaux d'individus appartenant à diverses espèces, la démarche conduit à décrire des mécanismes qui expliquent que les êtres vivants sont le résultat de l'évolution.

I.1. Définition de l'évolution

L'évolution est définie comme étant le processus par lequel les espèces se transforment. Ainsi un arbre phylogénétique peut établir les relations d'apparentement entre tous les êtres vivants.

I.2. Fondements de l'évolution et son essence génétique

Les théories de l'évolution peuvent être classées en deux groupes ; Les théories ignorant la génétique et la géochronologie et les théories intégrant la génétique et la durée des temps géologiques

I.2.1. Théories ignorant la génétique et la géochronologie

Ces théories ont été fondées par plusieurs scientifiques et représentées par deux courants : le créationnisme et l'évolutionnisme.

I.2.1.1. Créationnisme

a. Carl von LINNE (1707-1778) et la création unique

C'est un botaniste dont le père est pasteur. Linné se heurte aux difficultés de classement des plantes et plus généralement des êtres vivants. Dans *Species plantarum*, en 1753, il commence par répartir tous les végétaux en 24 classes selon la disposition de leurs étamines. À l'intérieur de

chaque classe, il établit des subdivisions (familles) caractérisées par la structure du pistil. Linné est considéré comme le père de la systématique et l'inventeur de la nomenclature binominale. Il travaille sur le vivant et ignore les fossiles; Selon LINNE, les espèces sont le fruit d'une création unique et ces espèces sont définitivement fixées dès l'origine (la fixité des espèces).

b. Georges CUVIER (1769-1832) et les créations successives

Il étudie les Vertébrés fossiles (paléontologie), note des extinctions successives des espèces. Il envisage ainsi des extinctions suivies de nouvelles créations et attribue les extinctions à des cataclysmes à répétition ("catastrophisme"); chaque création est d'essence divine et les espèces sont ainsi définitivement fixées dès l'origine (fixisme).

1.2.1.2. Evolutionnisme

Ce courant de pensée scientifique recherche des explications qu'il soumet dans le meilleur des cas à l'épreuve des observations, de l'expérimentation et de la critique.

a. Jean-Baptiste LAMARCK (1744-1829) et le transformisme (" lamarckisme ")

Observations de base

LAMARCK a travaillé sur des séries phylétiques de Mollusques fossiles (figure 1) montrant des modifications graduelles. Lamarck remarque l'existence de variations individuelles au sein de l'espèce (variabilité intra-populationnelle).



Figure1 : séries phylétiques étudiée par Lamarck (<https://www.futura-sciences.com/sante/definitions/biologie-evolution-convergente-5897/>).

Principes de la théorie de Lamarck

- les êtres vivants sont adaptés à leur milieu;
- la transformation est réalisée par modification du corps; - les organes utiles se développent, les organes inutiles s'atrophient - principe exprimé par " la fonction crée l'organe ";
- les transformations acquises par l'individu sont transmises à la descendance (principe de l'hérédité des caractères acquis)
- l'évolution va dans le sens d'une complexification des êtres vivants.

Exemples célèbres

- * l'atrophie des yeux de la Taupe vivant dans l'obscurité;
- * l'allongement du cou de la Girafe forcée de brouter les feuilles à la cime des arbres lors des périodes de sécheresse. La girafe voulant allonger son cou pour atteindre la cime des arbres a provoqué le développement de celui-ci par l'usage et a transmis ce caractère à sa descendance.

Objections et critiques

- * les variations des individus au sein de l'espèce ne sont pas forcément adaptatives;
- * l'hérédité des caractères acquis n'a jamais été démontrée; il en est de même pour les mécanismes de la transformation sous l'influence déterminante du milieu;
- * l'argumentation de Lamarck, purement intellectuelle, ne repose sur aucune base expérimentale.

b. Etienne Geoffroy SAINT-HILAIRE (1772-1844)

le " finalisme " est un héritage du lamarckisme, pour Etienne Geoffroy SAINT-HILAIRE, l'idée de transformation des espèces s'appuie sur des observations d'anatomie comparée

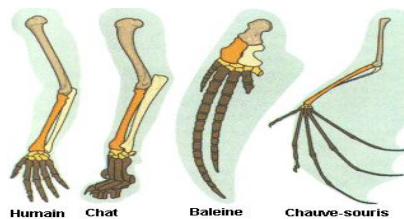


Figure 2 : Anatomie comparée des membres supérieurs (<https://www.futura-sciences.com/sante/definitions/biologie-evolution-convergente-5897/>).

L'anatomie comparée permet d'identifier que certaines structures corporelles d'espèces apparentées ont une origine commune. Par exemple, les membres antérieurs des humains, des chats, des baleines et des chauves-souris ont les mêmes types d'os bien qu'ils aient des utilités différentes (figure 2) (préhension, marche, nage et vol). Ce sont des structures homologues qui ont une origine commune.

c. Charles DARWIN (1809 -1882) et le darwinisme

Cette théorie est basée sur des observations réalisées durant son voyage autour du Monde sur le Beagle. En recensant la faune sur les îles Galápagos, Darwin a décrit 13 espèces apparentées de pinsons des Galapagos. Les plus importantes différences entre ces espèces se trouvent dans la taille et la forme du bec, l'analyse de ces différences en plus des travaux des cultivateurs et des éleveurs : par sélection artificielle ont poussé Darwin à fonder sa théorie.

Principes de la théorie de Darwin

- Les individus d'une même espèce manifestent une variabilité héréditaire;
- la variabilité au sein de l'espèce est le matériel du changement évolutif;
- les ressources du milieu de vie limitent l'excédent des populations ;

Selon Darwin l'évolution se fait donc grâce à des mécanismes évolutifs qui sont la compétition, la sélection naturelle des individus les plus aptes et la sélection sexuelle (choix des partenaires). Darwin nota que les espèces qui se ressemblent dérivent probablement d'un ancêtre commun.

Exemple célèbre :

L'allongement du cou de la Girafe. Selon Darwin au sein d'une population de girafes, le cou présente des variations de longueur. Quand les feuillages se font rares ou clairsemés (période de sécheresse), les individus aux cous les plus longs peuvent atteindre les feuilles inaccessibles aux autres et ont donc plus de chance de survivre et de transmettre leur caractère.

Objections et critiques sur 2 points fondamentaux :

- Darwin ne fournit aucune explication à la variation et à la variabilité (aptitude à varier) des espèces.
- pas d'explication de l'hérédité des caractères acquis.

I.2.2. Théories intégrant la génétique et la durée des temps géologiques

a- Génétique généalogique selon MORGAN (1900)

Avec la découverte d'un élément nouveau : Les gènes localisés dans le noyau sur les chromosomes comme supports de l'hérédité. Morgan étudie la transmission de différents caractères sur des élevages de Drosophile, à l'échelle du couple de géniteurs et de sa descendance, avec une attention particulière pour les mutations. Cette dernière est considérée comme le seul mécanisme de l'évolution, l'effet amplificateur du temps peut à lui seul expliquer l'apparition de nouvelles espèces.

Critiques :

-Les mutations sont des phénomènes aléatoires et brusques ne peuvent expliquer la variabilité du monde du vivant.

-La génétique généalogique est centrée sur la transmission des caractères existants et donc leur maintien ; les individus issus de mutations sont d'emblée adaptés à leur milieu et non soumis à la sélection.

b. Génétique évolutive des populations (# 1920): les généticiens néo-darwiniens

Fischer et Haldane (Grande Bretagne), Wright (U.S.A.), et Teissier (France) situent le problème à l'échelle des populations. Tout changement évolutif réside dans les variations de fréquence allélique et celles-ci sont influencées par l'environnement, par le biais de la sélection naturelle. Les variations des fréquences alléliques sont suivies au fil des générations successives.

Critiques

Quel lien entre les changements de fréquences alléliques au sein des populations et les faits évolutifs ?

c. Théorie synthétique (1940) ou néo-darwinisme

Julian Huxley, Ernst Mayr, Théodosius Dobzansky et George Gaylord Simpson

En 1942, J. Huxley crée le terme de **synthèse évolutive** pour cette théorie qui unit les connaissances de toute la biologie et l'analyse mathématique.

Principes de la Théorie synthétique

1) l'unité évolutive : la population

* l'espèce, constituée de populations, manifeste une variabilité héréditaire

2) le changement évolutif :

La théorie synthétique reconnaît l'existence de 2 modalités évolutives: **anagenèse et cladogénèse**.

Anagenèse : transformation progressive d'une seule lignée par la dérive morphologique de la population (transformation discrète, graduelle et irréversible) => 'pseudo-extinction' (figure 3).

Cladogénèse : éclatement d'une lignée ancestrale en plusieurs lignées descendantes (figure 3).

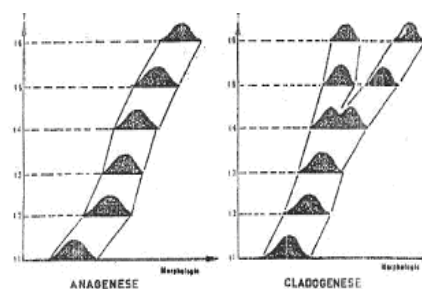


Figure 3 : illustration des processus anagenèse et cladogénèse (<https://www.biodeug.com/licence-3-paleontologie-chapitre-5-microevolution/>)

3) mécanisme du changement évolutif :

Les transformations évolutives sont dues à des mutations

4) orientation du changement évolutif : la sélection naturelle à toute mutation (et à la transformation correspondante) peut être attribuée une valeur adaptative et sélective

- valeur positive si elle procure un avantage par rapport aux individus non mutants;
- valeur négative dans le cas inverse.

Ces transformations sont triées par le milieu via la sélection naturelle: un allèle est retenu ou éliminé en fonction de sa valeur adaptative.

5) l'effet amplificateur du temps.

Si une (ou des) mutation(s) favorable(s) se produi(sen)t en grand nombre dans une population et si elles se répète(nt) sur plusieurs générations, alors les mutants favorisés envahissent la population qui évolue naturellement vers une autre espèce.

Les processus évolutifs sont ramenés à des changements de fréquence allélique au sein des populations.

Avec le temps, au fil des générations, et par addition / répétition de changements minimes (micro-évolution) se produisent des transformations majeures (macro-évolution) .

Critiques des paléontologistes, systématiciens embryologistes :

- quel est le lien entre les changements de fréquence allélique et les changements de phénotype?
- La théorie synthétique a ignoré la génétique du développement passant directement du gène à l'organisme soumis aux contraintes de l'environnement.
- Nous parlons de convergence lorsque la sélection naturelle conduit à l'acquisition de structures ou d'aptitudes similaires au sein de lignées évolutives différentes, mais soumises à des conditions environnementales équivalentes.

I.3. Critiques la théorie de l'évolution

a- Mutation

La mutation est un évènement rare. Le taux de mutations chez les êtres vivants varie en fonction de plusieurs facteurs, notamment l'espèce, le type de mutation, et les mécanismes de réparation de l'ADN. Voici quelques valeurs générales :

-Chez les bactéries et les virus : Le taux de mutation est relativement élevé. Par exemple, les virus à ARN, comme le VIH et le SARS-CoV-2, ont un taux de mutation d'environ 10^{-3} à 10^{-5} mutations par nucléotide et par réplication.

-Les bactéries, comme E. coli, ont un taux d'environ 10^{-9} mutations par nucléotide par génération.

-Chez les eucaryotes unicellulaires : Comme la levure *Saccharomyces cerevisiae*, le taux de mutation est d'environ 10^{-10} mutations par nucléotide par génération.

- Chez les mammifères (y compris l'humain) :

Le taux de mutation par nucléotide par génération est d'environ 10^{-8} . Chez l'humain, cela se traduit par environ 50 à 100 nouvelles mutations dans le génome d'un individu par génération. La majorité des mutations surviennent lors de la spermatogenèse, et le nombre de mutations augmente avec l'âge du père.

-Chez les plantes : Le taux de mutation varie également, mais il est généralement comparable à celui des animaux.

En plus de la rareté des mutations, La majorité des mutations sont neutres et une grandes partie sont délétères.

-La fixation d'une mutation peut nécessiter des millions d'années. En effet la fixation d'une mutation dans une population dépend de plusieurs facteurs, notamment la sélection naturelle, la dérive génétique, et le mode de reproduction. En moyenne, le temps nécessaire pour qu'une mutation se fixe (atteigne une fréquence de 100 % dans la population) peut être estimé selon son effet et sa fréquence initiale.

Cas d'une mutation neutre (sans effet sur la survie)

La fixation est due à la dérive génétique (fluctuations aléatoires des fréquences des allèles). Le temps moyen de fixation est environ $4 N_e$ générations, où N_e est la taille effective de la population. Exemple : Pour une population humaine avec une $N_e \approx 10\,000$, une mutation neutre mettrait en moyenne 40 000 générations (soit environ 1 million d'années) pour se fixer.

Cas d'une mutation bénéfique

Une mutation avantageuse est favorisée par la sélection naturelle, et sa fixation est beaucoup plus rapide. Le temps moyen de fixation est d'environ $(2/s) \ln(2N_e)$ générations, où s est le coefficient sélectif (force de la sélection).

Exemples :

Pour une mutation avec $s = 1\%$ (0.01) et $N_e = 10\,000$, la fixation prend environ 4 000 à 5 000 générations.

Avec $s = 10\%$ (0.1), la fixation peut se faire en quelques centaines de générations seulement.

b- Sélection

La sélection naturelle ne peut agir que sur la variation génétique existante. Les compromis limitent également la sélection naturelle. Les gènes voisins sur le même chromosome sont souvent liés et hérités ensemble. Les traits intermédiaires peuvent également limiter la sélection naturelle.

La sélection naturelle augmente généralement la capacité d'une population à survivre et à se reproduire, d'autres mécanismes évolutifs pourraient avoir l'effet inverse. Des allèles nocifs peuvent être introduits et des allèles utiles effacés par la migration ou dérive génétique.

À retenir

- L'évolution est un processus de transformation des espèces dans le temps.
- Deux grands groupes de théories : sans génétique (Lamarck, Darwin) et avec génétique (Morgan, théorie synthétique).
- La sélection naturelle est centrale dans l'évolution darwinienne, mais les mutations et la génétique expliquent la variabilité.

Glossaire

Fixisme : Théorie selon laquelle les espèces sont immuables.

Transformisme : Hypothèse d'un changement progressif des espèces.

Sélection naturelle : Survie et reproduction des individus les mieux adaptés.

Anagenèse : Transformation progressive d'une lignée.

Cladogenèse : Diversification d'une espèce en plusieurs lignées.

QCM

1. Quelle affirmation est correcte concernant la théorie de Darwin ?

- A. Elle repose sur la génétique mendélienne
- B. Elle explique l'origine des mutations
- ☐ C. Elle repose sur la sélection naturelle et la variabilité héréditaire
- D. Elle réfute l'existence d'un ancêtre commun

2. Quelle est la différence principale entre anagenèse et cladogenèse ?

- A. L'anagenèse implique plusieurs lignées, la cladogenèse une seule
- ☐ B. L'anagenèse est une transformation linéaire, la cladogenèse une diversification
- C. Les deux sont identiques
- D. L'anagenèse se produit uniquement chez les plantes

3. Qui est à l'origine de la théorie synthétique de l'évolution ?

- A. Linné et Lamarck
- B. Darwin et Wallace
- ☐ C. Huxley, Mayr, Dobzhansky, Simpson
- D. Morgan et Mendel

4. Qu'est-ce que la théorie fixiste soutient ?

- A. Les espèces changent en réponse à l'environnement
- B. Les espèces sont issues d'un ancêtre commun
- ☐ C. Les espèces ont été créées une fois pour toutes et sont immuables
- D. Les mutations sont à l'origine de toutes les espèces

II. Notion de l'espèce et les différents modes de spéciation

II.1. introduction

La classification évolutionniste traditionnelle reconnaît sept niveaux fondamentaux de classification de plus en plus globalisants, de l'espèce (considérée comme l'unité fondamentale et irréductible de l'évolution) au règne (la plus globale des catégories) :

- l'espèce ;
- le genre ;
- la famille ;
- l'ordre ;
- la classe ;
- l'embranchement (ou phylum) ;
- et le règne.

Chaque espèce est nommée par deux mots latins (*binomen*), le premier étant le nom du genre, et le second l'adjectif de l'espèce.

ex : *Canis lupus* (le loup)

Staphylococcus aureus (le staphylocoque doré).

Cette manière de nommer les espèces a été fixée par Carl Linné (1707-1778) et est utilisée depuis (nomenclature binominale ou linnéenne). La nomenclature linnéenne est utilisée depuis le XVIII^e siècle, mais la notion d'espèce a considérablement évolué depuis.

La notion d'espèce est d'autant plus importante en biologie qu'elle constitue l'unité taxinomique fondamentale évolutive la plus évidente pour tous. Son utilisation dans plusieurs domaines de la biologie en précise l'intérêt :

- Les taxinomies et les systématiques reposent souvent sur la notion d'espèce.
- Le raisonnement génétique est fondé sur la transmission d'allèles chez des individus d'une même espèce et rarement au niveau du genre ou de la famille.
- L'étude de l'évolution se rapporte à celle de l'espèce.
- L'écologie utilise abondamment la notion d'espèce.
- La recherche médicale, en parasitologie par exemple.

Le mot Espèce du latin « *species* » veut dire Catégorie, type ou encore apparence

II.2. Critères de classification d'une espèce

Le concept d'espèce a fait l'objet de nombreuses discussions, avec une intensification du débat depuis une vingtaine d'années. Coyne et Orr (2004) ont ainsi répertorié 25 concepts d'espèces différents, bien que leur liste ne soit pas exhaustive. Ce vieux débat persiste car il n'existe pas de concept applicable à tous les taxa sans ambiguïté (Hey, 2001), qui prenne en compte les particularités des systèmes de reproduction de certaines espèces - par exemple, la reproduction clonale des bactéries - et l'aspect dynamique de la formation des espèces.

Dans ce cours quatre concepts seront présentés :

- Morphologique (*phénétique*)
- Biologique
- Ecologique
- Cladiste

II.2.1. Concept morphologique

À partir de leurs caractéristiques, les individus sont classés en catégories distinctes : la ressemblance morphologique semble un critère suffisant pour déterminer l'appartenance d'un organisme à une espèce.

En paléontologie, ce concept est fondamental ; les fossiles ne livrent, en effet, que des informations morpho-anatomiques utilisées pour distinguer les espèces entre elles.

- **Phénétique**

Elle correspond à une méthode de classification du vivant reposant sur le nombre de caractères communs à deux ou plusieurs espèces. Elle s'appuie sur le principe que des individus partagent d'autant plus de similitudes que leur dernier ancêtre commun est récent. C'est une méthode numérique et mathématique, permettant d'établir une matrice des distances, représentée à l'échelle par un phénogramme (figure 4).

L'Espèce est donc définie comme étant un ensemble d'organismes qui se ressemblent plus entre eux qu'ils ne ressemblent à des organismes appartenant à d'autres ensembles équivalents.

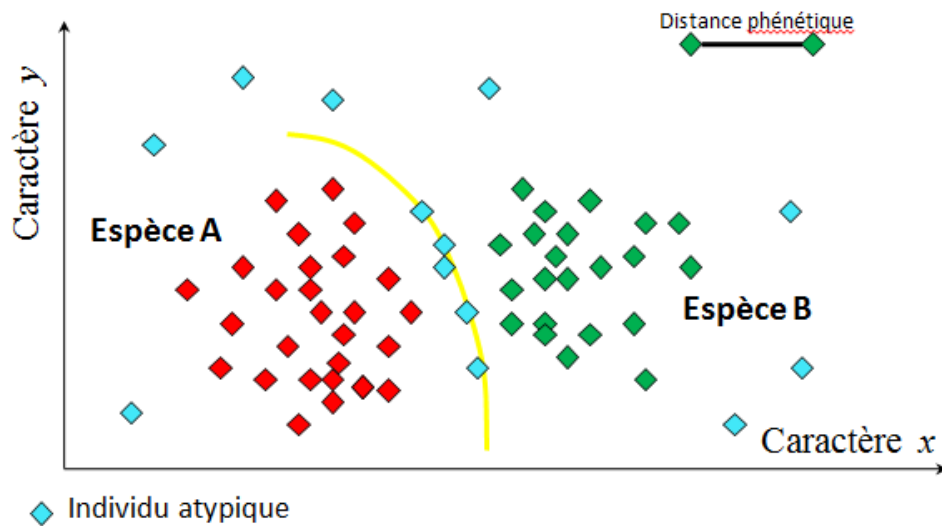


Figure 4 : exemple d'une classification phénétique (<https://essicolo.github.io/ecologie-mathematique-R/chapitre-ordination.html>).

