



الجمهورية الجزائرية الديمقراطية الشعبية

République Algérienne Démocratique et Populaire

وزارة التعليم العالي والبحث العلمي

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

جامعة محمد الصديق بن يحيى جيجل

Université Mohammed Seddik Benyahia-Jijel



Faculté des Sciences de la Nature et de la Vie
Département des Sciences de l'Environnement et des
Sciences Agronomiques
Master 1 Semestre 1 Ecologie Fondamentale et
Appliquée

كلية علوم الطبيعة والحياة
قسم علوم المحيط والعلوم الفلاحية
السنة الاولى ماستر علم البيئة الأساسي والتطبيقي

Génétique des Populations

Polycopié de cours

Première année Master Ecologie Fondamentale et Appliquée

U E fondamental Code : UEF1

Préparé par Dr YOUNSI Salah Eddine
Grade de l'auteur MCA

2025/2026

Ref.....

Sommaire

INTRODUCTION	01
Chapitre I: Structure génétique d'une population	03
1. Concept de population biologique	03
2. Définition et objectifs de la génétique des populations	04
3. Notion de la variation	05
4. Composition génétique d'une population	05
4.1. Estimation de la composition génétique d'une population	05
4.2. Fréquence phénotypique	06
4.3. Fréquence génotypique et allélique	07
5. Exercice d'application	09
Chapitre II: Model de Hardy-weinberg	12
1. Etude théorique d'une population idéale	12
1.1. Notion de population théorique idéale	12
1.2. Conditions de la loi de Hardy-Weinberg	12
1.3. Démonstration de la loi de Hardy-Weinberg	13
2. Application de la loi de Hardy-Weinberg	14
2.1. Gènes autosomiques	14
2.2. Gènes liés au sexe	15
2.3. Système multi-allélique (Codominance, Dominance partielle, Dominance totale)	15
3. Conformité au modèle de Hardy-Weinberg : Test d'équilibre	17
3.1. Test de vérification de la conformité d'une population a la structure de Hardy-weinberg	17
3.2. Exercice d'application	19
Chapitre III. Influence du régime de reproduction : Consanguinité	21
1- Notion de régime de reproduction	21
1.1. Les régimes fermés	21
1.2. Les régimes ouverts	21
2- Cas particulier de la consanguinité	21
2.1. Introduction :	21
2.2. Définition	22
2.3. Causes de la consanguinité	22
2.4. Principaux paramètres mesurant la consanguinité	24
2.4.1. Coefficient de consanguinité :	
2.5. Conséquences de la consanguinité sur les fréquences génotypiques	26
2.6. Conséquences de la consanguinité sur les fréquences alléliques	28
Chapitre IV. Influence des pression évolutives	30
1. Introduction	30
2. Présentation des pressions évolutives	30
2.1. Les mutations ; source de variabilité	30
2.2. Migration ou Flux Géniques	32
2.3. Sélection	32
REFERENCES BIBLIOGRAPHIQUES	35

INTRODUCTION

Tout organisme vivant dans la biosphère appartient nécessairement à une population, c'est-à-dire un ensemble d'individus de la même espèce génétiquement apparentés ou présentant une certaine similarité génétique. L'une des caractéristiques principales des populations naturelles est leur polymorphisme, c'est-à-dire la coexistence de plusieurs allèles pour un gène donné au sein d'une même population, ce qui reflète leur variabilité génétique [1, 2 et 3].

La génétique initiée par Gregor Mendel, classiquement appelée génétique mendélienne, a pour objectif de comprendre le déterminisme et la transmission des caractères héréditaires en analysant la descendance issue de croisements contrôlés entre individus de génotypes différents, notamment par l'étude des proportions des diverses catégories de descendants. Après la découverte de l'ADN comme support de l'information génétique, la génétique moléculaire a émergé pour étudier les mécanismes fins du déterminisme, de l'expression et de la transmission des caractères. Elle connaît aujourd'hui de nombreuses extensions avec les programmes de génomique (séquençage des génomes et identification des gènes) et de protéomique (inventaire et fonction des protéines d'un organisme) [2].

La compréhension du déterminisme et de la transmission des caractères nécessite également d'étudier les individus dans leurs conditions naturelles, où ils sont génétiquement uniques et libres de se reproduire avec tout autre individu de la même espèce. Cette partie de la génétique, qui considère les individus en interaction avec leur environnement, est la génétique des populations. Cette dernière étudie les conséquences des lois fondamentales de la transmission du matériel génétique non plus au niveau de l'individu mais à l'échelle globale des populations. Discipline à l'interface de la zoologie, de la botanique systématique, de l'écologie et de l'éthologie, elle s'appuie sur des formulations et un raisonnement mathématique [1].

La génétique des populations est née dans les années 1920-1930 de la volonté de concilier la théorie darwinienne de l'évolution avec les données de plus en plus précises acquises depuis le début du XXe siècle sur la transmission du matériel héréditaire. Cette synthèse a été principalement réalisée par des mathématiciens et des biologistes – notamment R.A. Fisher, J.B.S. Haldane et S. Wright – qui ont élaboré les premiers modèles formels décrivant l'évolution des populations. L'objectif de la génétique des populations est, grâce à des modèles de plus en plus sophistiqués, de rendre compte des mécanismes et des dynamiques de l'évolution [1, 2, 4, 5 et 6].

Dans les conditions naturelles, une population constitue donc une communauté génétique formée par l'ensemble des génotypes des individus qui la composent. Elle se caractérise par un

génome collectif, ou patrimoine génétique, également appelé pool génétique (ou pool génique), qui correspond à la somme des allèles présents dans l'ensemble des individus pour chacun des gènes.

Bien que chaque génotype individuel soit définitivement déterminé à la naissance et disparaisse à la mort de l'individu, le pool génétique d'une population présente une continuité à travers les générations tout en pouvant varier au cours du temps. C'est cette dynamique évolutive que la génétique des populations cherche à comprendre.

Différentes notions de génétique des populations sont détaillées dans ce présent document, qui correspond à un cours conçu en respectant le programme pédagogique officiel, à l'intention des étudiants de première année de Master Écologie Fondamentale et Appliquée, filière Écologie et Environnement, domaine Sciences de la Nature et de la Vie. Ce cours est effectivement enseigné à l'université Mohamed Seddik Ben Yahia de Jijel.

L'objectif principal du cours de génétique des populations, en Master, est de comprendre comment et pourquoi la diversité génétique d'une population change au cours du temps et dans l'espace. En résumé, il s'agit d'étudier l'évolution au fil des générations, mais d'un point de vue mathématique et statistique. À ce niveau d'études, le cours dépasse la simple théorie pour fournir les outils nécessaires à la recherche ou à la gestion de la biodiversité.

Le cours est divisé en quatre chapitres : le premier expose la structure génétique d'une population ; le deuxième présente le modèle de Hardy-Weinberg ; le troisième traite l'influence de la consanguinité en tant que régime de reproduction ; le dernier chapitre aborde l'influence des pressions évolutives, qualifiées de forces motrices du vivant, qui sont au nombre de quatre : les mutations, la sélection naturelle, la dérive génétique et la migration.

On note ainsi que ces différents chapitres sont accompagnés de séries de travaux dirigés (TD) permettant aux étudiants de réaliser des applications.

Chapitre I: Structure génétique d'une population

1. Concept de population biologique

Une population est un ensemble d'individus appartenant à la même espèce, susceptibles d'interagir entre eux au moment de la reproduction, occupant une aire géographique commune et jouant un rôle particulier dans un écosystème. C'est l'unité d'étude dans de nombreux domaines des Sciences de la Vie (épidémiologie, évolution, écologie, biogéographie, biologie de la conservation).

Simple au plan théorique, cette définition est souvent difficile à appliquer aux situations naturelles. Les limites d'une population sont incertaines et dépendent des caractéristiques intrinsèques des espèces (répartition spatiale et temporelle des individus, mobilité, mode de reproduction, durée de vie, socialité, etc). Lorsqu'une espèce présente de très grands effectifs et occupe un vaste territoire apparemment homogène, seule l'étude détaillée de la distribution des individus, de leurs comportements, de leurs déplacements et de leurs génotypes peut permettre de déceler d'éventuelles discontinuités correspondant à des limites de populations.

La notion de population fait donc appel à des critères fondamentaux suivants :

1.1. Critère d'ordre systématique : Provient de l'appartenance à la même espèce, qui désigne un ensemble d'individus interféconds, isolés quant à la reproduction d'ensembles similaires.

La population est à distinguer de la notion d'espèce qui rassemble tous les individus interfertiles même si ceux-ci n'ont jamais la possibilité de se croiser.

1.2. Critère d'ordre biologique : Ce deuxième critère complète le précédent. Il résulte de la possibilité de s'apparier ou de se croiser, et donc d'assurer le brassage des gènes d'une génération à la suivante, faisant des populations naturelles des entités adaptable et évolutives ; les individus appartenant à la même population participent au même pool génétique

1.3. Critère d'ordre géographique : La possibilité de se croiser implique très généralement la présence des individus dans la même aire géographique, au moins à l'époque de la reproduction.

1.4. Critère d'ordre écologique : La présence dans la même aire géographique et dans l'écosystème où l'espèce trouve les éléments nécessaires à son existence implique que les individus utilisent les mêmes ressources, aient les mêmes prédateurs, parasites, etc. ; en un mot, les individus d'une population ont la même fonction dans l'écosystème auquel ils participent, vivant sous les mêmes contraintes générales, sont soumis aux mêmes pressions de sélection. Ce critère permet de reconnaître éventuellement comme appartenant à des

populations différentes, des espèces jumelles sympatriques ou des semi-espèces occupant des niches écologiques différentes.

2. Définition et objectifs de la génétique des populations

La génétique des populations est une branche de la génétique qui étudie la variabilité des gènes au sein et entre des populations d'être vivants soumis à des pressions évolutives variées, suivant trois objectifs principaux :

- 1) **Mesurer la variabilité génétique** (appelée aussi diversité génétique) par l'estimation des fréquences alléliques pour différents gènes.
- 2) **Comprendre comment** la variabilité génétique se transmet d'une génération à l'autre.
- 3) **Comprendre comment et pourquoi** la variabilité génétique évolue au fil des générations.

