

COURS DE GENETIQUE DES POPULATIONS

Préparé par OURARI Malika

**Adressé aux étudiants de troisième année licence et de première année de
Master du système LMD en biologie et aux étudiants en médecine.**

PREAMBULE

Public cible : Ce polycopié de cours de génétique des populations naturelles est destiné aux étudiants de troisième année licence et de première année de Master du système LMD en biologie, aux étudiants en médecine et aux étudiants des filières concernées par l'étude des populations naturelles eucaryotes.

Volume horaire globale : 45 heures de cours 20 heures de travaux dirigés

Prérequis : Connaissances de base de biologie générale (division cellulaire, mitose, méiose, modes de reproduction), notion d'écologie (biogéographie, habitat, bioclimat), notions de base de génétique (modes de transmission des caractères héréditaires, lois de Mendel, concepts de gènes, allèles, génotypes, phénotypes), de biologie moléculaire (structure et fonction de l'ADN, de l'ARN, protéines, replication traduction etc.) et les concepts de bases en statistiques et en probabilités.

Objectifs du cours : A l'issue de ce cours l'étudiant doit être en mesure de :

- Calculer les fréquences alléliques et génotypiques dans les populations naturelles en équilibre et dans les populations en déséquilibre en fonction des modes de croisement (panmixie, consanguinité, homogamie) et en fonction des forces évolutives (mécanismes qui modifient la structure génétique des population : mutations, sélection, dérive génétique, migration).
- Identifier les différents types de sélection et de spéciation.
- Modéliser les processus génétiques : utiliser des modèles mathématiques pour prédire des changements dans les fréquences génétiques des populations.
- Etudier la variabilité génétique en explorant les causes et les conséquences de la variabilité dans les populations qui interviennent dans l'adaptation, la spéciation et la survie des populations.
- Appliquer la génétique des populations en médecine (résistances aux antibiotiques), en agronomie (résistance en insecticides, herbicides).

GENETIQUE DES POPULATIONS

CONTENU DU COURS

1. VARIATIONS DES POPULATIONS NATURELLES

1.1. Nature de la variation et notion de caractère : morphologique, anatomique, histologique, caryologique, physiologique, écologique, biochimique et moléculaires.

1.2. Les composantes écologique et génétique de la variation : écotype, écophène, écocline.

1.3. Origine de la variation génétique : mutations, recombinaisons, polyploïdie, hybridation et introgression.

2. STRUCTURE GENETIQUE, EVOLUTION ET SPECIATION

2.1. Concepts, définition et caractères généraux de la population : Dimensions, isolement, mode de reproduction.

2.2. Panmixie et équilibre de Hardy et Weinberg

- Données théoriques, fréquences alléliques et fréquences génotypiques
- les écarts à la panmixie : consanguinité et homogamie

2.2. Les populations en déséquilibres et la sélection naturelle :

- mutations, migration et flux génétiques
- valeur adaptatives et coefficients de sélection.
- Polymorphisme et avantages alternés des homozygotes et des hétérozygotes.
- Dérive génétique et différents types de sélection.

2.3. La spéciation.

- Concepts et définitions de l'espèce.
- Mécanismes d'isolement reproducteur.
- Spéciation allopatrique et spéciation sympatrique

INTRODUCTION

La génétique des populations est une discipline qui est née de la synthèse des théories de Mendel, de Darwin et des biométriciens du début du XXème siècle (Fisher, Wright, Haldane). Ils ont introduit les premières méthodes d'analyse (de mesure) de la variabilité génétique des populations naturelles. L'analyse de cette variabilité génétique dans et entre les populations nous permet de retracer leur histoire évolutive. La génétique des populations explique les modifications observées dans les populations en fonction du temps et de l'espace en étudiant les forces évolutives (reproduction sexuée, mutation, sélection naturelle, migration, dérive génétique..). Elle permet d'expliquer, voire de prédire l'évolution des populations.

RAPPELS

La génétique est l'étude du mode de transmission des **caractères héréditaires**. Les caractères héréditaires sont gouvernés par des éléments dénommés **gènes** arrangés linéairement sur un support : le **chromosome**. Un gène est une séquence d'ADN qui permet d'assurer la synthèse d'un **polypeptide**. Chaque gène sur un chromosome, occupe un emplacement défini appelé **locus**. Chaque gène se trouve en double exemplaire chez les diploïdes. Chaque exemplaire constitue un **allèle**. Un gène peut être représenté par deux allèles (diallèle) ou plusieurs allèles (polyallèles, ex. des groupes sanguins A, B, AB, O). Un allèle peut être soit **dominant** soit **récessif**. Selon Mendel, l'allèle dominant est l'allèle qui s'exprime chez l'hybride de première génération, l'autre étant l'allèle récessif. Pour un caractère donné, si les deux allèles du couple sont identiques, l'individu est dit **homozygote**. Par contre, si les deux allèles ne sont pas identiques, l'individu est dit **hétérozygote**. Un gène contient une information codée permettant d'assurer la synthèse d'un **polypeptide**. Toutes les formes alléliques d'un gène se trouvent à des positions semblables sur des chromosomes **homologues** (sauf dans le cas de translocations). Les chromosomes homologues sont des chromosomes portant les mêmes loci. Ils sont réunis par paires appelés **bivalents**, lors de la prophase I de la méiose. Chaque bivalent est constitué d'un chromosome d'origine maternelle et d'un chromosome d'origine paternelle. La combinaison des deux allèles situés face à face sur les chromosomes homologues forme un **génotype**.

La **génétique mendélienne** (1865) étudie les mécanismes héréditaires en observant les dissemblances et les ressemblances (caractères phénotypiques) chez les parents et chez la descendance. La **génétique moléculaire** analyse les mécanismes qui mènent du gène vers le caractère. Elle étudie la biochimie des acides nucléiques (réplication transcription, traduction, enzymes de restriction, séquençage...).

Sur un troisième plan, la **génétique des populations** étudie **statistiquement** les fréquences génotypiques et des fréquences alléliques dans les populations. Une population est définie comme étant 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.

OBJECTIFS DE LA GENETIQUE DES POPULATIONS

L'individu détient pour une **courte période de temps** une **petite partie du pool de gènes** ; un pool de gènes étant l'ensemble de gènes et de leurs différents allèles présents dans une population. La contribution d'un individu dans les phénomènes d'adaptation, d'évolution et de spéciation est donc minime par rapport à l'ensemble du pool de gènes de la population. C'est dans la population que s'organisent les gènes en combinaisons multiples (en différents génotypes) grâce à la reproduction sexuée (ségrégation, recombinaison, crossing over). C'est également, dans la population que les gènes peuvent se transformer par mutations. Cette variabilité génétique est à l'origine de la pérennité des populations naturelles en leur permettant de s'adapter à un environnement constamment changeant, de résister aux parasites et aux nouvelles maladies. **La population a donc deux caractéristiques qui sont absentes chez l'individu : la continuité dans le temps et l'aptitude au changement.** De ce fait, la population est considérée comme une unité de base d'adaptation, d'évolution, et de spéciation (formation de nouvelles espèces). L'évolution est définie par le généticien Dobszhansky comme toute variation dans les fréquences alléliques et/ ou génotypiques dans les populations.

La génétique des populations a pour objectifs :

- **d'identifier la variabilité génétique au sein des populations naturelles en estimant les fréquences alléliques et les fréquences génotypiques à l'aide de modèles mathématiques et statistiques ;**

- d'évaluer la variation de ces fréquences entre les populations et au cours du temps
- de relever l'influence des régimes de reproduction (panmixie, consanguinité, autogamie, homogamie) et des forces évolutives (mutations, dérive génétique, sélection naturelles, migration) dans la variation des fréquences génotypiques et des fréquences alléliques dans l'espace (selon les populations) et dans le temps (en fonction des générations successives).

DOMAINES D'APPLICATION DE LA GENETIQUE DES POPULATIONS

La génétique des populations trouve un intérêt

- Dans le domaine des sciences fondamentales, en permettant d'expliquer les mécanismes de l'évolution biologique et de retracer l'histoire évolutive des espèces.
- dans le domaine de l'écologie en permettant de montrer l'impact sur l'environnement des espèces invasives lors de l'introduction d'espèces dans des milieux nouveaux, dans les études de la conservation de la biodiversité, dans l'étude des OGM (dissémination, impact sur l'environnement), dans les programmes d'aménagement des pêches et des chasses ;
- en médecine, dans la recherche des tares génétiques, en épidémiologie, dans l'étude des effets de consanguinité ;
- dans le domaine agronomique dans l'amélioration génétique animale et végétale, par l'utilisation de croisements consanguins ou homogamiques et de la sélection artificielle.

CHAPITRE I. LA VARIABILITE GENETIQUE DANS LES POPULATIONS NATURELLES

1. Introduction
2. Notion de caractère
3. Principaux types de variation
4. composante écologique et génétique de la variation
5. origine de la variation génétique
6. Détection de la variation : notion de polymorphisme
7. Quantification du polymorphisme

1. Introduction

Une particularité du monde vivant est la variabilité des phénotypes des individus. A l'intérieur d'une espèce sexuée, il n'existe pas deux individus ayant exactement les mêmes caractéristiques phénotypiques même les jumeaux monozygotes : **l'individu est unique**. Certaines de ces variations s'expriment au niveau phénotypique (morphologie, physiologie, comportement) mais d'autres peuvent être mis en évidence par l'utilisation de techniques adaptées (variabilité des protéines ou des séquences d'ADN).

2. Notion de caractère

Selon Bidault (1968), un caractère est l'ensemble ou une partie des traits ou propriétés d'un individu susceptibles d'être soumis à une **mesure** ou de recevoir un **qualificatif** permettant des comparaisons avec les mêmes traits ou propriétés d'un autre individu. La mesure ou la qualification d'un caractère permet d'indiquer les **valeurs** ou les **états**. Pour les grains de pollen d'une plante par exemple, le diamètre, la forme et la couleur sont des caractères. Le caractère 'diamètre' présente une certaine valeur alors que la 'forme' s'observe sous un état particulier.

Pendant longtemps, les seuls caractères qui ont retenu l'attention des analystes de la variabilité sont ceux relatifs à la morphologie c'est à dire ceux qui concernent la forme et les dimensions des différentes parties d'un individu. Actuellement, l'analyse de la variation porte sur un très grand nombre de caractères de nature diverse.

- Caractères morphologiques

La plupart des caractères morphologiques qui étaient autrefois utilisés sous forme qualitative ont de plus en plus tendance à être exprimés quantitativement. C'est ainsi, par exemple, que les différentes parties (feuille, pétales, sépales) des organismes sont les plus fréquemment exprimés par des **rapports numériques** au lieu d'être mentionnés en termes subjectifs comme lancéolée, étroitement lancéolée etc.

- Caractères anatomiques et histologiques

Chez les plantes, parmi les caractères anatomiques les plus fréquemment usités, on peut citer les structures des feuilles et des tiges et celles de certains épidermes ainsi que les particularités des stomates. La structure des tiges (particularité de l'endoderme) distingue les différentes espèces d'*Equisetum*. Dans le genre *Festuca* (F. Graminées), l'anatomie foliaire permet de distinguer de nombreux taxons de rang spécifique ou infraspécifique. La structure des épidermes caractérise différentes tribus des graminées par la forme et la disposition des éléments épidermiques. Les dimensions des stomates distinguent des unités infraspécifiques chez *Poa annua* L.

- Caractères cytologiques (caryologiques)

Ils concernent le nombre, la forme et la structure interne du chromosome. C'est ce qu'on nomme le **caryotype**. L'étude morphologique est définie par la position du centromère (terminal, subterminal, médian), l'**indice centromérique**, par la longueur des bras et par la présence éventuelle des constriction secondaires (satellites). Tenant compte de la variation de ces caractères cités, on établit à partir d'une moyenne des caryotypes un **caryogramme**. Un caryogramme est une représentation schématique du caryotype moyen établi à partir de plusieurs cellules. L'analyse des caryotypes est reconnue avoir une grande importance dans l'explication du phénomène de l'évolution et celle de la délimitation des espèces. Ex. 1 : la sous famille des *Aveneae* et des *Hordeae* des graminées ont un nombre chromosomique de base égal à 7 et des chromosomes longs alors que d'autres sous familles comme les *Paniceae* ont un nombre de base de 9 ou 10 et des chromosomes courts. Ex. 2 : Le caryotype de l'homme à 46 chromosomes est obtenu après remaniement du caryotype à 48 chromosomes du chimpanzé.

- Caractères physiologiques et écologiques

Ces caractères se traduisent par des adaptations écologiques particulières liées à la capacité photosynthétique, à la résistance au froid, à la chaleur, à la sécheresse ou à différents

constituants du sol. Ex. *Festuca glauca* (Graminées) comprend deux variétés morphologiquement voisines dont l'une, tétraploïde, se trouve sur des sols calcaires, l'autre diploïde, se localise dans les sols siliceux (Bidault, 1968).

- Caractères biochimiques

Du point de vue technique la chimiotaxonomie est relativement récente. Certaines molécules telles que les alcaloïdes, les protéines, les enzymes, les polysaccharides, les anthocyanes sont utilisées par chromatographie ou électrophorèse. Cependant, c'est depuis le progrès de la biologie moléculaire, que l'on peut constituer une base de classification naturelle (qui reflète l'arbre phylogénétique de l'espèce). Les séquences d'acides aminés dans les protéines ou les séquences de nucléotides dans l'ADN ou l'ARN fournissent des données discontinues (comparaison de sites) alors que le **taux d'hybridation** plus ou moins élevé entre les brins d'ADN appartenant à des groupes différents, **les fréquences alléliques** s'expriment sous forme quantitative.

Des méthodes statistiques avec l'utilisation de l'informatique permettent de traiter simultanément un grand nombre de caractères (analyses multivariées) et de mieux cerner la variabilité des populations. Diverses méthodes d'analyse de données sont communément utilisées (Analyse en Composantes Principales ACP, Analyse Factorielle des Correspondances AFC, Classification Automatique Hiérarchique CAH).

3. Principaux types de variation

Selon le niveau d'étude on distingue la variation individuelle intra-population et inter-population.

3.1. Variation individuelle

Elle affecte les différentes parties d'un individu à un moment donné ou les mêmes parties à des moments différents.

- **Variation suivant l'âge** : c'est la différence entre les stades immatures ou les stades larvaires et adultes (chenilles, papillons). Selon l'âge physiologique, nous observons un polymorphisme foliaire chez le genre *Plantago*.