La notion morphologique est la plus utilisée y compris actuellement mais elle présente ses limites qui sont :

- Subjectivité des critères retenus pour la définition d'une espèce
- Variations intrapopulation: dimorphisme sexuel, écophénotype...

a. Dimorphisme sexuel

Femelle et mâle du canard colvert présentent des colorations très différentes; ils appartiennent cependant à la même espèce.

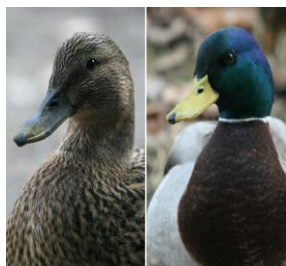


Figure 5 : exemple de dimorphisme sexuel chez le canard (<https://www.zoo-servion.ch/canard-carolin>).

Jusqu'à très récemment (2008), le mâle adulte (à gauche), long de quelques centimètres et la femelle adulte (à droite), dépassant les 15 centimètres (figure 6), d'une même espèce pélagique de poissons osseux, étaient considérés comme des représentants de deux espèces distinctes appartenant à deux familles différentes.

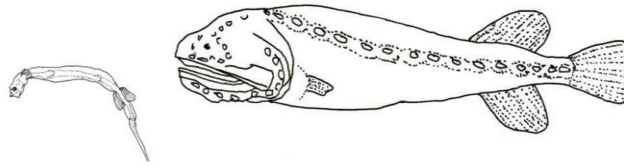


Figure 6 : exemple de dimorphisme sexuel chez une espèce pélagique de poissons osseux (Futuyma, D. J. 2017).

b. Ecophénotype : variations locales

C'est la variabilité morphologique qui s'exprime suite à une adaptation aux conditions locales et momentanées du milieu (figure 7).

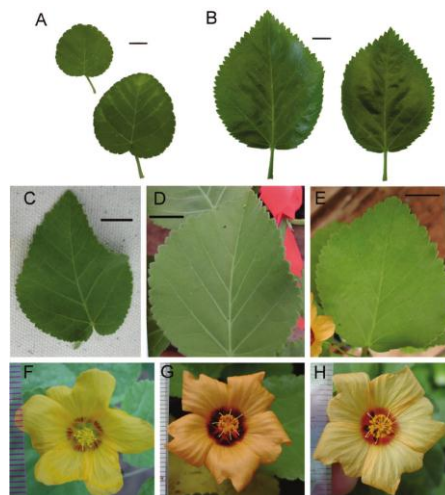


Figure 7 : Ecophénotypes chez *Sida fallax*. Photos représentatives de (A) feuilles d'écotype de plage, (B) feuilles d'écotype de montagne, (C , D , E) feuilles d'hybrides F1, (F) fleur d'écotype de plage, (G) fleur hybride (B 3 M) et (H) fleur d'écotype de montagne (Mitsuko, 2006).

- **espèces jumelles**

Des espèces différentes peuvent être morphologiquement très ressemblantes, on les appelle espèces jumelles. C'est l'exemple des Sturnelle (figure 8) dont la différenciation est faite à partir du comportement et des habitudes de reproduction.



Figure 8 : Sturnelle de l'ouest à gauche et Sturnelle des prés à droite (<http://www.evolution-biologique.org/mecanismes/speciation/les-especes-jumelles.html>)

Conclusion

- Le concept morphologique ne constitue pas réellement une définition de l'espèce
- Concept indispensable pour les espèces fossiles

II.2.2. Concept biologique

Georges BUFFON (1707-1788), dans son Histoire naturelle (1749), reconnaît que l'espèce est constituée d'individus interféconds. Ce concept a été réactualisé, principalement, par Ernst MAYR (1904- 2005) : l'espèce est un ensemble de populations naturelles véritablement ou potentiellement interfécondes, isolées d'autres groupes comparables avec lesquels elle ne se reproduit pas.

Le concept biologique de l'espèce identifie les espèces comme des groupes de populations naturelles interféconds qui sont reproductivement isolés d'autres groupes semblables (Mayr, 1942, cité par Coyne et Orr, 2004).

II.2.2.1. Barrières reproductives

L'utilisation du concept biologique de l'espèce permet d'étudier l'évolution de barrières reproductives, définies comme des traits biologiques des organismes limitant les flux de gènes entre individus de différents taxa (Coyne et Orr, 2004). Il faut souligner que la formation de nouvelles espèces est un processus dynamique, se traduisant par la possibilité de flux de gènes comme étape intermédiaire avant l'isolement reproducteur complet (Figure 9). Lorsque la spéciation est en cours,

il est plus difficile de se baser uniquement sur le critère d'isolement reproducteur pour définir à partir de quelle étape les taxa étudiés peuvent être reconnus comme de nouvelles espèces (Coyne et Orr, 2004).

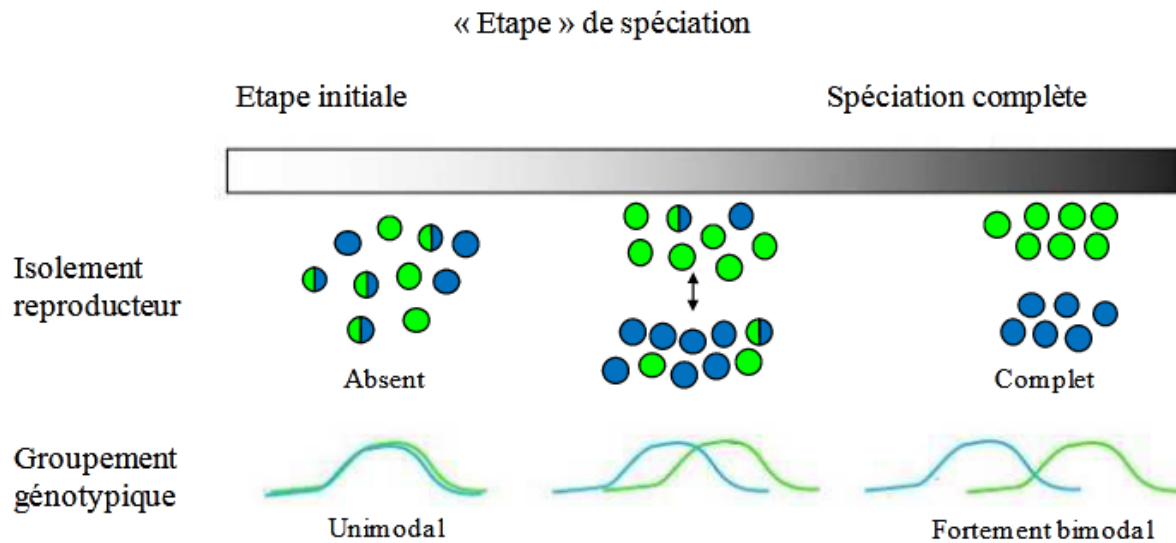


Figure 9 : Représentation de l'isolement reproductif assuré par des barrières biologiques qui s'opposent aux mélanges entre espèces différentes (<https://planet-vie.ens.fr/media/2207>).

Il existe deux types de barrières biologiques (figure 10):

- a. Les barrières prézygotiques, qui interviennent en faveur d'un isolement reproductif, empêchent les croisements,
- b. les barrières postzygotiques, mises en place après la fécondation ou la formation du zygote, empêchent la survie ou le développement ultérieur de l'embryon.

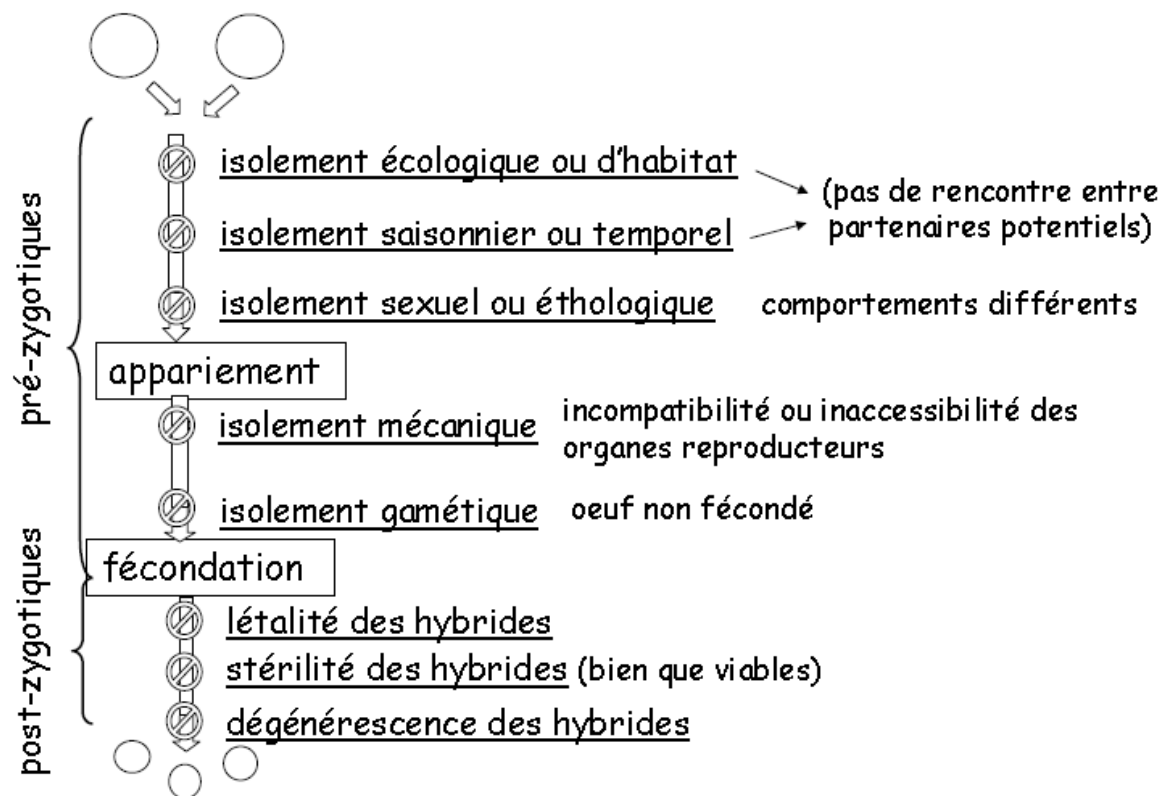


Figure 10 : différents types de barrières biologiques (<https://www.slideserve.com/chico/bases-biologiques-du-comportement-introduction>).

II.2.2.2. Exemples de barrières reproductibles

Le nombre et le type de barrières reproductives sont très variables entre taxa, mais des tendances générales ont pu être observées. Ainsi, chez les animaux, les barrières prézygotiques sont généralement contrôlées par quelques gènes à effets majeurs dans les premières étapes du processus de spéciation, tandis que les barrières postzygotiques sont plutôt contrôlées par de nombreux gènes à effets mineurs qui s'accumulent graduellement (Coyne et Orr, 2004). Les taxa de plantes sont typiquement séparés par de multiples barrières reproductives avec des effets individuels limités, plutôt qu'une seule barrière reproductrice d'effet majeur sur la réduction des flux de gènes entre taxa, suggérant que l'isolement reproducteur complet est atteint après un nombre important de générations (Rieseberg et Willis, 2007 ; Widmer et al., 2009). Cependant, la spéciation par polyploïdie, avec une augmentation du nombre de chromosomes, entraîne un isolement reproducteur quasi instantané et est fréquemment observée chez les plantes supérieures (Stebbins, 1940 ; voir pour revue Soltis et al., 2004 et Rieseberg et Willis, 2007, et notamment le cas de *Spartina anglica*, Ainouche et al., 2004).

Parmi les barrières prézygotiques on a les barrières reproductibles comportementales (figure 11), l'isolement écogéographique (figure 12), l'isolement temporel (figure 13) et l'isolement gamétique (figure 14). En ce qui concerne les barrières postzygotiques, plusieurs mécanismes y contribuent : non viabilité des hybrides, stérilité des hybrides et déchéance des hybrides (figure 15).

Avec la présence d'incompatibilité, malgré une capacité de dispersion importante pendant la phase larvaire planctonique, une forte structuration en mosaïque de la zone hybride a été décrite chez ces taxa. Chez les macroalgues marines, la majorité des études de spéciation se sont cependant concentrées sur la description d'espèces cryptiques, à l'aide de phylogénies et parfois de croisements testant l'isolement reproducteur. Des études montrent cependant le rôle dans l'isolement reproducteur du système de reproduction chez les espèces de *Fucus* tandis que d'autres suggèrent la présence d'isolement temporel chez des algues vertes d'eaux tropicales. La présence de protéines de reconnaissance entre gamètes jouant un rôle dans l'isolement reproducteur a également été montrée chez les algues rouges *Aglaothamnion*. Ces études d'isolement reproducteur chez les macroalgues marines sont donc actuellement peu nombreuses, ce qui ne permet pas de déterminer les contributions relatives de l'isolement prézygotique et postzygotique et d'inférer l'importance des traits du cycle de vie sur les barrières reproductives. Une connaissance détaillée de l'importance des différentes barrières reproductives est cependant une première étape nécessaire pour ensuite identifier les forces évolutives qui favorisent la spéciation et prédire l'issue d'une possible hybridation en cours.



Figure 11 : Barrières reproductibles comportementales
(https://slideplayer.fr/slide/8071444/#google_vignette)

Isolement
écologique
Pas dans le
même habitat.



Le campagnol des champs
vit dans les champs.



Le campagnol des rochers
vit sur les rochers.

Figure 12 : Barrière géographique (https://slideplayer.fr/slide/8071444/#google_vignette)

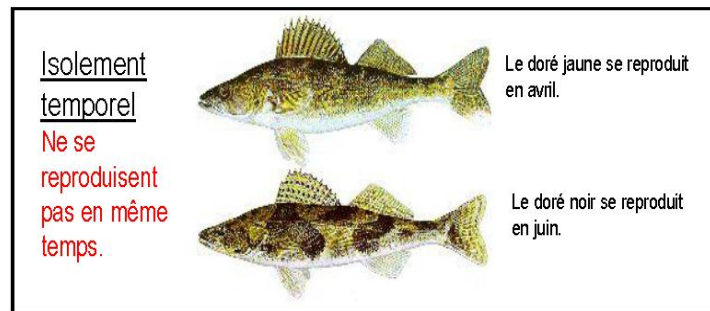


Figure 13 : Barrière chronologique (https://slideplayer.fr/slide/8071444/#google_vignette)



Les gamètes
libérés par les
mâles des
poissons ne
fécondent que
les œufs de
leur espèce.

Figure 14 : Barrière anatomique isolement gamétique
(https://slideplayer.fr/slide/8071444/#google_vignette)



Figures 15 : exemples de barrières postzygotiques
(https://slideplayer.fr/slide/8071444/#google_vignette)

Difficultés d'application du concept biologique

- ✓ Difficultés liées à des modalités particulières de reproduction
 - ➔ Espèces « uniparentales » dites apomictiques
- ✓ Difficultés liées à une méconnaissance des systèmes de reconnaissance
 - ➔ Notamment pour les groupes « inférieurs » : lichens, « algues », bryophytes
- ✓ Depuis les années 1950, découverte des espèces voisines capables de s'hybrider !!!
- ✓ G. COUSIN (1968) a découvert un grand nombre d'hybrides fertiles chez les Grillons, dont il a étudié 31 espèces, réparties en 3 genres.

Le concept reproductif est loin d'être universel : plus de la moitié de l'histoire de la vie sur Terre est due à des organismes asexués ou uniparentaux.

- Le flux génique

L'isolement reproductif permet le maintien de l'identité de l'espèce. Lors des croisements, les gènes circulent constamment parmi les membres de l'espèce, de génération en génération. Le flux génique, ainsi formé, réduit les différences locales et il suffit à maintenir les caractères spécifiques autour d'un type moyen.

Selon E. MAYR, le flux génique est le principal facteur de cohésion spécifique et il assure aussi la stabilité relative des frontières de l'espèce. À la suite d'une migration, seuls quelques pionniers frontaliers peuvent acquérir de nouvelles caractéristiques.

II.3. Concept écologique

Selon ce concept l'espèce est un groupe d'individus partageant une même niche écologique. ce concept n'est pas très utilisé car deux espèces distinctes peuvent partager la même niche et une même espèce peut occuper différents niches.

II.4. Concept cladiste (phylogénétique)

Les généalogies de gènes permettent d'identifier des groupes d'organismes partageant un ancêtre commun et donc une histoire commune. Différents concepts phylogénétiques de l'espèce ont alors été proposés en se focalisant sur l'histoire commune des individus. Ainsi, le premier concept définit l'espèce phylogénétique comme un groupe irréductible, basal, d'organismes qui partagent une relation d'ascendance et de descendance, et qui sont distincts d'autres groupes semblables.

Le processus de spéciation du point de vue phylogénétique peut être vu comme le processus évolutif par lequel les généalogies de gènes passent de peu structurées à résolues, avec des groupes phylogénétiques dont la séparation est importante. Pendant ce processus, les barrières reproductives sont essentielles car ce sont elles qui favorisent la séparation au cours du temps des taxa phylogénétiques, en limitant les flux de gènes. Lorsque deux taxa sont reproductivement isolés, les généalogies de gènes vont passer par des étapes successives de polyphylie, puis possiblement par une étape de paraphylie avant de montrer finalement une monophylie réciproque : à ce stade, les copies de gènes dans chaque taxon sont plus proches phylogénétiquement entre elles qu'avec les copies de gènes de l'autre taxon. Cependant, la monophylie réciproque est un stade qui ne peut être atteint qu'après une longue période de temps, qui dépend de la taille efficace des populations: même en l'absence de flux de gènes entre taxa, les processus de dérive génétique et de sélection naturelle n'éliminent que lentement le polymorphisme ancestral.

Le concept phylogénétique de l'espèce est critiqué en particulier car il est impossible d'atteindre la monophylie réciproque entre taxa pour tous les gènes du génome : d'autres événements de spéciation auront probablement lieu avant, et de plus, la sélection peut préserver des allèles communs entre taxa pour certains loci. Des critères subjectifs, ont alors été définis, comme la présence de 50% de gènes montrant une monophylie réciproque.