A la différence de la génétique mendélienne, la génétique des populations étudie les proportions (ou fréquences) des allèles, des génotypes et des phénotypes au sein d'un ensemble d'individus issus de croisements non contrôlés entre de nombreux parents (fig.01).

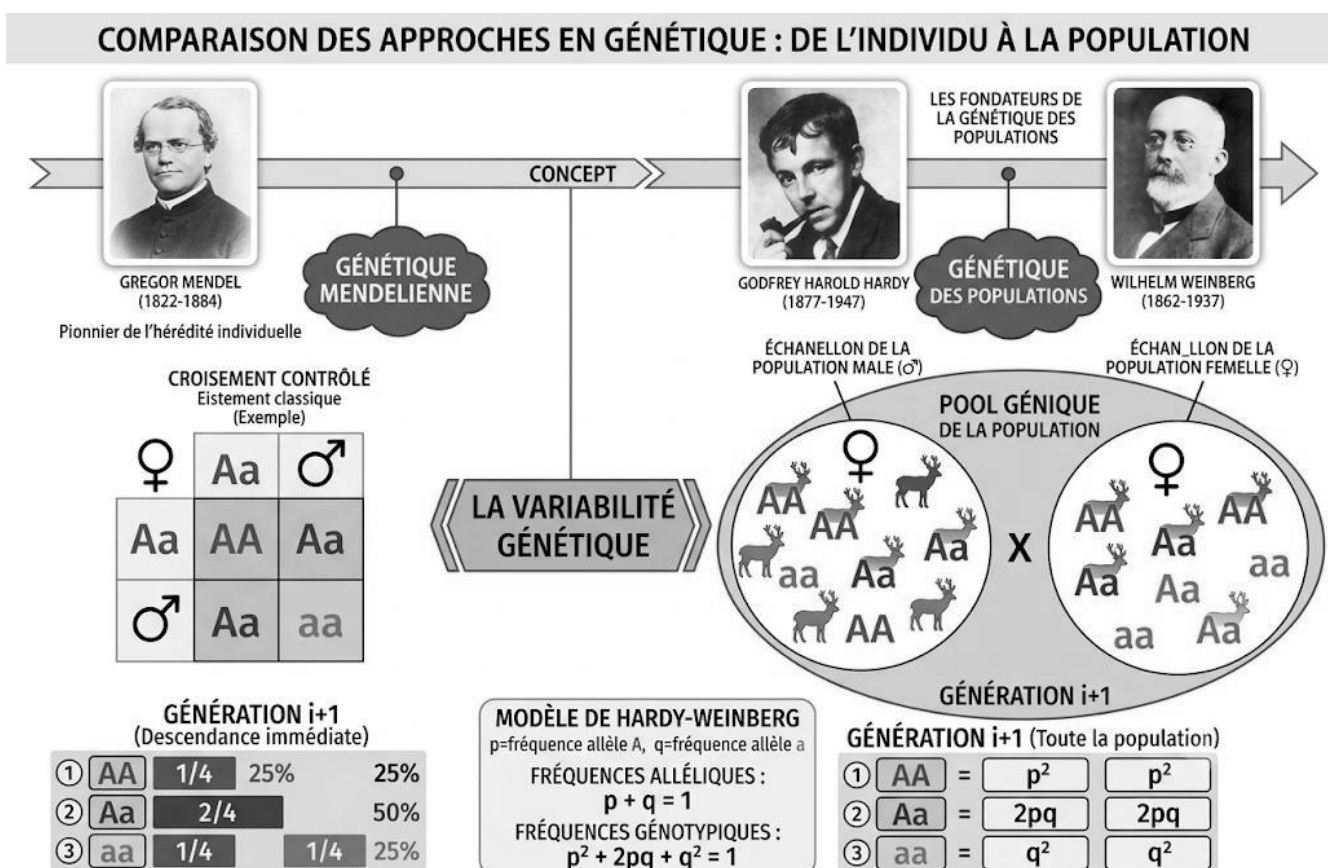


Figure 1 : Transition conceptuelle entre la génétique mendélienne (croisement individuel) et la génétique des populations (modèle d'Hardy-Weinberg). Portraits : G. Mendel, G.H. Hardy et W. Weinberg (Archives historiques, domaine public) [7]

3. Notion de la variation

La variation est l'une des propriétés générales des populations naturelles (animales et végétales) qui peut être perçue par des caractères morphologiques et physiologiques, témoignant de la diversité phénotypique.

Dans toute population, les individus présentent, pour la plupart des caractères, des phénotypes très différents. Même la plus simple des observations révèle une variabilité phénotypique impressionnante. Dans l'espèce humaine, on connaît une grande diversité de taille, de poids, de la forme du corps, de la couleur et de la texture des cheveux, de la couleur de la peau, des yeux, et de bien d'autres traits ou aptitudes physique et diversité phénotypiques, et spécialement la part qui résulte de différences entre génotypes. En particulier la génétique des populations comme discipline se repose de déterminer l'ampleur de la variabilité génétique dans les populations naturelles, d'en expliquer l'origine, le maintien et l'importance évolutive.

Dans la plupart des populations naturelles, cette variabilité se manifeste par la présence de multiples allèles sur de nombreux gènes ; diverses techniques d'étude révèlent cette variabilité, parmi lesquelles l'électrophorèse de protéines dont l'usage s'est répandu à la fin des années soixante pour révéler la variabilité génétique des enzymes et d'autres protéines.

4. Composition génétique d'une population

La composition génétique des populations est présentée par l'ensemble des fréquences des gènes qui la constitue. Elle est décrite donc par les fréquences phénotypiques, génotypiques et les fréquences alléliques.

4.1. Estimation de la composition génétique d'une population

L'ensemble des contraintes qui s'exercent sur la population va déterminer sa constitution génétique et l'importance des variations génétiques qu'elle présente constitue son potentiel évolutif. Les différents croisements qui se reproduisent dans la population vont déterminer les fréquences relatives des génotypes. La constitution génétique de la population peut être définie par la distribution des gènes sur les différents loci. Elle peut être appréhendée à plusieurs niveaux :

- a) Phénotypique si l'on considère juste l'apparence de l'organisme ;
- b) Génotypique si l'on calcule la fréquence des différents génotypes dans la population ;
- c) Allélique si l'on tente d'estimer la fréquence d'un allèle pour chaque locus étudié dans la population.

Les techniques statistiques sont utilisées pour calculer la constitution génétique à partir d'un échantillon de la population étudiée. Pour les caractères mendéliens, les calculs de fréquence restent

relativement simples tandis que les caractères quantitatifs à déterminisme plus complexe comme la taille ou le poids doivent être évalués par des méthodes plus sophistiquées.

❖ **Exemple :**

- si un gène d'un locus n'existe qu'en un seul état allélique, la proportion de cet allèle dans la population sera de 1,0.
- Pour un gène diallélique de type A/a et autosomique (non porté par un hétérochromosome) il existe six possibilités de croisement des génotypes parentaux (Tableau suivante)

Tableau 01 : Fréquence des génotypes pour un gène diallélique A et a

Type de croisements parentaux	Fréquence des génotypes des descendants
AA x AA	1,0 AA
aa x aa	1,0 aa
AA x aa	1,0 Aa
AA x Aa	0,5 AA, 0,5 Aa
Aa x aa	0,5 Aa, 0,5Aa
Aa x Aa	0,25 AA, 0,5 Aa et 0,25aa

4.2. Fréquence phénotypique

La fréquence phénotypique est la proportion d'un phénotype particulier (caractère observable) dans une population. Elle s'exprime souvent en fraction, en décimale ou en pourcentage. Elle dépend des fréquences alléliques et des relations de dominance entre allèles.

On calcul la fréquence phénotypique de la manière suivante :

$$f(\text{phénotype}) = \frac{\text{Nombre d'individus présentant ce phénotype}}{\text{Nombre total d'individus dans la population}}$$

Pour un locus dialléliques A (dominant) et a (récessif), les fréquences phénotypiques seront :

$$f([A]) = \frac{\text{nombre d'individus porteurs de ce phénotype [A]}}{\text{le nombre total des individus (N)}}$$

$$f[a] = \frac{\text{nombre d'individus porteurs de ce phénotype [a]}}{\text{le nombre total des individus (N)}}$$

4.3. Fréquence génotypique et allélique

4.3.1. Cas de gène autosomique

- ❖ **Fréquence génotypique** : Elle est décrite par l'ensemble des fréquences dans la population des différents génotypes au locus considéré. Sa connaissance directe n'est possible que si l'on peut reconnaître les différents génotypes à ce locus, par leur phénotype (variation discontinue, pas de dominance) ou par l'examen de leurs descendance.

Soit : $n(AA)$ = nombre d'individus ayant le génotype AA

$n(Aa)$ = nombre d'individus ayant le génotype Aa

$n(aa)$ = nombre d'individus ayant le génotype aa

N = le nombre total des individus ($n(AA) + n(Aa) + n(aa)$)

Les fréquences génotypiques pour cette population seront :

$$f(AA) = \frac{n(AA)}{N} = D$$

$$f(Aa) = \frac{n(Aa)}{N} = H$$

$$f(aa) = \frac{n(aa)}{N} = R$$

- ❖ **Fréquence allélique** : Ce sont les fréquences dans la population des différents allèles au locus considéré. Leur connaissance donne la structure allélique (ou dite aussi génique) de cette population.

- Dans le cas du même locus biallélique (A, a) précédent on aura :

- Fréquence de l'allèle A :

$$f(A) = p = \frac{n(Aa) + 2 \cdot n(AA)}{2 \cdot N} = \frac{H}{2} + D$$

- Fréquence de l'allèle a :

$$f(a) = q = \frac{n(Aa) + 2 \cdot n(aa)}{2 \cdot N} = \frac{H}{2} + R$$

Ces formules sont toujours vraies : elles n'impliquent aucune supposition sur le régime de reproduction ou l'existence ou non de pressions évolutives.

