

## Chapitre IV. Cause de la variabilité de la structure génétique des populations naturelles.

### 1. Définition de la variation de la fréquence allélique dans les populations

Dans la nature, les populations s'écartent toujours des conditions théoriques idéales. Les effectifs sont limités, fluctuent au cours du temps et les différents génotypes se distinguent par des capacités reproductives différentes.

De plus, les populations ne sont jamais génétiquement closes, car elles subissent l'apport de nouveaux allèles soit par migration soit par mutation.

Ces facteurs ont pour conséquence une variation de la fréquence des allèles au cours des générations successives.

Ces changements, appelés changements micro évolutifs ou microévolution, modifient la structure génétique de la population qui peut ou non évoluer vers une meilleure adaptation à son environnement.

L'analyse des conséquences génétiques de ces différents facteurs nécessite l'emploi de modèles mathématiques ou de simulations qui permettent de déterminer comment et sous quels facteurs les fréquences alléliques varient au cours du temps.

On recherche généralement les conditions pour lesquelles un équilibre peut être atteint, c'est-à-dire où les fréquences alléliques ne varient plus, bien que les populations restent sous l'action de ces différents facteurs (mutations, migrations, sélections).

Ce sont donc les conditions qui permettent un maintien de la variabilité génétique qui sont étudiées. L'étude de l'effet de ces différents facteurs sur la variation des fréquences alléliques et la recherche de cet équilibre se fait généralement de la façon suivante :

a. Mesure de la variation des fréquences alléliques entre 2 générations successives

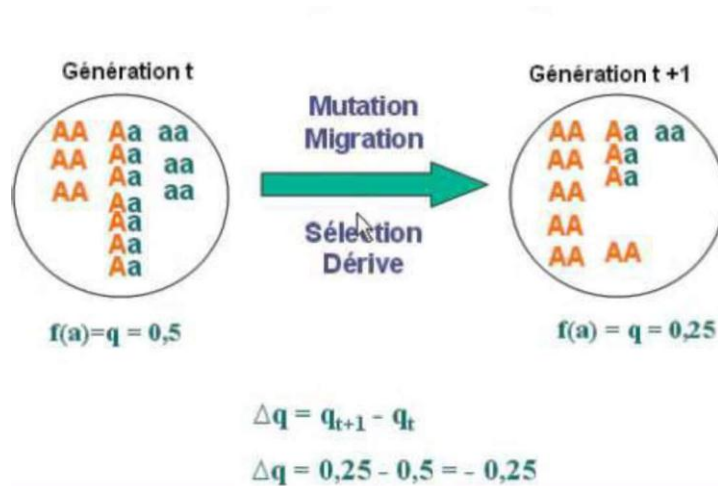
Pour l'allèle A  $\Delta p = p_{t+1} - p_t$

pour l'allèle a  $\Delta q = q_{t+1} - q_t$  et  $p_{t+1}$  et  $q_{t+1}$  étant respectivement la fréquence de l'allèle A aux générations t et t+1.

b. La recherche d'un équilibre, c'est-à-dire les conditions qui permettent d'annuler  $\Delta q$  ou  $\Delta p$

Pour l'allèle A  $\Delta p = 0$

Pour l'allèle a  $\Delta q = 0$



## 2. Les mutations

Les mutations sont à l'origine des variations génétiques observées dans les populations naturelles. Elles font apparaître de nouveaux allèles ou de nouvelles structures chromosomiques. Les mutations sont dans la plupart des cas récurrentes, car elles se produisent à chaque génération, mais leur fréquence est généralement faible de l'ordre de  $10^{-7}$  (Jean-Pierre, 2008). La fréquence avec laquelle les allèles d'un locus donné subissent une mutation à une génération donnée est appelée taux de mutation (Exemple tableau IV)

Tableaux IV : Les taux de mutation pour différentes espèces et

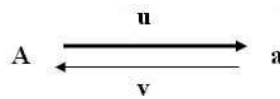
Espèce	Caractères Maladies	Taux de mutation	Mutation réverse
Homme	Achondroplasie	$1 \cdot 10^{-5}$	-
	Achondroplasie	$7 \cdot 10^{-6}$	-
	Hémophilie A	$5.7 \cdot 10^{-5}$	-
	Hémophilie B	$3 \cdot 10^{-6}$	-
	Myopathie de Duchène	$10.5 \cdot 10^{-5}$	-
	Myopathie de Duchène	$4.6 \cdot 10^{-5}$	-
Souris	Non agouti	$4.5 \cdot 10^{-5}$	$4.2 \cdot 10^{-6}$
	Albinos	$3.3 \cdot 10^{-5}$	-

Lorsque les différents états alléliques produits par mutation ne sont pas soumis à sélection (allèles neutres), l'évolution de la fréquence de ces allèles dépendra uniquement de leurs taux de mutation respectifs.