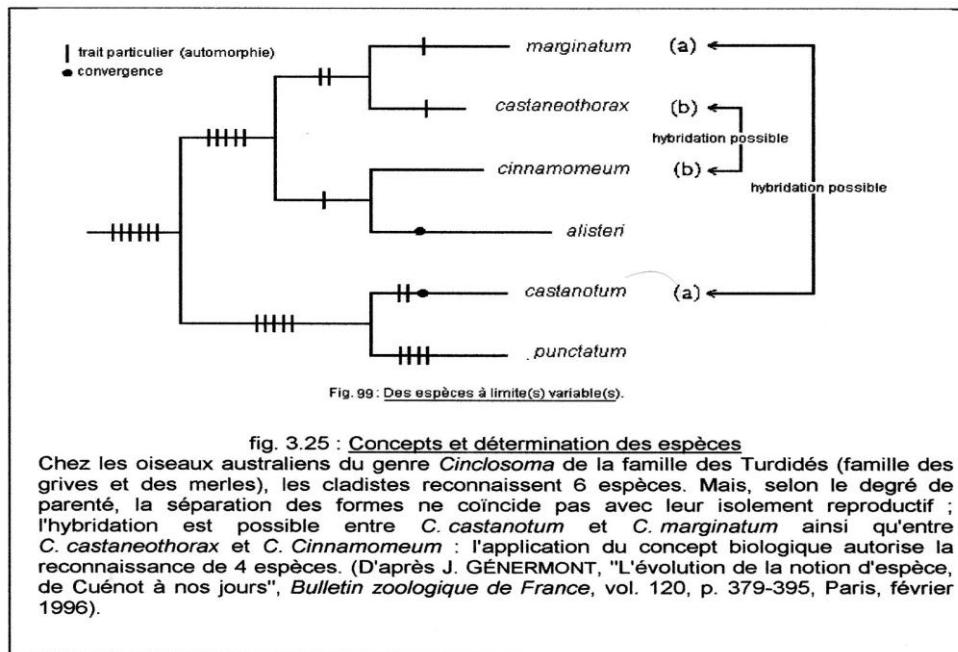


Figure 16 : exemples de classification cladiste

(<https://www.aquaportail.com/dictionnaire/definition/5073/phenetique>)

Conclusion

Depuis quelques années, la taxonomie traditionnelle, basée sur les critères morphologiques, tend à être complétée par l'utilisation de courtes séquences d'ADN d'un gène particulier pour caractériser les espèces et identifier des espèces cryptiques. L'utilisation d'un seul critère (un seul gène) pose cependant le problème de la généralisation à l'ensemble du génome : le gène considéré, généralement mitochondrial, peut par exemple être soumis à sélection ou présenter des copies nucléaires et ne pas refléter l'histoire de l'espèce, mais celle du gène. Quoiqu'il en soit, l'utilisation de phylogénies de plusieurs gènes a l'avantage de permettre d'étudier les aspects historiques, géographiques et démographiques des processus de spéciation. En particulier, les phylogénies de gènes nous renseignent sur l'ancienneté de la divergence entre taxa.

La définition la plus communément admise est énoncé par Ernst Mayr en 1942 : « une espèce est une population ou un ensemble de populations dont les individus peuvent effectivement ou potentiellement se reproduire entre eux et engendrer une descendance viable et féconde, dans des conditions naturelles. Ils sont génétiquement isolés d'autres groupes similaires ».

II.5. Notion de population

Selon les propositions de la théorie synthétique de l'évolution énoncées par E. Mayer (1942), la notion de population représente l'unité qui évolue dans le temps. « Une population est un ensemble d'individus d'une même espèce vivant dans une zone géographique suffisamment petite pour que tous les partenaires sexuels présentent la même probabilité de rencontre »

II.6. Spéciation

La spéciation représente l'ensemble des processus qui conduisent à l'apparition de nouvelles espèces.

II.6.1. Mécanismes de spéciation

La spéciation passe par trois étapes :

- 1-Arret des échanges génétiques
- 2-Evolution des mécanismes d'isolement reproductif
- 3-Accumulation de différences génétiques, comportementales et morphologiques.

Trois forces évolutives interviennent lors de la spéciation ; la dérive génétique, les mutations et la sélection.

Processus de spéciation

1- Anagenèse :

se rapporte à l'évolution graduelle d'une espèce qui se transforme progressivement pour donner naissance à une nouvelle espèce, sans qu'il y ait ramification ; on parle également d'une évolution phylétique.

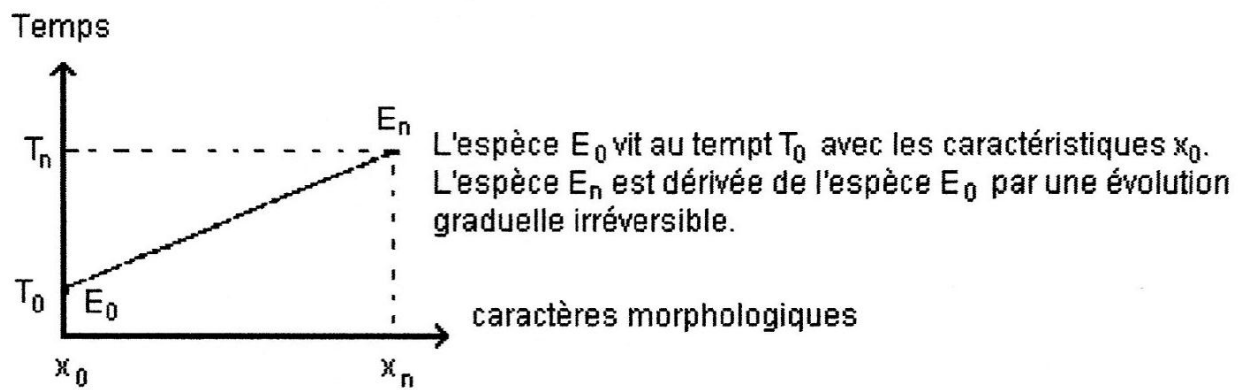


Figure 17 : Evolution d'une espèce dans le temps selon le processus Anagenèse (https://v-assets.cdnsnw.com/fs/Root/cktan-Sortie_MNHN_correction_CR_Chapitre_6.pdf)

2- Cladogenèse

C'est la séparation d'une lignée évolutive en deux rameaux sous l'effet d'une cause quelconque : glaciation, orogénèse, ouverture d'un isthme, cours d'eau....Chacun des deux rameaux ainsi isolés va évoluer irréversiblement pour son propre compte et donner peut-être une nouvelle espèce.

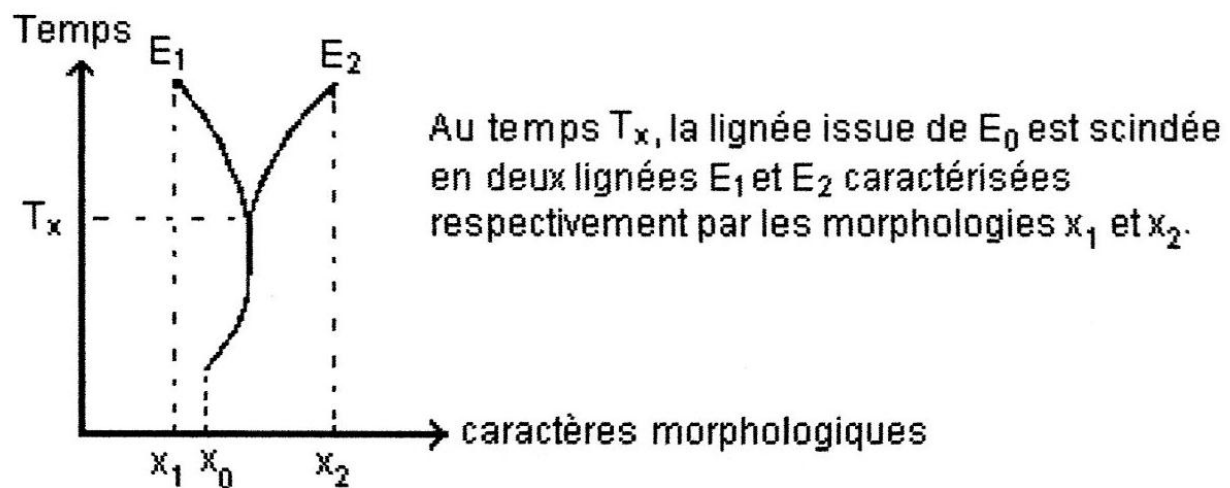


Figure 18 : Evolution d'une espèce dans le temps selon par Cladogenèse (https://v-assets.cdnsnw.com/fs/Root/cktan-Sortie_MNHN_correction_CR_Chapitre_6.pdf).

II.6.1. Modes de spéciation

Les modes de spéciation sont traditionnellement classés selon le degré de divergence géographique des aires de distribution des taxa en formation.

1- Spéciation allopatrique, où l'origine des différentes espèces résulte de l'évolution divergente de populations qui sont géographiquement isolées l'une de l'autre. Dans le processus de spéciation allopatrique, on distingue souvent selon la taille relative des populations impliquées la spéciation par vicariance de la spéciation par « bourgeonnement » (également appelée spéciation péripatrique). Dans le premier cas, l'isolement reproducteur évolue après la division de l'aire de distribution d'une espèce en deux nouveaux groupes isolés mais de tailles importantes. Dans le second cas, l'un des nouveaux groupes est de taille bien plus réduite que l'autre : le plus petit groupe peut être la conséquence de l'isolement géographique d'une petite population ou de la migration de quelques individus vers un habitat isolé.

- Pour une **spéciation allopatrique il faut des populations séparées géographiquement, ce qui provoque une nette diminution ou un** arrêt des échanges génétiques et l'accumulation progressive de différences entre les populations.
- Ce mode de spéciation présuppose aussi que les deux populations sont confrontées à des milieux différents permettant ou provoquant des adaptations différentielles.
- Apparition d'une barrière géographique
- Scission de la Population en plusieurs groupes (rupture des flux de gènes
- souvent **faible** effectif dérive génétique
- Conditions écologiques intergroupe différentes : Pression sélective différentes
- Différenciation génétique : même en cas de la disparition de la barrière écologique on a apparition d'une barrière biologique

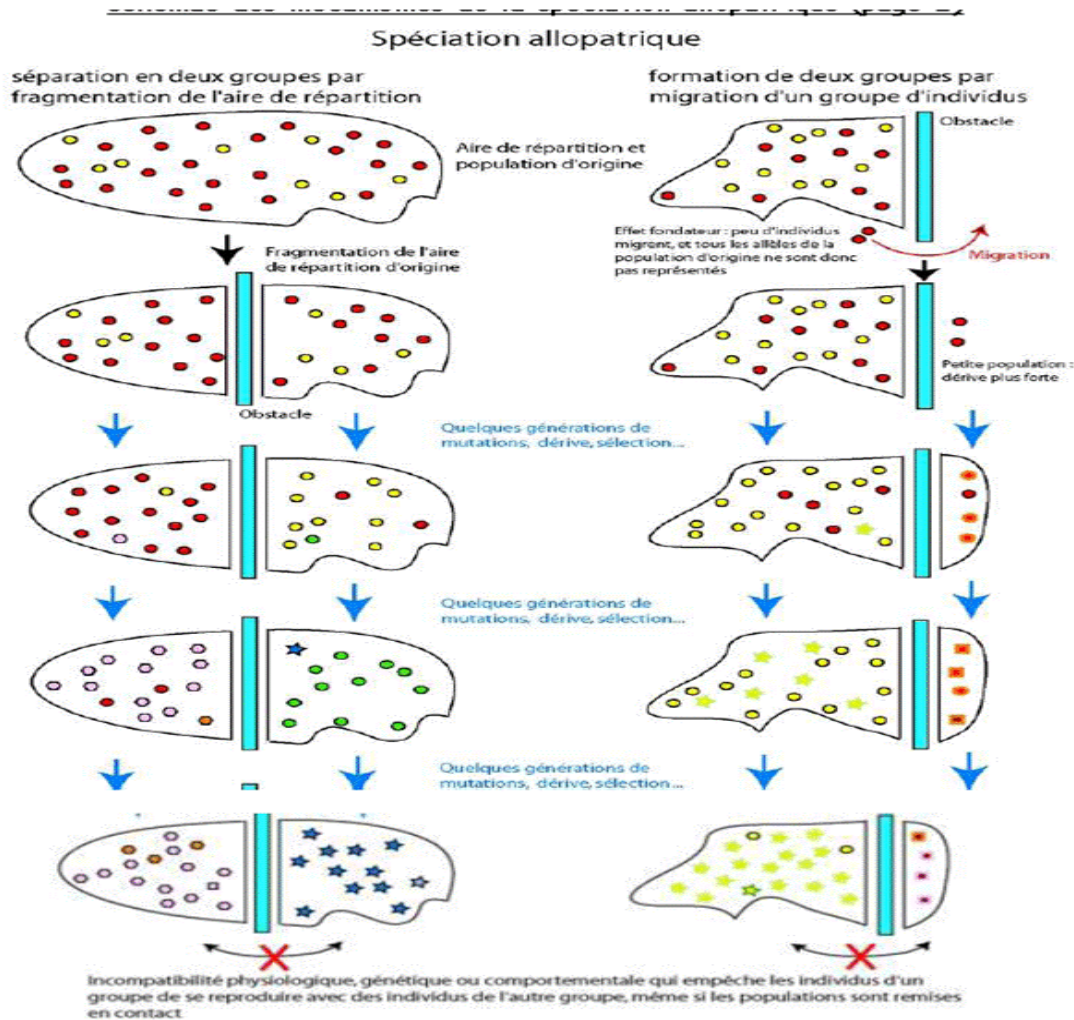


Figure 19 : Spéciation allopatrique (<https://planet-vie.ens.fr/media/2207>)

2- Spéciation sympatrique, où des barrières reproductives entre individus évoluent à l'intérieur d'une seule et même région géographique. Ce mode de spéciation se caractérise par la géographie et par la possibilité de flux de gènes entre les taxa durant le processus de spéciation.

- La spéciation est sympatrique si deux populations en voie de séparation vivent **côte à côte sur le même territoire**
- Un processus particulièrement fréquent chez les plantes.
- Apparition d'un isolement reproductif au sein d'une même population (coexistence de deux espèces sympatriques).
- Mode de spéciation rapide = quelques générations
- Populations non isolées géographiquement qui évoluent en espèces distinctes
- Apparition d'un polymorphisme génétique au sein de la population
- Accroissement de l'isolement reproductif à l'intérieur de la population
- Apparition d'une deuxième espèce en sympatrie avec la première

Causes de la spéciation sympatrique :

- Remaniements chromosomiques
- Hybridation et/ ou polyploidie

a- Hybridation interspécifique



Helianthus annuus



Helianthus petiolaris



Espèce nouvelle avec sa propre
niche écologique (milieu désertique)

Helianthus anomalus

Figure 20 : Exemple de spéciation par hybridation interspécifique (<https://www.futura-sciences.com/planete/dossiers/developpement-durable-mali-oasis-biodiversite-570/page/4/>)

b- Spéciation homoploïde

Isolement reproductif entre l'hybride et ses parents : formation d'une nouvelle espèce du même nombre chromosomique (homoploïde).

c-Remaniements chromosomiques

- L'évolution des caryotypes peut se faire sous l'influence de délétions, inversions, translocations, duplications, affectant des portions de chromosomes, fusions ou fissions (cassure au niveau du centromère).
- Le caryotype restructuré donne naissance à des individus plus fertiles entre eux qu'avec ceux ayant gardé le caryotype initial----- barrière de reproduction.

Exemples: lilas, Volail, poule et dinde

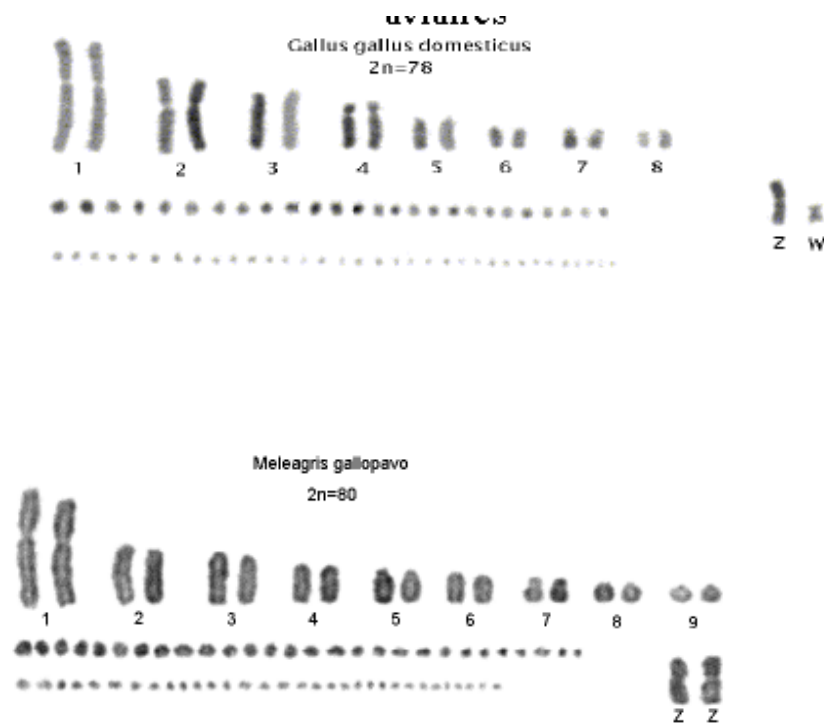


Figure 21 : Comparaison des caryotypes de la poule et la dinde (Vignal et al., 2013).

Tableau 1 : Comparaison des caryotypes de la poule, la dinde, le canard et la caille (Vignal et al., 2013).

Macrochromosomes de poule	Macrochromosomes de caille	Macrochromosomes de dinde	Macrochromosomes de canard
1	1 (inversion)	1 (réarrangements)	1
2	2 (inversion)	3 + 6	2
3	3	2	3
4	4 (inversion)	4 (sans le bras p)	4 (réarrangements et sans bras p)
5	5	5 (avec un bras p)	5 (réarrangements)
6	6	8	6 (réarrangements)
7	7 (réarrangements)	7	7 (réarrangements)
8	8 (inversion)	9	8 (réarrangements)
Z	Z	Z (inversion)	Z (réarrangements)
Nombre de sondes	28/29	43/48	25/26

d- Spéciation par polyploidie

- Mode de spéciation très répandue chez les plantes (70% des angiospermes et 95% des fougères)
- après le séquençage du génôme de plantes qu'on croyait diploïde, (ex: *Arabidopsis thaliana*= première plante supérieure séquencée en 2000) on a découvert que la plupart des angiospermes avaient subi plusieurs cycles de polyploidisation et diploidisation au cours de leur histoire. L'ancêtre d'*Arabidopsis* était polyploïde on parle de paleopolyploidisation.
- On sait aujourd'hui que la polyploidie a marqué toutes les lignées eucaryotes (champignons...) mais ça fonctionne toujours bien dans le système végétal.

Exemple :

Jusqu'en 1870, on ne rencontrait que deux espèces de spartine : *Spartina maritima* sur les côtes européennes et *Spartina alterniflora* sur les côtes américaines. En 1880, sur les côtes anglaises, est recensée une troisième espèce jusqu'alors inconnue, qui est nommée *Spartina anglica*.

Ces trois espèces possèdent des caryotypes différents :

_ *Spartina maritima* : $2n = 60$

_ *Spartina alterniflora* : $2n = 62$

_ *Spartina anglica* : $2n = 122$

Les scientifiques expliquent que, dans un premier temps, *Spartina maritima* et *Spartina alterniflora* se sont croisées formant un **hybride viable** mais **stérile** avant qu'un phénomène de **polyploïdisation** ne restaure la reproduction sexuée de l'hybride. Ce dernier devenant alors une nouvelle espèce, *Spartina anglica*.

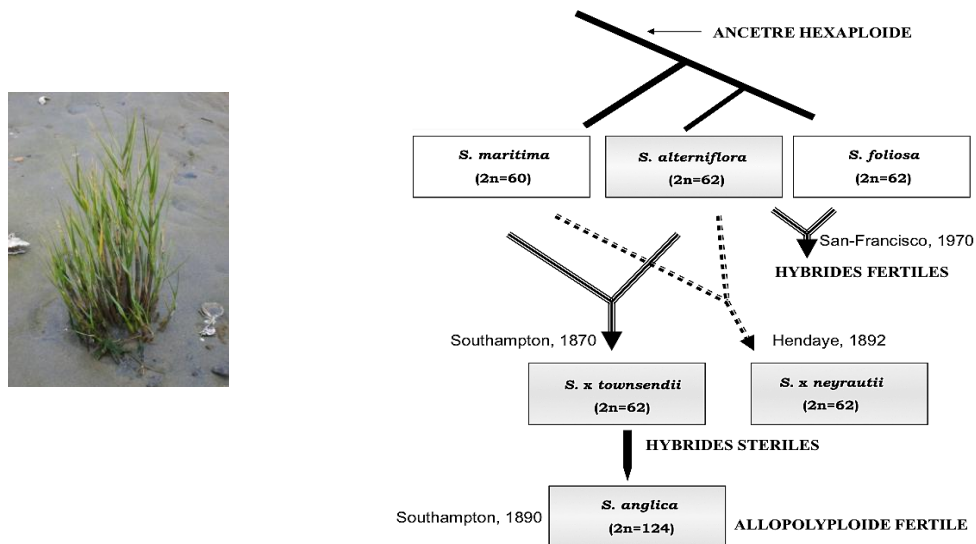


Figure 22 : spéciation des espèces de spartine (Ainouche et *al.*, 2004).