- **Variation saisonnière** : les mammifères des régions tempérées et froides sont capables de transformer leur pelage en fourrure d'hivers. De nombreux oiseaux ont, durant une partie de l'année un plumage obscur qui est remplacé par un plumage nuptial brillant avant le début de la saison des amours.

- **Variation des structures** des épidermes des graminées en fonction de la partie de la plante ou d'organe considéré. De nombreux caractères floraux de *Plantago coronopus* varient d'une façon nette en fonction de leur position le long de l'épi.

3.2. Variation à l'intérieur d'une population et entre les populations

Les approches biologiques modernes reconnaissent l'importance de la variation intra spécifique et cherchent à en déterminer la signification évolutive. L'étude de la variation intra et inter populations permet d'expliquer les mécanismes de spéciation. Les avantages que présente une réserve de variation sont :

- **Plus le nombre de types génétiques d'une population est important, plus celle-ci aura la possibilité de posséder des génotypes capables de survivre aux changements saisonniers et autres transformations temporelles et particulièrement à celles d'un environnement hostile. S'il existe dans une population vivant normalement dans un environnement humide des génotypes résistants à la sécheresse, cette population aura des chances de survivre au cours d'une période de sécheresse.**

- **Cette variation autorise également une plus grande utilisation du milieu car elle permet la colonisation d'habitats marginaux.**

L'étude de la variabilité inter et intra population nécessite de procéder d'abord à un échantillonnage adapté par lequel, à l'aide de méthodes statistiques (loi de distribution, moyenne, mode, écart-type, coefficient de corrélation), les propriétés de l'ensemble peuvent être déduites. L'échantillonnage doit porter sur toute l'aire de distribution. L'analyse doit se faire sur des caractères homologues au même stade phénologique sur des échantillons de terrain et sur leurs échantillons transplantés dans les mêmes conditions de culture (pour éliminer la part de la variation génétique).

4. Composantes écologique et génétique de la variation

4.1. Composante génétique : notions d'écotypes et d'écoclines

Les différents facteurs de l'environnement agissent fortement, par l'intermédiaire de la sélection naturelle, sur la variabilité génétique des populations. Le milieu sélectionne les individus qui lui sont génétiquement les mieux adaptés. Les populations, si elles prospèrent dans des conditions écologiques différentes, peuvent se différencier en plusieurs sous unités adaptées chacune à ses propres conditions stationnelles. Turesson a appelé **écotype** ces populations dont les caractéristiques héréditaires sont le produit de la réaction entre le génotype et le milieu.

La variation écotypique est plus fréquemment **continue** que **discontinue** et cette continuité est en rapport avec la variation graduelle de l'environnement. Ce résultat a conduit Gregor et al. à adopter le concept d'**écocline** pour qualifier cette **variation écotypique continue**. Exp. Dobzhansky (1948) montre que la fréquence de certaines variations dans la structure des chromosomes augmente régulièrement en fonction de l'altitude.

4.2. Composante écologique : notion d'écophène

Le milieu peut agir directement sur le phénotype d'un individu qui développe des modifications non héréditaires appelées **accommodations**. Les populations obtenues sont des **écophènes**. **Exp. 1.** Deux plantes dont l'une pousse dans de bonnes conditions et l'autre sur un sol pauvre. **Exp. 2.** Une souris élevée à 28 °C. a une queue qui mesure 93.1 mm et à 6°C, elle diminue à 75.9mm. **Exp. 3.** L'homme de la plaine qui va vivre en montagne voit croître le nombre de ses hématies : la diminution de la quantité d'O₂ consécutive à la diminution de la pression atmosphérique est compensée par la fabrication plus intense d'hématies mais le phénomène disparaît lors du retour en plaine.

On appelle **homéostasie**, l'aptitude d'un individu à changer son phénotype de façon qu'il puisse survivre et s'accommoder aux variations du milieu.

5. Origine de la variation génétique

Plusieurs facteurs sont à l'origine de la variabilité génétique : reproduction sexuée (ségrégation, recombinaison), mutations, sélection...

5.1. Recombinaison et ségrégation

Au cours de la méiose, les chromosomes homologues paternels et maternels s'apparient et se recombinent (crossing-over) ensuite ils se disjoignent (ségrégation) de façon aléatoire dans des cellules reproductrices haploïdes (spores ou gamètes).

5.2. Mutations

Ce sont des changements brusques et héréditaires qui affectent une partie ou l'ensemble du génotype. On distingue les mutations géniques et les mutations chromosomiques.

- Les mutations géniques

On parle de mutation génique lorsqu'elle affecte la séquence du gène. Si un faible nombre de nucléotides est touché, il s'agit d'une **mutation ponctuelle**. Ce sont des changements localisés à l'intérieur d'un cistron (unité fonctionnelle d'un gène).

- Mutations chromosomiques

On distingue les changements structuraux des chromosomes et les changements numériques des chromosomes.

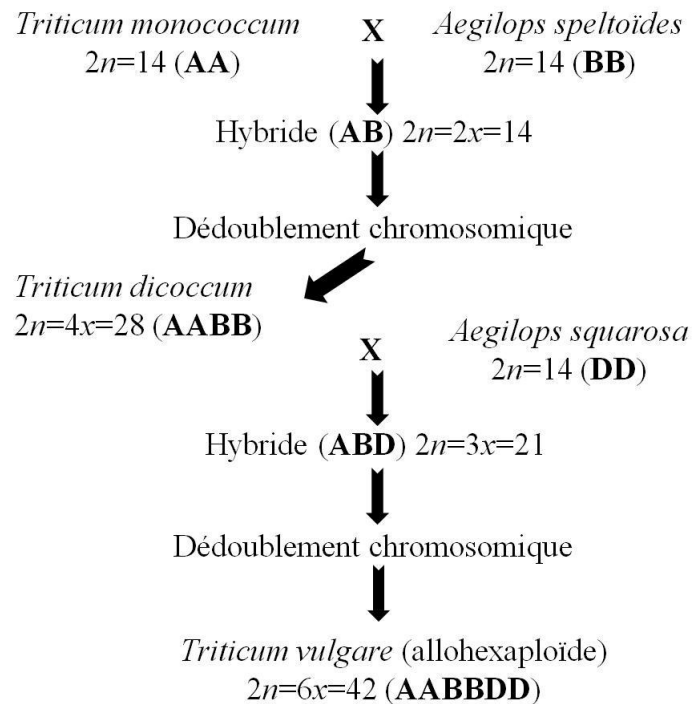
Parmi les **changements structuraux** des chromosomes on peut citer : la **fragmentation** (par rupture d'un chromosome), les **délétions**, (perte d'un segment de chromosome), les **duplications** (ajout d'un segment de chromosome à un chromosome originel), les **inversions**, les **translocations**.

Pour les **changements numériques des chromosomes**, nous constatons que chez un grand nombre d'organismes, le nombre chromosomique est, en général, constant. Cependant, certains organismes appartenant à des mêmes unités systématiques ont parfois des nombres chromosomiques variés. Exp. chez *Festuca ovina*, on observe des formes ayant 14, 28, 42, 49, 56 et 70 chromosomes. Pour ces nombres, le **plus petit multiple commun est 7**. Il représente le **nombre de base** de l'espèce. Il est symbolisé par la lettre '**x**'. le nombre le plus faible de la série représente deux fois le nombre de base. C'est le nombre diploïde de l'espèce. Le nombre haploïde de chromosomes contenus dans les gamètes est représenté par la lettre '**n**'. Le stock chromosomique de l'espèce diploïde sera **$2n=2x=14$** . Les autres taxons dont les nombres chromosomiques sont des multiples de 7 sont qualifiés de **polyploïdes**. Ils sont des tétraploïdes

lorsque $2n=4x=28$, **hexaploïdes** lorsque $2n=6x=42$, **octaploïdes** lorsque $2n=8x=56$, etc. La polyploïdie est plus fréquente chez les plantes. Les polyploïdes sont relativement courant dans certains groupes de vertébrés (amphibiens, reptiles). Dans la majorité des cas c'est une polyploïdie paire ($4x$, $6x$). Il a été mis en évidence l'existence d'une espèce de crapaud triploïde Batura (*Bufo pseudoraddei baturae*) qui vit dans les montagnes du Pakistan qui se reproduit d'une façon sexuée. La méiose produit alors des gamètes mâles à ' x ' chromosomes après réduction de cellules mères de $3x$ à $2x$ et des gamètes femelles à $2x$ chromosomes après augmentation des cellules mères de $3x$ à $4x$. La fécondation aboutit alors à la formation d'individus à $3x$ ($2x \times x$). Les nombres chromosomiques peuvent également varier par la présence ou l'absence d'un ou de plusieurs chromosomes. Ce qui définit l'**aneuploïdie** par opposition à l'**euploïdie** envisagée précédemment. Lorsqu'il y a perte d'un chromosome, on parle de **monosomie** ($2n-1$). Il y a **polysomie**, quand un chromosome du lot normal est présent à plus de deux exemplaires. Les **trisomiques** ont $2n+1$, les **tétrasomiques** ont $2n+2$ etc. chez les plantes, des chromosomes surnuméraires peuvent s'observer. On les désigne par les **chromosomes 'B'** par opposition aux **chromosomes 'A'** qui forment la garniture chromosomique normale. Leur origine est inconnue. Ils sont petits par rapport aux chromosomes A. ils ne présentent pas d'homologie avec les chromosomes normaux et ne s'apparient pas avec eux au moment de la méiose. Beaucoup d'espèces ont également plusieurs nombres de base et peuvent présenter plusieurs **séries polyploïdes** (*Centaurea* L., *Crepis* L.)

- Hybridation, introgression

Une hybridation est un croisement entre deux individus différents **génétiquement** et **taxonomiquement**. Elle permet, donc, de mettre en contact des génomes différents. Elle constitue une source de variabilité pour les populations naturelles. Elle est fréquente chez les plantes. Elle s'accompagne, en général, de la réduction de la fertilité. Exp. le croisement entre l'âne et le cheval donne le mulet qui est vigoureux mais stérile. Chez les plantes, dans le cas d'hybrides interspécifiques stériles, le passage de la forme diploïde à la forme polyploïde permet l'obtention d'**allopyploïdes** fertiles (polyploïdes après hybridation). Exp. histoire des blés actuels.



Triticum vulgare représentent l'ensemble des blés cultivés actuellement.

Etant donné leur très forte stérilité, ces hybrides ont tendance à se croiser avec l'un des parents. Il s'en suit une production de nombreux hybrides en retour. Ces croisements en retour avec l'un ou l'autre des types originels aboutissent, par suite des recombinaisons et ségrégations génétiques, au transfert d'une partie du matériel génétique d'une sous espèce dans une autre. Ce phénomène est connu sous le terme **d'hybridation introgressive** ou **introgression**.

6. Détection de la variation : notion de polymorphisme

La génétique des populations s'intéresse principalement à la **variabilité d'origine génétique** et que l'on désigne sous le nom de **polymorphisme**. Le polymorphisme correspond aux variations de la séquence nucléotidique de l'ADN d'un gène dans une population. Un gène est dit polymorphe s'il existe au moins deux allèles à une fréquence égale ou supérieure à 5% ou à 1%. Plusieurs types de polymorphisme sont étudiés dans l'analyse de la variation des populations :

6.1. Polymorphismes de l'ADN :

- **Polymorphisme nucléotidique** ou **polymorphisme d'un seul nucléotide (SNP, single-nucleotide polymorphism)** est la substitution d'une seule paire de base du génome entre individus d'une même espèce. Les SNP sont à la base des différences dans la résistance aux maladies (drépanocytose, β -thalassémie). Les SNP peuvent se retrouver au sein de régions codantes de gènes (**exon**), de régions non codantes de gènes (**intron**), ou de régions intergéniques, entre les gènes. On parlera de formes alléliques *synonymes* dans le cas où plusieurs formes d'un SNP mènent à la même séquence polypeptidique, et de formes *non synonymes* dans le cas où les séquences produites diffèrent. Les SNP permettent d'identifier des génotypes (reconnaître des personnes, par exemple), ou permettant de contribuer à la construction d'arbres généalogiques (individus) ou phylogénétiques (espèces).

- **Polymorphisme de longueur des fragments de restriction** (ou **RFLP, restriction fragment length polymorphism**) est caractéristique des molécules d'ADN permettant de les distinguer dans deux génomes différents (empreintes génétiques, tests de paternité). Après extraction et purification, l'ADN est coupé en fragments de restriction par une enzyme de restriction au niveau d'une séquence qui lui est spécifique (site de restriction). Les fragments de restriction obtenus, sont ensuite séparés selon leur longueur par électrophorèse.

- **Microsatellites** : une à six paires de bases répétées n fois ($5 \leq n \leq 40$). Ex : TGTGTGTGTG..... $\geq 50\ 000$ microsatellites (TG) $_n$ présents dans tout le génome humain. On les met en évidence par PCR (Polymerase Chain Reaction ou réaction de polymérisation en chaîne) exclusivement. Elle permet d'amplifier *in vitro* une région un fragment d'ADN donné afin d'en obtenir une quantité suffisante pour le détecter et l'étudier.

6. 2. Les polymorphismes chromosomiques

- Remaniements chromosomiques / mutations chromosomiques (gain, perte, réarrangement de segments)
- Changement du nombre de chromosomes sans changement de structure : hétéroploïdie => aneuploïdie et euploïdie, mutations génomiques (polyploïdie)

6.3. Les polymorphismes enzymatiques et protéiques

C'est grâce à l'électrophorèse que ces polymorphismes ont été mis en évidence. Elle trouve une application dans la détection de maladies héréditaires graves dues à l'existence de gènes ou d'associations de gènes défavorables. L'électrophorèse est une technique utile pour analyser la variabilité des populations, les différencier géographiquement, suivre leur évolution sous l'action de facteurs divers. L'électrophorèse suivie de révélation spécifique des activités enzymatiques permet d'obtenir des zymogrammes à partir d'extraits bruts. Les autres protéines peuvent être séparées en fonction de leur charge électrique, ou de leur conformation spatiale.

6.4. Les polymorphismes sérologiques

Les groupes sanguins A, B, O chez l'homme, (Landsteiner, 1900) est la forme la plus anciennement reconnue du polymorphisme. Ces groupes sanguins sont au nombre de quatre (A, B, AB et O) se différencient par la présence, l'absence ou la combinaison des antigènes A ou B à la surface des globules rouges. Le deuxième polymorphisme correspond aux facteurs *Rhesus* qui sont des facteurs présents à la surface des globules rouges. Si ces facteurs apparaissent sur les globules rouges d'une personne, elle est dite " *Rhesus* (Rh) positif ". Si ces facteurs n'apparaissent pas sur ses globules rouges, elle est dite " Rh négatif ".