- Cas d'un locus multiallélique: $A_1, A_2, \dots, A_j, \dots, A_n$
 - Fréquence alléliques P : $P_1, P_2, \dots, P_j, \dots, P_n$
 - $P_i = 2n A_i A_i + \sum n A_i A_j / 2N = Fr A_i A_i + \frac{1}{2} \sum F A_i A_j$

Remarque :

En général on n'échantillonne qu'une fraction de la population (entité elle-même généralement mal définie dans ses limites). Les fréquences calculées comme ci-dessus à partir d'un échantillon sont des estimateurs sans biais des fréquences de la population.

4.3.2. Cas de gène lié au sexe :

Pour les **gènes autosomiques** (portés par les chromosomes non sexuels), les formules données précédemment s'appliquent directement. Pour les **gènes liés au sexe** (portés par les chromosomes sexuels, typiquement le chromosome X chez les mammifères), le calcul des fréquences doit tenir compte du fait que :

- Les femelles (XX) ont deux copies du gène.
- Les mâles (XY) n'ont qu'une seule copie (ils sont **hémizygotes**).

On traite séparément les effectifs mâles et femelles. Les formules diffèrent selon que l'on calcule des fréquences génotypiques ou alléliques.

❖ **Exemple de gène lié au sexe pour un locus sur le chromosome X****a) Effectifs et génotypes possibles**

- **Femelles (XX)** : trois génotypes possibles : $X^A X^A$, $X^A X^a$, $X^a X^a$.
- **Mâles (XY)** : deux génotypes possibles : $X^A Y$ (hémizygote A), $X^a Y$ (hémizygote a).

Soient :

- N_f = nombre de femelles, N_m = nombre de mâles.
- n_{AA} , n_{Aa} , n_{aa} : effectifs des femelles pour chaque génotype.
- m_A , m_a : effectifs des mâles pour chaque allèle sur X.

b) Fréquences génotypiques

On calcule séparément chez les femelles et chez les mâles (car les mâles n'ont pas de « génotype diploïde » au sens classique).

❖ Chez les femelles :

$$f_{fem}(AA) = \frac{n_{AA}}{N_f}, \quad f_{fem}(Aa) = \frac{n_{Aa}}{N_f}, \quad f_{fem}(aa) = \frac{n_{aa}}{N_f}$$

❖ Chez les mâles :

$$f_{mâles}(X^A Y) = \frac{m_A}{N_m}, \quad f_{mâles}(X^a Y) = \frac{m_a}{N_m}$$

c) Fréquence allélique (du gène lié à X)

Il faut compter le nombre total d'allèles X^A et X^a sur l'ensemble des chromosomes X dans la population.

- Nombre total de chromosomes X = $2N_f + N_m$ (deux X par femelle, un X par mâle).
- Nombre d'allèles $X^A = 2 \times n_{AA} + n_{Aa} + m_A$.
- Nombre d'allèles $X^a = 2 \times n_{aa} + n_{Aa} + m_a$.

D'où :

$$f(X^A) = \frac{2n_{AA} + n_{Aa} + m_A}{2N_f + N_m}$$

$$f(X^a) = \frac{2n_{aa} + n_{Aa} + m_a}{2N_f + N_m}$$

Vérification : $f(X^A) + f(X^a) = 1$.

5. Exercice d'application

- ❖ **Exercice 1 :** Le schéma ci-après, représente le pool génique correspondant à un gène autosomal diallélique (l'allèle A est dominant alors que l'allèle a est récessif), dans une population P diploïde de 13 individus. Calculez les fréquences phénotypiques, génotypiques et allélique de la population (fig.02).

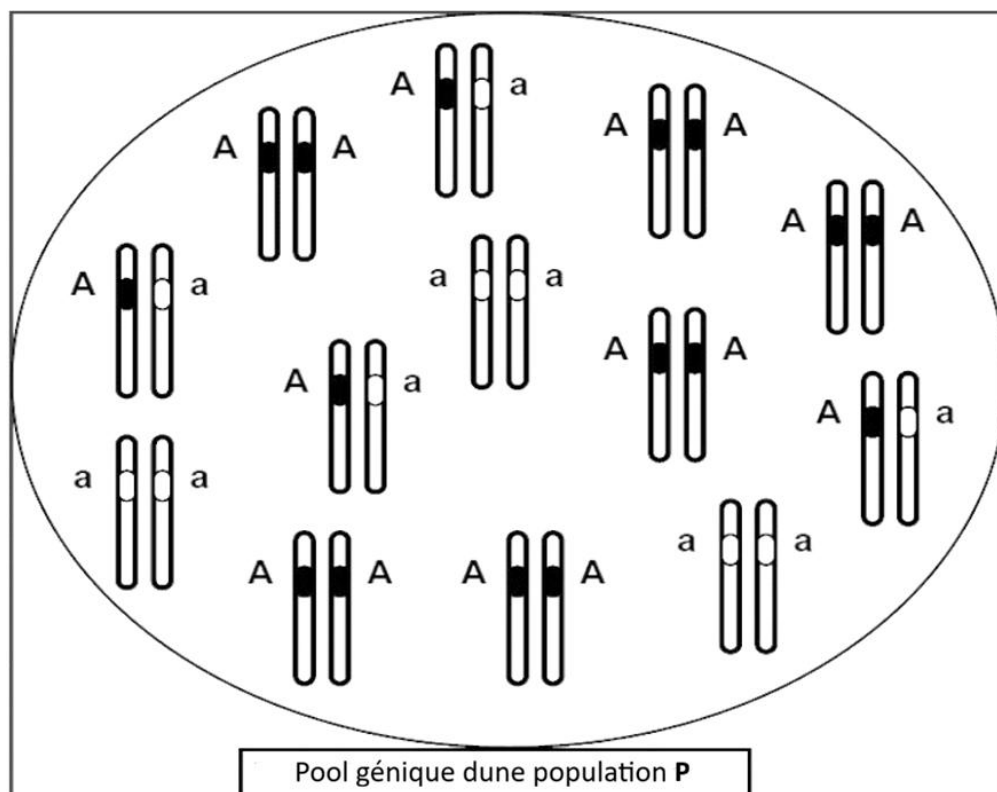


Figure 2 : Pool génique d'une population P

❖ **Solution de l'exercice 1**▪ **Calcul des fréquences phénotypiques :**

Soit N : l'effectif de la population, $n [A]$: le nombre d'individus porteurs du phénotype $[A]$ et $n [a]$: le nombre d'individus porteurs du phénotype $[a]$.

On calcule les fréquences phénotypiques suivant les formules suivantes :

$$F [A] = n[A]/N ; F [a] = n[a]/N \text{ avec } F [A] + F [a] = 1$$

Application numérique : $N = 13$, $n [A] = 10$ et $n [a] = 3$

$$F [A] = [10]/13 = 0.77;$$

$$F [a] = 3/13 = 0.23$$

▪ **Calcul des fréquences génotypiques :**

On considère N : l'effectif de la population, $n (AA)$: le nombre d'individus porteurs du génotype AA , $n (aa)$: le nombre d'individus porteurs du génotype aa et $n (Aa)$: le nombre d'individus porteurs du génotype Aa .

On calcule les fréquences génotypiques suivant les formules suivantes :

$$F (AA) = n(AA)/N = D ; F (Aa) = n(Aa)/N = H ; F (aa) = n(aa)/N = R$$

On a : $D + H + R = 1$, ainsi $F [A] = f (AA) + F [Aa]$ et $F [a] = f (aa)$

Application numérique : $N = 13$, $n (AA) = 6$, $n (Aa) = 4$ et $n (aa) = 3$

$$F (AA) = 6/13 = 0.46 = D ;$$

$$F (Aa) = 4/13 = 0.31 = H ;$$

$$F (aa) = 3/13 = 0.23 = R$$

▪ **Calcul des fréquences alléliques :**

- La fréquence de l'allèle $A = f (A) = P$;

$$p = \frac{n(Aa) + 2 \cdot n(AA)}{2 \cdot N} = \frac{H}{2} + D$$

$$P = (4 + 2 \times 6) / 2 \times 13 = 0.31/2 + 0.46 = 0.615$$

- La fréquence de l'allèle $a = F (a) = q$

$$q = \frac{n(Aa) + 2 \cdot n(aa)}{2 \cdot N} = \frac{H}{2} + R$$

$$q = (4 + 2 \times 3) / 2 \times 13 = 0.31/2 + 0.23 = 0.385$$

On peut vérifier ici que $p + q = 1$

❖ **Exercice 2** : Un exemple concret pour un locus lié à X. Soit la population suivante :

- 200 femelles : 90 $X^A X^A$, 80 $X^A X^a$, 30 $X^a X^a$
- 100 mâles : 60 $X^A Y$, 40 $X^a Y$

• **Calculs des fréquences génotypiques et alléliques :**

- **Fréquences génotypiques**

Femelles : $f(AA) = 90/200 = 0,45$, $f(Aa) = 80/200 = 0,40$, $f(aa) = 30/200 = 0,15$

Mâles : $f(X^A Y) = 60/100 = 0,60$, $f(X^a Y) = 40/100 = 0,40$

- **Fréquence alléliques**

$$\text{Nb total de X} = 2 \times 200 + 100 = 500$$

$$\text{Nb de } X^A = 2 \times 90 + 80 + 60 = 180 + 80 + 60 = 320$$

$$\text{Nb de } X^a = 2 \times 30 + 80 + 40 = 60 + 80 + 40 = 180$$

$$f(X^A) = 320/500 = 0,64;$$

$$f(X^a) = 180/500 = 0,36$$

Chapitre II : Model de Hardy-Weinberg

1. Etude théorique d'une population idéale

1.1. Notion de population théorique idéale

La prévision de la variabilité génétique d'une population est très difficile à réaliser du fait des mutations, de la transmission simultanée de plusieurs gènes qui peut entraîner des interactions, etc. Pour éviter ces problèmes, on peut tenter de formuler des hypothèses par la définition d'un modèle où les croisements seraient réellement aléatoires (effectif infini), sans migration et mutation, en faisant abstraction de la sélection (tous les individus ont les mêmes chances de se reproduire).