- Les hybrides et allopolyploïdes européens ont hérité du génome chloroplastique de *S. alterniflora*.
- Les nombres chromosomiques sont ceux rapportés par Marchant (1968), qui observa une variation des nombres de chromosomes pour *S. anglica* ($2n = 120, 122, 124$).

3- Spéciation parapatricque est un intermédiaire entre les deux premiers modes de spéciation : dans ce cas, l'isolement reproducteur évolue entre des populations dont les aires de distribution sont adjacentes et entre lesquelles les échanges de gènes sont possibles mais limités.

- Evolution rapide d'un ensemble d'espèces à partir d'un ancêtre commun afin de coloniser des niches écologiques vacantes
- radiation adaptative
- Formation rapide de plusieurs espèces à partir d'une population ancestrale ayant colonisé plusieurs milieux différents et isolés (ex : archipels)

- Suite à des pressions sélectives différentes, on a une divergence génétique, une différenciation morphologique et un isolement reproducteur.

Exemple: les astéracées des îles de Hawaï, les descendants d'une seule espèce ont évolué en 28 espèces occupant différents habitats.

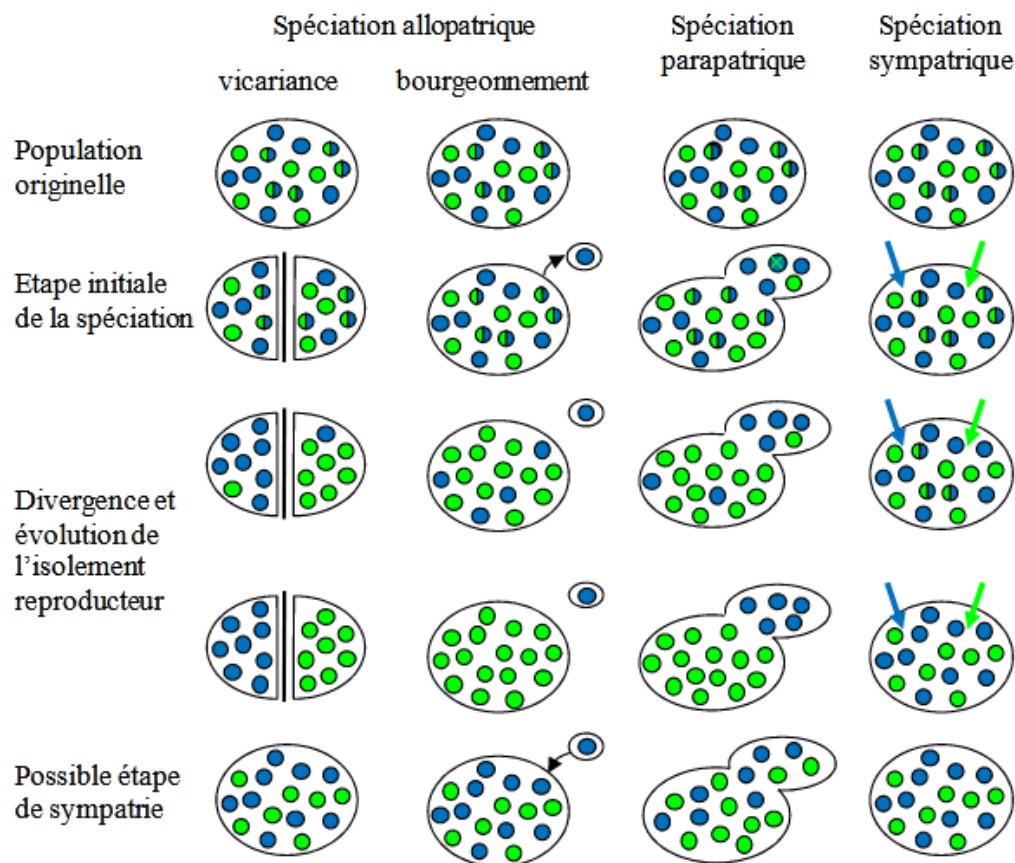


Figure 23 : Différents modes de spéciation (<https://www.futura-sciences.com/planete/dossiers/developpement-durable-mali-oasis-biodiversite-570/page/4/>)

1.6.2. Isolement reproductif : quand l'épigénétique s'en mêle

Les modifications génétiques sont à l'origine de la diversité biologique, sur laquelle agit la sélection naturelle. Mais pour faire émerger de nouvelles espèces, des barrières reproductives doivent être mises en place, comme par exemple l'impossibilité de fécondation ou le caractère délétère de certaines combinaisons de gènes telle que l'incompatibilité hybride. Une étude publiée le 7 mars 2017 dans la revue *PNAS* suggère que les allèles épigénétiques, ayant hérité d'un état de méthylation de l'ADN, contribuent à l'isolement reproductif chez les plantes.

L'accumulation de mutations dans une sous-population peut favoriser l'isolement reproductif et ainsi contribuer à la spéciation. Chez les plantes, cet isolement peut être du à l'échec de la pollinisation entre deux sous-espèces émergentes, à la létalité des hybrides mais aussi à leur stérilité ou la létalité de leur descendance. Ce sont des combinaisons délétères d'allèles provenant des parents qui engendrent cette incompatibilité hybride.

Des chercheurs de l'Institut de biologie moléculaire des plantes à Strasbourg et du Howard Hughes Medical Institute (Indiana University, USA) se sont intéressés à l'incompatibilité hybride entre deux écotypes d'*Arabidopsis*. Le cas étudié implique un gène essentiel (nécessaire à la synthèse de l'histidine) dupliqué dans chacun des écotypes parentaux (Figure 25). Dans l'un des écotypes, une des deux copies, altérée par une mutation, est un pseudo gène. Dans l'autre écotype, les deux gènes paralogues sont conservés mais une seule copie est exprimée. Cette fois, c'est une modification épigénétique, n'affectant pas la séquence d'ADN, mais néanmoins transmise à la descendance qui intervient dans cette perte de fonction. C'est l'hyper méthylation de la région du promoteur de ce gène qui est responsable de cette extinction du gène qu'on désigne alors comme un épiallèle silencieux. L'incompatibilité entre les deux écotypes se manifeste par la létalité au stade embryonnaire d'environ 1/16 de la descendance des hybrides F1.

Par une approche génétique, les chercheurs ont montré que l'élimination de la méthylation de l'ADN dans le promoteur rétablit l'expression de l'épiallèle et lève l'incompatibilité hybride entre les deux écotypes d'*Arabidopsis thaliana*. Comme la variabilité génétique, la variabilité épigénétique naturelle contribue donc à l'incompatibilité hybride et limite les génotypes possibles issus de centaines de croisements entre écotypes. Le caractère réversible de l'épigénétique en fait cependant un acteur singulier. En effet, au sein d'une population, ces incompatibilités épigénétiques peuvent fluctuer plus rapidement que les incompatibilités dues à des mutations génétiques. Le débat reste ouvert sur l'impact de l'épigénétique sur la spéciation.

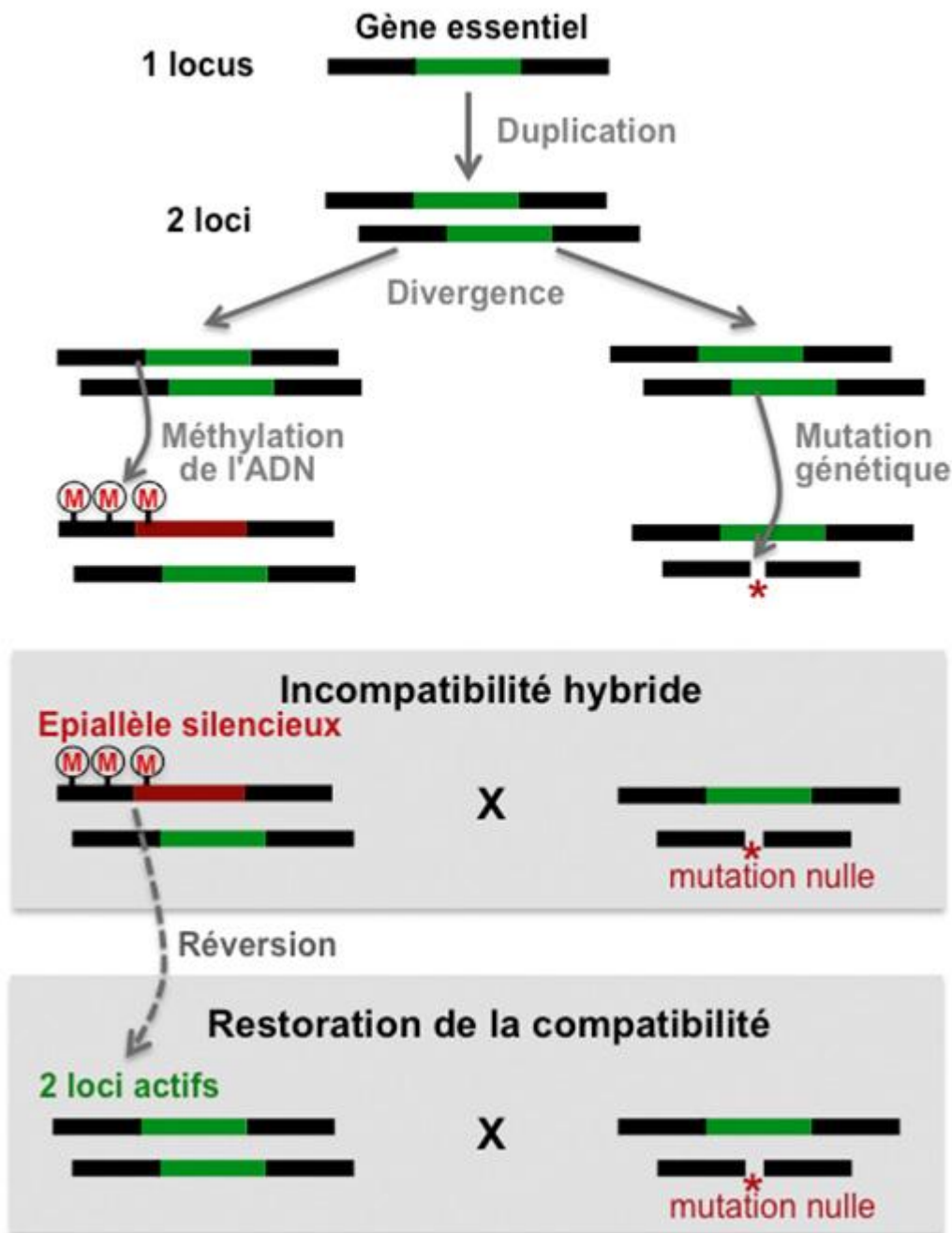


Figure 24 : Duplication d'un gène essentiel génère deux loci paralogues indépendants. Si la méthylation de l'ADN éteint une copie dans un écotyle d'Arabidopsis (locus rouge à gauche), et si une mutation nulle détruit l'autre dans un autre écotyle (astérisque rouge à droite), les croisements entre ces écotypes présenteront une incompatibilité hybride (1/16 de létalité embryonnaire) dans la descendance de la génération F1. Contrairement à l'incompatibilité hybride uniquement due à des mutations, la perte de méthylation de l'ADN (réversion) peut rapidement restaurer l'expression des deux copies dans le premier écotyle et la viabilité de toutes les combinaisons possibles de ces gènes dupliqués (Todd Blevins. CNRS-IBMP).

II.6.3. Cospéciation

Dans ce cas, lorsqu'il y a spéciation de l'hôte, le parasite hébergé par la nouvelle espèce va effectuer une spéciation à son tour. La phylogénie des parasites reflèterait ainsi la phylogénie de l'hôte.

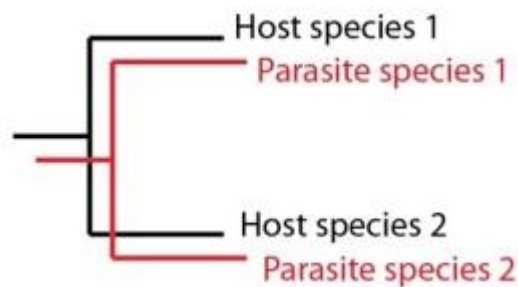


Figure 25 : exemple de cospéciation (<https://www.futura-sciences.com/planete/dossiers/developpement-durable-mali-oasis-biodiversite-570/page/4/>)

Modes de spéciation, d'après Butlin *et al.* (2008)

Etape dans le processus de spéciation :

- 1) Initiation de l'isolement reproducteur intrinsèque
- 2) Augmentation de l'étendue génomique de l'isolement
- 3) Isolement reproducteur complet

Pour chaque étape :

Quel est le contexte spatial ?

Position dans un continuum, pouvant être variable au cours du temps avec m = taux d'échange de gènes

- Sympatrie ($m = 0,5$)
- Parapatricie ($0 < m < 0,5$)
- Allopatricie ($m \approx 0$)

Quelle est la force ou quelles sont les forces conduisant à la divergence ?

Catégories non-exclusives et pouvant interagir entre elles

- Hasard (mutation, hybridation, dérive)
- Sélection naturelle (écologie)
- Sélection sexuelle ou conflit sexuel

Quelle sont les bases génétiques de l'isolement reproducteur ?

Catégories non-exclusives et pouvant interagir entre elles

- Bases génétiques ou non-génétiques (plasticité, apprentissage)
- Incompatibilité intrinsèque *versus* mal-adaptation
- Mécanisme à un-gène, un-allèle et deux-allèles
- Avec ou sans restriction de la recombinaison

À retenir

- L'espèce est l'unité fondamentale en biologie évolutive.
- Il existe plusieurs concepts d'espèce : morphologique, biologique, écologique et phylogénétique.
- La spéciation peut être allopatrique, sympatrique ou parapatric, et dépend des barrières reproductives.

Glossaire

Espèce : Groupe d'individus interféconds et isolés reproductivement.

Barrière reproductive : Mécanisme empêchant l'hybridation entre espèces.

Spéciation : Apparition de nouvelles espèces.

Flux génique : Échange d'allèles entre populations.

Cospéciation : Spéciation parallèle de deux espèces en interaction (ex : hôte-parasite).

QCM

1. Quel concept d'espèce repose sur l'interfécondité et l'isolement reproductif ?

A. Morphologique

☐ B. Biologique

C. Phylogénétique

D. Écologique

2. La spéciation allopatrique nécessite :

A. Un changement chromosomique

B. Une barrière écologique temporaire

☐ C. Une séparation géographique

D. Une niche écologique différente

3. Quel mode de spéciation est le plus fréquent chez les plantes ?

- ☐ A. Sympatrique par polyploïdie
- B. Allopatrique par vicariance
- C. Parapatricque
- D. Homoploïde

4. Une espèce cryptique est :

- A. Morphologiquement distincte mais génétiquement identique
- B. Une espèce disparue
- ☐ C. Identique morphologiquement à une autre espèce
- D. Apparue par mutation spontanée

5. Quelle affirmation est correcte à propos de l'épigénétique ?

- A. Elle modifie la séquence d'ADN
- B. Elle est toujours irréversible
- ☐ C. Elle peut provoquer des isollements reproductifs
- D. Elle n'a aucun effet évolutif

III. 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
- La migration (ou flux de gènes)
- La sélection naturelle et adaptation

III.1. Dérive Génétique

La dérive génétique est le changement de la composition génétique d'une population de petite taille causé par des effets d'échantillonnage aléatoire. Il peut s'agir aussi bien de la perte, de la fixation de certains allèles ou du changement de leur fréquence (Falconer & Mackay 1996, Frankam 2002).

Elle se produit de façon plus marquée lorsque l'effectif de la population est faible. Elle peut conduire à la disparition de certains allèles.

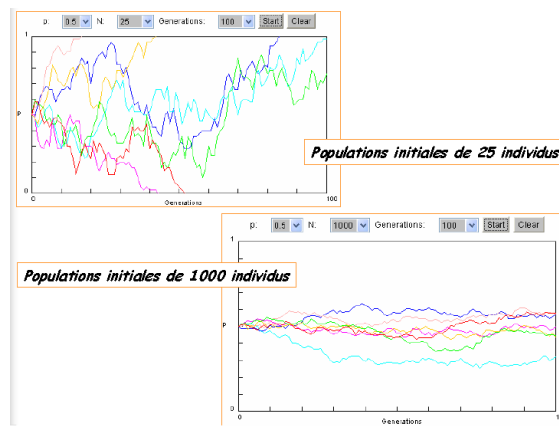


Figure 26 : Modèle de dérive génétique (p est la fréquence initiale de l'allèle, N le nombre d'individus de la population) (<https://planet-vie.ens.fr/thematiques/evolution/mecanismes-theories-et-concepts-de-l-evolution/la-derive-genetique>).

Effet fondateur

La fondation d'une population se fait généralement par l'installation d'un petit nombre d'individus dans une nouvelle localisation. Accompagnés de nombreux changements génétiques tels que la disparition d'allèles rares ou la réduction de la diversité génétique, les effets fondateurs jouent un grand rôle écologique et évolutif.

Par ailleurs, les effets fondateurs sont également supposés modifier l'architecture génétique des caractères phénotypiques de grande importance écologique et évolutive et être impliqués dans les processus de spéciation et d'invasion biologique. La présence de dépression de consanguinité, suite à un goulot d'étranglement, pourrait avoir de sérieuses conséquences sur les chances de maintien d'une population. Cependant, les résultats de certaines études suggèrent l'existence de mécanismes de purge des effets délétères dus à la consanguinité. En l'absence de consensus sur l'existence et la généralité du phénomène de purge, plus d'études sont nécessaires sur les mécanismes génétiques associés à des changements démographiques (taille de la population notamment). A l'heure où l'accroissement des pressions anthropiques entraîne la fragmentation et la dégradation de l'habitat de nombreuses espèces, et réduit leurs effectifs de manière dramatique, la conservation et la gestion des petites populations isolées passent par une meilleure compréhension des mécanismes génétiques qui peuvent affecter leur maintien. D'autre part, Frankham (2005) rappelle le paradoxe des espèces envahissantes. Ces espèces, qui généralement subissent de forts goulots d'étranglement, ont une diversité génétique réduite et devraient donc subir les effets de la consanguinité et voir leur valeur sélective diminuée. Pourtant, on observe chez ces espèces une grande capacité d'invasion. Parmi les causes du succès de colonisation des espèces envahissantes, on peut citer: une forte fécondité, la purge des allèles délétères du génome, un fort taux de migration ou encore la possibilité de reproduction asexuée. Une meilleure compréhension des phénomènes génétiques liés à l'invasion pourrait permettre une meilleure gestion et un meilleur contrôle de ces espèces envahissantes.

Effets fondateurs et fréquence allélique Lors de la fondation d'une population par un faible nombre d'individus, la diversité allélique (N_A) est réduite de manière proportionnelle au nombre d'individus fondateurs. Les allèles rares de la population sont perdus les plus rapidement. Parallèlement à cette perte, augmente la fréquence des allèles caractérisés au départ par des fréquences moyennes. La distorsion de la distribution des fréquences des allèles permet la détection du goulot d'étranglement créé par l'effet fondateur. La diminution de N_A entraîne aussi un changement du niveau d'hétérozygotie de la population. L'hétérozygotie diminue cependant moins

rapidement que NA et peut apparaître en excès par rapport à l'hétérozygotie attendue pour un équilibre mutation dérive.