7. Quantification du polymorphisme

Le polymorphisme pour un gène dans une population peut se mesurer par différents indices :

- Nombre moyen d'allèles par locus (A)

Le nombre moyen d'allèles par *locus* (A), appelé également taux d'allélisme ou richesse allélique, est défini pour n_i allèles au *locus* i et pour L *loci* comme :

$$A = 1/L \sum_{i=1}^L n_i$$

Exemple : pour 3 *loci* numérotés 1, 2 et 3 ayant respectivement 2, 3 et 2 allèles, $A = (2 + 3 + 2)/3 = 2,33$. La mesure de ce paramètre est particulièrement importante pour les stratégies de conservation.

- **Proportions de *loci* polymorphes (p)**

- **Fréquences alléliques pour les *loci* polymorphes**

- **Indices de diversité basés sur les fréquences des allèles** (ex. indice de Nei, $H_e = 2pq$)

* **Indice de diversité génétique de Nei**. Il se calcule de la manière suivante :

$$H_e = \text{Nombre de loci polymorphes} / \text{Nombre de loci étudiés}$$

~ S'il n'y a qu'un seul allèle, $H_e = 0$.

~ Plus il y a d'allèles plus l'indice sera grand.

~ Plus les allèles auront des fréquences proches les unes des autres et plus l'indice sera grand.

Ex1. Le polymorphisme des allozymes (Hamrick et Godt, 1989)

Forme de vie	P	He	
annuelles	30,2%	0,105%	
Pérennes herbacées	28%	0,096	
Pérennes ligneuses	50%	0,149	

D'après ces résultats, **les arbres se révèlent comme étant les plus variables**

Ex.2. Chez le guépard ($p = 0,02$ et $H_e = 0,0004$) (Goulot d'étranglement, perte de la variation génétique donc perte de sa capacité évolutive).

- **Indice de Fixation (Fis)**

Le paramètre Fis de Wright, dénommé aussi indice de fixation et appelé auparavant coefficient de consanguinité (Wright, 1969) est calculé selon la formule :

$$Fis = (H_e - H_o) / H_e = 1 - (H_o / H_e);$$

H_o : hétérozygotie observée ; H_e : hétérozygotie attendue dans l'hypothèse de Hardy-Weinberg. Fis = 1 signifie fixation complète (cas d'autofécondation), Fis inférieur à 1 : hétérozygotie excédentaire, Fis = 0 : population en équilibre de Hardy-Weinberg.

CHAPITRE II. STRUCTURE GENETIQUE EVOLUTION ET SPECIATION

1. Populations et leurs caractéristiques

- 1.1. Isolement**
- 1. 2. Mécanismes d'isolement**
- 1.3. Taille des populations**
- 1.4. Modes de reproduction**

2. Panmixie et équilibre de Hardy-Weinberg

- 2.1. Fréquences génotypiques**
- 2.2. Fréquences alléliques**
- 2.3. Cas de plusieurs allèles**
- 2.4. Cas de gènes liés aux chromosomes sexuels**

3. Modes de croisement

- 3.1. Croisements maintenant les fréquences alléliques**
 - 3.1.1. Consanguinité et autofécondation**
 - 3.1.2. Homogamie**

3.2. Croisements modifiant les fréquences alléliques (cas des mâles dominants, voir chapitre sélection)

4. populations en déséquilibre

- 4.1. Mutations**
- 4.2. Sélection naturelle**
- 4.3. Effet combiné des mutations et de la sélection**
- 4.4. Migration**

5. Spéciation

- 5.1. Concepts et définition de l'espèce**
- 5.2. Mécanismes d'isolement reproducteur**
- 5.3. Mécanismes de spéciation**

1. Populations et leurs caractéristiques

1.1. Isolement reproducteur

Les moyens divers qui empêchent le croisement d'une espèce avec une autre se nomment mécanismes d'isolement. La spéciation est liée à deux types de processus biologiques plus ou moins indépendants et pouvant intervenir successivement ou simultanément. Les uns contrôlent la diversification des populations (mutation, recombinaison, ségrégation...) alors que les autres mettent en place des barrières d'isolement qui constituent le terme ultime de l'individualisation des unités spécifiques (c.à.d : espèces). La diversification peut se faire graduellement sous l'influence prépondérante de la sélection naturelle ou d'une façon brusque par l'intermédiaire de mutations chromosomiques importantes ou d'hybridation. On distingue donc, une **spéciation progressive** et une **spéciation abrupte**. Dans le phénomène de la spéciation abrupte, les nouveaux types sont d'emblée isolés de ceux qui leurs ont donné naissance. La **diversification** et l'**isolement** sont donc **simultanés**. L'étude des populations a montré que les populations évoluent selon un ou plusieurs des trois éléments structuraux suivants : 1) séries de populations contigües à changements graduels (variation en cline) ; 2) populations géographiquement séparées de la partie principale du domaine de l'espèce (isolats géographiques) ; 3) ceintures étroites à variabilité croissante, souvent très vite bordées de chaque côté par de groupes de populations ou de sous espèces stables et très uniformes (ceintures hybrides).

- Variation en cline :

A travers une série de populations contigües, on constate généralement que les changements suivent une progression régulière, en raison de la variation graduelle de l'environnement et du courant de gènes entre populations adjacentes qui tend à adoucir les différences existantes entre les populations.

- Isolats géographiques

Un isolat est défini comme une population ou un groupe de populations séparées par une barrière extrinsèque et ne pouvant échanger librement leurs gènes avec d'autres populations de l'espèce. La caractéristique essentielle d'un isolat géographique est d'être séparé du reste de l'espèce par une discontinuité. Cependant, l'isolement n'est jamais complet car il se produit toujours un certain courant génique (même dans les îles océaniques, sinon, elles n'auraient pu

être colonisées à l'origine). Ces isolats ont trois destinées possibles : ils deviennent des espèces distinctes, ils disparaissent totalement ou ils rétablissent le contact avec le corps principal de l'espèce.

- Ceintures hybrides

Lorsque deux populations sont isolées l'une de l'autre, leurs pools de gènes deviennent indépendants et leurs constitutions génétiques divergent constamment. Lorsque l'isolement géographique disparaît et les populations rétablissent le contact, on peut constater l'existence de certains phénomènes dans la zone de contact qui indiquent le degré de différenciation génétique atteint durant l'isolement.

1. 2. Mécanismes d'isolement

Les mécanismes d'isolement sont de deux types : les **barrières internes** et les **barrières externes**.

- **Les barrières externes** : ce sont des mécanismes pré-copulatoires ou prépollinisation qui empêchent les croisements interspécifiques tels que :

* Isolement saisonnier : les espèces ne se rencontrent pas. Ex. deux espèces qui ne fleurissent pas à la même période de l'année.

* Isolement éthologique : (ethos=habitude, coutumes, comportement). Les espèces se rencontrent mais n'échangent pas de gènes (pas de fécondation). Il y a restriction au croisement au hasard.

* Isolement mécanique : il est très efficace chez les plantes. Les loges du stigmate ont des tailles et des formes qui varient selon l'espèce. Ce qui constitue des barrières à la pollinisation par les insectes.

* Isolement écologique du type sympatrique : niches écologiques sont différentes.

- Les barrières internes :

Ce sont des mécanismes postcopulatoires ou postpollinisation. Ils réduisent le succès des croisements interspécifiques.

* Mortalité des gamètes : elle empêche la fécondation. Les tubes polliniques sont incapables d'atteindre les ovules. Ex.1 : chez les polyploïdes, les tubes polliniques sont de grands diamètres et cheminent difficilement dans les styles des espèces diploïdes. Ex. 2 : les spermatozoïdes peuvent se heurter dans le tractus génital de la femelle à une réaction antigénique.

* Mortalité du zygote : il y a fécondation. Dans le genre *Datura*, l'embryon cesse son développement quand il atteint le stade huit cellules. Différents mécanismes sont à l'origine de la mortalité du zygote tels que :

~ L'albumen qui peut se comporter comme un inhibiteur de développement de l'embryon.

~ Présence d'un gène létal qui n'a aucune influence sur les individus de l'espèce mais entraîne la mort de l'hybride.

~ L'embryon n'arrive jamais à maturité à cause d'une dysharmonie génétique entre les génomes parentaux.

~ L'inviabilité des hybrides résulte de l'interaction de certains génotypes hybrides avec la présence de l'un de ces espèces parentes.

~ Stérilité des hybrides : il y a deux types de stérilité hybride (génique et chromosomique). La stérilité génique se traduit par le fait que les organes sexuels n'arrivent pas à la méiose ou par la présence d'anomalies méiotiques d'origine génétique. La stérilité chromosomique résulte de la non homologie entre les chromosomes maternels et paternels de l'hybride.

~ Inviabilité ou faiblesse de descendants de la F1 ou des générations suivantes. Il arrive que l'isolement ne se manifeste qu'au cours de la 2^e génération des hybrides.

1.3. Taille des populations

L'évaluation de la taille des populations est un problème techniquement difficile. Il faut en effet, considérer les variations de la taille dans le temps, les déplacements d'individus avec les échanges interpopulations qu'ils entraînent, la durée des générations et leurs chevauchements éventuels. Il faut tenir compte également de la répartition réelle des individus

d'une population dans la nature car on s'est contenté souvent des hypothèses d'uniformité et de continuité qui conduisent à la notion très insuffisante de **densité**. A cause de ces difficultés matérielles et conceptuelles, la taille d'une population est souvent mal estimée. **Ex. 1.** Une population d'escargots *Cepea nemoralis* rassemble de l'ordre de 5000 à 20 000 individus. **Ex.2.** Les populations de phalènes varient de 500 à 20 000. Même les groupes pourtant très organisés de vertébrés supérieurs n'ont pas une taille fixe. **Ex.** Les groupes de babouins (*Papio*) peuvent aller de quelques uns à près de 200 individus. Ce qui importe en génétique des populations, c'est l'impact de ces fluctuations numériques sur la biologie des populations correspondantes, sur leurs structures et sur leur composition génétique.

1.4. Modes de reproduction

Il existe deux modes de reproduction : sexuée et asexuée. La fonction biologique de la reproduction sexuée est de reproduire une grande variété de génotypes différents. Elle permet donc de recombinaison les facteurs génétiques des individus parents en de nombreux zygotes génétiquement uniques. Il existe différents types de reproduction sexuée :

- L'autogamie : est la fécondation entre deux gamètes provenant du même individu. L'hermaphrodisme est très répandu dans le règne animal (protozoaire, poissons ; un individu produit les gamètes mâles et femelles). Ce mode n'aboutit pas nécessairement à une consanguinité. Il existe, chez la plupart des espèces hermaphrodites, un nombre de mécanismes qui réduisent ou éliminent les possibilités d'autogamie (production à un moment donné que des gamètes mâles ou des gamètes femelles, les uns après les autres : protandrie ou protérogynie).

- l'allogamie : c'est la fécondation entre deux gamètes provenant d'individus différents. Chez les animaux, on parle de gonogorisme (séparation des sexes). La présence d'un système génétique d'incompatibilité ou encore divers mécanismes floraux peuvent intervenir pour favoriser l'allogamie. Un système d'incompatibilité est un mécanisme génétique et physiologique qui favorise l'allogamie. Il est caractérisé par la présence d'allèles *S*. Deux individus portant cet allèle sont incompatibles (empêche l'autogamie). L'hétérostylie est un mécanisme d'incompatibilité qui prévient également l'autogamie.

- L'apomixie : c'est la formation d'une graine sans fécondation. Etant donné que ce mode de reproduction empêche les recombinaisons et les ségrégations génétiques, on pourrait le considérer comme un facteur limitant ou négatif de l'évolution. Cependant, l'apomixie est

toujours partielle, c'est-à-dire, qu'une plante donnée peut fort bien se reproduire en même temps par les voies sexuées et apomictiques (apomixie facultative). La reproduction sexuée, par suite des combinaisons et ségrégations, donnent lieu à des génotypes variés qui peuvent, à leur tour être sexués ou apomictiques ou les deux à la fois. Ainsi, certaines recombinaisons génétiques favorables peuvent être conservées sans modification par apomixie. Le premier avantage de l'apomixie est donc la possibilité d'une rétention de combinaisons génétiques favorables. Il arrive également que la descendance d'une plante apomictique comprenne des individus ayant un ou plusieurs chromosomes en plus ou en moins en raison des anomalies qui surviennent au moment de la méiose. Ces nouveaux types ne se maintiennent que par la reproduction asexuée.

- Parthénogénèse : est un mode de reproduction dans lequel un nouvel individu est produit sans fécondation. Elle est fréquente chez plantes et chez certains arthropodes, les insectes et les crustacées. Les gamètes mâles et femelles peuvent se développer en embryons sans qu'il y ait fécondation ni méiose. L'absence de méiose permet donc à des individus triploïdes ($3x$), pentaploïdes ($5x$) de se reproduire sans problème.

2. Panmixie et équilibre de Hardy-Weinberg

On considère :

- Des gènes autosomaux
- Des gènes existant sous deux formes alléliques. La plupart des gènes se présentent sous plusieurs formes alléliques. Certaines estérases sont déterminées par six allèles. Pour notre étude, nous considérons le cas d'une paire d'allèles ;
- Une population panmictique (croisement au hasard, indépendamment du génotype et du lien de parenté).
- Un effectif très grand pour pouvoir négliger les fluctuations de la fréquence de gènes.
- Stabilité de l'effectif. L'étude des populations naturelles a montré qu'il existe une stabilité relative de l'effectif. Lorsqu'une population colonise un milieu, son effectif augmente d'une génération à l'autre jusqu'à une limite déterminée par certains facteurs (nourriture, place disponibles). Lorsque le maximum est atteint, la population se trouve en équilibre avec son milieu. L'effectif demeure pratiquement constant (invariable) d'une génération à l'autre. On dit que la population est à l'équilibre ou stationnaire.

La loi de Hardy-Weinberg permet de calculer la fréquence génique dans une population stationnaire. Trouvée simultanément par Hardy (mathématicien anglais) et Weinberg (médecin allemand), cette loi s'énonce ainsi : dans une population panmictique à l'équilibre où il n'y a ni mutation, ni sélection, ni migration et d'effectif élevé, où tous les œufs arrivent à l'état adulte, où tous les individus sont également fertiles, où tous les gamètes participent à la fécondation, la proportion des allèles est des génotypes d'une génération à l'autre est absolument constante. Quelque soit la fréquence de départ, cet état est atteint dès la première génération.