Le premier modèle publié concernant la structure génotypique d'une population l'a été simultanément et indépendamment par le mathématicien anglais G.H. Hardy et le médecin allemand W. Weinberg en 1908.

Dans ce modèle, les fréquences des allèles et des génotypes suivent une loi, la loi de Hardy-Weinberg, qui est le modèle de référence de la génétique des populations. Cette loi énonce que les fréquences alléliques et les fréquences génotypiques restent stables de génération en génération [8].

1.2. Conditions de la loi de Hardy-Weinberg.

Dans leurs modèles, ces auteurs ont adopté les six hypothèses ci-après permettant de simplifier les calculs et le raisonnement :

- 1.2.1. Diploïdie de l'organisme :** ça concerne les organismes ayant des cellules diploïdes ($2n$), avec deux jeux complets de chromosomes homologues (un jeu de chromosomes venant du parent mâle et un du parent femelle).
- 1.2.2. Reproduction sexuée :** C'est un mode de reproduction qui implique la fusion de deux cellules spécialisées appelées gamètes (le spermatozoïde chez le mâle et l'ovule chez la femelle). Cette fusion, appelée fécondation, donne naissance à une cellule-œuf ou zygote, qui se développe ensuite pour former un nouvel individu.
- 1.2.3. Système de reproduction panmictique :** Un système de reproduction panmictique (du grec pan = tout, mixis = mélange) signifie que tous les individus d'une population, quel que soit leur génotype, ont la même probabilité de se reproduire entre eux. En d'autres termes, il s'agit d'une reproduction aléatoire (croisements au hasard) sans aucune contrainte de sélection, de parenté, de distance géographique, ni de préférence sexuelle.
- 1.2.4. Effectif infini de la population :** Implique de la taille de la population est très grande, de sorte que la fréquence d'un événement soit égale à sa probabilité (la loi des grands nombres)

1.2.5. Des générations non chevauchantes : Aucun croisement entre individus de générations différentes

1.2.6. Absence des pressions évolutives :

- Absence de migration :** La population doit être close génétiquement (absence de flux migratoires). En effet, un individu étranger a peu de chance de présenter la même composition génétique p^2 , $2pq$ et q^2 que les individus de la population étudiée
- Absence de mutations :** Tout changements dans la séquence des nucléotides (A, T, C, G) de l'ADN. Elle peut toucher un gène (mutation ponctuelle), un chromosome (délétion, duplication, inversion, translocation), ou même l'ensemble du génome (polyploidie).
- Absence de sélection :** il faut considérer que tous les couples sont fertiles dont chacun des adultes reproducteurs va donner un nombre égal des gamètes, en plus de l'absence de la sélection gamétique, soit la probabilité de survie de chaque gamète jusqu'à la fécondation est indépendante de son génotype.

1.3. Démonstration de la loi de Hardy-Weinberg

Considérant une population diploïde dans laquelle coexistent deux allèles en un locus, **A** et **a**, en fréquences respectives **p** et **q**, ($p + q = 1$). La population gamétique est donc composée de gamètes porteurs de l'allèle **A** en fréquence **p**, et de gamètes porteurs de **a** en fréquence **q**. En appliquant les hypothèses **H₄** et **H₅**, la génération diploïde suivante s'obtient par le tableau ci-dessous où la fréquence de chaque case est le produit des fréquences de la ligne et de la colonne correspondantes :

Tableau 2: Fréquences des genotypes pour la generation suivante

Gamètes mâles Gamètes femelles	A	a
	p	q
A	AA p ²	Aa pq
a	Aa pq	aa q ²

La structure génotypique de la nouvelle génération produite est donc :

p² individus AA

2pq individus Aa

q² individus aa

Cette structure génotypique est célèbre sous le nom de structure de **Hardy-Weinberg**.

Selon l'hypothèse H_0 les gamètes que produira cette nouvelle génération contiendront :

- Le gène A en fréquence $p_1 = p_2 + \frac{1}{2} (2pq) = p (p+q) = p$
- Le gène a en fréquence $q_1 = q_2 + \frac{1}{2} (2pq) = q (p+q) = q$

En résumé, la loi de **Hardy-Weinberg** est actuellement énoncée sous la forme suivante :

- Dans une population isolée d'effectif illimité, non soumise à la sélection, et dans laquelle il n'y a pas de mutation, les fréquences alléliques restent constantes.
- Si les accouplements sont panmixiques, les fréquences génotypiques se déduisent directement des fréquences alléliques selon les relations ci-dessous ; elles restent donc aussi constantes.

$$\text{Fr}(\mathbf{AA}) = p^2$$

$$\text{Fr}(\mathbf{Aa}) = 2pq$$

$$\text{Fr}(\mathbf{aa}) = q^2$$

2. Application de la loi de Hardy-Weinberg

2.1. Gènes autosomiques

2.1.1. Codominance $A=B$:

Soit une population en équilibre (soumise à la loi de H-W), si les fréquences alléliques à un locus donné A/a sont respectivement p et q :

- Les fréquences génotypiques seront : $f(AA) = p^2$, $f(Aa) = 2pq$ et $f(aa) = q^2$ (tab. ...).
- Les fréquences phénotypiques seront : $f[A] = p^2$, $f[Aa] = 2pq$ et $f[a] = q^2$ (tab. ...).

Tableau 3 : fréquences des phénotypes et génotype dans le cas de codominance

Génotypes	AA	Aa	aa
Phénotypes	[A]	[Aa]	[a]
Fréquences	p^2	$2pq$	q^2

2.1.2. Dominance et Récessivité $A > a$:

Dans une population théorique-idéale, on considère un gène représenté par deux allèles: l'allèle **A** dominant et l'allèle **a** récessif.

On suppose qu'à une génération initiale G_0 on a :

Tableau 3 :

Tableau 4 : fréquences des phénotypes et génotype dans le cas de dominance A sur a

Génotypes	AA	Aa	aa
Phénotypes	[A]	[A]	[a]
Fréquences	p^2	$2pq$	q^2

2.2. Gènes liés au sexe

2.2.1. Calcul des fréquences

Dans une population équilibrée, on considère un gène à deux allèles (A et a) porté sur le chromosome sexuel X. Soit p la fréquence de l'allèle A et q la fréquence de l'allèle a

> Chez les femelles on a : $f(X^A X^A) = p^2$; $f(X^A X^a) = 2pq$; $f(X^a X^a) = q^2$

$$f(\varphi[A]) = f(X^A X^A) + f(X^A X^a) = p^2 + 2pq ; f(\varphi[a]) = f(X^a X^a) = q^2$$

> Chez les mâles on a : $f(\sigma[A]) = f(X^A Y) = f(A) = p$; $f(\sigma[a]) = f(X^a Y) = f(a) = q$

Conclusion : chez les femelles, les fréquences des génotypes sont conformes à la loi de H-W, alors que chez les mâles, les fréquences des génotypes sont celles des allèles.

2.2.2. Application aux maladies héréditaires liées au chromosome X

a) Fréquence d'une maladie récessive

> On considère une maladie héréditaire récessive liée au chromosome X. La fréquence de l'allèle (a), responsable de la maladie est égale à $q = 0,1$. Dans ce cas :

- La fréquence de la maladie chez les femelles est : $f(\varphi[a]) = f(X^a X^a) = q^2 = 0,01$
- La fréquence de la maladie chez les mâles est : $f(\sigma[a]) = f(X^a Y) = f(a) = q = 0,1$

> On a : $q > q^2$, on déduit que les mâles sont plus touchés que les femelles.

b) Fréquence d'une maladie dominante

> On considère une maladie héréditaire dominante liée au chromosome X. La fréquence de l'allèle (A), responsable de la maladie est égale à $p = 0,01$. Dans ce cas :

- La fréquence de la maladie chez les femelles est : $f(\varphi[A]) = f(X^A X^A) + f(X^A X^a) = p^2 + 2pq = 0,0001 + 0,0198 = 0,0199$
- La fréquence de la maladie chez les mâles est : $f(\sigma[A]) = f(X^A Y) = f(A) = p = 0,01$ On a : $p^2 + 2pq > p$, on déduit que les femelles sont plus touchés que les mâles.

2.3. Système multi-allélique (Codominance, Dominance partielle, Dominance totale)

La loi de Hardy-Weinberg s'applique également à des gènes multiallélique. Pour un locus multiallélique, la loi de Hardy-Weinberg s'étend naturellement en remplaçant les carrés des fréquences alléliques pour les homozygotes et le double produit pour les hétérozygotes. Le principe fondamental reste le même.

Soit un gène avec **k allèles** $A_1, A_2, \dots, A_k$ de fréquences : $p_1, p_2, \dots, p_k$ avec

$$\sum_{i=1}^k p_i = 1$$

À l'équilibre de Hardy-Weinberg :

- **Fréquence des homozygotes** $A_i A_i$:

$$p_i^2$$

- **Fréquence des hétérozygotes** $A_i A_j$ (pour $i \neq j$) :

$$2p_i p_j$$

La somme de toutes les fréquences génotypiques vaut 1 :

$$\sum_{i=1}^k p_i^2 + \sum_{i < j} 2p_i p_j = \left(\sum_{i=1}^k p_i \right)^2 = 1$$

➤ **Exemple de gène à 3 allèles (système ABO, par exemple)**

Prenons le système **ABO** (allèles A,B,O) avec fréquences p, q, r de sorte que ($p+q+r=1$).