Le déséquilibre de liaison Le déséquilibre de liaison, l'association non aléatoire de deux gènes au sein du génome, devrait augmenter lors de la fondation de nouvelle population. Le déséquilibre de liaison est affecté par différents paramètres dont certains sont génomiques. Par exemple, le taux de recombinaisons méiotiques s'accroît avec la distance qui sépare les loci et se traduit par une diminution du DL, de même, un taux de mutation élevé aux loci étudiés réduit la probabilité de mettre en évidence le déséquilibre. Sur le plan démographique, plusieurs facteurs affectent le DL par exemple: le nombre d'individus fondateurs de la population: plus le nombre d'individus est important plus le DL est faible; la vitesse d'accroissement de la population: un accroissement rapide diminue le déséquilibre car les événements de recombinaison sont plus nombreux. Le déséquilibre de liaison peut donc également être utilisé pour retracer l'histoire des populations.

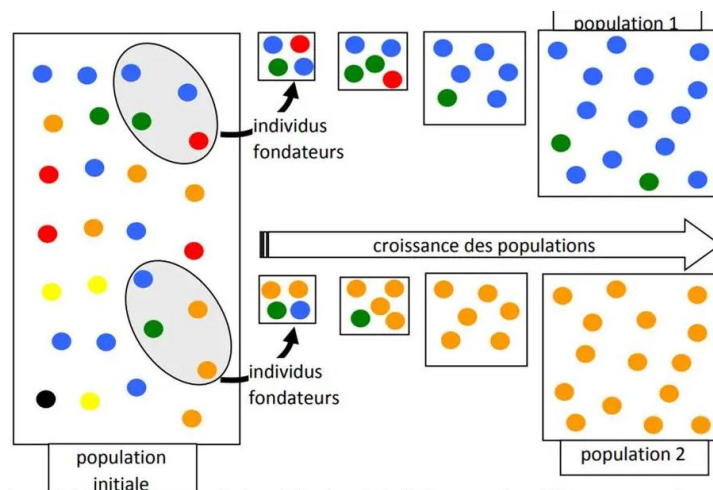


Figure 27 : principe de l'effet fondateur (<https://planet-vie.ens.fr/thematiques/evolution/mecanismes-theories-et-concepts-de-l-evolution/la-derive-genetique>).

Exemple 1:

Les Huttérites sont les membres d'une secte qui, persécutés en Europe, se sont installés en Amérique du Nord au dix-neuvième siècle. Comme les Amish, ils ont établi une série de colonies d'une centaine d'individus dans lesquelles ils vivent en autarcie sans se marier avec les personnes étrangères. Mc Lellan et ses collaborateurs y ont étudié la fréquence de certains allèles des groupes sanguins et du système de compatibilité tissulaire HLA .

Tableau 2 : fréquences des allèles du complexe HLA et des groupes sanguins chez les Huttérites, les Amish et la population européenne et américaine (<https://planet-vie.ens.fr/thematiques/evolution/mecanismes-theories-et-concepts-de-l-evolution/la-derivee-genetique>) .

	HLA					Groupes sanguins		
	A10	A11	A28	B5	B12	A	B	K
Europe, USA	3 à 5	4,5 à 7,4	2 à 4	4 à 8	9 à 18	25 à 32	6 à 14.5	3 à 5
Huttérites	14	0	0	14	8	35	2	14
Amish	7	14	0,7	6	19	66	6,5	0,2

Exemple 2 : la vigne marronne de la Réunion



Figure 28 : profil électrophorétique des marqueurs AFLP des vignes marronne de la Réunion et du Vietnam <https://planet-vie.ens.fr/thematiques/evolution/mecanismes-theories-et-concepts-de-l-evolution>).

Absence de variabilité génétique sur l'île de la Réunion sans doute la conséquence de l'introduction de quelques plans seulement.

Exemple 3 :

Cet effet fondateur a été observé au Saguenay-Lac-Saint-Jean. Parmi les colonisateurs de cette région, certains individus portaient les allèles de la dystrophie myotonique (dont la fréquence était plus élevée que dans la population initiale). C'est pourquoi, aujourd'hui, la fréquence de cette maladie dans cette région est près de 50 fois plus élevée que dans la population québécoise.

Goulot d'étranglement

Le goulot d'étranglement, qui est une réduction temporaire et brutale des effectifs d'une population, peut provoquer une augmentation du taux d'individus homozygotes issus des accouplements entre apparentés et aboutir à une diminution de leur aptitude adaptative (dépression de consanguinité). De plus, cette réduction des effectifs s'accompagne d'une diminution de la diversité génétique et par conséquent d'une limitation du potentiel évolutif de la population.

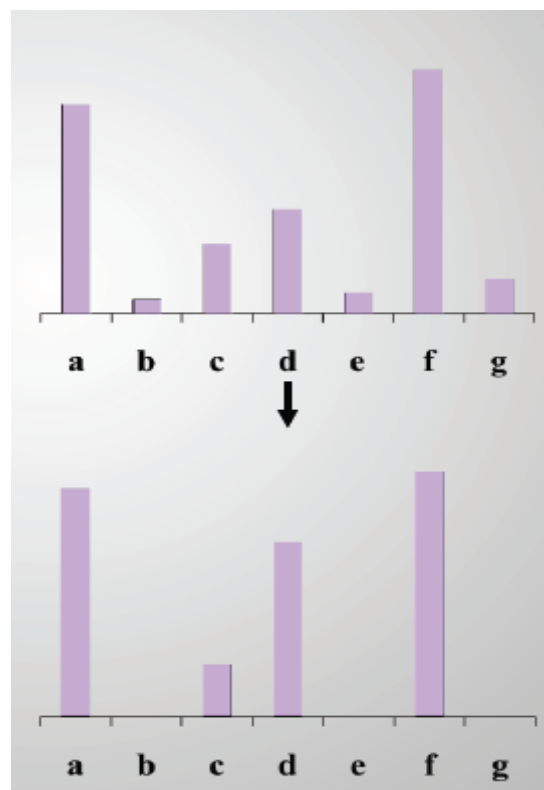


Figure 29 : principe du goulot d'étranglement.

Les allèles rares sont éliminés en premier (<https://planet-vie.ens.fr/thematiques/evolution/mecanismes-theories-et-concepts-de-l-evolution>)

Un goulot d'étranglement, associé par exemple à la fondation d'une nouvelle population, provoque une très forte dérive, à savoir une composition génétique nouvelle due à l'échantillonnage aléatoire d'allèles de la population ancestrale. Frankham (1995), suivant les propositions de Souté (1976), a montré plus généralement de quelle manière la taille de la population influence la diversité génétique. Frankham propose dix prédictions sur la composition génétique d'une population en relation avec sa taille. Il montre notamment que les populations insulaires ont en moyenne une diversité plus faible que les populations continentales. Les fluctuations des effectifs de la population peuvent aussi avoir des effets importants sur l'érosion de la diversité génétique et seront variables selon la durée et le nombre d'individus résiduels dans la population et la dynamique de la population. Plus que l'effectif, la « taille efficace » de la population est importante pour expliquer ces effets sur la composition génétique de la population. La taille efficace représente le nombre d'individus d'une population qui contribuent à la reproduction et au pool génétique de la population.

Exemple de dérive suite à un goulot d'étranglement (Autre exemple d'effet fondateur) : L'éléphant de mer au Mexique et en Californie. *Mirounga angustirostris*, fut très chassé, et passa par un goulot d'étranglement vers 1892 (20 à 100 individus contre 175 000 actuellement). L'étude du polymorphisme de la région contrôle de l'ADN mitochondrial d'animaux fossiles et de peaux et os anciens révèle une perte de diversité génétique

Frequencies of northern elephant seal mtDNA genotypes and diversity statistics (standard errors).

	Seal bones ca. 1000–1800	Seal skins 1892	Seal bones 1914–1980	Modern seals, southern California	Modern seals, central California
	This study	This study	This study	This study	Hoelzel <i>et al.</i> [2]
Genotype 1	1	0	4	78	29
Genotype 2	0	2	2	31	11
Genotype 3	2	0	0	0	0
Genotype 4	1	0	0	0	0
Genotype 5	1	0	0	0	0
Haplotype diversity, <i>h</i>	0.90 (0.16)	0 (0)	0.53 (0.17)	0.411 (0.038)	0.409 (0.065)
Nucleotide diversity, π	0.0065 (0.014)	0 (0)	0.0086 (0.0028)	0.0067 (0.0006)	0.0066 (0.0011)

Genotypes are based on nucleotide sequences from a 179 bp segment (C) of the control region (Figure 1).



Figure 30 : diminution de la diversité génétique chez l'éléphant de mer (<https://planet-vie.ens.fr/thematiques/evolution/mecanismes-theories-et-concepts-de-l-evolution>)

III.2. Mutation

La mutation est considérée en génétique évolutive comme la source ultime de variation. Il est possible de distinguer 3 grandes classes de mutations : les substitutions nucléotidiques, les insertions/ délétions de quelques nucléotides et les remaniements géniques de grande taille. Des mécanismes mutationnels plus rares seront également présentés.

La cause proximale des mutations peut être : une erreur dans la réplication de l'ADN, l'insertion d'un élément transposable, la cassure physique d'un chromosome, la non disjonction au moment de la méiose, ou encore agents mutagènes, chimiques ou radioactifs

Remplacements d'une base :

Substitutions nucléotidiques constituent près de 70 % des mutations : on distingue les transitions (remplacement d'une base pyrimidique (C ou T) ou purique (A ou G) par une autre base de même nature) et les transversions (remplacement d'une base purique par une base pyrimidique, ou inversement). Les substitutions nucléotidiques peuvent être induites par des agents environnementaux mutagènes (substances chimiques, rayonnements...) ou par le métabolisme endogène, mais également être la conséquence d'erreurs spontanées intervenant lors de la réplication de l'ADN, et non détectées par les systèmes de réparation.

Délétions ou insertions de quelques nucléotides :

Les délétions ou insertions d'un ou plusieurs nucléotides (moins de 20) sont, après les substitutions nucléotidiques, les anomalies de séquences nucléotidiques les plus fréquentes, représentant près de 25 % des anomalies répertoriées dans la base de données HGMD. Des délétions ou insertions non multiples de trois bases entraînent au niveau des séquences codantes un décalage du cadre de lecture (frame shift) qui aboutit à l'apparition d'un codon Stop prématuré et à l'éventuelle présence d'une protéine incomplète. Ces anomalies moléculaires surviennent souvent au niveau de courtes répétitions en tandem, très probablement par un mécanisme de glissement (slippage) de l'ADN polymérase en raison de l'appariement décalé de séquences répétées lors de la réplication de l'ADN. Selon la façon dont le mésappariement est résolu, ce dérapage peut être à l'origine d'insertions ou de délétions d'un ou plusieurs motifs répétés.

MUTATIONS PONCTUELLES										
SUBSTITUTIONS										
Hémoglobine normale β^A										
Site	Init	1	2	3	4	5	6	7	8	9
Codon	AUG	GUG	CAC	GUG	ACU	CCU	GAG	GAG	AGG	UCC
Acide aminé	Val	His	Leu	Thr	Pro	Glu	Glu	Lys	Ser	Ala
Mutation silencieuse β^A^+										
Site	Init	1	2	3	4	5	6	7	8	9
Codon	AUG	GUG	CAC	GUG	ACU	CCU	GAG	GAG	AGG	UCC
Acide aminé	Val	His	Leu	Thr	Pro	Glu	Glu	Lys	Ser	Ala
Drépanocytose β^S										
Site	Init	1	2	3	4	5	6	7	8	9
Codon	AUG	GUG	CAC	GUG	ACU	CCU	GUG	GAG	AGG	UCC
Acide aminé	Val	His	Leu	Thr	Pro	Val	Glu	Lys	Ser	Ala
β^0 thalassémie										
Site	...	33	34	35	36	37	38	39	40	41
Codon	...	UUG	UAC	UAC	UUG	ACC	UAG	AGG	UUC	...
Acide aminé	...	Val	Val	Tyr	Pro	Try	Thr	STOP	...	...
DÉLÉTIONS										
β^0 thalassémie										
Site	...	5	6	7	8	9	10	11	12	13
Codon	...	CCU	G-G	ACA	AGU	GUG	GCG	UUA	GUG	GCG
Acide aminé	...	Pro	Gly	Arg	Ser	Leu	Pro	Leu	Pro	Cys
β^0 thalassémie										
Site	...	38	39	40	41	42	43	44	45	...
Codon	...	ACC	CAG	ACG	UU	G	AGU	CCU	UUG	GGG
Acide aminé	...	Thr	Gln	Arg	Leu	Ser	Pro	Try	Gly	...

A l'échelle moléculaire :

- . Mutation silencieuse
- . Mutation faux-sens
- . Mutation non-sens

Document 6 :
Les mutations, exemple du gène β de l'hémoglobine

Figure 31 : exemple de substitutions dans le gène β de l'hémoglobine (Verneuil Hubert De., et al 2003).

Mutation chromosomiques

Les mutations chromosomiques impliquent le remaniement de nombre ou de la structure d'un ou de plusieurs chromosomes. Il peut s'agir d'une délétion, d'un échange de matériel génétique entre deux chromosomes (translocation), d'une duplication ou d'une inversion d'une région, ou encore, d'une insertion. Dans ce dernier cas, plusieurs gènes peuvent être atteints simultanément

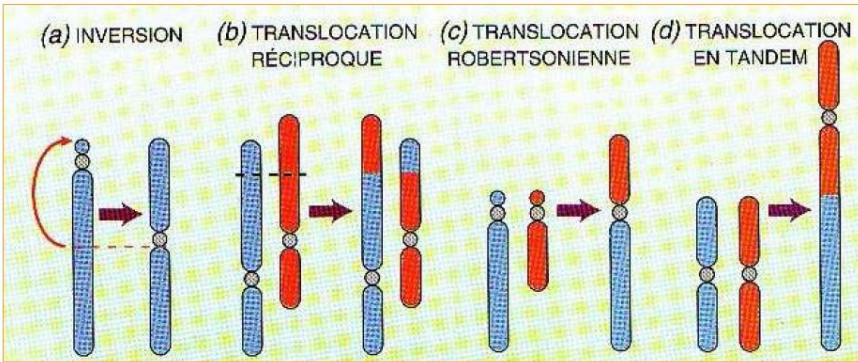


Figure 32 : Remaniements chromosomique (Verneuil Hubert De., et al 2003).

Mécanismes rares

a-insertions d'éléments transposables

Insertion d'éléments transposables Les éléments transposables représentent environ la moitié du génome humain et sont divisés en deux types : les transposons (environ 3 % du génome humain), qui se mobilisent sous forme d'ADN selon un mécanisme de type « couper/ coller » gouverné par une transposase, et les rétrotransposons (plus de 40 % du génome humain), qui utilisent un intermédiaire ARN lors de leur mobilisation selon un mécanisme de type copier/coller gouverné en particulier par une transcriptase inverse. Le déplacement des éléments transposables d'un point à un autre du génome peut engendrer des mutations soit directement par leur insertion à côté ou à l'intérieur d'un gène, ce qui peut en altérer la séquence codante et/ou l'expression, soit indirectement par la recombinaison homologue entre deux copies d'éléments transposables de séquences similaires et situés à des positions non orthologues, ce qui peut amener à des délétions ou des duplications (voir ci-dessus). Chez l'homme, les insertions responsables de maladies génétiques concernent exclusivement les rétrotransposons sans LTR (long terminal repeat) autonomes (LINE) et non autonomes (SINE) [10]. Depuis les premières descriptions par les groupes de H.H. Kazazian en 1988 [14] et de F. Collins en 1990 [15], une cinquantaine d'insertions de novo de séquences LINE-1 (seul élément actif au sein des LINE) et de séquences Alu (SINE) ont été rapportées. Parmi les 500 000 séquences LINE-1 et le million de séquences Alu du génome humain, seule une faible proportion est encore active (moins de 100 séquences LINE-1 et quelques milliers de séquences Alu), et la fréquence des événements de transposition a été estimée chez l'homme à 1 événement pour 50 à 200 cellules germinales [16]. Dans tous les cas, les insertions d'éléments transposables sont des événements rares, dont l'identification a néanmoins permis d'enrichir les connaissances des rétrotransposons chez l'homme. Enfin, il est à noter qu'aucun cas d'insertion de novo de rétrotransposons à LTR ou HERV (human endogenous retrovirus) n'a été rapporté à ce jour.

b-Conversions géniques

Conversion génique La conversion génique est un mécanisme de recombinaison qui, contrairement au crossing-over, correspond à un transfert unidirectionnel d'information génétique [17]. Elle conduit au remplacement d'une séquence d'ADN par une autre, apparentée non allélique (conversion interlocus) ou allélique (conversion interallélique). Ce remplacement peut entraîner dans le gène receveur une série de changements nucléotidiques répartis sur une région assez courte (Figure 3). Dans certains cas, la séquence donneuse est un pseudogène, inactivé par l'accumulation

de mutations, et le transfert d'une de ces mutations inactive le gène receveur : c'est le cas de certaines mutations du gène de la stéroïde 21 hydroxylase dans l'hyperplasie surrénalienne, ou du gène de la β glucocérébroside dans la maladie de Gaucher.

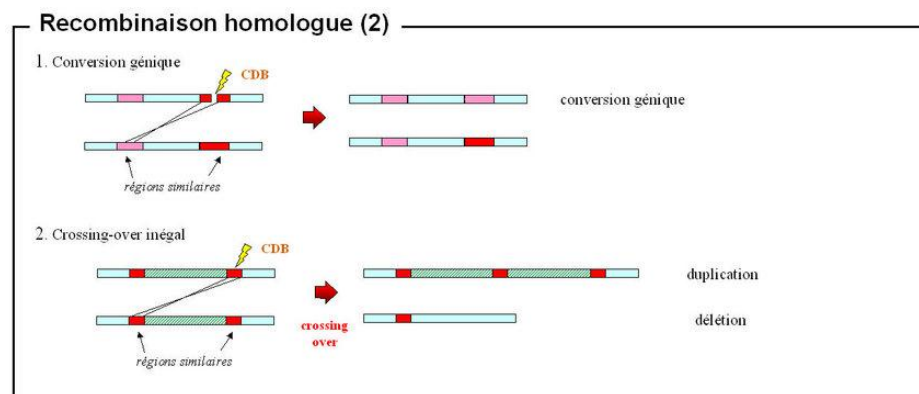


Figure 33 : principe de la conversion génique (Verneuil Hubert De., et al 2003).

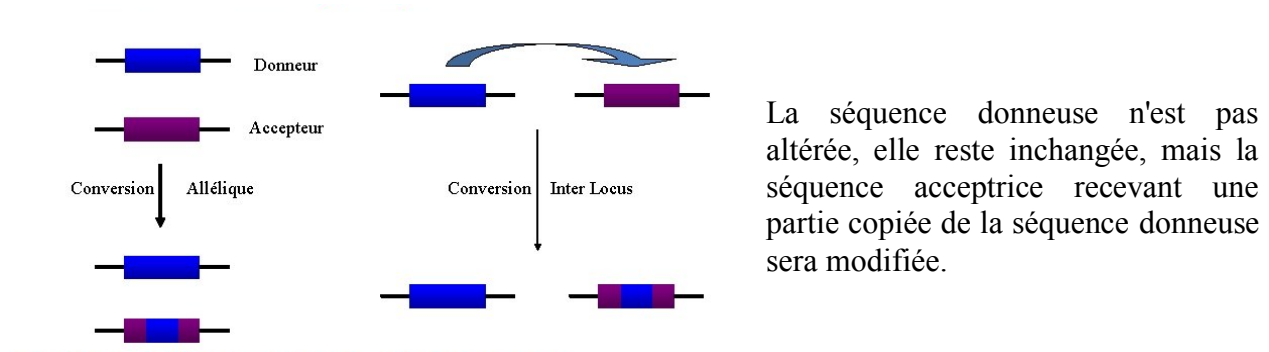


Figure 34 : conséquences de la conversion génique (Verneuil Hubert De., et al 2003).