2.1. Fréquences génotypiques

Soient deux allèles (A et a) à un même gène dans une population diploïde et d'effectif N. A étant dominant sur a. Il existe trois catégories de génotypes (AA, Aa, aa). Appelons x, le nombre d'individus ayant le génotype AA, y le nombre d'individus ayant le génotype Aa et z, le nombre d'individus ayant le génotype aa.

$$x+y+z=N$$

On se propose de déterminer la fréquence de chaque génotype. **La fréquence génotypique est le rapport du nombre d'individus ayant ce génotype au nombre total d'individus de la population. L'ensemble des fréquences de chacun des génotypes constitue la structure génotypique de la population.**

$\left. \begin{array}{l} \text{La fréquence du génotype AA} = \\ x/(x+y+z)=x/N \end{array} \right\}$
--

2.2. Fréquences alléliques

Même énoncé que précédemment mais cette fois ci, on se propose de calculer la fréquence de chaque allèle. Appelons 'p', la fréquence de l'allèle A et 'q' la fréquence de l'allèle a (p+q=1).

Chaque individu AA possède deux allèles A donc nous avons 2x allèles A issus des homozygotes dominants dans la population

Chaque individu Aa possède un allèle, donc nous avons 'y' allèles A issus des hétérozygotes Aa dans la population.

Le nombre total d'allèles 'A' dans toute la population est de 2x (issus des homozygotes) + y allèles (issus des hétérozygotes)

La fréquence de l'allèle A est $p = (2x+y)/2N = (x+\frac{1}{2}y)/N = x/N + \frac{1}{2}y/N$

La fréquence de l'allèle A est donc égale à la fréquence du génotype homozygote AA additionnée de la demi-fréquence des hétérozygotes Aa, soit $fr(A) = fr(AA) + \frac{1}{2} fr(Aa)$.

Il est donc possible de calculer les fréquences alléliques en fonction des effectifs partiels des génotypes et de l'effectif total.

Exercices d'application

Ex. 1. Considérons les effectifs des groupes sanguins M et N dans une population de 10694 individus.

Groupe sanguin	M	MN	N
Effectif	3356	5178	2160

Calculer les fréquences des différents groupes sanguins, les fréquences phénotypiques, les fréquences génotypiques et les fréquences alléliques.

Fréquences des groupes sanguins M : 0.3138 MN : 0.4842 N : 0.2020

Les fréquences des groupes sanguins M, NM, N sont des **fréquences phénotypiques**. Comme M et N sont co-dominants (ils ne sont pas dominants l'un par rapport à l'autre), les phénotypes et les génotypes se confondent. Ces fréquences sont aussi des fréquences génotypiques ; sachant que le groupe M correspond au génotype MM, le groupe N correspond au génotype NN et le groupe MN au génotype MN.

Les fréquences alléliques :

$$p = fr(M) = fr(MM) + \frac{1}{2} fr(MN) = 3356/10694 + \frac{1}{2} (5178/10694) = 0.5559$$

$$q = fr(N) = fr(NN) + \frac{1}{2} fr(MN) = 2160/10694 + \frac{1}{2} (5178/10694) = 0.4449$$

Montrons que, dans une population en équilibre de Hardy-Weinberg, les fréquences alléliques et les fréquences génotypiques sont constants d'une génération à l'autre.

Soit à la génération n, la fréquence de l'allèle A = p et la fréquence de l'allèle a = q avec $p+q=1$. Que deviennent ces fréquences à **la génération n+1** ?

Dans le cas où la répartition des allèles A et a est à égale entre les spermatozoïdes et les ovules, les individus de la génération n produisent p spermatozoïdes porteurs de l'allèle A, q spermatozoïdes porteurs de l'allèle a, p ovules porteurs de A et q ovules porteurs de a.

Le résultat des croisements aléatoires qui déterminent la composition génétique de la génération n+1 est donné dans le tableau suivant :

	A(p)	a (q)
A(p)	AA (p^2)	Aa (pq)
a (q)	Aa (pq)	aa (q^2)

Dans le cas de deux allèles, on a 'p' chances de tirer A et donc p^2 chances de réaliser AA. De façon symétrique, on a une probabilité q^2 pour réaliser aa. Pour réaliser le génotype Aa, il y a 2 fois pq possibilités.

La génération n+1 sera constituée de p^2 zygotes AA + $2pq$ zygotes Aa + q^2 zygotes aa

Les fréquences alléliques à cette génération seront égales à :

$$\left. \begin{array}{l} \text{fr (A)} = \text{fr (AA)} + \frac{1}{2}\text{fr (Aa)} \\ p(p+q) = p \end{array} \right\} \begin{array}{l} \text{fr (a)} = p^2 + pq = \\ q(p+q) = q \end{array} \quad \begin{array}{l} \text{proportions} \\ \text{qui} \\ \text{sont} \end{array}$$

Il est facile de vérifier que la somme des différentes probabilités génotypiques est bien égale à 1 :

$$p^2 + 2pq + q^2 = (p+q)^2 = 1$$

Application à un gène récessif responsable de la phénylcétonurie, due à un déficit métabolique et provoquant un retard mental prononcé chez les individus homozygotes récessifs. Cette

maladie touche dans une population 01 individu sur 25000. Quelle est la fréquence q de cet allèle ?

La fréquence des malades (aa) est $1/25000 = q^2$

La fréquence de l'allèle a est $q = \sqrt{q^2} = \sqrt{1/25000} = 1/159$

$$p+q=1 \Rightarrow p=1-q=158/159=79/80$$

Les fréquences des différents génotypes sont donc :

- fr (aa) = $1/25000 \Rightarrow$ malades
- fr (Aa) = $1/80 \Rightarrow$ normaux
- fr (AA) = $79/80 \Rightarrow$ normaux

On constate ainsi, que la fréquence des hétérozygotes est 300 fois plus élevée que celle des homozygotes atteints de phénylcétonurie. **D'une façon générale, plus un allèle est rare, plus est grande la proportion des hétérozygotes par rapport à celle des homozygotes pour cet allèle.**

Remarque :

Au phénotype $[A]$ dû à l'allèle dominant (A) peuvent correspondre 2 génotypes (AA) et (Aa). Le seul phénotype dont le génotype est certain est celui qui est dû à l'allèle récessif (aa). Si la population est en équilibre, on peut obtenir une estimation de ' q ' à partir de q^2 (fréquence des individus homozygotes aa).

Ex. si 75% de la population a le phénotype $[A] \Rightarrow 25\%$ a le phénotype $[a]$. Si la population est en équilibre pour ce locus, on peut écrire :

$$q^2 = \text{fréquence des } (aa) \text{ d'où } q^2 = 0.25 \Rightarrow q=0.5 \Rightarrow p = 0.5$$

2.3. Cas de plusieurs allèles

Soient trois allèles A, a, a' avec une relation de dominance hiérarchisée $A > a > a'$ et dont les fréquences respectives sont p, q , et r . les croisements aléatoires donneront naissance à des zygotes dont les fréquences seront comme suit :

$(p+q+r)^2 =$	p^2	+	q^2	+	r^2	+	$2pq$	+	$2pr$	+	$2qr$
	(AA)		(aa)		(a'a')		(Aa)		(Aa')		(aa')

Les différentes fréquences : p^2 $2pq$ $2pr$ q^2 $2qr$ r^2

Les différents génotypes : (AA) (Aa) (Aa') (aa) (aa') $(a'a')$

Les différents phénotypes : $[A]$ $[a]$ $[a']$

Ex. Soient trois allèles déterminant la coloration du pelage chez le lapin : C=sauvage, ch=himalayan et c=albinos, en fonction de la relation de dominance $C > ch > c$ et les fréquences respectives suivantes p, q et r.

a. Si une population de lapins contenant des individus sauvages, himalayan et albinos est panmictique, quelle seraient, en fonction de p, q et r à la génération suivante les fréquences attendues des différents génotypes. En déduire les fréquences phénotypiques en fonction de p, q et r.

b. un échantillon de lapins contient 168 lapins sauvages, 30 himalayan et 2 albinos. Calculer les fréquences p, q et r.

c. Etant donné les fréquences $p=0.5$, $q=0.1$ et $r=0.4$, calculer les proportions des différents génotypes parmi les lapins sauvages, les fréquences des lapins sauvages et la fréquence des lapins himalayan.

Réponse :

a) :

Fréquences génotypiques: p^2 $2pq$ $2pr$ q^2 $2qr$ r^2

Les différents génotypes sont : (CC) (Cch) (Cc) $(chch)$ (chc) (cc)

Les différents phénotypes sont : $[C]$ $[ch]$ $[c]$

Les fréquences phénotypiques :

$$\text{fr } [C] = p^2 + 2pq + 2pr, \quad \text{fr } [ch] = q^2 + 2qr, \quad \text{fr } [c] = r^2$$

b) :

Fréquences phénotypiques

$$\text{fr [C]} = 168/200 = 0.84 = p^2 + 2pq + 2pr$$

$$\text{fr [ch]} = 30/200 = 0.15 = q^2 + 2qr$$

$$\text{fr [c]} = 2/200 = 0.01 = r^2$$

$$\text{fr [c]} = 2/200 = 0.01 = r^2 \Rightarrow \boxed{r = 0.1}$$

$$(q^2 + 2qr + r^2) = \text{fr [ch]} + \text{fr [c]} = 0.15 + 0.01 = 0.16 = (q+r)^2 = 0.16 \Rightarrow q+r=0.4 \Rightarrow q=0.4-0.1=0.3$$

$$\boxed{q=0.3} ; \text{ puisque } p+q+r=1 \Rightarrow p = 1-q-r = 1-0.3-0.1=0.6 \quad \boxed{p=0.6}$$

Les fréquences des différents génotypes parmi les lapins sauvages

$$\text{fr (CC)} = p^2 = (0.5)^2 = 0.25$$

$$\text{fr (Cch)} = 2pq = 2 \times 0.5 \times 0.1 = 0.1$$

$$\text{fr (Cc)} = 2pr = 2 \times 0.5 \times 0.4 = 0.4$$

Fréquences des lapins sauvages (phénotypes) :

$$\text{fr [C]} = \text{fr (CC)} + \text{fr (Cch)} + \text{fr (Cc)} = 0.25 + 0.1 + 0.4 = 0.75$$

Fréquences des lapins himalayan (phénotypes) :

$$\text{fr [ch]} = \text{fr (chch)} + \text{fr (chc)} = q^2 + 2qr = (0.1)^2 + 2 \times 0.1 \times 0.4 = 0.09$$

2.4. Cas des gènes liés aux chromosomes sexuels

Un gène est lié au sexe lorsque'il est porté par le cromosomes X. pour deux allèles A et a, il y a :

[- Trois génotypes pour les femelles : $X^A X^A$, $X^A X^a$, $X^a X^a$

- et deux génotypes pour les mâles : $X^A Y$, $X^a Y$

L'hétérozygotie n'existe que chez les femelles. Les mâles sont **hémizygotes** pour les gènes liés au sexe.

On divise la population de départ en deux souspopulations l'une femelle et l'autre mâle pour calculer les fréquences alléliques et les fréquences génotypiques. On peut donc calculer séparément les fréquences alléliques chez les femelles comme nous l'avons fait précédemment dans le cas des gènes autosomaux. De même, on peut calculer séparément les fréquences alléliques chez les mâles.

Chez les mâles, la fréquence des mâles $X^A Y$ correspond à la fréquence p de l'allèle A et la fréquence des mâles $X^a Y$ correspond à la fréquence q de l'allèle a . **Chez les mâles, la fréquence génotypique est à l'égal de la fréquence allélique.**

Remarque : Le phénotype récessif est plus fréquent chez les mâles ($q_{\text{♂}}=0.5$, $q_{\text{♀}}^2=0.0025$).

Théoriquement, **lorsque la population est en équilibre, les fréquences alléliques calculées chez les femelles et chez les mâles sont identiques.** Il suffit donc de calculer les fréquences chez les mâles du fait de la simplicité des calculs pour avoir les fréquences alléliques de la population.

Bien souvent, les fréquences calculées chez les femelles et chez les mâles ne sont pas identiques. La fréquence p de A dans l'ensemble de la population est donnée par la formule :

$$p = \left(\frac{2}{3} \right) p_{\text{♀}} + \left(\frac{1}{3} \right) p_{\text{♂}}$$

Ex.1. Soit un locus A/a situé sur le chromosome X . dans une population, on dénombre :

90 ♂ $X^A Y$

10 ♂ $X^a Y$

77 ♀ $X^A X^A$

21 ♀ $X^A X^a$

2 ♀ $X^a X^a$

- Cette population est-elle en équilibre ?
- Calculer la fréquence p de A dans la population totale.

Chez les ♀ : $p_{\text{♀}} = \left(\frac{77}{100} \right) + \frac{1}{2} \left(\frac{21}{100} \right) = 0.875$; $q_{\text{♀}} = \left(\frac{2}{100} \right) + \frac{1}{2} \left(\frac{21}{100} \right) = 0.125$

Chez les ♂ : $p_{\text{♂}} = 90/100 = 0.9$

$q_{\text{♂}} = 10/100 = 0.1$

Considérons la population en équilibre donc les fréquences alléliques chez les mâles sont à égales aux fréquences alléliques chez les femelles et considérons les fréquences chez les mâles $p=0.9$ et $q=0.1$

Nb de ♀ $X^A X^A = p^2 \times 100 = (0.9)^2 \times 100 = \mathbf{81}$

Nb de ♀ $X^A X^a = 2pq \times 100 = 2 \times 0.1 \times 0.9 \times 100 = \mathbf{18}$

Nb de ♀ $X^a X^a = q^2 \times 100 = (0.1)^2 \times 100 = \mathbf{1}$

Calculons le χ^2 :

$$\chi^2_{\text{cal}} = \sum (n_i \text{ obs} - n_i \text{ cal})^2 / n_i \text{ cal} = (77-81)^2/81 + (21-18)^2/18 + (2-1)^2/1 = 1.6$$

Nombre de degré de liberté $3-1=2$, $\chi^2_{\text{cal}} < \chi^2_{\text{tab}} (6) \Rightarrow$ la population est en équilibre

Fréquence p de l'allèle A dans l'ensemble de la population :

$$p = (2/3)p_{\text{♀}} + (1/3)p_{\text{♂}} = (2/3) \times 0.875 + (1/3) \times 0.9 = 0.883$$

Ex.2 : chez les chats domestiques, la couleur du pelage est déterminée par un couple d'allèles codominants liés au sexe $X^b X^y$

	Phénotype		
	noir	mosaïque	jaune
♀	$X^b X^b$	$X^b X^y$	$X^y X^y$
♂	$X^b Y$	-----	$X^y Y$

Soit p, la fréquence de b et q celle de y

$$p = [2 \times (\text{nb de ♀ noires}) + (\text{nb de ♀ mosaïque}) + (\text{nb de ♂ noirs})] / (2 \text{ nb de ♀} + \text{nb de ♂})$$

$$q = [2 \times (\text{nb de ♀ jaunes}) + (\text{nb de ♀ mosaïques}) + (\text{nb de ♂ jaunes})] / (2 \text{ nb de ♀} + \text{nb de ♂})$$

3. Modes de croisement

La loi de Hardy et Weinberg est une loi essentiellement théorique. Elle n'est applicable que si les croisements se font au hasard. C'est un cas idéal qui ne se rencontre jamais dans la nature où l'une au moins et généralement plusieurs de ses conditions d'application sont absentes. Toujours intervient dans le choix des conjoints un certain nombre de facteurs qui rendent tel type de croisement plus probable que tel autre. Ces croisements non panmictiques peuvent se diviser en deux catégories :

- Croisements modifiant les fréquences génotypiques mais maintenant les fréquences alléliques ;
- Croisements modifiant les fréquences génotypiques et les fréquences alléliques.