En respectant les conditions d'équilibre on aura des résultats comme le montre le tableau suivant :

Tableau 5: Fréquences génotypiques attendues à la génération suivante pour le locus triallélique du groupe sanguin ABO

Gamètes mâles Gamètes femelles	A	B	O
	p	q	r
A	AA p^2	Aa pq	AO pr
B	AB pq	BB q^2	BO qr
O	AO pr	BO qr	OO r^2

> **La fréquence des génotypes : → ça donne six génotypes possibles**

- 1) $F(AA) = p^2$
- 2) $F(BB) = q^2$
- 3) $F(OO) = r^2$
- 4) $F(AB) = 2pq$
- 5) $F(AO) = 2pr$
- 6) $F(BO) = 2qr$

> **La fréquence des phénotypes : → ça donne quatre génotypes possibles**

- 1) $F[A] = p^2 + 2pr$
- 2) $F[B] = q^2 + 2qr$
- 3) $F[O] = r^2$
- 4) $F[AB] = 2pq$

3. Conformité au modèle de Hardy-Weinberg : Test d'équilibre

Les facteurs qui modifient la structure génétique de la population alternant l'équilibre d'Hardy-Weinberg peuvent dès lors être considérés comme des **processus évolutifs**. Ces facteurs peuvent venir perturber les fréquences génotypiques ou même modifier les fréquences alléliques. Les changements des fréquences alléliques constituent des événements majeurs d'évolution mais des modifications importantes des fréquences génotypiques peuvent également avoir des implications considérables comme c'est le cas lors des accouplements non aléatoire par exemple.

L'évolution peut donc être définie comme un changement durable dans une population provoqué par une altération des conditions d'équilibre.

3.1. Test de vérification de la conformité d'une population a la structure de Hardy-weinberg.

Lorsqu'un échantillon issu d'une population est étudié à un locus donné, et que sa composition génotypique peut être établie, il est possible de tester si les fréquences génotypiques observées correspondent ou non à une population ayant la structure de Hardy-Weinberg.

Si un échantillon présente pour un locus biallélique la composition suivante ;

Tableau 6 : Effectifs observés pour chaque génotype

Génotypes	AA	Aa	aa	Total
Effectifs observés	N_1	N_2	N_3	N

Les allèles A et a sont par conséquent représentés en fréquences p et q telle que :

$$p = (2N_1 + N_2) / 2N$$

$$q = (2N_3 + N_2) / 2N$$

Ces fréquences sont des estimateurs des fréquences réelles de la population possédant la structure de Hardy-Weinberg.

Tableau 7 : Effectifs estimés dit attendus ou théorique pour chaque génotype

Génotypes	AA	Aa	aa
Effectifs estimés (théoriques)	Np^2	$2Npq$	Nq^2

Afin d'évaluer la validité de l'hypothèse génétique, nous utiliserons un test de la distribution statistique appelé ; **test de chi-deux** (χ^2). Ce test mesure les écarts entre la répartition observée dite expérimentale et la répartition attendue dite théorique. Il puisse convertir les déviations par rapport aux valeurs attendus, en probabilité que telles inégalité soit obtenues par hasard.

$$\chi^2 = \sum_{\text{génotype}} \frac{(\text{Effectifs observés} - \text{Effectifs théoriques})^2}{\text{Effectifs théoriques}}$$

$$\chi^2 = \frac{(N_1 - Np^2)^2}{Np^2} + \frac{(N_2 - 2Npq)^2}{2Npq} + \frac{(N_3 - Nq^2)^2}{Nq^2}$$

Par ailleurs ce test doit prendre en considération la taille de l'échantillon et le nombre de variables ou degrés de liberté. A tout valeur de **chi-deux**, on associe un second nombre appelé le nombre de degré de liberté (ddl).

$$\text{ddl} = \text{nombre de génotypes possible (k)} - \text{nombre des allèles au locus considéré (r)} \rightarrow \text{ddl} = k - r$$

Tableau 8 : des nombres de degrés de liberté selon le nombre d'allèles étudiées.

Allèle	Génotypes	ddl
2	3	1
3	6	3
4	10	6
5	15	10
6	21	15

Selon la valeur de **chi-deux** calculée et le nombre de degré de liberté associés, on obtient un nombre correspondant qui témoigne de la qualité de conformité (de validité). Ce nombre est la valeur correspondante de la probabilité qui se détermine d'après la table de **chi-deux** ou de la figure appropriée (annexe). C'est la probabilité que le hasard seul puisse produire entre les valeurs observées et les valeurs attendus un écart au moins aussi grand que l'écart réellement obtenu.

Si la probabilité est inférieure à **0,05**, la qualité de conformité est considérée comme suffisamment faible pour que le modèle soit jugé non conforme (**rejet de l'hypothèse de conformité**), il y a des différences significatives entre les effectifs observés et les effectifs attendus. Donc la population ne suit pas la structure d'équilibre d'Hardy-Weinberg.

Inversement, si la probabilité est supérieure à **0,05** l'adéquation (ou la conformité) est considérée suffisamment correcte pour que le modèle ne soit pas rejeté (**accepte de l'hypothèse de conformité**). Il n'y a pas de différences significatives entre les effectifs observés et les effectifs attendus. Donc la population considérée suit bien la structure d'équilibre d'Hardy-Weinberg.

3.2. Exercice d'application

On peut illustrer une utilisation de la loi de Hardy-Weinberg avec les données réelles concernant les phénotypes du groupe sanguin MN parmi un échantillon de 1000 britanniques [9]. Le groupe sanguin MN est sous la dépendance de deux allèles d'un seul gène. Les deux allèles sont appelés M et N, et chaque allèle contrôle la production de molécules situées à l'aide de réactifs chimiques appropriés. Les substances correspondant aux allèles M et N sont appelées M et N, respectivement, avec :

- a) Les individus MM produisent seulement la substance M
- b) Les individus NN produisent seulement la substance N
- c) Les individus MN produisent seulement la substance MN

Parmi 1000 individus il y avait : 298 phénotypes M, 489 MN et 213 N. Puisque chaque génotype correspond à un phénotype unique dans ce système nous savons immédiatement que notre échantillon contient : 298 génotypes MM, 489 génotypes MN et 213 génotypes NN.

- 1) Estimez les fréquences des allèles M et N?
- 2) Déterminez si les fréquences génotypiques suivent la loi de Hardy-Weinberg ?

Réponse

Soit p est la fréquence de l'allèle M et q la fréquence de l'allèle N

$$p = (2n(MM) + n(MN)) / 2n = (496 + 489) / 2000 = 0,5425$$

$$q = (2n(NN) + n(MN)) / 2n = (418 + 489) / 2000 = 0,4575$$

(Encore une fois remarquez que $p+q = 1$)

Si la population est en équilibre de Hardy-Weinberg, on s'attend à ce que les fréquences génotypiques de MM, MN, et NN soient respectivement p^2 , $2pq$ et q^2 :

$$p^2 = (0.5425)^2 = 0,2943,$$

$$2pq = 2(0.5425)(0,4575) = 0,4964$$

$$q^2 = (0.4575)^2 = 0,2093$$

Puisque le nombre total d'individus étudiés est égale à 1000, les effectifs attendus des génotypes MM, MN et NN sont:

$$MM = 0,2943 \times 1000 = 294,3$$

$$MN = 0,4964 \times 1000 = 496,4$$

$$NN = 0,2093 \times 1000 = 209,3$$

A cette étape il est pratique de présenter les données dans un tableau, comme suivant:

Tableau 9 : Effectifs des génotypes attendus et observés

Génotypes	Nombre individus observés	Nombre effectifs attendus
MM	298	294,3
MN	489	496,4
NN	213	209,3

Avec cette façon de présenter les données, il est évident que l'adéquation entre les nombres observés et les nombres attendus est très bonne ; elle n'est pas parfaite à cause des fluctuations statistiques aléatoires des effectifs des divers génotypes qui peuvent constituer un échantillon donné.

- Pour vérifier cette conclusion (adéquation ou conformité), nous utiliserons le test statistique classique (**test de chi-deux**) pour quantifier l'ampleur de l'adéquation :
- Le degré de liberté pour ce chi-deux : $ddl = k - r = 3 - 2 = 1$

À partir de la figure ou d'un tableau approprié on lit la valeur de la probabilité correspondante au nombre de degré de liberté (1) pour la valeur de chi-deux.

$p = 0,67 > 0,05$ donc l'adéquation est considérée suffisamment correcte pour que le modèle ne soit pas rejeté. Les fréquences génotypiques sont dans les proportions de Hardy-Weinberg pour ce gène (MN)

Chapitre III. Influence du régime de reproduction : Consanguinité

1- Notion de régime de reproduction

Le régime de reproduction décrit la manière dont les gamètes s'assemblent vis-à-vis du (ou des) gène(s) considéré(s), pour former la génération suivante. Ils sont classés selon deux types :

1.1. Les régimes fermés : ce sont d'une part, la consanguinité, dont la forme la plus extrême est l'autogamie, et d'autre part, l'homogamie. Dans une population un régime fermé donne un taux d'hétérozygotes inférieur à **2pq**.

1.2. Les régimes ouverts : ils sont représentés essentiellement par l'hétérogamie. L'hétérogamie est un régime de reproduction où les unions se font entre individus phénotypiquement dissemblables (on parle parfois aussi d'homogamie négative). Cela entraîne un taux d'hétérozygotes supérieur à 2pq.

Ce régime concerne notamment les gènes sexuels (ou parasexuels). Chez les mammifères, le sexe mâle est déterminé par la formule chromosomique YX, et le sexe femelle par la formule XX.

On considérant X et Y comme des allèles, la fréquence de X est égale à 3/4, celle de Y à 1/4; le taux d'hétérozygotes en panmixie serait de:

$$2pq = 2(1/4)(3/4) = 3/8$$

L'hétérozygotie modifie non seulement la structure génotypique, mais aussi la fréquence d'allèles dans la population. Ainsi chez les mammifères, la fréquence des chromosomes X et Y est maintenue à (3/4 : 1/4). Cette particularité tient à ceci : si pour une raison ou une autre, les individus reproducteurs d'un sexe sont moins fréquents que ceux de l'autre sexe à une génération, ce sont eux qui, en moyenne, se reproduisent le plus. L'hétérogamie entraîne donc une forme spéciale de sélection qui favorise des gènes de la forme plus rare, ce qui les maintient à une fréquence d'équilibre précise.

2- Cas particulier de la consanguinité

2.1. Introduction :

L'interdiction de l'inceste est l'un des rares tabous universels de l'humanité. De l'Amazonie aux grandes métropoles modernes, presque toutes les sociétés ont tracé une frontière morale et légale interdisant les unions au sein du noyau familial strict. Ce phénomène fascinant ne s'explique pas par une seule raison, mais par la convergence de trois forces : la biologie évolutionniste, la psychologie du développement et la nécessité sociologique [10, 11, 12, 13].