Lorsqu'un gène subit de manière récurrente une mutation le faisant passer de l'état allélique A vers l'état allélique a sans que la mutation inverse soit possible (mutation de a vers A), l'allèle a va progressivement remplacer l'allèle A dans la population.

Cette évolution sera cependant extrêmement lente, voire indétectable, compte tenu des taux faibles de mutation. L'allèle a sera par contre rapidement éliminé si il est à l'origine d'un désavantage pour l'individu qui le porte. Lorsque les mutations sont récurrentes et que les mutations réverses sont possibles, l'évolution des fréquences alléliques sera fonction des valeurs respectives des différents taux de mutations.

Dans le cas d'un gène à deux allèles A et a, les fréquences respectives sont p et q, et les taux de mutation sont respectivement u pour le passage de A vers a et v pour le passage de a vers A.



Si  $p_0$  et  $q_0$  sont les fréquences alléliques de A et a à la génération  $G_0$ , les fréquences  $p_1$  et  $q_1$  de ces allèles à la génération  $G_1$  seront les suivantes :

$$f(A) \quad p_1 = p_0 - u p_0 + v q_0$$

$$f(a) \quad q_1 = q_0 - v q_0 + u p_0$$

La variation  $\Delta q$  de la fréquence de l'allèle a entre les générations  $G_0$  et  $G_1$  est :

$$\Delta q = q_1 - q_0$$

$$\Delta q = -v q_0 + u p_0$$

A l'équilibre, les fréquences des allèles A et a ne varieront plus entre deux générations successives ( $p_{n+1} = p_n$  et  $q_{n+1} = q_n$ ). Ces valeurs d'équilibre, appelées  $p_e$  pour A et  $q_e$  pour a, seront obtenues lorsque  $\Delta q = 0$  c'est à dire :

$$\Delta q = -v q_e + u p_e = 0 \quad \Delta q = -v q_e + u (1 - q_e) = 0 \quad \Delta q = -v q_e + u - u q_e = 0 \quad \Delta q = u - q_e (u + v) = 0 \quad q_e = u / (u + v)$$

Et de la même façon, pour l'allèle A :

$$p_e = v / (u + v)$$

La valeur des fréquences alléliques à l'équilibre dépend donc uniquement des taux de mutation.

Si  $u = 10^{-6}$  et  $v = 10^{-7}$  (taux de mutation réalistes chez les eucaryotes), la fréquence de l'allèle A à l'équilibre sera  $p_e = 0.09$  donc  $q_e = 0.91$ .

Ainsi, si seul l'allèle A est présent à la première génération ( $p_0 = 1$ ), la fréquence de cet allèle diminuera au cours des générations pour atteindre la fréquence de 0.09 à l'équilibre.

Cependant, cette évolution ne se fera que très lentement puisqu'il faut à peu près 1000 générations pour faire passer  $p$  de 1 à 0.99.

Ainsi, le processus de mutation seul n'a pas d'effet important sur la structure génétique des populations car la variation des fréquences alléliques qu'il induit est extrêmement faible voire indétectable à l'échelle de quelques dizaines voire quelques centaines de générations. On peut démontrer que l'évolution des fréquences alléliques au cours du temps est :  $p_{n+1} = p_e + (p_0 - p_e)(1 - u - v)^n$

Les mutations sont cependant d'une importance capitale puisqu'elles génèrent des formes alléliques nouvelles dont les fréquences vont pouvoir varier rapidement sous l'action d'autres processus en particulier la sélection naturelle.