3.1. Croisements maintenant les fréquences alléliques

Le maintien des fréquences alléliques est assuré dans les croisements **consanguins** c-à-d les croisement entre individus apparentés et dans le cas extrême de consanguinité qu'est l'**autofécondation**. Il est assuré également dans les croisements **homogamiques** où les individus phénotypiquement semblables ont tendance à s'accoupler (choix du partenaire selon la taille, la couleur, les ornements).

3.1.1. Consanguinité et autofécondation

La consanguinité est l'union entre individus présentant un certain degré de parenté. La consanguinité peut avoir une origine sociale : ex. classique est donné par les Dynasties de l'Egypte antique : dans le but de conserver le sang royal il y avait des mariages entre frères et sœurs, ou une origine géographique tribus, villages isolés. Chez les oiseaux, il est très fréquent que les couples se forment à partir d'une même couvée. Un autre exemple est celui des blastophages du figuier qui se développent et qui se reproduisent dans une même figue. L'autofécondation est très répandue chez les plantes à fleurs bisexuées et chez les mollusques hermaphrodites. Néanmoins l'autofécondation n'est pas une règle absolue car il arrive que des dispositions spéciales tendent à l'empêcher telles que des gènes stériles, l'hétérostylie, la protandrie ou la protogynie.

Les croisements consanguins sont utilisés pour améliorer les plantes cultivées ou les animaux domestiques. Nous allons voir pourquoi en examinant le cas de l'autofécondation.

Supposons panmictique à l'origine la population dans laquelle nous allons suivre le devenir des allèles A et a qui sont présents aux fréquences p et q respectivement. La composition génotypique de la population à la génération de départ F_0 est :

$$AA : p^2 \qquad Aa : 2pq \qquad aa : q^2$$

Les individus AA d'une part, aa d'autre part s'autofécondent, redonnant pour les premiers des AA, pour les seconds des aa. Dans le cas des Aa, le résultat de l'autofécondation est différent puisqu'en F_1 :

$$2pq (Aa \times Aa) \text{ donnent } 2pq \left(\frac{1}{4}AA + \frac{1}{2} Aa + \frac{1}{4} aa \right) = pq \left(\frac{1}{2}AA + Aa + \frac{1}{2}aa \right)$$

dont les AA et aa vont rejoindre leurs semblables issus des croisements homogamiques pour former la troisième génération.

Le nombre d'hétérozygotes diminue donc de moitié à chaque génération ;

$$2pq \text{ à la génération } F_0$$

$$pq \text{ à la génération } F_1$$

$$pq/2 \text{ à la génération } F_2$$

$$pq/2^2 \text{ à la génération } F_3$$

$$pq/2^{n-1} \text{ à la génération } F_n$$

Le nombre d'homozygotes s'accroît simultanément de la quantité correspondante par addition des termes successifs d'une progression géométrique de raison $\frac{1}{2}$, soit à la génération n, une augmentation pour chaque type d'homozygote (il y a deux types) est à égale à :

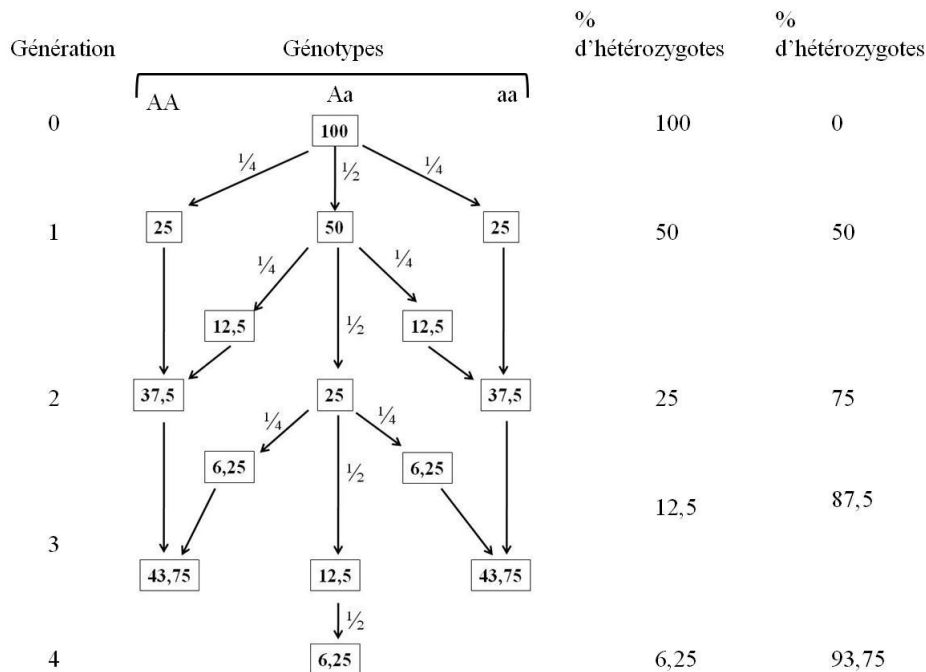
$$\frac{1}{2} (pq + pq/2 + pq/4 + \dots + pq/2^{n-1}) = pq/2 + pq/4 + \dots + pq/2^n \simeq pq$$

La fréquence des individus homozygotes AA est au bout de n générations est à égale à :

$$p^2 \text{ (fréquence de départ) } + pq \text{ (augmentation) } = p(p+q) = \mathbf{p}$$

$$\text{Celle des homozygotes aa est à égale } pq + q^2 = q(p+q) = \mathbf{q}$$

Donc les fréquences génotypiques sont à égales aux fréquences alléliques (car il n'y a plus d'hétérozygote).



Les fréquences alléliques sont donc :

$$\text{fr (A)} = \text{fr (AA)} + \frac{1}{2} \text{fr (Aa)} = \text{fr (A)} + 0 = \mathbf{p}$$

$$\text{fr (a)} = \text{fr (aa)} + \frac{1}{2} \text{fr (Aa)} = \text{fr (a)} + 0 = \mathbf{q}$$

Les fréquences alléliques restent donc constantes avec ce type de croisement mais les fréquences génotypiques varient. Les hétérozygotes sont assez rapidement éliminés (au bout de neuf générations, environ).

Coefficient de consanguinité

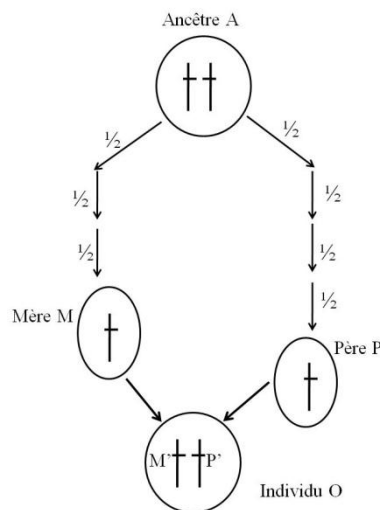
Le degré de consanguinité **d'une population** est mesuré par le coefficient α qui varie de zéro (absence d'homozygote) à un (homozygotie complète) et **qui exprime le taux d'hétérozygotie perdue à chaque génération**. Le coefficient de consanguinité α **d'un individu** est la probabilité pour que deux allèles qu'il possède en un locus désigné au

hasard soient identiques. Au niveau d'un individu, ce coefficient indique la probabilité qu'ont les deux allèles d'un locus d'être deux répliques d'un même allèle d'un ancêtre commun.

Calcul du coefficient de consanguinité

a. Cas d'un ancêtre commun

Un individu O reçoit un allèle P' de son père P et un allèle M' de sa mère M. Si cette mère et ce père ont un ancêtre commun A, la probabilité pour que ces deux allèles P' et M' soient identiques s'obtient comme suit :



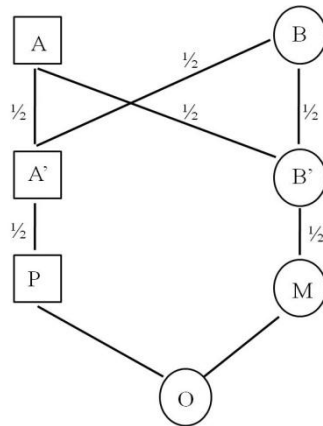
- La probabilité pour que l'allèle transmis par le père vienne de A est à égale à $(\frac{1}{2})^n$, (n : nombre de générations entre A et P).
- La probabilité pour que l'allèle transmis par la mère vienne de A est à égale à $(\frac{1}{2})^{n'}$, (n' : nombre de générations entre A et M).
- La probabilité pour que tous les deux viennent de A est $(\frac{1}{2})^{n+n'}$
- Dans ce cas, avec une probabilité $(\frac{1}{2})$, c'est le même allèle de A qui a donné P' et M'. ce qui donne une probabilité : $p_1 = (\frac{1}{2})^{n+n'+1}$

Si l'ancêtre A est consanguin nous avons une probabilité :

$$p_2 = a_O = (\frac{1}{2})^{n+n'+1} \times (a_A + 1)$$

b. Cas de deux ancêtres communs

Exemple d'un mariage entre deux cousins germains. La généalogie d'un tel mariage est donnée dans la figure qui suit :



O a forcément reçu un des exemplaires qu'elle possède de son père P. Celui-ci a une chance sur deux de l'avoir reçu de A' qui a une chance sur deux de l'avoir reçu de A. De même, O a forcément reçu l'autre exemplaire de sa mère M, laquelle a une chance sur deux de l'avoir reçu de B' qui a une chance sur deux de l'avoir reçu de A. Donc, la probabilité pour que les deux exemplaires portés par O proviennent l'un et l'autre de A est à égale à $(\frac{1}{2})^4$. Il y a en outre une chance sur deux pour que A ait communiqué le même exemplaire à A' et à B'. la probabilité pour que O possède deux copies identiques d'un allèle ou de l'autre porté par A est : $\frac{1}{2} \times (\frac{1}{2})^4 = (\frac{1}{2})^5$.

Le même raisonnement montre qu'il y a aussi $(\frac{1}{2})^5$ chance pour que O possède deux copies identiques de l'un ou l'autre allèle ancêtre porté par B.

Le calcul du coefficient de consanguinité revient donc à additionner les probabilités pour que le descendant possède deux copies identiques provenant d'un ancêtre commun, ici :

$$\alpha_O = (\frac{1}{2})^{n+m+1} (A) + (\frac{1}{2})^{n+m+1} (B)$$

D'une façon générale :

$$\alpha = \sum (\frac{1}{2})^{n_i + m_i + 1} \times (1 + \alpha_i)$$

i : ancêtre commun. **ni** : nombre de générations entre l'ancêtre i et la mère. **mi** : nombre de générations entre l'ancêtre i et le père. **α_i** =coefficient de consanguinité de l'ancêtre i.

La formule ($AA p^2 + Aa 2pq + aa q^2$) appliquée dans le cas de l'équilibre de Hardy-Weinberg ne convient plus dans le cas de consanguinité. Elle se transforme en :

$$AA (p^2 + \alpha pq) + Aa (2pq - 2\alpha pq) + aa (q^2 + \alpha pq)$$

c. Consanguinité et homozygotie

La consanguinité a des conséquences importantes dans le cas des tares récessives. Elle élève la probabilité pour que le descendant soit homozygote et exprime le trait correspondant. Dans le cas de **consanguinité**, la probabilité pour que l'enfant exprime la tare est égale à $q^2 + \alpha pq$, alors que dans le cas de **panmixie** (équilibre de Hardy Weinberg) elle est égale à q^2 .

Prenons pour exemple la probabilité pour que le descendant d'un mariage entre cousins germains soit phénylcétonurique. Dans ce cas $q=1/159$ et $\alpha= 1/16$ soit

$$q^2 + \alpha pq = (1/159)^2 + [(1/16) \times (158/159) \times (1/159)] = 11/25000$$

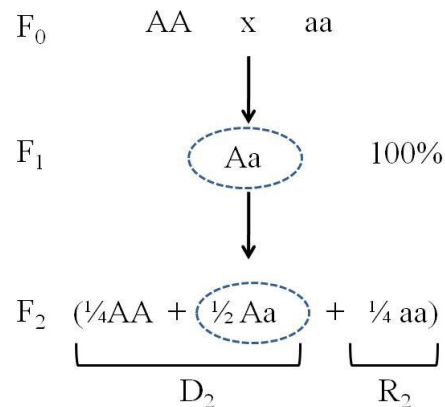
La probabilité est 11 fois supérieure à celle attendue dans un mariage au hasard.

3.1.2. Homogamie

Les couples peuvent ne pas se former au hasard comme nous l'avons vu dans les populations panmictiques. Il peut y avoir un choix en fonction du phénotype. Un système de croisement est de type homogamique, s'il existe une tendance à la ressemblance (homogamie positive) ou au contraire à la dissemblance (homogamie négative) entre les partenaires d'un même couple. L'homogamie ne change pas les fréquences des allèles mais affecte celles des génotypes. L'homogamie que pratiquent les éleveurs ou les agriculteurs est l'homogamie phénotypique : lorsqu'un caractère morphologique ou physiologique leur semble suffisamment intéressant pour être reproduit à grande échelle, ils croisent entre eux les individus porteurs de ce caractère. Certains de ces caractères sont le fait de gènes dominants qui se manifestent donc à l'état hétérozygote. D'autres sont dus à des gènes récessifs visibles seulement à l'état homozygote.