L'omniprésence de la consanguinité chez les souverains a modifié l'Histoire à trois niveaux :

- **Une mortalité infantile record** : Dans les familles royales consanguines, le taux de survie des enfants était parfois inférieur à celui des familles paysannes de la même époque. Chez les Habsbourg d'Espagne, la moitié des enfants nés mouraient avant l'âge de 1 an.
- **Le déclenchement de guerres de succession** : L'incapacité biologique à produire un héritier mâle viable (due à la stérilité ou à la mortalité précoce) créait de soudains vides de pouvoir que les pays voisins s'empressaient de contester par les armes.
- **L'instabilité politique** : Des dirigeants physiquement ou mentalement diminués par des maladies génétiques récessives ont souvent dû laisser la régence à des ministres ou à des favoris, déstabilisant la gouvernance des empires.

La consanguinité désigne le résultat d'une reproduction entre deux individus apparentés, c'est-à-dire qui partagent au moins un ancêtre commun récent. Plus les parents sont proches biologiquement (comme des cousins germains, ou un frère et une sœur), plus le degré de consanguinité de leur descendance est élevé.

C'est un concept central en génétique, en biologie de l'évolution, mais aussi en anthropologie en raison des tabous et des lois qui l'entourent dans les sociétés humaines.

2.2. Définition

La consanguinité est un régime de reproduction où les unions se font entre individus apparentés. Deux individus sont dits apparentés si et seulement s'ils ont au moins un ancêtre commun. Un individu est dit consanguin si et seulement si ses deux parents sont apparentés.

2.3. Causes de la consanguinité

La consanguinité n'apparaît jamais par hasard. Qu'elle soit observée chez les humains, les animaux ou les plantes, ses causes se séparent en deux grandes catégories : les facteurs subis (liés à la géographie ou à l'histoire) et les facteurs choisis (liés à la culture, la politique ou la science).

Voici les causes principales qui provoquent des unions consanguines.

2.3.1. L'isolement géographique et les petites populations (L'Endogamie)

Lorsqu'une population est physiquement coupée du reste du monde, le choix des partenaires devient mécaniquement limité. Au fil des générations, tous les membres du groupe finissent par partager des ancêtres communs [14].

- **Les îles et les vallées isolées** : Les communautés vivant sur des îles lointaines ou dans des villages de haute montagne n'ont pas la possibilité d'intégrer du sang neuf facilement.
- **L'effet fondateur** : C'est le cas lorsqu'un très petit groupe d'individus (parfois quelques dizaines de personnes) quitte une population pour coloniser un nouveau territoire. La

diversité génétique de départ est si faible que la consanguinité devient inévitable au bout de quelques décennies.

2.3.2. Les causes culturelles, religieuses et sociales

Chez l'être humain, la consanguinité est souvent le résultat de traditions ou de règles sociales qui encouragent le mariage à l'intérieur d'un même groupe (ce qu'on appelle l'**endogamie**) [15].

- **La préservation du patrimoine et des terres** : Dans de nombreuses sociétés agraires historiques, épouser un cousin germain permettait de ne pas diviser les terres familiales lors des héritages et de garder les richesses au sein du même clan.
- **Les barrières religieuses ou de castes** : Lorsque la loi ou la tradition interdit d'épouser quelqu'un en dehors de sa religion, de son ethnie ou de sa classe sociale, et que ce groupe est minoritaire, le cercle des partenaires se réduit considérablement, augmentant le risque de consanguinité.

2.3.3. Les stratégies politiques et dynastiques

Comme évoqué avec l'exemple des familles royales européennes, la consanguinité a souvent été utilisée comme une arme géopolitique.

Pour sceller des traités de paix, éviter de céder des morceaux de royaume à des étrangers ou maintenir la "pureté" d'une lignée sacrée (comme chez les Pharaons d'Égypte), les mariages incestueux ou très consanguins étaient activement encouragés au sommet de l'État.

2.3.4. Les causes environnementales et biologiques (Chez les animaux)

Dans la nature, les animaux ne choisissent pas la consanguinité par culture, mais y sont contraints par les crises environnementales [16].

- **Le goulot d'étranglement génétique** : Lorsqu'une espèce est massivement chassée, victime d'une épidémie ou que son habitat est détruit, sa population s'effondre. Les quelques survivants n'ont d'autre choix que de se reproduire entre parents proches.
- **La fragmentation des habitats** : La construction de routes, de villes ou de clôtures sépare les forêts en "îlots" isolés. Les animaux sauvages (comme les grands félins ou les ours) se retrouvent coincés dans des zones trop petites et ne peuvent plus rencontrer de partenaires extérieurs à leur groupe d'origine.

2.3.5. La sélection artificielle par l'Homme

Enfin, l'Homme est la cause directe de la consanguinité chez les animaux domestiques et les plantes à travers l'**élevage sélectif** [17].

Pour créer et stabiliser une race de chien (par exemple, obtenir des traits physiques bien précis comme le nez écrasé des bouledogues), les éleveurs font s'accoupler entre eux les individus les plus

proches qui possèdent ce trait. Si cela permet de "fixer" une race ou d'améliorer le rendement d'une variété de blé, cela affaiblit aussi considérablement leur patrimoine génétique.

Tableau 10 : Tableau récapitulatif synthétisant les concepts fondamentaux qui proviennent de plusieurs disciplines.

Type de cause	Exemple type	Moteur principal
Géographique	Îles isolées, tribus reculées	Absence physique de partenaires extérieurs
Sociale / Culturelle	Protection des héritages, traditions	Volonté de garder les biens ou l'identité dans le clan
Politique	Familles royales, dynasties	Consolidation du pouvoir et des alliances
Écologique	Espèces menacées, parcs fragmentés	Destruction de la faune et baisse de population
Scientifique / Agricole	Races de chiens, élevage intensif	Recherche de caractéristiques spécifiques

Source : [16]

2.4. Principaux paramètres mesurant la consanguinité

Les principaux paramètres utilisés pour mesurer la consanguinité sont le coefficient de consanguinité et de parenté :

2.4.1. Coefficient de consanguinité :

La consanguinité ou le coefficient de consanguinité (F_x) d'un individu I est la probabilité que deux gènes homologues soit identique ; la probabilité de trouver à un locus donné deux gènes identiques. Généralement ce coefficient s'applique à un seul individu qui mesure l'impact biologique des unions familiales sur son ADN.

> **Formule** : Le généticien Sewall Wright (1922) a mis au point la formule officielle pour calculer le coefficient de consanguinité d'un individu X :

$$F_X = \sum \left(\frac{1}{2}\right)^{n+1} \cdot (1 + F_A)$$

N : est le nombre de pas qui séparent le père et la mère en remontant vers leur ancêtre commun A.

F_A : est le coefficient de consanguinité de l'ancêtre commun lui-même (si cet ancêtre n'est pas consanguin, $F_A = 0$ et le terme $(1 + F_A)$ devient égal à 1).

Σ : signifie qu'on additionne les résultats si les parents partagent plusieurs ancêtres communs (par exemple, un grand-père et une grand-mère).

> Valeurs de la consanguinité

- $F=0$: aucun ancêtre commun (individu non consanguin)
- $0 < F < 10$: degré variable de consanguinité
- $F=1$: individu complètement homozygote (autofécondation complète)

❖ Exemple : cas d'enfant de cousins germains

Prenons l'enfant de deux cousins germains. Les parents partagent deux grands-parents communs (A1 et A2). On suppose que ces grands-parents ne sont pas consanguins ($F_A = 0$).

> Chemin via le Grand-père (A1) : Pour relier le père et la mère en passant par le grand-père, il y a 4 pas (Père --> Grand-parent --> Oncle/Tante --> Mère). Donc $n = 4$.

$$\left(\frac{1}{2}\right)^{4+1} = \left(\frac{1}{2}\right)^5 = \frac{1}{32}$$

> Chemin via la Grand-mère (A2) : Le chemin est identique, soit $n = 4$.

$$\left(\frac{1}{2}\right)^{4+1} = \left(\frac{1}{2}\right)^5 = \frac{1}{32}$$

> Somme des chemins :

$$F_X = \frac{1}{32} + \frac{1}{32} = \frac{2}{32} = \frac{1}{16} = 0,0625 \text{ (soit 6,25\%)}$$

L'enfant de cousins germains a 6,25% de chances, pour chaque gène, de posséder deux copies strictement identiques héritées de ses grands-parents.

2.4.2. Coefficient de parenté :

Le **coefficient de parenté** (Phi ou ϕ) est une mesure génétique fondamentale. Alors que le *coefficient de consanguinité* s'applique à **un seul individu** (pour savoir à quel point ses gènes sont identiques), le coefficient de parenté s'applique à **deux individus distincts**. Il mesure la probabilité que deux allèles (versions d'un gène), tirés au hasard chez deux individus différents, soient strictement identiques parce qu'ils ont été hérités d'un ancêtre commun.

La parenté ou le coefficient de parenté ϕ_{ij} de deux individus I et j est la probabilité que deux gènes homologues tirés au hasard l'un chez I, l'autre chez J, soit identique ; c'est la probabilité de trouver à un même locus deux gènes identiques.

D'après leurs définitions, le coefficient- de consanguinité d'un individu I est égal au coefficient de parenté de ses deux parents (P et M)

$$f_I = \phi_{PM} = \sum_i \left(\frac{1}{2}\right)^{ni+1} (1+f_{Ai})$$

A un locus donné, il aura donc pu recevoir de ses deux parents des copies d'un même gène de l'ancêtre commun ; un individu consanguin est, par conséquent, plus souvent homozygote (et moins hétérozygote) qu'un individu issu de l'union de deux individus non apparentés. La consanguinité entraîne donc un déficit d'hétérozygotes par rapport à $2pq$. Ce régime de reproduction, comme pour l'autogamie, concerne l'ensemble du génome.

2.5. Conséquences de la consanguinité sur les fréquences génotypiques

La consanguinité a un impact majeur et très précis sur la structure génétique d'une population : **elle modifie radicalement les fréquences génotypiques, mais elle ne change pas les fréquences alléliques** [8].