En résumé, les mutations font apparaître à chaque génération des allèles nouveaux, et en dernier ressort, ce sont elles les principales responsables de l'existence du polymorphisme. De ce point de vue, leur rôle est capital car elles déterminent le potentiel adaptatif des populations donc leur maintien à long terme. Par contre, les taux de mutation sont très faibles et la répétition des mutations à chaque génération ne modifie que très lentement les fréquences de ces allèles. A une échelle de temps de quelques dizaines à quelques centaines de générations, l'effet des mutations sur l'évolution des fréquences alléliques est donc négligeable

**3. Les migrations Dans la nature, les populations d'une même espèce ne sont pas génétiquement isolées.** Il se produit à chaque génération des échanges d'adultes, mais aussi d'embryons (graines chez les plantes), de pollen, ou de gamètes (gamètes des animaux aquatiques) qui sont autant d'échanges de gènes, appelés flux géniques.

Ces flux géniques sont généralement d'autant plus importants que les populations sont proches géographiquement. C'est le cas lorsqu'une population est subdivisée en sous-populations à cause d'une discontinuité de l'habitat ou des ressources. Ces échanges de gènes provoqués par la dispersion des individus entre sous populations limitent alors leur divergence génétique (Jean-Pierre, 2008).

Les conséquences génétiques de ces phénomènes migratoires dépendent de l'organisation spatiale des populations, de leur densité en individus et des nombreux facteurs

environnementaux responsables de leur dispersion, l'ensemble devenant rapidement complexe à analyser (Figure 7)

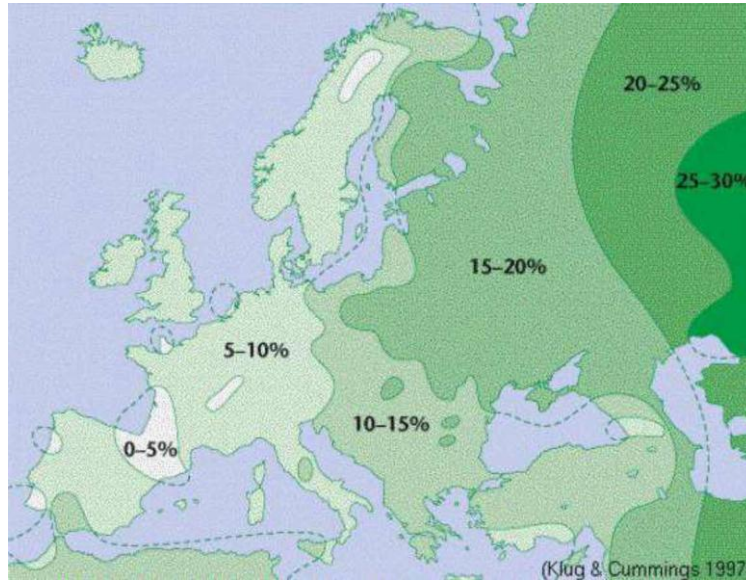


Figure: **Migration chez l'homme. Diffusion de l'allèle B du système ABO en Europe de l'Ouest suite au flux migratoire venu de l'Est.**

### 3.1. Le modèle insulaire

Le modèle de migration le plus simple, appelé modèle insulaire (Figure 8), est celui où les échanges d'individus s'effectuent dans un seul sens avec un flux de gènes unidirectionnel entre une population 1 de grande taille (le continent) et une population 2 de petit effectif (l'île). C'est le cas d'une île proche d'un continent, d'une étendue d'eau située en aval d'un lac, ou d'une minorité isolée du reste de la population sur une base ethnique ou socio-culturelle.

Dans tous ces cas, la population 2 (l'île) est constituée de résidents dont le taux d'émigration peut être négligé alors que la population 1 (le continent) envoie à chaque génération une proportion donnée de migrants. Si à chaque génération, la population 2 d'effectif  $N$  reçoit  $n$  individus migrants de la population 1, le flux migratoire  $m$  correspond à la proportion de migrants qui arrivent dans la population résidente :  $m = n / (N + n)$