Imaginons un cas simple d'homogamie phénotypique. Une population renferme des allèles A dominants et des allèles a récessifs dans laquelle l'accouplement ne se produit qu'entre

phénotype [A] (AA, Aa) ou entre phénotype [a] (aa). Calculons le résultat de ce mode de croisement en prenant comme population de départ une F₂ panmictique.



Dans F₂, les R₂ se croisent entre eux donnant exclusivement des aa et les D₂ se croisent entre eux. L'évaluation du résultat de ce dernier croisement est aisée en considérant les fréquences alléliques : les fréquences des allèles A et a dans la fraction D₂ de la population sont respectivement 2/3 et 1/3.

$$\frac{(1/4 AA + 1/2 Aa)}{D_2} \text{ soit } AA + 2Aa \Rightarrow AA + Aa + Aa$$

Dans la fraction D₂, la fréquence de l'allèle A = fr (A) = 4 allèles A sur 6 allèles au total soit : 2/3 ;

La fréquence de l'allèle a = 1/3

Les génotypes résultants de leurs croisements fournis par l'expression [(2/3)A + (1/3)a]² se trouvent dans les proportions :

	A (2/3)	a(1/3)
A (2/3)	AA(4/9)	Aa (2/9)
a(1/3)	Aa (2/9)	aa(1/9)

Soit : $(4/9)AA$, $(4/9) Aa$ et $(1/9) aa$

Mais D_2 ne représente que les $3/4$ de la population, les proportions par rapport à l'ensemble de la F_2 sont ramenés à :

$$[(4/9)AA+(1/3)Aa+(1/9)aa] \times (3/4) = (1/3)AA+(1/3)Aa+(1/12)aa$$

La F_3 est constituée par :

$$\left[\begin{array}{c} (1/3)AA+(1/3)Aa+(1/12)aa \\ \text{Issus de } D_2 \end{array} \right] + \left[\begin{array}{c} 1/4aa \\ \text{issus de } R_2 \end{array} \right] = \left[\begin{array}{c} (1/3)AA + (1/3)Aa \\ D_3 \end{array} \right] + \left[\begin{array}{c} (1/3)aa \\ R_3 \end{array} \right]$$

Un calcul identique peut être mené pour calculer la F_4 à partir de cette F_3

D_3 contient 3A pour 1 a ; il fournira donc en F_4 :

$$[(3/4)A+(1/4)a]^2 = (9/16)AA + (6/16)Aa + (1/16)aa$$

Et comme D_3 représente $2/3$ de F_3 , ces proportions ramenées à la population totale deviennent :

$$[(9/16)AA + (6/16)Aa + (1/16)aa] \times (2/3) = (3/8) AA + (1/4)Aa + (1/24)aa$$

$$\left[\begin{array}{c} (3/8) AA + (1/4)Aa + (1/24)aa \\ \text{Issus de } D_3 \end{array} \right] + \left[\begin{array}{c} (1/3) aa \\ \text{Issus de } R_3 \end{array} \right] = 3/8AA + 1/4Aa + 3/8aa$$

L'examen des générations successives montre que le nombre d'allèles dominants demeure égal au nombre d'allèles récessifs. Les fréquences des génotypes homozygotes dominants et récessifs demeurent elles aussi égales d'une génération à l'autre.

$$\boxed{\text{fr}(AA)_n = \text{fr}(aa)_n}$$

La fréquence des hétérozygotes est obtenue aisément par récurrence :

En F_1 , elle est à égale à 1 ;

en F_2 , à $1/2$

en F_3 , à $1/3$

en F_4 , à $1/4$

en F_n , à $1/n$

et nous pouvons écrire :

$$\text{fr (Aa)}_n = 1/n \quad \text{fr (AA)}_n = \text{fr(aa)}_n = [1 - (1/n)] \times 1/2 = (n-1)/2n$$

La réduction à 1/1000 de la fréquence des hétérozygotes exige donc 1000 générations alors qu'il suffisait de 09 générations dans le cas de la consanguinité absolue. Il est donc plus rapide et plus facile pour sélectionner une variété intéressante d'utiliser les croisements consanguins (autofécondation cas idéal) des individus porteurs du caractère intéressant.

Il est important de signaler que ce mode de croisement présente des inconvénients du fait de la perte de polymorphisme génétique qu'il entraîne.

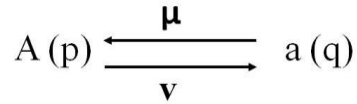
4. Populations en déséquilibre

4.1. Mutations

La mutation est la cause fondamentale de la **variabilité génétique**. Le mot 'mutation' exprime tout changement génétique (inversion, translocation, polyploïdie). L'effet de la mutation sur la structure de la population est différent suivant qu'il s'agit d'un événement rare ou d'un événement qui se répète de génération en génération avec un certain taux. La mutation rare ne confère pas un avantage particulier et ne peut produire de grands changements dans une population et risque fort bien de disparaître à moins que la dérive génétique abordée plus loin, ne la perde à une fréquence élevée mais par pur hasard. A la différence du cas de la mutation rare, la mutation répétitive est régulière ainsi que sa fréquence. Le gène normal peut disparaître et seul le gène muté restera.

En général, la mutation est réversible c'est-à-dire qu'elle s'effectue dans deux directions généralement avec des fréquences différentes.

Appelons ' μ ', la probabilité de mutation de l'allèle dominant 'A' vers l'allèle 'a' et ' ν ', la probabilité de mutation de l'allèle récessif 'a' vers l'allèle dominant 'A'.



Si la fréquence de l'allèle 'A' est à égale à (p) à la génération (n), on peut prévoir la fréquence (p') à la génération (n+1). La fraction des allèles a transformés par mutation en 'A' (vq) s'ajoutera à 'p' et la fraction des allèles 'A' transformés en 'a' par la mutation inverse (μp) s'en retranchera.

$$p' = p + vq - \mu p = (1 - \mu)p + vq$$

De manière générale, Δp correspond au changement de la fréquence de l'allèle A à chaque génération.

$$\Delta p = p' - p = p + vq - \mu p - p = vq - \mu p$$

$$\Delta p = vq - \mu p$$

Un état d'équilibre sera atteint lorsque $\Delta p = 0$ c-à-d lorsque le nombre de mutations de 'A' vers 'a' compensera exactement celui des mutations de 'a' vers 'A'

$$\Delta p = vq - \mu p \Rightarrow vq = \mu p \Rightarrow v(1-p) = \mu p \Rightarrow v - vp - \mu p = 0 \Rightarrow v - p(v + \mu) = 0$$

$$\bar{p} = \nu / (\mu + \nu) \quad \bar{q} = \mu / (\mu + \nu)$$

En conclusion, les fréquences des allèles à l'équilibre ne dépendent que des fréquences mutationnelles. Connaissant les valeurs de μ et de ν , il est très simple de calculer $\bar{p}$ et $\bar{q}$

Cas où la mutation est irréversible

Supposons qu'un allèle A soit à la fréquence p_0 et que μ soit la fréquence mutationnelle de A vers a. Après intervention de la mutation, la nouvelle fréquence de A est $p_1 = p_0 (1-\mu)$. La fréquence de A va donc décroître de génération en génération.

Par récurrence :

$$p_1 = p_0 - \mu p_0 = p_0 (1-\mu)$$

$$p_2 = p_1 - \mu p_1 = p_1 (1-\mu) = p_0 (1-\mu) (1-\mu) = p_0 (1-\mu)^2$$

Après n générations : $p_n = p_0 (1-\mu)^n$.

Exercice d'application

Combien faudra-t-il de générations pour faire passer la fréquence $p = 0,9$ à $p_n = 0,3$ si $\mu = 10^{-6}$?

$$p_n = p_0 (1-\mu)^n \Rightarrow p_n/p_0 = (1-\mu)^n \Rightarrow \log(p_n/p_0) = \log (1-\mu)^n \Rightarrow \log p_n - \log p_0 = n \log (1-\mu) \Rightarrow n = (\log p_n - \log p_0) / \log (1-\mu)$$

$$n = (\log p_n - \log p_0) / \log (1-\mu) = 1098623 \text{ générations}$$

4.2. Sélection

La loi de Hardy et Weinberg suppose que tous les génotypes sont équivalents et qu'ils participent de la même façon à la constitution de la génération suivante. Dans les populations naturelles, les différents génotypes n'ont pas la même viabilité ni la même fertilité. On parle de sélection naturelle lorsque différents génotypes ne sont pas également viables ou également fertiles. Il en résulte que certains génotypes sont retirés de la population soit radicalement soit partiellement. La sélection s'opère également au niveau des gamètes. Dans ce cas, on aboutit toujours, à plus ou moins long terme, à l'élimination de l'allèle considéré. La sélection naturelle est la force directrice de l'évolution. C'est le processus qui conduit à une meilleure adaptation des organismes à leur environnement quelque peu différents. Si les conditions sont différentes, la sélection naturelle va favoriser la divergence génétique entre les populations. Si les

conditions sont similaires, la sélection naturelle va empêcher les populations de diverger génétiquement. La sélection agit sur les phénotypes et non sur les génotypes. On définit la **valeur adaptative** (ou **valeur sélective**) comme le rapport de l'effectif trouvé à une génération à l'effectif théorique attendu par application de la loi de Hardy-Weinberg. Ex. Lorsque 100 œufs de drosophiles 'sauvages' donnent naissance à 100 adultes, la valeur adaptative du phénotype considéré est à égale à $100/100=1$. Mais lorsque 100 œufs 'Bar' de la même espèce ne donnent naissance qu'à 70 adultes, la valeur sélective de 'Bar' est de 0,70.

Plusieurs facteurs sont susceptibles d'intervenir dans cette valeur sélective : fertilité, longévité, résistance à la sous alimentation, aux facteurs climatiques, rapidité de développement etc. Les différents constituants de la valeur sélective n'ont pas une valeur absolue. Elles varient avec les conditions du milieu.

Considérons une population renfermant les allèles A et a aux fréquences p et q respectivement. En absence de sélection gamétique, les différents gamètes donneraient p^2 œufs (AA), $2pq$ œufs (Aa) et q^2 œufs (aa). Mais ces œufs ne donneront pas tous des adultes car certains mourront faute de pouvoir résister aux mauvaises conditions du milieu. Il y aura donc sélection de génotypes résistants ; si nous appelons δ_1 , δ_2 et δ_3 les valeurs sélectives globales de chacun des génotypes :

Les p^2 œufs (AA) donneront $\delta_1 p^2$ adultes (AA)

Les $2pq$ œufs (Aa) donneront $2\delta_2 pq$ adultes (Aa)

Les q^2 œufs (aa) donneront $\delta_3 q^2$ adultes (aa)

La fréquence des allèles A produits par ces adultes-en supposant toujours que n'existe aucune sélection gamétique- sera donc :

$$p_1 = (\delta_1 p^2 + \delta_2 pq) / (\delta_1 p^2 + 2\delta_2 pq + \delta_3 q^2) = p [(\delta_1 p + \delta_2 q) / (\delta_1 p^2 + 2\delta_2 pq + \delta_3 q^2)]$$

$$\text{avec : } \delta_1 p^2 + 2\delta_2 pq + \delta_3 q^2 = w = \text{valeur sélective moyenne de la population}$$

la fréquence de l'allèle a sera

$$q_1 = (\delta_3 q^2 + \delta_2 pq) / (\delta_1 p^2 + 2\delta_2 pq + \delta_3 q^2) = q [(\delta_2 p + \delta_3 q) / (\delta_1 p^2 + 2\delta_2 pq + \delta_3 q^2)]$$

La variation de fréquence de l'allèle A entre deux générations est donc égale à :

$$\Delta p = p_1 - p = [(\delta_1 p^2 + \delta_2 pq) / w] - p = pq [p(\delta_1 - \delta_2) + q(\delta_2 - \delta_3)] / w$$

Dans cette formule de Δp , la population est en équilibre si $p=0$ ou $p=1$ ou encore si $p = (\delta_3 - \delta_2)/(\delta_1 - 2\delta_2 + \delta_3)$.

Lorsque le nombre de générations n augmente indéfiniment la population évolue vers l'un des trois états d'équilibre.

- Si $\delta_2 \leq$ au plus grand des deux valeurs δ_1 et δ_3

$$p \rightarrow 0, \text{ si } \delta_1 < \delta_3$$

$$p \rightarrow 1, \text{ si } \delta_1 > \delta_3$$

- Si $\delta_2 \geq$ au plus grand des deux valeurs δ_1 et δ_3 , $p \rightarrow (\delta_3 - \delta_2)/(\delta_1 - 2\delta_2 + \delta_3)$.

L'effet de la sélection est donc soit l'élimination du polymorphisme, si l'un au moins des deux génotypes homozygote est favorisé par rapport au génotype hétérozygote, soit le maintien du polymorphisme, si le génotype hétérozygote est le plus favorisé.

Δp permet de prévoir l'évolution de la fréquence des allèles en fonction des valeurs relatives de δ_1 , δ_2 et δ_3 . Dans la formule de Δp , le dénominateur est toujours positif. A chaque fois que p et q sont différents de 1 (c-à-d les deux allèles coexistent dans la population), le signe de Δp dépend uniquement du signe du numérateur M . Si $M > 0 \Rightarrow \Delta p > 0 \Rightarrow$ la fréquence de l'allèle A augmente. Si M est $< 0 \Rightarrow \Delta p < 0$ et la fréquence de l'allèle A diminue. Enfin si M est nul, Δp est nul et la population est en équilibre.

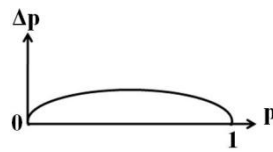
Nous allons examiner successivement les trois situations possibles : Un homozygote est avantagé et l'autre défavorisé, l'hétérozygote est avantagé, l'hétérozygote est défavorisé.

- Cas où un homozygote est avantagé et l'autre défavorisé (valeur adaptative de l'hétérozygote est intermédiaire).

Cette situation peut être symbolisée par la relation $\delta_1 \geq \delta_2 \geq \delta_3$ ou $\delta_1 \leq \delta_2 \leq \delta_3$

Si on reporte sur un graphique la valeur de M en fonction de p , on constate que cette valeur est toujours positive. Ceci signifie que **Δp est toujours positif. La fréquence de A augmente à chaque génération jusqu'à atteindre la valeur 1. L'allèle a sera éliminé. La population tend à perdre sa diversité génétique puisqu'elle ne comportera que des homozygotes AA .**

On dit que la population a fixé l'allèle A. l'état d'équilibre n'existe que lorsque p sera égal à 1.