En clair, cela signifie que la consanguinité redistribue les cartes (les gènes) en détruisant les combinaisons mixtes au profit des combinaisons identiques, sans pour autant faire apparaître ou disparaître de nouveaux gènes.

Dans une population idéale sans consanguinité (loi de Hardy-Weinberg), pour un locus dialléliques (les allèles A et a avec les fréquences respectives p et q), les proportions de la population sont :

- Homozygotes AA : p^2
- Hétérozygotes Aa : $2pq$
- Homozygotes aa : q^2

Lorsque la consanguinité s'installe, elle agit comme un aimant qui sépare les allèles hétérogènes. Les fréquences génotypiques sont alors modifiées selon les formules de Wright :

Tableau 11 : Influence de la consanguinité sur les fréquences génotypiques

Génotype	Fréquence sans consanguinité (F=0)	Fréquence avec consanguinité (F)	Effet de la consanguinité
Homozygotes dominants (AA)	p^2	$p^2 + Fpq$	Augmentation
Hétérozygotes (Aa)	$2pq$	$2pq - 2Fpq = 2pq(1-F)$	Diminution
Homozygotes récessifs (aa)	q^2	$q^2 + Fpq$	Augmentation

Source : [17]

Ce qu'il faut retenir : La fraction $2Fpq$ d'individus hétérozygotes (Aa) "disparaît" et se retrouve divisée équitablement entre les deux groupes d'homozygotes ($+Fpq$ chez les AA et $+Fpq$ chez les aa).

2.5.1. Exemple :

Imaginons une maladie génétique rare causée par l'allèle récessif a, dont sa fréquence dans la population est de 1% ($q = 0,01$). L'allèle sain A a une fréquence de 99% ($p = 0,99$).

Si l'on compare une population normale ($F = 0$) et une population très consanguine issue systématiquement d'unions entre cousins germains ($F = 0,0625$) :

1^{er} Cas : Population sans consanguinité ($F = 0$)

- Fréquence des malades (aa) = $q^2 = (0,01)^2 = 0,0001$ (soit 1 personne sur 10 000).

2^{ème} Cas : Population consanguine ($F = 0,0625$)

- Fréquence des malades (aa) = $q^2 + Fpq$
- Calcul : $0,0001 + [0,0625 \times 0,99 \times 0,01] = 0,0001 + 0,00061875 = 0,00071875$ (soit environ 7,2 personnes sur 10 000).

On constate dans cet exemple que le risque de voir naître des enfants malades (aa) a été multiplié par plus de 7, uniquement à cause de la modification des fréquences génotypiques induite par F.

2.5.2. Conséquences biologiques à long terme

Cette modification des fréquences génotypiques entraîne trois effets biologiques majeurs :

- La mise en lumière des tares cachées : Les allèles récessifs délétères (maladies, malformations) "dorment" généralement chez les individus hétérozygotes (Aa) sans faire de dégâts. La consanguinité, en créant des homozygotes aa, rend ces maladies visibles et actives.
- L'action renforcée de la sélection naturelle : En rendant les gènes défectueux visibles (sous forme aa), la consanguinité permet à la sélection naturelle de les éliminer plus rapidement (si les individus malades meurent ou ne peuvent pas se reproduire). C'est ce qu'on appelle la "purge" du fardeau génétique.
- La dépression consanguine : La baisse globale des hétérozygotes (Aa) réduit la "vigueur" générale de la population. Les hétérozygotes ont souvent un avantage sélectif (meilleur système immunitaire, meilleure adaptabilité). Leur raréfaction affaiblit la population face aux maladies ou aux changements climatiques.

2.6. Conséquences de la consanguinité sur les fréquences alléliques

Contrairement aux fréquences génotypiques qu'elle bouleverse, la consanguinité seule n'a **strictement aucun impact immédiat sur les fréquences alléliques**.

En clair : elle ne crée pas de nouveaux allèles, elle n'en supprime aucun directement, et elle ne modifie pas les proportions globales des gènes présents dans la population. Elle se contente de les **réorganiser**.

Pour comprendre pourquoi les fréquences alléliques (les proportions p et q) restent inchangées après une génération de consanguinité, il suffit de recalculer la fréquence de l'allèle A (notée p') à partir des nouvelles fréquences génotypiques de Wright :

Rappelons que la fréquence d'un allèle s'obtient en prenant tous les homozygotes de cet allèle + la moitié des hétérozygotes.

$$p' = \text{Fréquence}(AA) + \frac{1}{2}\text{Fréquence}(Aa)$$

Remplaçons par les valeurs modifiées par le coefficient de consanguinité F :

$$p' = (p^2 + Fpq) + \frac{1}{2}[2pq(1 - F)]$$

$$p' = p^2 + Fpq + pq(1 - F)$$

$$p' = p^2 + Fpq + pq - Fpq$$

En simplifiant, les termes + Fpq et - Fpq s'annulent :

$$p' = p^2 + pq$$

Si l'on met p en facteur :

$$p' = p(p + q)$$

Comme la somme des fréquences alléliques $p + q = 1$:

$$p' = p \times 1 = p$$

Le résultat est sans appel : La fréquence de l'allèle après l'union consanguine (p') est exactement la même qu'avant (p).

❖ Cas d'un effet indirect à long terme

Si la consanguinité seule ne change pas les fréquences alléliques à la génération 1, elle va indirectement provoquer leur modification à long terme en servant de carburant à deux autres mécanismes évolutifs :

- 1) **La Sélection Naturelle** : Comme la consanguinité augmente le nombre d'homozygotes récessifs (aa), elle expose les défauts génétiques au grand jour [Crnokrak, P., & Barrett, S. C. (2002) ; Hedrick, P. W. (1994)].
 - > Si ces individus atteints d'une maladie récessive meurent avant de se reproduire ou sont stériles, ces allèles défectueux (a) sont éliminés de la population.
 - > À long terme, la fréquence de l'allèle délétère q va donc baisser drastiquement. La consanguinité a permis à la sélection d'épurer le génome de la population.
- 2) **La Dérive Génétique** : La consanguinité survient souvent dans de très petites populations (îles, parcs naturels fragmentés). Dans ces petits groupes, la reproduction consanguine s'accompagne d'un phénomène de loterie génétique appelé la **dérive génétique**. Au fil des générations, certains allèles vont être transmis par pur hasard beaucoup plus souvent que d'autres. À terme, cela conduit inévitablement à la **fixation** d'un allèle ($p = 1$) et à la **perte** définitive de l'autre ($q = 0$), détruisant ainsi la diversité génétique de la population [Wright, S. (1931) ; Kimura, M., & Ohta, T. (1971)].

Chapitre IV. Influence des pressions évolutives

1. Introduction

Si l'on observe un groupe d'êtres vivants à un instant donné – qu'il s'agisse d'une colonie de bactéries, d'un banc de saumons ou d'une population humaine –, on constate l'évidence que tous les individus ne sont pas identiques. Ils partagent des traits communs, mais possèdent des variations de couleur, de taille, de résistance aux maladies ou d'adaptabilité. Cette diversité visible (le phénotype) n'est que le reflet d'une diversité invisible : la composition génétique de la population, c'est-à-dire la nature et la fréquence des allèles qu'elle abrite.

En 1908, les scientifiques Hardy et Weinberg ont démontré une loi théorique fondamentale. Selon eux, dans un monde idéal, la composition génétique de cette population resterait strictement stable d'une génération à l'autre. C'est l'état de stase, ou l'équilibre de Hardy-Weinberg.

Pourtant, dans la nature, cet équilibre n'existe jamais. Le vivant est en perpétuel mouvement. Les populations changent, s'adaptent, s'appauvrissent ou se diversifient. Pourquoi ? Parce qu'elles sont soumises en permanence à des forces de rupture que l'on appelle les pressions évolutives.

Dans cette partie, nous allons voir que la composition génétique d'une population résulte toujours d'un compromis entre plusieurs forces. Une seule pression est rarement maîtresse : ce sont leurs interactions qui produisent la diversité du vivant et la différenciation entre populations.

2. Présentation des pressions évolutives

Les pressions évolutives sont au nombre de quatre. Elles sont qualifiées de forces motrices du vivant qui agissent comme des modeleurs sur le pool de gènes de la population :

2.1. Les mutations ; source de variabilité

La variabilité génétique est le résultat des mutations qui font apparaître de nouveaux allèles, auxquelles il faut ajouter les phénomènes de recombinaison (notamment pour les caractères quantitatifs). Les mutations peuvent affecter une portion plus ou moins grande d'ADN et, en fonction de leur localisation dans le génome, peuvent avoir ou non des effets phénotypiques. Il existe ainsi tous les intermédiaires entre les mutations neutres qui n'ont aucun effet sur l'organisme et les mutations létales, qui réduisent l'espérance de vie des individus.

2.1.1. Définition : Une altération permanente et (généralement) transmissible de la séquence nucléotidique de l'ADN d'un gène ou d'un chromosome. C'est la **source ultime de toute nouveauté génétique**.

2.1.2. Les causes des mutations : Elles sont décrites de la manière suivante :

❖ Mutations spontanées :

- **Erreurs de réplication :** L'ADN polymérase (l'enzyme qui copie l'ADN) commet une erreur rare non corrigée.

- **Lésions chimiques** : Instabilité chimique des bases (ex. désamination : une cytosine se transforme en uracile).

❖ **Mutations induites (par des mutagènes) :**

- **Agents physiques** : Rayons UV (crée des liaisons entre thymine adjacentes), rayons ionisants (X, gamma) qui cassent l'ADN.
- **Agents chimiques** : Composés qui altèrent la structure des bases (ex. nitrites, benzopyrène de la fumée de cigarette).
- **Agents biologiques** : Certains virus peuvent intégrer leur ADN dans le nôtre.