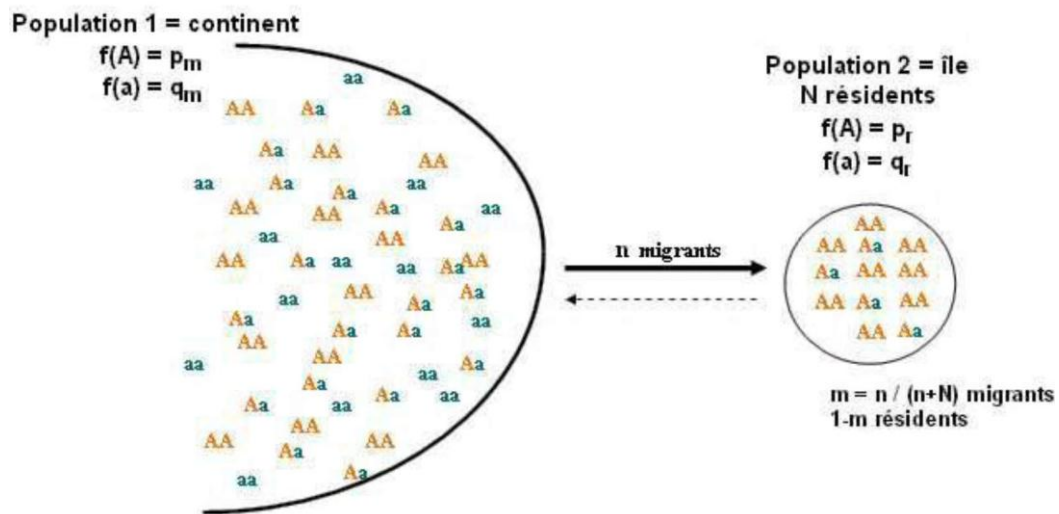


Figure : Modèle insulaire (île-continent).

A chaque génération, il y a une proportion  $m$  d'individus migrants et une proportion  $1-m$  d'individus résidents.

Si à une génération  $G_0$  donnée les deux populations ont des fréquences différentes pour les allèles d'un même gène, par exemple  $p_m$  pour un allèle  $A$  dans la population de migrants et  $p_0$  pour ce même allèle dans la population qui reçoit ces individus (île), la nouvelle fréquence de  $A$  à la génération suivante  $G_1$  après l'arrivée de  $m$  migrants sera:

$$p_1 = (1 - m) p_0 + m p_m$$

La variation de la fréquence allélique  $\Delta p$  de l'allèle  $A$  entre ces 2 générations sera alors :

$$p_1 - p_0 = m (p_m - p_0)$$

Explication :

$$p_1 - p_0 = (1 - m) p_0 + m p_m - p_0$$

$$p_1 - p_0 = p_0 - m p_0 + m p_m - p_0$$

$$p_1 - p_0 = m p_m - m p_0$$

$$p_1 - p_0 = m (p_m - p_0)$$

Soit entre les générations  $t$  et  $t+1$  :

$$p_{t+1} - p_t = m (p_m - p_t)$$

La variation des fréquences alléliques est donc proportionnelle au flux migratoire  $m$  et à la différence de fréquence entre les deux populations. L'équilibre  $p_e$  (ou  $q_e$ ) sera atteint lorsque  $\Delta p = 0$  c'est-à-dire lorsque les fréquences alléliques dans la population qui reçoit de nouveaux individus (l'île) seront les mêmes que celles de la population immigrante (le continent).

$$p_e = p_m$$

$$q_e = q_m$$

Après  $t$  générations de migrations, il y a une proportion  $(1-m)^t$  d'individus résidents, où la fréquence de A est  $p_0$ , et  $1-(1-m)^t$  migrants, où la fréquence de A est  $p_m$ . La fréquence allélique à la génération  $t$  dans l'ensemble de la population sera :

$$p_t = (1-m)^t p_0 + (1 - (1-m)^t) p_m$$

C'est-à-dire :

$$p_t = (1-m)^t (p_0 - p_m) + p_m$$

Lorsque les deux sens de migrations existent, les fréquences des allèles dans les deux populations seront fonction des fréquences initiales et de leurs taux de migration respectifs.

### 3.2. Autres modèles de migration

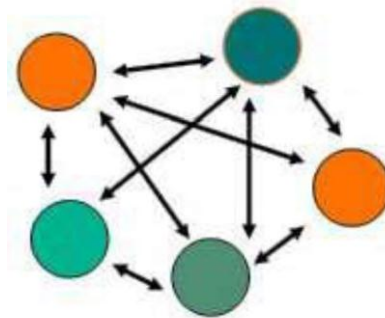
Lorsque plus de deux populations sont partiellement isolées, les échanges de migrants sont multidirectionnels et l'analyse des conséquences des flux géniques se complexifie rapidement. Plusieurs types de modèles de migration ont été proposés.