Un exemple de cette situation est fourni par l'acquisition de la résistance aux insecticides DDT (DichloroDiphénylTrichloroéthane). Après avoir anéanti les poux du corps vecteur du typhus, en Italie, puis au Japon et en Corée de 1944 à 1946, le DDT s'est montré incapable d'enrayer une pandémie de typhus en Espagne. De 1950 à 1951, les soldats coréens traités avec le DDT ont vu leurs poux se multiplier au lieu de disparaître. En 1957, le pou est devenu résistant au DDT dans presque tous les pays du monde. Des observations similaires sont faites sur de nombreuses autres espèces (Anophèles vecteurs du paludisme). Une interprétation possible de ces phénomènes est la sélection des génotypes les plus résistants.

Génotypes	(RR)	(Rs)	(ss)
Phénotypes	[R]	[R]	[s]
$\delta_1=1$	[R]	(RR)	résistant
$\delta_2=1$	[R]	(Rs)	résistant
$\delta_3=0$	[s]	(ss)	sensible

Au bout d'un certain nombre générations, l'allèle 's' est éliminé par contre l'allèle 'R' est fixé. On aboutit donc à la fixation de R. Ce qui pose le problème de résistance au DDT.

Nous allons envisager un cas particulier de cette situation où un allèle dominant est génétiquement létal. La létalité génétique signifie l'absence totale de descendance. Elle peut être due à la mort de l'individu avant l'âge de la reproduction ou bien encore à sa stérilité. Si la létalité est due à un allèle dominant, on a la relation $\delta_1 \geq \delta_2 = \delta_3=0$. Dans ce cas, on constate que $\Delta p = q$. En effet, aucun individu portant l'allèle défavorisé n'a de descendant et cet allèle est entièrement éliminé à chaque génération. Une situation s'approche beaucoup de ce cas chez l'homme et correspond au rétinoblastome. Cette tare est due à un allèle dominant qui provoque, très tôt chez l'enfant, une tumeur de la rétine généralement mortelle. La chirurgie permet

maintenant, par l'ablation de l'œil de prévenir la mort de l'enfant et la transmission de l'allèle est donc possible.

Cette sélection qui favorise les phénotypes extrêmes est appelée **sélection directionnelle**. Elle intervient lorsque le milieu change progressivement dans une direction particulière (bioclimat chaud → froid par ex.). Elle change la population dans une direction.

ou $\delta_1 \leq \delta_2 \leq \delta_3$: Les homozygotes récessifs sont éliminés et la moyenne de la population sera graduellement déplacée vers le phénotype dominant

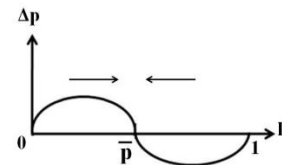
$\delta_1 \geq \delta_2 \geq \delta_3$: Ce sont les dominants qui sont éliminés et la moyenne est déplacée vers le phénotype récessif.

- **Cas où l'hétérozygote est avantage** (la valeur adaptative des hétérozygotes est supérieure à celles des homozygotes, $\delta_3 < \delta_1 < \delta_2$).

Dans ce cas, il existe une valeur d'équilibre $\bar{p}$ pour laquelle p est différent de zéro et différent de 1. Cet équilibre est stable.

Si $p < \bar{p}$, $\Delta p > 0 \Rightarrow p$ augmente à chaque génération

Si $p > \bar{p}$, $\Delta p < 0 \Rightarrow p$ diminue à chaque génération



Quelque soit la valeur de p , la sélection tend à la modifier dans la direction de la valeur d'équilibre $\bar{p}$.

L'anémie falciforme (drépanocytose) est un exemple classique de cette situation. Cette maladie est déterminée par un allèle semi-dominant. Il y a modification des acides aminés de la chaîne β de l'hémoglobine.

AA → normaux ($Hb^S Hb^S$)

aa → anémie létale ($Hb^S Hb^S$)

Le génotype Aa ($Hb^S Hb^s$) présente des hématies qui changent de forme lorsque la pression en O_2 diminue. Ils prennent la forme en croissant, en faucille. L'étude des populations humaines a montré que la fréquence de l'allèle défavorable très faible dans les populations nord-

européennes (les homozygotes sont contre sélectionnés) devient relativement élevée dans certaines populations d'Afrique noire, des Indes et du pourtour de la méditerranée. On note en outre, une forte corrélation entre la fréquence de cet allèle et la distribution géographique de la malaria. On a montré que les hétérozygotes sont moins sensibles à la malaria que les homozygotes normaux AA. Ainsi, dans les régions où sévit la malaria, les hétérozygotes résistent à cette maladie alors que les homozygotes ($Hb^S Hb^S$) et ($Hb^S Hb^S$) ne sont pas résistants. Il y a là, avantage sélectif des hétérozygotes, avantage qui assure la conservation de l'allèle nocif Hb^S .

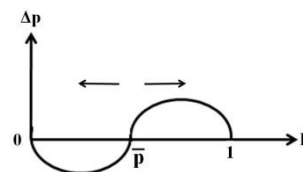
Parmi la population noire d'origine africaine des USA, on trouve la fréquence de l'allèle responsable de l'anémie falciforme dix fois moins élevée ($q=0.04$). La malaria n'existant pas dans cette région, les hétérozygotes ont perdu leur avantage de sélection. L'avantage sélectif dépend fortement des conditions du milieu. Selon que la malaria sévit ou non, les hétérozygotes sont avantagés ou légèrement désavantagés par rapport aux individus normaux. Ce type de sélection maintient dans la population l'allèle létal. C'est la **sélection stabilisatrice**. Elle agit lorsque le milieu change peu ou pas. Elle élimine les mutants mal adaptés et s'oppose à leur diffusion. Elle a pour effet de maintenir une population génétique constante parce qu'elle favorise les phénotypes moyens ou normaux et élimine les individus extrêmes. Elle favorise le polymorphisme en **équilibre balancé**.

- Cas où l'hétérozygote est désavantagé ($\delta_3 > \delta_1 > \delta_2$)

Dans ce cas, on observe également un changement du signe du numérateur M donc de Δp en fonction de p. il existe un état d'équilibre où $\Delta p=0$ et p et q sont différents de 1. Cet **équilibre est instable**.

$$p < \bar{p} \Rightarrow \Delta p = p' - p < 0 \Rightarrow p \text{ diminue}$$

$$p > \bar{p} \Rightarrow \Delta p = p' - p > 0 \Rightarrow p \text{ augmente}$$



Dès que la valeur de p s'écarte du point d'équilibre, la pression de sélection va s'écarter de plus en plus et cela entraîne la fixation de l'un des deux allèles ; il y aura diminution de la variabilité génétique. Cette situation peut s'appliquer aux systèmes d'incompatibilité des groupes sanguins ABO et *Rhesus*. Le facteur *Rhesus* est déterminé par deux allèles, (A) responsable de la synthèse d'un antigène de surface de l'hématie dit (Rh^+) et (a) absence de cet antigène. L'allèle A est

dominant et l'allèle a est récessif. Les génotypes (AA) et (aa) correspondent au phénotype $[Rh^+]$. Les individus $[Rh^-]$ n'ont pas d'anticorps contre les antigènes *Rhesus* mais ils peuvent les synthétiser dès qu'ils sont en contact avec les hématies $[Rh^+]$. Ainsi, une femme au génotype (aa) c-à-d au phénotype $[Rh^-]$ avec un homme au génotype (Aa) c-à-d au phénotype $[Rh^+]$ peut engendrer un enfant (Aa) au phénotype $[Rh^+]$ à la première grossesse qui se passe normalement. Une petite quantité d'Ag $[Rh^+]$ passe du fœtus vers la mère au moment de l'accouchement et stimule la production d'anticorps. A la deuxième grossesse, ces anticorps augmentent et provoque chez le bébé une maladie hémolytique du nouveau-né, anémie souvent fatale. Actuellement, ces accidents peuvent être prévus. C'est un mécanisme de sélection contre les hétérozygotes. Mais on n'a pas noté de fixation d'un des deux allèles. Cette sélection est dite **sélection diversifiante**. Elle fractionne une population homogène en plusieurs types d'adaptations différentes. Elle contribue à maintenir la variation dans la population. L'établissement et la survie de ces différents types requièrent un milieu hétérogène c-à-d formé de plusieurs niches écologiques.

La sélection naturelle est donc une force évolutive puissante. L'efficacité de la sélection dépend du milieu en ce sens que c'est le milieu qui fait la différence entre les génotypes appropriés et les non adaptés. C'est également le milieu qui oriente la sélection vers la forme qui convient :

- ~ elle est stabilisatrice lorsque le milieu ne change pas
- ~ si le milieu change, elle devient directionnelle
- ~ enfin, s'il offre une variété de niches écologiques, elle peut devenir diversifiante

Coefficient de sélection

Ce coefficient indique l'intensité de la sélection qui s'exerce sur un génotype. C'est l'inverse de la valeur sélective ou valeur adaptative. Il s'exprime par la relation $S = 1 - \delta$. Si la valeur sélective exprime le degré d'avantage d'un génotype, le coefficient de sélection exprime le degré de désavantage.

Si :

	AA	Aa	aa
δ	1	0.85	0.60
S	0	0.15	0.40

Examinons divers cas possibles pour un locus A/a

Premier cas : l'allèle A dominant, l'allèle a récessif et la sélection est dirigée contre le récessif

	AA	Aa	aa
Valeur sélective (δ)	1	1	1
Fréquence	p^2	$2pq$	q^2
δ après sélection	1	1	1-S
Fréquence après sélection	p^2	$2pq$	$q^2 (1-S)$

Après sélection, la fréquence q est :

$$q_1 = [q^2 (1-S) + pq] / [p^2 + 2pq + q^2 (1-S)] = [q^2 (1-S) + pq] / [1 - Sq^2]$$

La variation de fréquences $\Delta q = q_1 - q = [-Sq^2 (1-q)] / [1 - Sq^2]$

La variation de fréquence (Δq) dépend du coefficient de sélection (S) et de la fréquence initiale de l'allèle a (q). Ce fait, nous le constatons dans tous les cas examinés.

Deuxième cas : sélection contre le dominant

	AA	Aa	aa
Valeur sélective (δ)	1	1	1
Fréquence	p^2	$2pq$	q^2
δ après sélection	$1-S$	$1-S$	1
Fréquence après sélection	$(1-S)p^2$	$(1-S)2pq$	q^2

En procédant comme précédemment, on trouve

$$\Delta q = q_1 - q = [Sp^2 (1-q)] / [(1-S) (1-p^2)]$$

L'efficacité de la sélection est déterminée par Δq . Si Δq est grand, la sélection est efficace. Δq est fonction de deux paramètres : fréquence initiale de l'allèle a (q) et la valeur du coefficient de sélection (S).

Application :

Nous allons examiner cette efficacité dans trois populations avec deux valeurs du coefficient de sélection (S) différentes.

Population I : $q=0.9$, population II : $q= 0.5$, population III : $q=0.1$

Avec dans le premier cas $S=0.2$, et dans le deuxième cas $S=0.8$, et dans le cas de dominance avec S appliqué aux récessifs.

$$\Delta q = q_1 - q = [-Sq^2 (1-q)] / [1-Sq^2]$$

1^{er} Cas : $S=0.2$ pop I, $\Delta q= -0.0193$
 Pop II, $\Delta q= -0.025$
 Pop III, $\Delta q=-0.0018$

2^e Cas : $S=0.8$ pop I, $\Delta q= -0.184$
 Pop II, $\Delta q= -0.125$
 Pop III, $\Delta q=-0.007$

Remarques

> Lorsque S est faible (premier cas), la sélection est plus efficace lorsque la fréquence d'un allèle est moyenne. Δq est le plus grand lorsque $q=0.5$

> Lorsque le coefficient de sélection (S) a une valeur importante, la sélection est très efficace quand la fréquence allélique (q) est élevée mais elle devient très rapidement de plus en plus faible. Ce dernier type de sélection est celui que l'on peut appliquer à une population que l'on veut améliorer en retirant les récessifs par exemple. Très rapidement, cette opération n'aura qu'un intérêt très limité.

Nombre de générations, nécessaires pour obtenir une variation de fréquence donnée

Mise au point pour améliorer une population animale ou végétale. Supposons que l'on applique un coefficient de sélection (S) aux récessifs. Au départ, la population comporte $p^2 + 2pq + q^2$ génotypes avec (q) comme fréquence initiale de l'allèle a. Après sélection totale des récessifs :

$$q_1 = pq / (p^2 + 2pq) = q / (1+q)$$

$$q_2 = p_1 q_1 / (p_1^2 + 2p_1 q_1) = q_1 / (1+q_1) = [q / (1+q)] / [1 + q / (1+q)] = q / (1+2q)$$

$$\text{par récurrence : } q_n = q / (1+nq)$$

la même formule permet de déterminer le nombre de générations nécessaires pour passer d'une valeur q à une valeur q_n : $n = 1/q_n - 1/q$

Ex. La fréquence d'un allèle récessif étant de 0.001 dans une population ; combien de générations seront nécessaires pour obtenir une fréquence égale à 0.0005 si $S=1$?

$$n = 1/0.0005 - 1/0.001 = 1000 \text{ générations}$$

NB :

Les allèles récessifs étant souvent à des fréquences de l'ordre de 10^{-5} , il devient impossible d'envisager d'extirper un allèle récessif d'une population.