- **Mécanismes épigénétiques**

Les mécanismes épigénétiques des mutations dans certaines circonstances, l'expression phénotypique de la maladie dépend du sexe du parent qui transmet l'anomalie moléculaire. Cette différence traduit le phénomène d'empreinte génomique parentale, une modification épigénétique, temporaire et réversible du génome nucléaire. Cette empreinte apposée durant la gamétogenèse conduit à l'expression différentielle du matériel génétique en fonction de son origine parentale. Chez l'homme, l'empreinte génomique a été impliquée dans un nombre croissant d'anomalies du développement, de maladies héréditaires et de cancers, une soixantaine de gènes soumis à

empreinte étant identifiés à ce jour. Les pathologies liées à l'empreinte sont dues à une modification, pour un locus ou un gène donné, de l'équilibre entre allèle paternel et allèle maternel. Ce déséquilibre peut résulter d'une disomie uniparentale, c'est-à-dire de la présence dans une cellule diploïde de 2 séquences homologues héritées d'un seul parent, voire de 2 chromosomes entiers. Chez l'homme, la première observation de ce type date de 1988 : chez un enfant à caryotype normal, la présence en double copie d'un chromosome 7 d'origine maternelle porteur d'une mutation p. Phe508del du gène CFTR était associée à une mucoviscidose et une petite taille, suggérant l'existence au niveau du chromosome 7 d'un gène impliqué dans la croissance et soumis à empreinte, le gène maternel n'étant pas exprimé. Cet exemple illustre les deux types d'effets délétères de la disomie uniparentale : transmission d'une maladie récessive par homozygotie (mucoviscidose) et absence d'expression d'un gène par empreinte parentale (petite taille). L'un des meilleurs exemples est aujourd'hui fourni par les syndromes de Prader-Willi et d'Angelman, liés à une disomie maternelle et paternelle du chromosome 15, respectivement. Un déséquilibre de l'empreinte peut également être observé malgré une transmission biparentale, de petites délétions emportant, dans certains cas, les « centres de l'empreinte », éléments intervenant en cis dans la régulation de l'empreinte. Enfin, il a également été rapporté chez certaines femmes des cas de môles hydatiformes récurrents d'origine biparentale, conséquence de défauts globaux d'empreinte de la lignée germinale femelle, les allèles maternels des gènes présentant alors un profil épigénétique paternel. Cette impossibilité d'imprimer l'empreinte maternelle, phénomène baptisé immaculate misconception par Surani [19], doit donc être le fait de mutations dans un gène essentiel à l'empreinte maternelle, encore inconnu, mais localisé par analyse de liaison au sein d'une région de 1,1Mb sur le chromosome 19q13.4 [20]. Il est aujourd'hui clairement établi que la méthylation de l'ADN ainsi que les modifications post-traductionnelles des histones participent à l'établissement de l'expression mono-allélique des gènes soumis à empreinte. Cependant, la nature précise de la marque primaire de l'empreinte et son devenir pendant le développement restent en partie mystérieux ; leur identification devrait permettre d'envisager une meilleure compréhension des maladies liées à des modifications épigénétiques.

Mutations génomiques

Polyploïdie, un mécanisme important sur le plan évolutif

La polyploïdie est un phénomène très répandu chez les végétaux qui a contribué à l'évolution des plantes à fleurs. Ainsi, le génome de l'ancêtre commun aux plantes à fleurs résulterait de trois fusions de génomes. Par la suite, de nombreuses autres fusions de génomes serait à l'origine de la diversification des plantes à fleurs. La polyploïdie est peu répandue dans le monde animal, ce qui pourrait s'expliquer par la production d'individus stériles. En effet, chez de nombreuses espèces, le sexe est déterminé par la garniture chromosomique de la cellule oeuf.

Plusieurs mécanismes peuvent être à l'origine du phénomène de polyploïdie. Le plus courant est une anomalie survenant lors d'une mitose. Les chromosomes dupliqués ne sont pas répartis dans les deux cellules filles car la cytokinèse, c'est-à-dire la division du cytoplasme, ne se réalise pas.

À l'issue de la mitose, le noyau renferme donc deux fois plus de chromosomes que le noyau de la cellule mère. Ce phénomène est nommé endoréplication et peut se produire spontanément ou être induit par un stress thermique ou l'emploi de colchicine. La colchicine est un alcaloïde extrait de la colchique. Cette molécule empêche la dépolymérisation des microtubules au moment de la métaphase.

La polyploïdie peut également avoir pour origine l'hybridation de deux génomes appartenant à des espèces proches d'un point de vue génétique mais néanmoins différentes. Les chromosomes sont en partie homologues, ce qui pose des problèmes au moment de la méiose et induit une stérilité des hybrides. La fertilité peut être restaurée par un mécanisme d'endoréplication. Dans certains cas, la polyploïdie a pour origine la fusion entre un gamète diploïde et un gamète haploïde conduisant à la formation d'individus triploïde stériles. Le fait que la polyploïdie soit répandue chez les végétaux est certainement lié à leur capacité à se reproduire par multiplication végétative, ce qui évite les problèmes posés par l'appariement de chromosomes d'origine différente lors de la méiose.

Les polyploidies sont parfois créés et exploités par l'Homme à des fins commerciales et/ou agronomiques. Les truites triploïdes et les huîtres triploïdes sont stériles et peuvent donc être commercialisées toute l'année. De nombreuses espèces cultivées sont soit autopolyploïdes portant plusieurs exemplaires d'un même génome (ex. : luzerne, pomme de terre, trèfle, orangers...), soit allopolyploïdes présentant des génomes de plusieurs espèces apparentées au niveau diploïde (ex. : blé dur, blé tendre coton, colza, tabac, fraisier, prunier domestique...).

Les avantages sont divers : production de fruits sans pépins, obtention de fruits plus gros, de plantes plus grandes... Certains scientifiques ont proposé, pour expliquer l'origine de la diversification importante des vertébrés constatée il y a 500 Ma, qu'il y aurait eu deux duplications complètes du génome. Ce réservoir de nouveaux gènes aurait permis l'acquisition de nouvelles fonctions et l'accroissement de la complexité et de la diversification.

Il existe deux grandes catégories de polyploïdes :

- **Les autopolyploïdes : se forment par duplication du génome au sein de la même espèce.**

Dactylis glomerata = $2n = 4x = 28$

- **Les allopolyploïdes : se forment suite la duplication d'un génôme chez un hybride.**
L'hybride plus ou moins stérile peut donner des gamètes non réduits

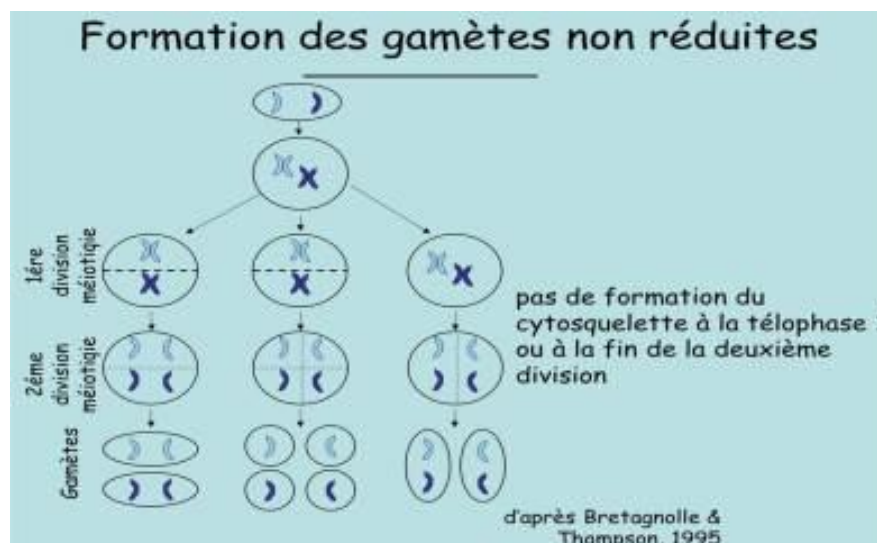


Figure 35 : illustration de la non disjonction méiotique (Verneuil Hubert De., et al 2003).

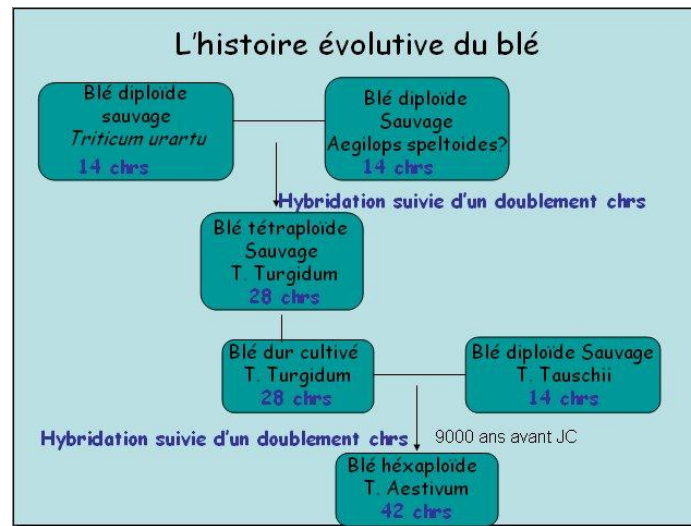


Figure 36 : Histoire évolutive du Blé (<https://accres.ens-lyon.fr/accres/thematiques/evolution/accompagnement-pedagogique>).

Les différentes espèces de blé ont été générées par des événements successifs de polyploïdisation intervenant après des croisements interspécifiques entre trois espèces ancestrales diploïdes

III.3. Sélection naturelle

La notion d'adaptation est un sujet important de la biologie évolutive. L'adaptation peut être décomposée en deux mécanismes évolutifs distincts : la plasticité phénotypique et l'adaptation génétique. La plasticité phénotypique correspond à la capacité d'un organisme à modifier sa physiologie, sa morphologie ou son développement en réponse à des changements environnementaux. Il s'agit essentiellement d'un mécanisme au niveau individuel qui s'exerce sur une durée relativement courte (vie de l'organisme). En revanche, l'adaptation génétique, est le résultat de l'action de la sélection naturelle qui augmente la fréquence des phénotypes les plus favorables à la survie et à la reproduction. On dit alors qu'ils qui ont une meilleure fitness, aussi appelée valeur sélective ou valeur adaptative. Par l'élimination des individus les moins adaptés à leur environnement, la fitness globale au niveau populationnel aura tendance à augmenter. Pour que l'adaptation puisse s'opérer, il faut que les caractères liés à la fitness soient au moins partiellement héréditaires, c'est-à-dire que la valeur de ces caractères soit contrôlée par des gènes transmissibles de génération en génération. D'autres caractères héréditaires peu liés à la valeur sélective peuvent évoluer de façon aléatoire par la dérive génétique. Il paraît donc essentiel d'étudier la part de l'évolution qui est due à la dérive génétique et celle due à la sélection naturelle, autrement dit à l'adaptation, voir même d'essayer d'identifier l'action de cette sélection et les régions qui y sont soumises.

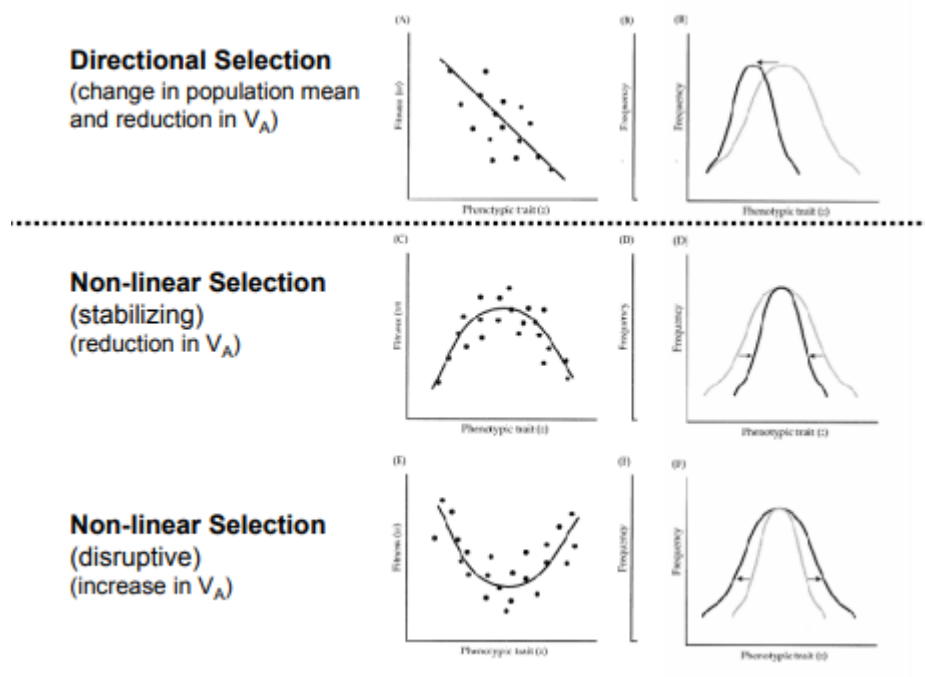


Figure 37 : Illustration des différentes formes de sélection au niveau phénotypique: relations entre fonctions de fitness et variation du caractère phénotypique (A, C et E) et conséquences sur la moyenne et la variance d'un caractère quantitatif (indiquées par des flèches) (B, D et F) pour les sélections directionnelle, stabilisatrice et divergente (Conner et Hartl 2004).

À retenir

- Les mutations génèrent la variation génétique.
- La sélection naturelle trie les individus selon leur adaptation.
- La dérive génétique agit de façon aléatoire, surtout dans les petites populations.
- Le flux génique tend à homogénéiser les populations.

Glossaire

Mutation : Changement dans l'ADN pouvant introduire de la nouveauté génétique.

Dérive génétique : Modification aléatoire des fréquences alléliques.

Sélection naturelle : Processus adaptatif favorisant certains génotypes.

Flux génique : Mouvements de gènes entre populations.

Effet fondateur : Nouvelle population créée par un petit nombre d'individus.

QCM

1. Quelle force évolutive est à l'origine première de la variation génétique ?

A. Sélection naturelle

B. Dérive génétique

☐ C. Mutation

D. Flux génique

2. La sélection naturelle agit sur :

A. Les mutations apparues pendant la vie de l'individu

☐ B. La variation héréditaire existante

C. Les fréquences géniques aléatoires

D. Les migrations interspécifiques

3. La dérive génétique est plus marquée :

A. Dans les grandes populations

B. Lors d'un isolement géographique

C. Lorsque la mutation est fréquente

☐ D. Dans les petites populations

4. Le flux génique a pour effet principal de :

A. Créer de nouvelles mutations

B. Renforcer l'isolement reproductif

☐ C. Réduire les différences génétiques entre populations

D. Fixer les allèles délétères

5. L'effet fondateur peut entraîner :

A. Une augmentation de la diversité génétique

☐ B. Une uniformité génétique au sein d'une nouvelle population

C. Une spéciation immédiate

D. Une sélection dirigée vers les caractères les plus favorables

IV. Diversité génétique

La diversité génétique correspond à la diversité des gènes au sein d'une même espèce. Tous les êtres vivants sont composés d'unités de base, les cellules, qui contiennent une carte d'identité génétique de l'individu, appelée le génome et qui est constitué de l'ensemble des gènes. Ce patrimoine génétique est contenu dans l'ADN (abréviation d'Acide DésoxyriboNucléique), molécule qui contient toutes les informations nécessaires au bon fonctionnement d'un organisme.

IV. 1. Notion de polymorphisme

Le polymorphisme génétique est l'existence, dans une population, de plusieurs états alternatifs de l'ADN, ou allèles, en une position définie du génome, ou *locus*.

Le polymorphisme génétique est à la base de la diversité génétique, il correspond à des variations de séquences d'ADN au sein d'un groupe d'individus. Ces variations naturelles sont dues à des mutations successives au cours de l'évolution qui permettent de caractériser la diversité génétique entre individus et populations. En général, la majorité des polymorphismes sont neutres, mais une partie de ces variations peut influencer les différences phénotypiques observées entre individus qui pourront leur permettre d'avoir une meilleure fitness dans leur environnement. Le polymorphisme peut nous renseigner sur les différents processus qui façonnent la variabilité génétique. Ils peuvent être de nature démographique (taille des populations, migrations) et sélectif (lié à l'environnement). L'étude de l'importance relative de ces différentes composantes façonnant le polymorphisme permet de mieux comprendre l'histoire évolutive des populations et des espèces.

IV.2. Déterminisme épigénétique

Mécanismes épigénétiques

L'ADN des cellules eucaryotes présente différents niveaux de compaction et l'un des premiers niveaux est la chromatine. Elle résulte de l'association entre la molécule d'ADN et des protéines. La chromatine permet la formation de chromosomes à l'intérieur du noyau des cellules eucaryotes. L'état de compaction de cette chromatine est dynamique et peut varier pour une même région génomique pendant le cycle cellulaire. Au-delà de son importance dans la compaction de l'information génétique, c'est aussi à ce niveau que sont régulés les processus cellulaires fondamentaux tels que la réplication, la réparation de l'ADN. Elle constitue le premier niveau de régulation de l'expression des gènes. La chromatine peut être dans un état dit « ouvert » qui va

favoriser l'expression génique ou au contraire « fermé » qui va réprimer l'expression génique. L'état « ouvert » et décondensé de la chromatine est appelé euchromatine, et la structure condensée et « fermée » de la chromatine est appelée hétérochromatine. Ces états de condensation sont contrôlés par des modifications biochimiques : la méthylation de l'ADN et/ou les HPTM (Figure 13) (Jenuwein and Allis, 2001). Ces mécanismes vont déterminer l'accessibilité de l'ADN à la machinerie de transcription et aux facteurs de transcription qui vont permettre la transcription l'ADN en ARNm, qui seront ensuite traduits en protéines (Felsenfeld, 2014).

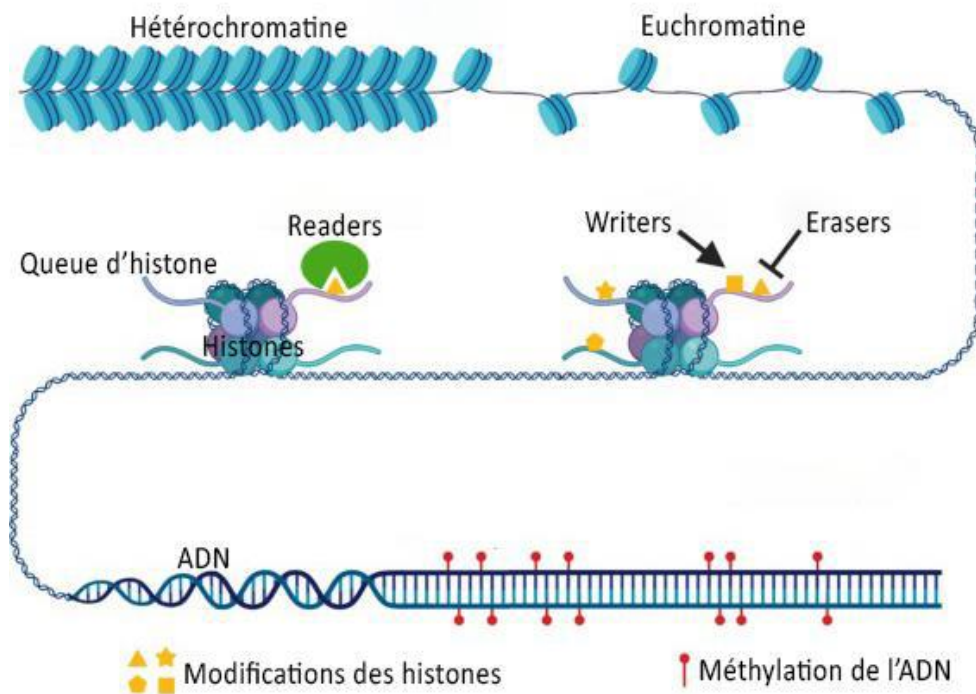


Figure 38: Mécanismes épigénétiques influant sur l'état de la chromatine (Cao and Yan, 2020).