#### 3.2.1. Le modèle de l'archipel

Correspond à un ensemble d'îles interconnectées (Figure 9) par des échanges de migrants dans toutes les directions possibles avec des conséquences sur l'ensemble des populations.

Sous l'effet de ces flux géniques multidirectionnels, les différences de fréquences alléliques entre populations diminuent progressivement, et les populations convergent vers une fréquence allélique commune qui correspond à la moyenne des fréquences alléliques dans les populations concernées, pondérée par leurs effectifs respectifs.





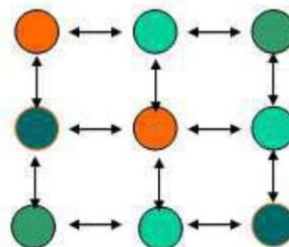
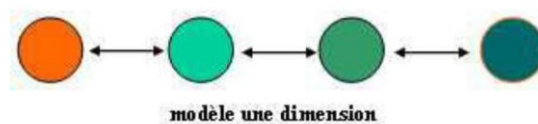
**Figure: Modèle archipel.**

### 3.2.2. Le modèle linéaire

Appelé aussi modèle en pas japonais (stepping stone model), correspond au cas où les échanges entre populations sont « canalisés » par une structure particulière de l'habitat.

C'est ce qui se passe dans une rivière, où les populations de poissons sont échelonnées de l'amont vers l'aval, les poissons de la population amont ne pouvant atteindre la population aval qu'en passant par les populations intermédiaires.

Dans ces conditions, l'intensité des échanges migratoires est directement corrélée à la distance géographique entre les populations

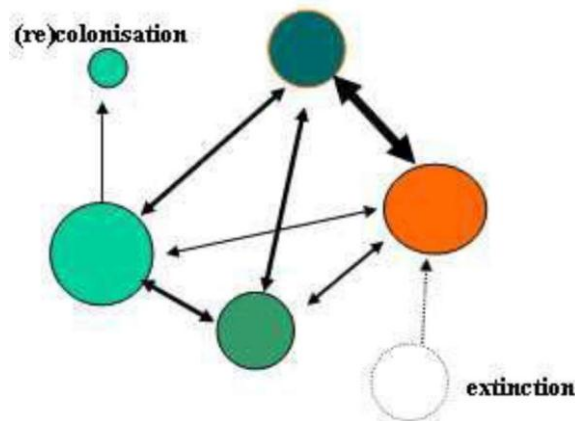


**Figure: Modèle de migration en pas japonais (stepping stone model).**

### 3.2.3. Le modèle en méta-populations

Est certainement le plus réaliste puisqu'il prend en compte une taille inégale des populations, des flux migratoires de différente intensité et surtout la possibilité que certaines populations disparaissent (extinction) alors que d'autres sont nouvellement créées par colonisation.

L'analyse du fonctionnement en métapopulation doit alors incorporer les effets de la variation des effectifs (dérive génétique) (Figure 11).



**Figure: Modèle en métapopulation.**

### 3.3. Estimation des taux de migrations

Contrairement aux taux de mutation qui sont très faibles, les taux de migration peuvent être élevés et leurs effets sont souvent importants et rapides. Les migrations tendent à uniformiser les fréquences alléliques des populations concernées, et ainsi s'opposent aux phénomènes qui tendent au contraire à les diversifier (dérive génétique ou sélection). La différenciation réelle que l'on observe entre les populations traduit donc le compromis qui s'établit entre ces processus. La mesure directe des flux migratoires n'est possible que par le repérage ou le

marquage des individus, et n'est réalisable que chez des animaux de grande taille (principalement les vertébrés y compris l'homme).

Chez la plupart des organismes, les flux géniques ne peuvent qu'être estimés par des mesures indirectes grâce à l'utilisation de marqueurs génétiques (biochimiques ou moléculaires) dont l'expression phénotypique est la plus faible possible (marqueurs neutres).

Plus les fréquences alléliques de ces marqueurs sont différentes entre populations, moins les flux géniques sont intenses et moins les flux migratoires sont importants.