4.3. Effet combiné des mutations et de la sélection

Dans la majorité des cas, un allèle mutant est soumis à un coefficient de sélection. De ce fait l'allèle sauvage demeure dans la population. La mutation et la sélection ayant des effets inverses, il est évident qu'un équilibre doit s'établir. Nous allons considérer trois cas :

Premier cas : $A > a$, (S) appliqué aux récessifs $A \rightarrow a$ avec un taux de mutation (μ) et $a \rightarrow A$ avec un taux de mutation (ν). Lorsque l'équilibre est atteint,

$$\Delta q = q_1 - q = \mu p - \nu q = [-Sq^2 (1-q)] / [1-Sq^2] = 0$$

$$\Rightarrow q = \sqrt{\mu/S}$$

Deuxième cas : $A = a$ (codominance, S appliqué aux homozygotes (aa).

$$\mu p - \nu q = [(-1/2 Sq) (1-q)] / (1-Sq)$$

après approximation, on obtient : $q = \mu/S$

Troisième cas : $A > a$, S appliqué à [A]

$$\text{A l'équilibre } \mu p = \nu q = (Sq^2 p) / [1-S(1-q^2)]$$

$$\bar{p} = 2\nu/S$$

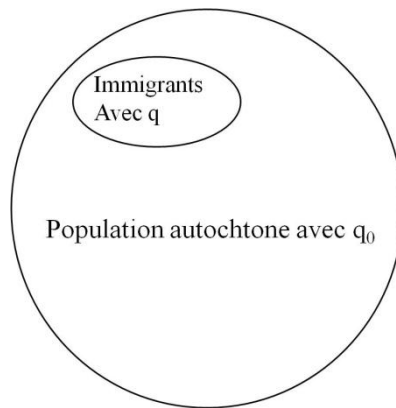
4.4. Migration

Dans une population, les individus peuvent migrer. Ils ont deux possibilités, soit constituer à eux seuls une population, soit s'intégrer dans une autre population. L'influence des migrations sur les fréquences alléliques dépend du taux d'immigration et des différences de fréquences alléliques entre la population d'origine et les immigrants. Dans le premier cas, les migrants emportent avec eux une partie du pool génétique. Les fréquences alléliques de ce groupe ne sont plus de ce fait identiques à celles de la population de départ : il y a dérive génétique (la variation imprévisible des fréquences liées à la faiblesse de l'effectif). De dérive génétique en dérive génétique, le pool de gènes se différencie de plus en plus d'autant que joue simultanément la mutation et la sélection. Ainsi, il peut arriver que les individus d'une population issue de la population de départ aient un génome tellement différent qu'ils ne peuvent plus se croiser avec les individus de la population mère. Il y a eu création d'une nouvelle espèce. C'est un phénomène de spéciation.

Taux d'immigration et fréquences alléliques

Dans le cas où les migrants s'intègrent dans une autre population, la fréquence allélique de cette population va varier. La variation dépend du taux d'immigration et de la différence entre les fréquences alléliques caractéristiques de la population immigrante et de celles de la population autochtone (accueillante).

Soit une population accueillant à chaque génération une proportion 'm' de nouveaux immigrants. Dans la population autochtone $m = \text{nombre d'immigrants} / \text{effectif total de la population} \Rightarrow$ le reste de la population accueillante est à égale à $1-m$. Soit 'q', la fréquence d'un certain allèle chez les immigrants et q_0 , sa fréquence chez les autochtones.



Après immigration, la fréquence de l'allèle considéré dans la nouvelle population est :

$$q_1 = mq + (1-m) q_0 = m (q - q_0) + q_0$$

$$\Delta q = q_1 - q_0 = m (q - q_0)$$

Ce modèle d'immigration correspond au peuplement d'une île. Si le taux d'immigration est constant de génération en génération, la population de l'île perd son individualité et devient comparable à celle du continent.

En comparant les fréquences $q_1, \dots, q_n$ de la population accueillante à celle des immigrants q , nous avons :

$$q_1 - q = m (q - q_0) + q_0 - q = m (q - q_0) - (q - q_0)$$

$$= (q - q_0) (m - 1) = (q_0 - q) (1 - m)$$

$$q_2 - q = (q_1 - q) (1 - m) = (q_0 - q) (1 - m) (1 - m) = (q_0 - q) (1 - m)^2$$

$$q_n - q = (q_0 - q) (1 - m)^n$$

Lorsque $q_n = q$, l'île perd son individualité.

Application :

Chez les Noirs américains, la fréquence de l'allèle R_0 du *Rhesus* est de 0.45. Chez les Africains, elle est de 0.63 et chez les Européens, elle est de 0.03.

- Calculer la fréquence théorique de R_0 chez les Noirs américains sachant que $m = 0.03$ depuis 10 générations.
- Quelle est la proportion des gènes d'origine africaine chez les Noirs américains ?
- Quelle est la proportion de gènes provenant des Européens ?

Soit la formule suivante

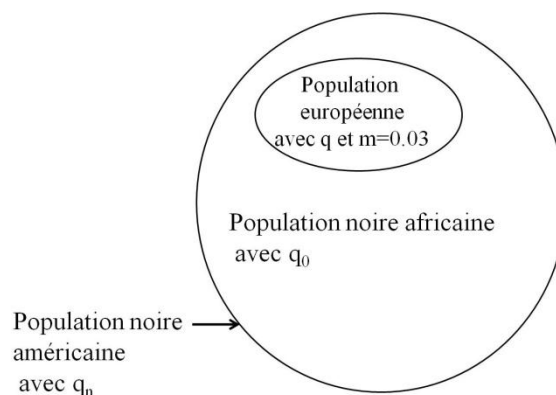
$q_n - q = (q_0 - q) (1 - m)^n$, où :

Noirs africains $\rightarrow q_0 =$ fréquence chez les autochtones,

Noirs américains $\rightarrow q_n$,

Européens $\rightarrow q =$ fréquence chez les immigrants. Ce sont les gènes européens qui ont migré vers la population noire.

$$q_n = (q_0 - q) (1 - m)^n + q = (0.63 - 0.03)(1 - 0.03)^{10} + 0.03 = 0.47$$



La proportion des gènes d'origine africaine chez les Noirs américains est de :

$$(1-m)^{10} = (0.93)^{10} = 0.73 = 73\%$$

La proportion des gènes provenant des Européens est égale à la différence soit 27%.

5. Spéciation

5.1. Concepts et définition de l'espèce

L'espèce est l'unité de base du système de la hiérarchie linnéenne de la taxonomie (science qui étudie les normes et les règles de la systématique). Les taxonomistes sont en permanence confrontés aux difficultés d'identification et de dénomination des espèces. D'une façon générale, le problème est de savoir si la catégorie 'espèce' peut être définie objectivement. Historiquement, trois conceptions générales de l'espèce ont successivement prévalu dans les milieux scientifiques :

- Le concept typologique de l'espèce :

Il remonte aux temps des philosophes Platon, Aristote et celui de Linné et de ses successeurs. La conception typologique de l'espèce correspond à une description effectuée sur un petit nombre d'individus prélevés en une même localité. Selon cette conception, les individus d'une même espèce sont considérés comme relativement **uniformes** et conformes au **type décrit**.

D'après ce concept, la diversité que l'on observe dans l'univers reflète l'existence d'un nombre limité de 'phénomènes universels' de base ou **types**. Les individus d'une même espèce n'ont pas de relation particulière les uns par rapport aux autres. Donc, la similitude morphologique est pour l'essentialiste, le critère de l'espèce. C'est ce qu'on nomme « concept morphologique de l'espèce ». Les caractéristiques morphologiques fournissent des faits valables pour la détermination du statut de l'espèce. Il est cependant, complètement différent d'utiliser le degré de différenciation morphologique comme critère essentiel pour déterminer le statut de l'espèce.

- Le concept nominaliste de l'espèce

Les nominalistes nient l'existence de phénomènes universels « réels ». Pour eux, seuls existent les individus, alors que les espèces sont des abstractions créées par l'homme. D'après eux les espèces ont été inventées pour pouvoir se référer collectivement à de grands nombres

d'individus. Ce concept nominaliste de l'espèce a été populaire en France, au dix-huitième siècle et conserve encore des adhérents de nos jours.

- **Le concept biologique de l'espèce**

Ce concept insiste sur le fait que :

~ Les membres d'une espèce forment une **communauté reproductive**. Les individus d'une espèce se comportent vis-à-vis des autres membres de la même espèce comme des conjoints potentiels et se recherchent les uns les autres dans le but de se reproduire.

~ L'espèce est également une **unité écologique** qui, quelque soit les individus qui la composent, interagissent en tant qu'unité avec les autres espèces qui partagent son environnement.

~ L'espèce est une **unité génétique** constituée d'un pool de gènes intercommuniquants, tandis qu'un individu n'est rien d'autre qu'un simple vaisseau temporaire transportant une petite fraction du contenu du pool de gènes pendant une courte période de temps.

La définition de l'espèce qui résulte de ce concept est la suivante : l'espèce est un groupe de populations naturelles capables d'inter croisements et qui sont reproductivement isolées d'autres groupes semblables. Une espèce est un pool de gènes protégé. Une espèce possède ses propres moyens (mécanismes d'isolement) pour se protéger contre le courant génique nuisible en provenance d'autres pools de gènes.

5.2. Mécanismes d'isolement reproducteur

Ce sont des mécanismes qui suppriment ou ne permettent, normalement, que très modérément les échanges génétiques entre les différents membres d'espèces différentes. De telles barrières ont pour résultats de maintenir l'intégrité spécifique. Ces mécanismes (barrières) sont de deux types : les barrières internes et les barrières externes.

- **Les barrières externes** : elles correspondent à des mécanismes qui empêchent les croisements interspécifiques

~ **Isolement géographique** : la simple séparation dans l'espace aboutit à une suppression des échanges entre espèces.

~ **Isolement par l'habitat** : ou isolement écologique. Il provient du fait que les individus appartenant à deux espèces différentes ont des préférences écologiques distincts. Un bon exemple est celui de deux espèces voisines de violette. *Viola arvensis* et *V. tricolor* qui peuvent donner expérimentalement des hybrides fertiles mais demeurent toutefois distinctes dans la nature à cause de leur préférence écologiques ; *V. arvensis* étant calcicole et *V. tricolor* silicole.

~ **Isolement saisonnier** : on parle d'I. S., lorsque les périodes de reproduction de deux espèces ne coïncident pas au cours de l'année. L'exemple classique est celui des deux espèces voisines de laitue (*Lactuca canadensis* et *L. graminiflora*) qui produisent des hybrides artificiels viables et fertiles mais ne fleurissent pas à la même époque de l'année. *L. graminiflora* ayant une floraison printanière précoce et l'autre fleurissant en été.

~ **Isolement éthologique** : de comportement (*ethos*=habitude, coutume). Il y a restriction à l'accouplement au hasard. Il est très répandu chez les animaux à psychisme développé. Il est fondé sur l'interaction entre *stimulus* provenant des partenaires des deux sexes ; ces stimuli étant visuels, auditifs, olfactifs ou tactiles. Exemple de deux espèces de grenouilles *Microhyla olivacea* et *M. carolinensis*. Leurs chants de cour sont très distincts. Et il n'existe pratiquement pas de mécanisme autre que la voix permettant la séparation des groupes. D'autres *stimuli* tels que des signaux chimiques peuvent être à l'origine des mécanismes d'isolement ; chaque signal étant plus ou moins spécifique.

~ **Isolement mécanique** : très efficace chez les plantes. Les loges du stigmate ont des tailles et des formes qui varient selon l'espèce. Ce qui constitue des barrières à la pollinisation par les insectes. Chez certaines fleurs la fécondation peut échouer parce que le pollen d'une espèce à style court ne peut développer un tube pollinique suffisamment long dans les espèces à longs styles ou parce que cette croissance y est ralentie ou empêchée.

- **Les barrières internes** : ce sont des mécanismes qui réduisent le succès de certains croisements interspécifiques.

~ **Empêchement de la fécondation** : il consiste en une inviabilité des gamètes d'une espèce au contact des voies génitales ou du style d'une autre espèce. Chez de nombreuses drosophiles, les spermatozoïdes d'une espèce étrangère provoquent dans le tractus génital de la femelle réceptrice une réaction antigénique. Chez les polyploïdes, les tubes polliniques sont de grand diamètre et cheminent difficilement dans le style des espèces diploïdes.

~ **Mortalité du zygote** : dans ce cas, il y a fécondation. Chez le genre *Rana* (grenouille), l'arrêt du développement des hybrides survient selon les croisements à différentes phases critiques au cours de la période embryonnaire. Dans le genre *Datura* (solanacées), l'embryon cesse son développement quand il atteint le stade 8 cellules. Les causes principales de la mortalité des hybrides sont multiples. Dans certains cas, l'albumen se comporte comme un inhibiteur du développement embryonnaire. Il peut y avoir un gène létal qui n'a aucune influence sur les individus de l'espèce mais entraîne la mort de l'hybride. L'embryon n'arrive jamais à maturité à cause d'une désharmonie génétique entre les génomes parentaux. L'inviabilité des hybrides résulte de l'interaction de certains génotypes hybrides avec le cytoplasme de l'une des espèces parentales.

~ **Stérilité des hybrides** : l'exemple le plus célèbre de stérilité des hybrides de première génération est celui du mulet. Cet animal est particulièrement résistant mais stérile. Cette stérilité des hybrides est un phénomène général. Il peut exister également une barrière aux échanges génétiques lors des générations ultérieures (cas du genre *Gossypium*). Il existe deux types de stérilité : génique et chromosomique. La stérilité génique se traduit par le fait que les organes reproducteurs n'arrivent pas à la méiose ou la présence d'anomalies méiotiques. La stérilité chromosomique résulte de la non homologie entre les chromosomes maternels et paternels de l'hybride.

5.2. Mécanismes de spéciation

Des mécanismes multiples peuvent intervenir et empêcher l'échange génique entre certains individus d'une même espèce. Ce qui permet au pool génétique harmonieusement équilibré (l'espèce) de se scinder en deux ensembles incapables d'échange génétique autrement dit en deux espèces différentes. La spéciation est donc la division d'une lignée phylétique. C'est un processus de multiplication des espèces. Il semble, à présent, que plusieurs modes de spéciation sont possibles selon Mayer (1942). On distingue l'évolution phylétique (échelle géologique), fusion, multiplication instantanée (par restructuration chromosomique, polyploïdie), multiplication graduelle (sympatrique, allopatrique) divergence progressive des populations jusqu'à la distinction spécifique.

- **Spéciation allopatrique ou géographique.** Selon Mayer (1942), une nouvelle espèce se forme quand une population est géographiquement isolée des autres populations de son espèce parente et acquiert durant la période d'isolement des caractères qui fournissent ou

garantissent l'isolement reproducteur (isolats marginaux, colonisation, différenciation des races géographiques).

- **Spéciation sympatrique** : d'après cette théorie, deux espèces peuvent s'individualiser à partir d'une seule sans passer par un stade d'isolement géographique. La division du pool de gènes est causée par des facteurs écologiques. Une spécificité poussée concernant l'alimentation augmenterait considérablement l'efficacité des mécanismes d'isolement reproducteur. Citons le cas de la mouche de la pomme (*Rhagoletis pomonella*). Fixés primitivement sur l'aubépine et éventuellement sur quelques rosacées, ces insectes envahirent les vergers 200 ans plus tard et durant les 100 années suivantes, une avance dans la période de maturité sexuelle des populations de mouches de la pomme par rapport aux populations de l'aubépine a pu être constatée.