L'ADN peut être méthylé directement au niveau des cytosines ou de l'adénine. L'ADN génomique s'enroule autour d'octamères d'histones, formant ainsi un nucléosome. Les histones coeurs peuvent être modifiées au niveau de leurs résidus par des liaisons covalentes tels que la méthylation, l'acétylation, la phosphorylation, etc. Les HPTM peuvent être ajoutées ou retirées par l'action d'enzymes spécifiques : les *writers* et les *erasers* respectivement. Les *readers* reconnaissent les HPTM en se fixant sur celles-ci.

IV.3. Déterminisme génétique

L'ensemble des gènes connus qui sont impliqués dans l'élaboration et l'expression du caractère. L'équivalence un gène = un caractère étant la plupart du temps fausse. Le déterminisme génétique des caractères peut être plus ou moins complexe en fonction des caractères :

- Caractères simples, qualitatifs : gènes majeurs, effets forts (expliquent un fort pourcentage de la variabilité)
- Caractères complexes, quantitatifs : gènes mineurs, grand nombre de gènes, expliquent une partie de la variabilité
- Caractères intermédiaires, à déterminisme oligogénique



Analyse du déterminisme génétique de plusieurs sources de résistance à l'oïdium (*Uncinula necator*)

Marquage d'une résistance monogénique à l'oïdium provenant de *Muscadinia rotundifolia*

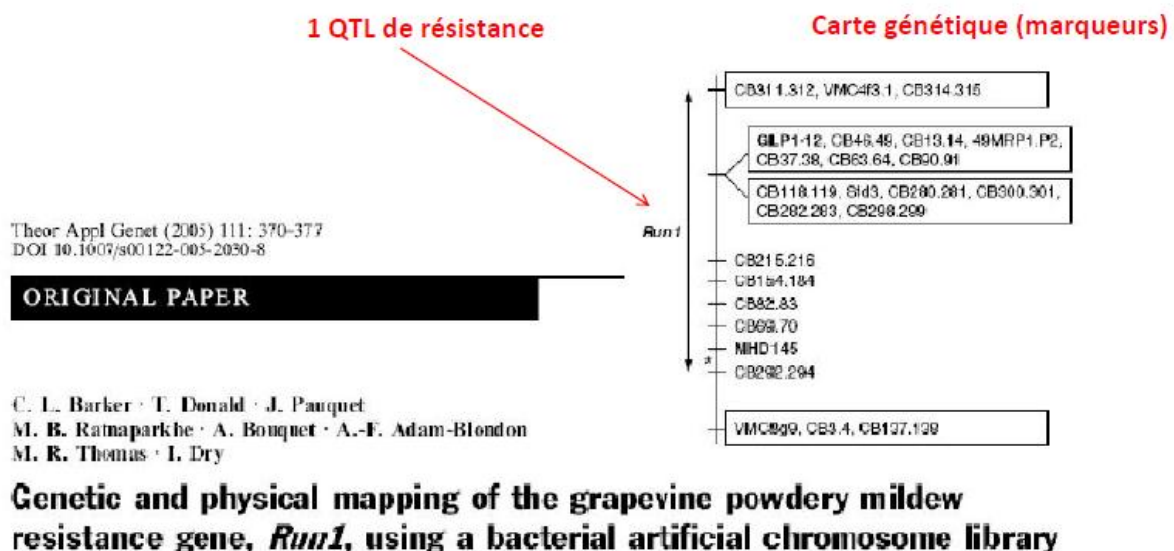


Figure 39 : Exemple d'étude du déterminisme génétique de la résistance à l'oïdium (Agnès D et al.,

2013).

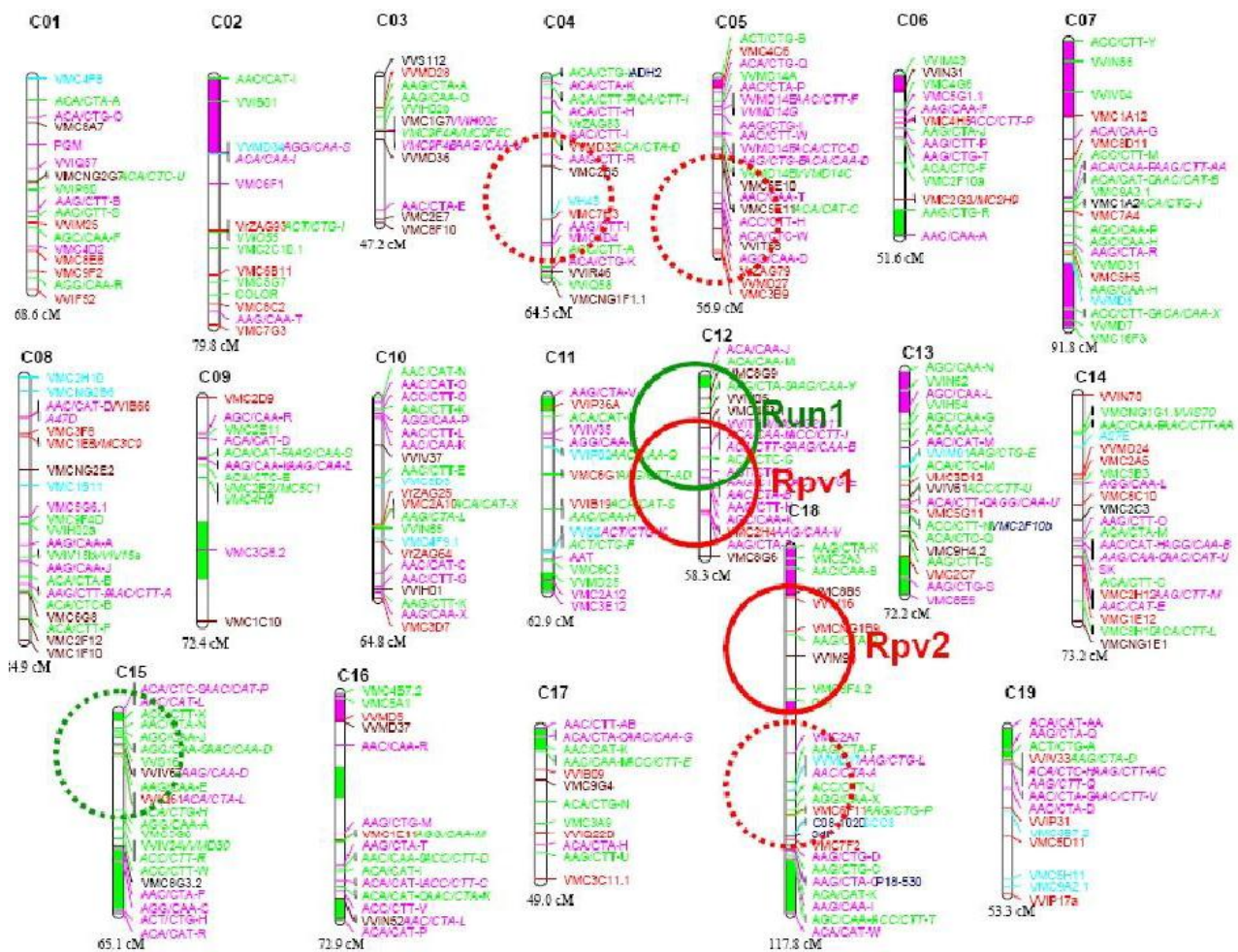


Figure 40 : Résultat de l'étude du déterminisme génétique de la résistance à l'oidium (Agnès D et al., 2013).

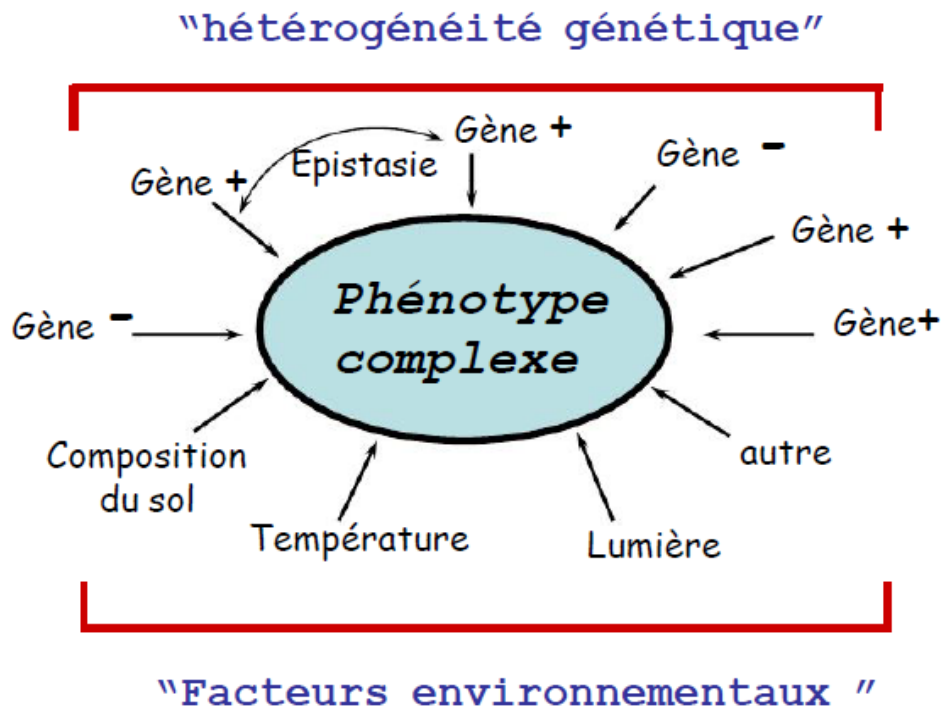


Figure 41 : Illustration du génotype au phénotype (<http://gec.sdv.univ-paris-diderot.fr/cours/genetique/chap/08.html>).

À retenir

- Le polymorphisme reflète la variabilité génétique au sein des populations.
- Le déterminisme génétique repose sur la variation allélique.
- L'épigénétique influence l'expression des gènes sans modifier leur séquence.
- Le phénotype résulte d'interactions entre génotype et environnement.

Glossaire

Polymorphisme : Présence de plusieurs allèles pour un gène.

Épigénétique : Modifications héréditaires de l'expression génique sans altération de la séquence.

Phénotype : Expression observable d'un génotype.

Plasticité phénotypique : Capacité à modifier son phénotype selon l'environnement.

Génotype : Ensemble des gènes d'un individu.

QCM

1. Qu'est-ce que le polymorphisme génétique
 - A. Une mutation spontanée
 - B. Une reproduction asexuée
 - ☐ C. La coexistence de plusieurs allèles pour un même gène
 - D. La capacité d'un individu à se reproduire

2. Le déterminisme épigénétique implique :
 - A. Un changement de la séquence d'ADN
 - B. Une transmission non héréditaire
 - ☐ C. Des modifications réversibles de l'expression des gènes
 - D. Une absence d'effet sur l'évolution

3. Parmi les mécanismes épigénétiques, on retrouve :
 - A. La duplication génique
 - ☐ B. La méthylation de l'ADN
 - C. La recombinaison homologue
 - D. L'épissage alternatif

4. Le passage du génotype au phénotype est influencé par :
 - A. Uniquement la séquence d'ADN
 - B. Le climat uniquement
 - ☐ C. Les interactions entre gènes et environnement
 - D. L'allèle dominant uniquement

5. Une espèce possédant une forte diversité génétique est :
 - A. Plus sensible aux épidémies
 - B. Moins adaptable à l'environnement
 - ☐ C. Plus apte à s'adapter aux changements
 - D. Génétiquement instable

V. Evolution des séquences nucléotidiques

De nombreuses métaphores se sont développées pour illustrer la motivation première de la génomique : « décrypter le code du vivant » ou encore « déchiffrer le livre de la vie ». Le génome désigne la séquence complète d'ADN disposée en chromosomes et contenue dans les noyaux de chaque cellule d'un organisme. Chez les Vertébrés, la taille de cette séquence peut aller de 350 millions à 130 milliards de bases (KAPUSTA, SUH et FESCHOTTE 2017). Les images intuitives du génome comme un livre ou un code mettent en avant la notion d'information, ici génétique, qu'il contient. Il s'agit de l'information nécessaire à la formation, survie et reproduction d'un individu. Cependant, le génome est une entité dynamique, soumis aux forces évolutives de mutations, dérive et sélection. Immédiatement, plusieurs questions théoriques peuvent se poser concernant le fonctionnement des génomes. A quel point sont-ils robustes, ou, pour user un peu plus des métaphores sus-citées, combien de modifications peuvent casser le programme de la vie ? Qu'est-ce qui différencie et/ou unit les génomes de différentes espèces ? Par exemple, combien de différences existe-t-il entre un génome humain, de souris, de poisson clown ? Ces premières questions entraînent une multitude d'autres, représentant des enjeux majeurs pour les domaines de la médecine, de l'agronomie et des biotechnologies. Cette sous-partie vise à introduire l'émergence et les principes généraux de la génomique comparative moderne.

V.1. Annotation des gènes : premier pas vers la fonction des génomes

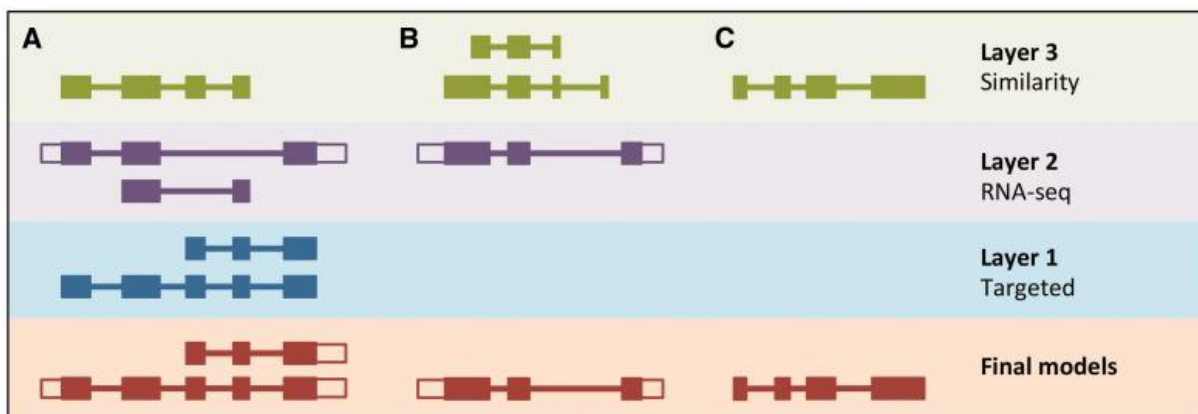


Figure 42: Annotation des gènes codants (AKEN et al. 2016).

Dans le processus d'annotation d'Ensembl, différentes couches hiérarchiques permettent l'annotation des gènes (1 *en bleu* : données protéiques issues de l'espèce annotée, 2 *en violet* : données RNA-seq issues de l'espèce annotée, 3 *en vert* : données protéiques issues d'espèces proches). La prédiction finale des gènes (*en rouge*) s'effectue en utilisant en priorité les modèles des couches 1, puis la 2, puis la 3.

V.2. Familles de gènes

L'étude des séquences génétiques montre que beaucoup de gènes sont présents à l'intérieur du génome non pas en une seule copie, mais en copies multiples, ce qui permet la synthèse d'une protéine par plusieurs gènes à la fois. Par exemple dans le génome humain, 15% de tous les gènes codant pour des protéines sont présents en plusieurs copies. Le processus évolutif qui permet de générer de nombreuses copies d'un même gène est la duplication.

Ce mécanisme peut avoir diverses origines. En particulier, il peut être dû à des erreurs de recombinaisons lors de la méiose (recombinaisons inégales), ce qui donne lieu à des duplications en tandem, i.e. des copies multiples adjacentes sur le chromosome. Un autre mécanisme donnant lieu à des duplications non adjacentes est la rétrotransposition, où l'ARN messager est inversement transcrit en ADN complémentaire puis inséré dans le génome. La présence de copies dupliquées peut également être due à des mécanismes de plus grande envergure entraînant la duplication d'un chromosome, ou même du génome entier. La duplication de gènes joue un rôle primordial dans l'évolution des espèces. En effet, c'est une source importante d'innovation génétique, et de création de nouvelles fonctions. Immédiatement après la duplication, les deux copies du gène ont exactement la même séquence. Cependant, au cours de l'évolution, elles accumulent des mutations, pouvant entraîner la perte de fonction de l'une des deux copies. Un tel gène est appelé "pseudogène". Lorsque trop de mutations se sont accumulées, le pseudogène n'est plus reconnaissable. On parle alors de la "perte" du gène. On appelle famille de gènes un ensemble de gènes dans un ou plusieurs génomes ayant évolué à partir d'un ancêtre commun. Les familles de gènes sont généralement identifiées par homologie de séquence. Par exemple, en utilisant la méthode de recherche BLAST, tous les gènes ayant un score de similarité supérieur à un certain seuil sont considérés "homologues" et regroupés dans une même famille. Il est important de distinguer entre deux types de gènes homologues : les orthologues et les paralogues. Deux gènes homologues dans deux génomes différents sont dits orthologues s'ils sont issus de leur dernier ancêtre commun par spéciation. D'autre part, deux gènes qui sont homologues dans le même génome ou dans deux génomes différents sont dits paralogues s'ils sont issus de leur dernier ancêtre commun par duplication. Alors que les gènes orthologues conservent généralement la même fonction, les gènes paralogues peuvent en développer des nouvelles. En effet la présence de deux copies dans un même génome permet à l'une des deux copies d'évoluer et d'acquérir éventuellement une nouvelle fonction, sans affecter la fonction initiale, assurée par l'autre copie. La figure 2.3 illustre l'évolution d'une famille de gènes dans trois génomes A, B et C, ayant évolué selon la phylogénie ((A,B),C) (Figure 44 (a)). Le gène g du génome ancestral G de A, B et C a subi une duplication précédant la première spéciation de la phylogénie. Deux descendants de chacune

des deux copies créées x et y sont présents dans chacune des espèces actuelles A , B et C . Les six gènes actuels 1,2,3,4,5,6 sont tous homologues, et forment donc une seule famille de gènes. Dans cette famille, tous les gènes de la sous-famille $F1 = \{1,2,3\}$ descendant de x sont paralogues aux gènes de la sous-famille $F2 = \{4,5,6\}$ descendant de y . Les gènes de chacune des sous-familles sont orthologues entre eux. En se référant à la figure 2.3, si on étudiait seulement les gènes 1,3 et 5, la phylogénie déduite serait $((A,C),B)$ au lieu de $((A,B),C)$ et ce malgré le fait que les espèces A et B sont plus proches l'une de l'autre que A et C . La duplication des gènes fait en sorte que des copies distinctes du même gène peuvent nous induire en erreur quant à la vraie phylogénie des espèces. Ainsi, pour représenter une histoire évolutive correcte il serait important d'avoir l'ensemble complet des gènes d'une famille. Une alternative plus réaliste est de sélectionner uniquement un ensemble de gènes orthologues. D'un point de vue pratique, il n'existe pas de méthode simple et universelle (homologie de séquence ou autre) permettant de distinguer entre gènes orthologues et gènes paralogues.