Des modèles permettent d'estimer les flux migratoires (en nombre d'individus par génération) à partir des structures génétiques des populations (variance des fréquences alléliques entre populations, statistique  $F_{st}$ ). En multipliant les locus examinés, on peut rechercher et analyser les déséquilibres gamétiques, ce qui constitue une méthode encore plus puissante. C'est grâce à ces méthodes qu'il est possible d'évaluer l'impact génétique des introductions volontaires sur les populations indigènes (alevinages sur les populations de poissons dans les rivières), ou d'étudier les déplacements naturels de populations sur de très grandes distances. En résumé, les migrations sont à l'origine de flux géniques importants entre populations. Quel que soit le modèle considéré (modèle insulaire, archipel, en pas japonais, métapopulations), ces flux géniques tendent à uniformiser les fréquences alléliques des populations interconnectées.

Ils s'opposent donc aux facteurs qui font diverger la composition génétique des populations (dérive génétique). Lorsque le repérage direct des individus n'est pas possible, la mesure des flux migratoires se fait indirectement en utilisant des modèles génétiques qui considèrent que les flux géniques entre populations sont d'autant plus faibles que la différenciation entre ces populations est plus forte.

#### **4. La sélection**

Le modèle de Hardy-Weinberg suppose que l'effet de la sélection est négligeable sur quelques générations, une condition évidemment irréaliste sur une grande échelle de temps puisque la sélection, avec d'autres forces y fut responsable de l'évolution du vivant.

Charles Darwin fut le premier des naturalistes à concevoir la sélection naturelle comme une force active de la transformation (l'évolution) des espèces. La pensée darwinienne est populationnelle car l'évolution des espèces est conçue non comme une adaptation ou une transformation des individus sous les contraintes de l'environnement, mais comme le tri, au sein des populations, par la sélection naturelle, des individus qui sont le mieux adaptés à ses contraintes, et qui, de ce fait, laissent un plus grand nombre de descendants. Cette fécondité différentielle induit une modification de la diversité (génétique) des populations vers une augmentation de la fréquence des types les mieux adaptés, et aboutit, après une longue période, à une transformation de l'espèce.

À la même époque, un autre naturaliste, Alfred Russel Wallace, vivant aux Indes, avait, indépendamment de Darwin, exprimé les mêmes conceptions. Il s'en était d'ailleurs ouvert à Darwin lui-même qui se dépêcha d'écrire un essai afin de publier ses conceptions conjointement avec celles de Wallace, dans le même numéro du journal of the Linnean society. Darwin poursuivit et approfondit sa réflexion et, par sa position en Angleterre, s'imposa comme le seul fondateur de la théorie de l'évolution par la sélection naturelle (Jean-Pierre, 2008).

Le fait que les concepts développés par Darwin aient pu surgir à la même époque chez un autre naturaliste signifie tout simplement que les esprits étaient mûrs pour reconsidérer complètement la problématique transformiste en dépassant les conceptions économiques, philosophiques et théologiques du XVIIIème siècle. Les conceptions darwiniennes apparaissent, pour tout historien des idées, comme un transfert, dans les sciences naturelles, de la pensée libérale, économique et philosophique, qui accompagna le développement

industriel et commercial du début du XIXème siècle. Que Darwin, à des passages clé de son ouvrage, fit explicitement référence à Malthus n'empêcha nullement Marx et Engels de considérer la théorie darwinienne comme de la plus grande importance sur le plan philosophique, révolutionnaire au sens Hégélien du terme. Cependant la sélection telle que l'a conçue Darwin, ne résout pas vraiment toute la question de l'évolution des espèces. D'une part la formulation mathématique

de la sélection, si elle rend compte de l'évolution

darwinienne de la diversité des populations par la « sélection du plus apte », apporte aussi des résultats inattendus dans le cadre strict du darwinisme, dont la possibilité de maintenir cette diversité. D'autre part cette formulation montre qu'une population soumise à une contrainte sélective de l'environnement peut évoluer dans plusieurs directions, parfois opposées, ce qui pose le problème de la force qui « exerce le choix évolutif ».