2.1.3. Les différents types de mutations : Il existe différents types moléculaires de mutations qui n'ont pas les mêmes conséquences phénotypiques. Elles sont classées selon l'échelle et l'impact sur la séquence

❖ **Mutation à l'échelle moléculaire** : Elle s'effectue au niveau de l'ADN selon les manières suivantes :

- **Substitution (ou mutation ponctuelle)** : Remplacement d'une base par une autre.
- **Transition** : Purine → Purine ($A \leftrightarrow G$) ou Pyrimidine → Pyrimidine ($C \leftrightarrow T$).
- **Transversion** : Purine ↔ Pyrimidine.

❖ **Mutation à l'échelle structurale** : Elle s'effectue au niveau des chromosomes selon les manières suivantes :

- **Duplication** : Une portion de chromosome est dupliquée.
- **Délétion** : Une portion est perdue.
- **Inversion** : Une portion est détachée et remise à l'envers.
- **Translocation** : Échange de portions entre chromosomes non homologues.
- **Réciproque** : Échange mutuel.
- **Non réciproque** (ou de type "Philadelphia", impliquée dans certaines leucémies).

❖ **Mutation à l'échelle génomique** : Elle s'effectue au niveau du nombre de chromosomes selon les manières suivantes :

- **Aneuploïdie** : Perte ou gain d'un ou plusieurs chromosomes.
Ex. : Trisomie 21 (3 copies du chromosome 21), Syndrome de Turner (monosomie X : 45,X).
- **Polyploïdie** : Multiplication de l'ensemble complet des chromosomes ($3n, 4n...$). Fréquent chez les plantes.

2.1.4. Impact des taux de mutation sur les fréquences alléliques

Dans la nature, la modification spontanée de l'ADN est un événement rare. On mesure la pression de mutation à l'aide de taux, généralement notés :

- u : le taux de mutation de l'allèle sain vers l'allèle muté ($A \rightarrow a$).
- v : le taux de mutation retour ($a \rightarrow A$).

Chez la plupart des espèces (des bactéries à l'être humain), ces taux oscillent entre 10^{-5} et 10^{-8} par gène et par génération [Haldane, J. B. S. (1927) ; Crow, J. F., & Kimura, M. (1970)].

2.1.5. Vitesse du processus

La vitesse à laquelle la mutation modifie la composition génétique d'une population est considérée comme **le processus le plus lent de toute la biologie de l'évolution**.

Pour donner une échelle temporelle à vos étudiants, on ne parle pas en années ni en siècles, mais en **dizaines de milliers, voire en millions de générations**.

Voici les éléments théoriques et mathématiques pour analyser cette cinétique (vitesse de changement) dans votre cours.

Pour mesurer la vitesse de ce processus, on cherche à calculer le temps t (en nombre de générations) nécessaire pour passer d'une fréquence initiale p_0 à une fréquence finale p_t . En transformant l'équation avec le logarithme népérien ($\ln$), on obtient la formule de la cinétique :

$$t = \frac{1}{u} \ln \left(\frac{p_0}{p_t} \right)$$

2.2. Migration ou Flux Géniques

Le flux génique désigne l'apport de gènes par l'arrivée d'individus provenant d'une autre population.

- **Effet homogénéisateur** : Il contrecarre les effets de la dérive génétique et de la sélection locale. Si le flux est continu, il fusionne la composition génétique des deux populations pour n'en former qu'une seule grande et unique.
- **Source d'allèles** : Pour une population locale isolée, l'arrivée d'un migrant équivaut à une "mutation" à grande échelle : elle réintroduit instantanément des allèles qui avaient été perdus par dérive ou par sélection.

2.3. Sélection

Dans la nature, les individus sont soumis à différentes conditions environnementales qui modulent le nombre de leurs descendants. De très nombreux caractères (pelage, couleur, taux d'hormones, ...) sont impliqués dans ce processus. Pour les caractères possédant un déterminisme

génétique, différents phénotypes (correspondant à différents génotypes) sont présents dans la population.

Dans un environnement donné, si les individus porteurs de ces phénotypes se reproduisent mieux que d'autres, ils auront tendance à transmettre plus fréquemment les gènes dont ils sont porteurs. C'est l'idée de base présentée par Darwin et Wallace (1858), qui proposent que les facteurs héréditaires avantageant certains individus soient ainsi mieux transmis au cours des générations. Ils associent l'idée d'adaptation à celle de plus forte reproduction. Il s'agit du processus de la **Sélection Naturelle**.

Selon les caractères favorisés dans une population, la sélection naturelle peut produire des résultats complètement différents. Elle peut :

- Préserver les caractéristiques d'une population en favorisant les individus moyens;
- Changer les caractéristiques d'une population en favorisant les individus qui varient dans une direction par rapport à la moyenne de la population;
- Changer les caractéristiques d'une population en favorisant les individus qui varient dans les deux directions.

La sélection naturelle est la seule pression évolutive qui **oriente** les changements génétiques en fonction de l'environnement. Elle élimine ou propage les gènes selon la valeur sélective (*fitness*) qu'ils confèrent à l'individu. Elle modifie la composition génétique de trois manières différentes :

Tableau 12 : Les modalités de la sélection naturelle et leurs dynamiques génétiques

Type de Sélection	Action sur la population	Impact sur la composition génétique
Directionnelle	Favorise un phénotype extrême (ex: résistance à un antibiotique).	Baisse de la diversité : Un allèle avantageux remplace progressivement tous les autres jusqu'à sa fixation ($p=1$).
Divergente (Disruptive)	Favorise les deux extrêmes au détriment des individus moyens.	Séparation en deux blocs : La population se divise génétiquement, ce qui peut mener à la création de deux espèces.
Stabilisante (Équilibrante)	Élimine les extrêmes et favorise les hétérozygotes (ex: drépanocytose/paludisme).	Maintien de la diversité : Les différents allèles sont activement conservés ensemble dans la population.

Source : [19 ; 20]

2.3.1. Coefficient de sélection

Le **coefficient de sélection** (noté s) mesure l'intensité de la sélection naturelle **contre** un génotype. Il représente la **réduction relative de la valeur sélective (fitness)** d'un génotype par rapport au génotype le plus avantage.

Plus précisément, si un génotype a une fitness W , et que le génotype de référence a une fitness $W_{\max}=1$, alors : $W=1-s$ où s est le coefficient de sélection.

2.3.2. Différentes formes de sélection naturelles

La sélection naturelle n'agit pas toujours de la même manière. Selon la relation entre le génotype et le fitness, on distingue plusieurs **formes de sélection** qui ont des conséquences très différentes sur la composition génétique des populations. Les **formes fondamentales selon l'effet sur le phénotype** sont les suivants :

- ❖ **Sélection directionnelle** : La sélection favorise un extrême du phénotype. La moyenne de la population se déplace dans une direction.
- ❖ **Sélection stabilisante** : La sélection favorise les phénotypes intermédiaires et élimine les extrêmes. La moyenne reste stable, mais la variance diminue.
- ❖ **Sélection disruptive (diversifiante)** : La sélection favorise les **deux extrêmes** du phénotype et défavorise l'intermédiaire. La distribution devient bimodale.

REFERENCES BIBLIOGRAPHIQUES

1. **Henry, Jean-Pierre, & Gouyon, Pierre-Henri. (2023).** *Précis de génétique des populations*. Éditions Dunod.
2. **Serre, J.-L. (2006).** *Génétique des populations : cours et exercices corrigés*. Dunod, collection Sciences sup. (présentation des concepts de base et de la mesure de la variabilité génétique).
3. **De Meeûs, T. (2015).** *Initiation à la génétique des populations naturelles*. IRD Éditions. (approche pratique et didactique).
4. **Fisher, R.A. (1930).** *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford. (ouvrage fondateur, réconcilie Mendel et Darwin).
5. **Haldane, J.B.S. (1932).** *The Causes of Evolution*. Longmans, Green and Co. (synthèse des travaux de Haldane sur l'évolution).
6. **Wright, S. (1931).** « Evolution in Mendelian Populations ». *Genetics*, 16(2), 97-159. (article majeur introduisant la dérive génétique)
7. **Griffiths, A. J., Wessler, S. R., Carroll, S. B., & Doebley, J. (2015).** Introduction à l'analyse génétique (6e éd.). De Boeck Supérieur.
8. **Hardy, G. H. (1908).** *Mendelian proportions in a mixed population*. *Science*, 28(706), 49-50.
9. **Race, R. R., & Sanger, R. (1975).** *Blood groups in man* (6e éd.). Blackwell Scientific Publications.
10. **Lévi-Strauss, Claude. (1949).** *Les Structures élémentaires de la parenté*. Paris : Mouton & Co.
11. **Lieberman, Debra, TOOBY, John, & Cosmides, Leda. (2007).** "The architecture of human kin detection." *Nature*, 445(7129), 730-733. (*Une étude moderne en psychologie évolutionniste analysant les mécanismes cognitifs humains de détection de la parenté pour éviter l'inceste*).
12. **Bittles, Alan H. (2012).** *Consanguinity in Context*. Cambridge University Press.
13. **Ember, Melvin. (1975).** "On the Origin and Extension of the Incest Taboo". *Behavior Science Research*, 10(4), 249-281.
14. **Mayr, E. (1942).** *Systematics and the Origin of Species*. Columbia University Press.
15. **Ourdieu, P. (1972).** *Les stratégies matrimoniales dans le système des stratégies de reproduction*. *Annales. Histoire, Sciences Sociales*, 27(4), 1105-1127.
16. **Soulé, M. E. (1980).** *Thresholds for survival: maintaining fitness and evolutionary potential*. *Conservation Biology: An Evolutionary-Ecological Perspective*, 151-169.
17. **Wright, S. (1921).** *Systems of mating*. *Genetics*, 6(2), 111-178.
18. **Wright, S. (1921).** *Systems of mating. II. The effects of inbreeding on a Mendelian population*. *Genetics*, 6(2), 124-143.

19. **Samadi, Sarah et DAVID, Patrice.** (2021). *La théorie de l'évolution : Une logique pour la biologie*. Éditions Flammarion.
20. **Futuyma, Douglas J., & Kirkpatrick, Mark.** (2017). *Evolution* (4th ed.). Sinauer Associates. (*Le manuel de biologie évolutive le plus utilisé au monde, détaillant parfaitement les graphiques de distribution phénotypique pour chaque type de sélection*).