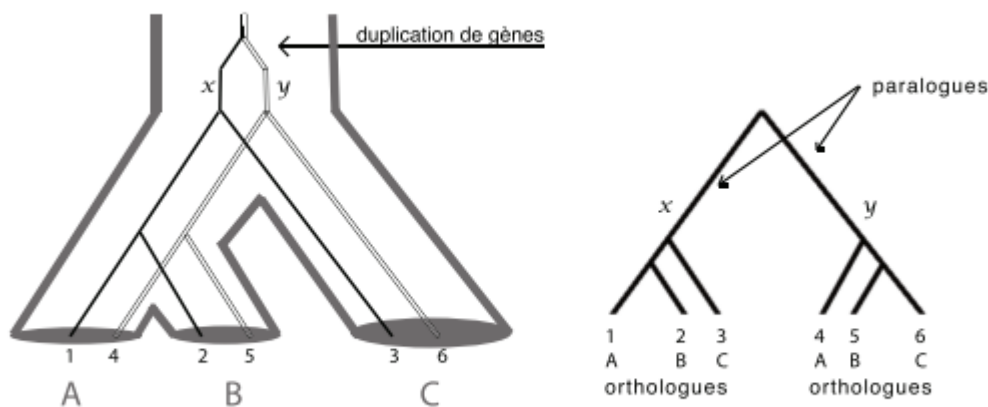


Figure 43 : Exemple de duplication de gènes qui donne lieu à deux lignées de gènes paralogues : le groupe des gènes $F1 = \{1,2,3\}$ descendants de x et le groupe de gènes $F2 = \{4,5,6\}$ descendants de y . Les gènes de chacune des familles $F1$ et $F2$ sont orthologues entre eux. Afin de retrouver la bonne phylogénie pour les espèces étudiées A, B et C , les copies de gènes sélectionnés doivent être des orthologues (Olivier D, 2022).

À retenir

- Les gènes évoluent par mutations, duplications et sélection.
- Les familles de gènes proviennent de duplications successives.
- L'évolution moléculaire peut être neutre, purificatrice ou adaptative.
- L'horloge moléculaire estime le temps de divergence entre espèces.

Glossaire

Annotation : Identification des régions fonctionnelles du génome.

Famille de gènes : Gènes issus de la duplication d'un ancêtre commun.

Mutation synonyme : Mutation sans effet sur la protéine.

dN/dS : Ratio utilisé pour estimer le type de sélection.

Horloge moléculaire : Méthode estimant le temps de divergence génétique.

QCM

1. Qu'est-ce qu'une famille de gènes ?

- A. Des gènes situés sur le même chromosome
- B. Des gènes hérités de parents différents
- ☐ C. Des gènes issus de la duplication d'un gène ancestral
- D. Des gènes responsables des relations de parenté

2. Une mutation silencieuse est :

- A. Toujours bénéfique
- B. Une substitution non-synonyme
- ☐ C. Une mutation sans effet sur la protéine
- D. Une inversion chromosomique

3. Le ratio $dN/dS > 1$ indique :

- A. Une forte dérive génétique
- B. Une mutation neutre
- ☐ C. Une sélection positive
- D. Une erreur de recombinaison

4. L'annotation d'un génome permet de :

A. Créer de nouveaux gènes

☐ B. Identifier les régions fonctionnelles d'un ADN

C. Éliminer les séquences non codantes

D. Rendre l'ADN plus stable

5. L'horloge moléculaire est basée sur :

A. La durée de vie des organismes

☐ B. Le taux constant de mutation dans le temps

C. La taille des génomes

D. La vitesse de réplication de l'ADN

VI. Phylogénie Moléculaire

VI.1. Reconstruction des relations de parenté de séquences homologues

Des séquences similaires peuvent être homologues et descendre d'un ancêtre commun, en reconstruire l'évolution permet une meilleure compréhension. La reconstruction de l'évolution de séquences homologues appartient au domaine de la phylogénie moléculaire, dont l'objectif principal est de déterminer l'arbre phylogénétique qui décrit au mieux les relations de parenté évolutives des séquences homologues. Les séquences homologues actuelles sont les seules observations disponibles. Elles sont issues de différents événements évolutifs de duplications et de spéciations, ont accumulé des substitutions et ont été sujettes à des pressions de sélection. Le problème est alors de retracer l'histoire de ces processus sous forme d'un arbre phylogénétique, avec comme seule information de départ les séquences des homologues qui coexistent actuellement.

VI.2. Arbre phylogénétique

Un arbre phylogénétique représente les liens de parenté entre des entités (e.g. espèces, séquences) et sa topologie représente leur histoire évolutive (Figure 45). Il est composé de nœuds et de branches, où chaque nœud correspond à une instance de l'entité observée (ex., un gène), les entités actuelles sur lesquelles a été basée l'inférence de l'arbre sont les feuilles de l'arbre (ex., gène actuel dont la séquence a permis l'inférence), les autres nœuds sont internes et correspondent à des nœuds ancestraux (ex., gène ancestral). Dans le cas d'un arbre binaire, chaque nœud est lié à trois branches, et dans le cas où cet arbre est enraciné (i.e., on en connaît l'origine), une branche le lie à son ancêtre, et les deux autres branches à ses descendants. Si l'arbre n'est pas binaire, il possède au minimum un nœud qui n'est pas résolu et son nombre de descendants est alors supérieur à deux. La topologie d'un arbre peut être décrite par l'ensemble de ces bipartitions : chaque branche sépare l'arbre en deux parties, avec un sous-ensemble de taxons dans chacune (ex., seq1 seq2 | seq3 seq4 sur la Figure 45). La longueur d'une branche représente la quantité d'évolution qui s'est produite entre deux nœuds (ex., le nombre de substitutions par site). Une branche peut également être associée à des valeurs de supports statistiques (ex., *bootstrap*).

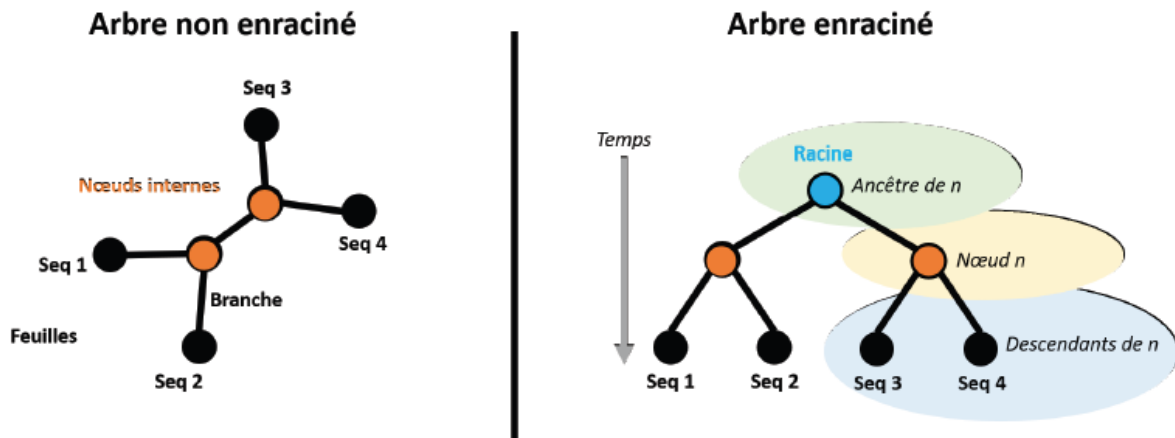


Figure 44: Arbre phylogénétique (Olivier D, 2022).

Exemple d'un arbre phylogénétique enraciné et non enraciné représentant les liens de parenté de quatre séquences homologues : seq1, seq2, seq3 et seq4 (feuilles de l'arbre). Chaque nœud interne (correspondant à une entité ancestrale) est lié à trois autres nœuds afin de représenter les relations de parenté. Si l'arbre est enraciné (l'origine en est connue), chaque nœud n possède un ancêtre et deux descendants. La topologie de cet arbre peut être décrite par la bipartition : seq1 seq2 | seq3 seq4.

VI.3. Construction d'un arbre phylogénétique

La première étape pour construire un arbre phylogénétique à partir de séquences homologues est de construire un alignement multiple de ces séquences (MSA, détaillé en gris sur la Figure 45), afin d'identifier des caractères homologues (issu d'un ancêtre commun et ayant pu subir des substitutions [WM17]) sur la base de leur similarité. Il existe une multitude de logiciels permettant de construire un MSA à partir de séquences homologues, chacun basé sur des principes différents et étant plus ou moins efficace et adapté aux types de séquence étudiés ou à l'utilisation que l'on veut en faire. La seconde étape (qui peut être optionnelle) est le nettoyage de l'alignement, elle consiste à ne garder que les colonnes du MSA les plus informatives. Les *indels* n'étant généralement pas considérés dans les modèles d'évolution, une colonne qui en possède une grande quantité est jugée peu informative et donc porteuse de peu de signal phylogénétique.

Cependant, une colonne peu informative reste informative, le dilemme est de savoir si le nettoyage du MSA a pour effet de s'enlever du bruit ou de l'information. Les principaux programmes utilisés pour nettoyer automatiquement un MSA sont Blocks et Trima. De plus, le nettoyage d'un MSA peut se compléter ou se faire exclusivement manuellement.

La troisième étape est l'inférence d'un arbre phylogénétique à partir d'un MSA, nettoyé ou non (en bleu, en jaune, en violet et en vert sur la Figure 2.2), à partir de différentes heuristiques qui vont permettre d'identifier la topologie de l'arbre sachant sa $\hat{z}$ vraisemblance : une fonction du MSA et du modèle de substitution.

La quatrième et dernière étape consiste à enraciner la topologie obtenue (en orange sur la Figure 45), en déterminant son origine dans le temps au milieu d'une branche, afin d'éviter toute erreur de raisonnement. L'enracinement d'un arbre se fait majoritairement par l'utilisation d'un groupe externe, groupe proche mais n'appartenant pas au groupe d'intérêt (un alignement doit être possible). L'enracinement de l'arbre peut également être estimé arbitrairement sur la base de la topologie de l'arbre (ex., enraciner au milieu, enraciner en fonction de la divergence la plus récente) ou par l'utilisation d'informations extérieures (ex., utilisation de l'arbre des espèces pour enraciner un arbre de gènes).

Dans le cas des méthodes de maximum de vraisemblance, l'arbre phylogénétique inféré est alors le plus vraisemblable, telle que la probabilité des données (MSA) soit maximale, sachant l'arbre et le modèle de substitution. Cependant, la probabilité que l'arbre soit correct (i.e., représente strictement la réalité biologique) est très faible. Différentes méthodes peuvent alors permettre de mesurer la robustesse de l'arbre et de ses différents sous-barbes, par exemple par échantillonnage qui consiste à faire varier les caractères donnés en entrée (enlever des colonnes, enlever des lignes du MSA) permet d'inférer des scénarios alternatifs à partir des mêmes données. Itérer ce processus avec différents tirages de caractères permet alors d'observer les fréquences auxquelles les différentes bipartitions sont retrouvées. Ces fréquences indiquent pour chaque bipartition de l'arbre à quel point elle est supportée par les données (ex., *boots rap*). Généralement pas considérées dans les modèles d'évolution, une colonne qui en possède une grande quantité est jugée peu informative et donc porteuse de peu de signal phylogénétique.

À retenir

- La phylogénie moléculaire permet de reconstruire les relations évolutives à partir des séquences d'ADN ou de protéines.
- Les séquences homologues sont comparées pour établir des arbres phylogénétiques.
- Un arbre phylogénétique montre les relations de parenté entre espèces ou gènes.
- La distance génétique et les méthodes d'alignement sont essentielles à l'inférence phylogénétique.
- Plusieurs méthodes existent pour construire un arbre : distance, maximum de vraisemblance, parcimonie.

Glossaire

Phylogénie : Étude des relations évolutives entre espèces ou gènes.

Séquence homologue : Séquences dérivant d'un ancêtre commun.

Arbre phylogénétique : Représentation graphique des relations évolutives.

Distance génétique : Mesure de divergence entre séquences.

Parcimonie : Méthode de construction d'arbre favorisant le scénario avec le moins de changements.

QCM

1. Quel est l'objectif de la phylogénie moléculaire ?

- A. Étudier la reproduction
- B. Identifier les maladies
- C. Reconstituer les relations évolutives à partir de séquences
- D. Analyser les mutations uniquement

Réponse : C

2. Une séquence homologue :

- A. A une fonction identique
- B. Provient d'un ancêtre commun
- C. Est identique chez toutes les espèces
- D. N'évolue pas

Réponse : B

3. Un arbre phylogénétique :

- A. Représente l'ADN mitochondrial uniquement
- B. Représente les relations de parenté évolutive
- C. Est toujours linéaire
- D. Ne contient pas de racine

Réponse : B

4. Quelle méthode privilégie le scénario le plus simple ?

- A. Distance génétique
- B. Maximum de vraisemblance
- C. Parcimonie
- D. Phénétique

Réponse : C

5. Plus deux espèces sont proches génétiquement, plus :

- A. Leur divergence est ancienne
- B. Elles ont d'ADN mitochondrial
- C. Elles partagent un ancêtre récent
- D. Elles ont des environnements similaires

Réponse : C

VII. Références :

- Abbas Ak, Murphy KM, Sher A. Functional diversity of helper T lymphocytes. *Nature*. 1996; 383:787-93.
- Agnès Doligez et al. New stable QTLs for berry weight do not colocalize with QTLs for seed traits in cultivated grapevine (*Vitis vinifera* L.). December 2013.
- Agnès Doligez et al. New stable QTLs for berry weight do not colocalize with QTLs for seed traits in cultivated grapevine (*Vitis vinifera* L.). December 2013. *BMC Plant Biology* 13(1):217.
- Ainouche M L. Alex. Baumel Alex. Baumel Armel. Salmon Armel. Salmon Glenn. Yannic Glenn. Hybridization, polyploidy and speciation in *Spartina* (Poaceae) January 2004. *New Phytologist* 161:165-172
- AKEN, Bronwen L. et al. "The Ensembl gene annotation system". In : *Database: The Journal of Biological Databases and Curation*. 2016.
- Blattner FR, Plunkett III G, Bloch CA, Perna NT, Burland V, Riley M. The complete genome sequence of *Escherichia coli* K-12. *Science*. 1997;277:1453–74.
- Butterfield ER, Abbott JC, Field MC. Automated Phylogenetic Analysis Using Best Reciprocal BLAST. *Methods Mol Biol*. 2021; 2369:41-63. doi: 10.1007/978-1-0716-1681-9_4. PMID: 34313983.
- Chen Y, Hong T, Wang S, Mo J, Tian T, Zhou X. Epigenetic modification of nucleic acids: from basic studies to medical applications. *Chem Soc Rev*. 2017 May 22; 46(10):2844-2872. doi: 10.1039/c6cs00599c. PMID: 28352906.
- Cornwell W, Nakagawa S. Phylogenetic comparative methods. *Curr Biol*. 2017 May 8;27(9): 333-336. doi: 10.1016/j.cub.2017.03.049. PMID: 28486113.
- Didelot X. Phylogenetic Analysis of Bacterial Pathogen Genomes. *Methods Mol Biol*. 2023;2674:87-99. doi: 10.1007/978-1-0716-3243-7_6. PMID: 37258962.
- Edgar RC, Batzoglou S. Multiple sequence alignment. *Curr Opin Struct Biol*. 2006 Jun;16(3):368-73. doi: 10.1016/j.sbi.2006.04.004. Epub 2006 May 5. PMID: 16679011.
- Emmanuel J.P. DOUZERY. *Phylogénie moléculaire*. Institut des Sciences de l'Evolution de Montpellier (ISE-M : UMR 5554 / CNRS) ; Université Montpellier II France. 2010 ; 183-240.
- Emmanuelle EVENO. L'adaptation à la sécheresse chez le pin maritime (*Pinus pinaster* Ait.) : patrons de diversité et différenciation nucléotidiques de gènes candidats et variabilité de caractères phénotypiques. Thèse de doctorat. Ecole doctorale : Sciences du Vivant, Géosciences, Sciences de l'Environnement. 2008.
- Frederique Pitel, Juliette Riquet. Les marqueurs anonymes et la détection de leur polymorphisme. *Productions Animales*, 2000, HS 2000, pp.45-53. hal-02694953.
- Futuyma, D. J. (2017). *Evolution* (4th edition). Sinauer Associates/Oxford University Press. (Illustration inspirée de la figure 16.4, p. 525.)

HARRY Myriam. Génétique moléculaire et évolutive. 2^{ème} édition Maloine. 2008

Jean Générmont. Les mécanismes de la spéciation. Société Française de Génétique. Medecine et Science. n ° 6 vol. 7. 1991 ; I-V.

Katoh K, Rozewicki J, Yamada KD. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform.* 2019 Jul 19;20(4):1160-1166. doi: 10.1093/bib/bbx108. PMID: 28968734; PMCID: PMC6781576.

Kuhls K, Mauricio I. Phylogenetic Studies. *Methods Mol Biol.* 2019;1971:9-68. doi: 10.1007/978-1-4939-9210-2_2. PMID: 30980297.

Li Y. Modern epigenetics methods in biological research. *Methods.* 2021 Mar;187:104-113. doi: 10.1016/j.ymeth.2020.06.022. Epub 2020 Jul 6. PMID: 32645449; PMCID: PMC7785612.

Mitsuko Yorkston and Curtis C. Daehler Interfertility between Hawaiian Ecotypes of *Sida fallax* (Malvaceae) and Evidence of a Hybrid Disadvantage. *International Journal of Plant Sciences.* 2006. Volume 167, Number 2.

Nadine Hanna, Béatrice Parfait, Dominique Vidaud, Michel Vidaud. Mécanismes et conséquences des mutations Médecine et Sciences 2005 ; 21 : 969-80

Olivier DENNLER. Caractérisation en modules fonctionnels des protéines ADAMTS- TSL, par approches de phylogénies. Thèse de doctorat présentée et soutenue à Rennes, le 19 Décembre 2022

Pirovano W, Heringa J. Multiple sequence alignment. *Methods Mol Biol.* 2008;452:143-61. doi: 10.1007/978-1-60327-159-2_7. PMID: 18566763.

Salipante SJ, Horwitz MS. Phylogenetic fate mapping. *Proc Natl Acad Sci U S A.* 2006 Apr 4;103(14):5448-53. doi: 10.1073/pnas.0601265103. Epub 2006 Mar 28. PMID: 16569691; PMCID: PMC1414797.

Sean B Carroll, John Doebley, Anthony J F Griffiths, Susan Wessler. Introduction à l'analyse génétique. 6^{ème} édition. De boeck. 2013.

Silva-Rocha R, Martínez-García E, Calles B, Chavarría M, Arce-Rodríguez A, de Las Heras A and Platero R.. The Standard European Vector Architecture (SEVA): a coherent platform for the analysis and deployment of complex prokaryotic phenotypes. *Nucleic acids research.* 41(Database issue):D666-75. 2013. doi: 10.1093/nar/gks1119.

Sophie Laget et Pierre-Antoine Defossez. Le double jeu de l'épigénétique. Médecine et Sciences 2008 ; 8-9(24) : 725 – 730.

Tang H, Finn RD, Thomas PD. TreeGrafter: phylogenetic tree-based annotation of proteins with Gene Ontology terms and other annotations. *Bioinformatics.* 2019 Feb 1;35(3):518-520. doi: 10.1093/bioinformatics/bty625. PMID: 30032202; PMCID: PMC6361231.

Ugo Bastolla, Markus Porto, H. Eduardo Roman, Michele Vendruscolo. Structural Approaches to Sequence Evolution. Springer. 2007

Verneuil Hubert De et al. Biochimie et biologie moléculaire (Conforme au programme UE1/UE2 1re et 2e années) : Conforme au programme UE1/UE2 1re et 2e années. 2003. Lavoisier medecine sciences.

Vidaurre V, Chen X. Epigenetic regulation of drosophila germline stem cell maintenance and differentiation. Dev Biol. 2021 May;473:105-118. doi: 10.1016/j.ydbio.2021.02.003. Epub 2021 Feb 18. PMID: 33610541; PMCID: PMC7992187.