Enfin le coût de la sélection est incompatible avec l'importance de la diversité observée dans les populations naturelles, ce qui implique qu'une grande part de cette diversité est sélectivement neutre, mais que sa variation participe obligatoirement à la définition des espèces nouvelles ; la question se pose alors de définir les forces qui participent à la variation de la diversité sélectivement neutre (Serre, 2006).

## 5. La drive génétique

Est l'évolution d'une population ou d'une espèce causée par des phénomènes aléatoires, impossible à prévoir. Du point de vue génétique, c'est la modification de la fréquence d'un allèle, ou d'un génotype, au sein d'une population, indépendamment des mutations, de la sélection naturelle et des migrations. La drive génétique est causée par des phénomènes aléatoires et imprévisibles, comme le hasard des rencontres des spermatozoïdes et des ovules, dans le cas d'une reproduction sexuée. Les effets de la drive génétique sont d'autant plus importants que la population est petite, car les écarts observés par rapport aux fréquences alléliques y seront d'autant plus perceptibles. Cette situation peut se produire au moment de l'apparition d'une espèce, ou après un goulot d'étranglement (quand une grande partie d'une espèce a disparu, à la suite de phénomènes épidémiques ou d'une crise climatique ou d'une catastrophe par exemple). Elle peut aussi survenir dans une situation d'insularisation écologique vraie (insularisation causée par une montée de la mer) ou due à la fragmentation écologique. La drive génétique concerne tous les allèles même si l'impact sur les allèles neutres (c'est-à-dire qui ne confèrent ni avantages ni inconvénients) est plus conséquent.

Malgré tout un allèle favorable peut disparaître ou un allèle défavorable se fixer dans une population par drive, ce qui est fréquent pour des populations aux tailles très réduites.

### 5.1. La notion d'effectif efficace

Dans les grandes populations, les variations (liées au hasard) du nombre d'enfants produits par des individus de génotypes différents, n'ont pas d'effet significatif sur la fréquence des gènes.

Dans les petites populations, ces variations peuvent avoir un effet considérable :

-Si un gène particulier n'est retrouvé que chez un petit nombre d'individus, si ces individus n'ont pas d'enfants ou, que par chance (hasard), ces enfants n'héritent pas de ce gène, le gène en question va complètement disparaître de la population (éteint : fréquence = 0) et son allèle va devenir fixé (fréquence = 1).

La part de la dérive génique aléatoire dépend de la taille de la population.

Elle est plus grande dans les petites populations où les variations dans la fréquence des gènes peuvent être considérables d'une génération à l'autre.

#### Démonstration :

Influence de la dérive aléatoire sur les fréquences géniques, soit  $N_e$  l'effectif de la population des reproducteurs. La nouvelle génération est formée à partir d'un échantillon de  $2N_e$  gamètes. Soient  $p$  et  $q$  la proportion des 2 allèles à la première génération.

La loi de probabilité de  $p$  est une loi binomiale de variance

$V_p = pq/2N_e$  et  $\sigma$  (écart-type) = racine carrée de  $V_p$ .

Exemple :

$N_e = 50$  et  $p = 0.5$

$\sigma^2 = 0.5 \times 0.5 / 100 = 1/400$

$\sigma = 1/20$

A la génération suivante,  $p$  sera compris entre 0.40 et 0.60 (+/- 2 écart-types) avec une probabilité de 95%.

A la génération suivante, si  $N = 50$  et  $p = 0.4$   $\sigma^2 = 0.4 \times 0.6 / 100 = 0.24 / 100 = 1/400$   $\sigma = 0.05$   $p$  sera compris entre 0.3 et 0.5. Dans un tel système, il n'y a que 2 états stables :  $p = 0$  et  $p = 1$ . Sous l'influence de la dérive, une population évoluera vers l'un de ces deux états.

Dans une petite population, lorsqu'une mutation apporte un nouvel allèle, le destin ultime de celui-ci est :

-Soit de disparaître : probabilité =  $1 - 1/2 N_e$

-Soit d'éliminer les autres : probabilité =  $\frac{1}{2N}$   
Cours de Génétique et Dynamique des Populations Dr : bounamous A  
~~Dans une population de 500 individus, un nouvel allèle a 1 chance sur 1000 d'éliminer les~~

autres et 999 chances sur 1000 d'être éliminé.

La durée moyenne est de :

$2 \log(2N)$  génération pour la

disparition.  $4N$  générations pour la

fixation.