

Université de Djelfa

Faculté des Sciences de la nature et de la vie

Département de Biologie

Domaine : Sciences de la nature et de la vie

Filière : Sciences Biologiques

Spécialité : Biologie Moléculaire (Licence)

Année universitaire 2019 /2020

Module : Génétique des Populations

IV. Les écarts à la panmixie

Objectifs : Ce chapitre vise à apprendre aux étudiants :

- Les facteurs qui peuvent perturber le modèle de Hardy-Weinberg
- Le calcul de l'effet de ces différents facteurs sur la variation des fréquences alléliques au cours des générations successives.

IV. Les écarts à la panmixie

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 :

- 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 $\Delta q = q_{t+1} - q_t$
 - p_t et p_{t+1} étant respectivement la fréquence de l'allèle A aux générations t et $t+1$
- la recherche d'un équilibre, c'est à dire les conditions qui permettent d'annuler Δq ou Δp :
 - pour l'allèle A $\Delta p = 0$
 - pour l'allèle a $\Delta q = 0$

A. Mutations et migrations

1. 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^{-5} à 10^{-7} .

A) Mutations à sens unique

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.

Si la fréquence de **A**, pour une génération, est **p0**, la fréquence des gènes nouvellement mutés en **a** est le produit **u.p0** pour la génération suivante.

La nouvelle fréquence de **A** = **p0-u.p0** et la modification de la fréquence est
Dp=-u.p0

B) Mutations réciproques

Lorsque les mutations sont récurrentes et que les mutations reverses 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 **p0** et **q0** sont les fréquences alléliques de **A** et **a** à la génération **G0**, les fréquences **p1** et **q1** de ces allèles à la génération **G1** seront les suivantes :

$$f(A) \quad p1 = p0 - u p0 + v q0$$

$$f(a) \quad q1 = q0 - v q0 + u p0$$

La variation **Dq** de la fréquence de l'allèle **a** entre les générations **G0** et **G1** est:

$$Dq = q1 - q0$$

$$Dq = -v q0 + u p0$$

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

$$Dq = 0 \text{ c'est à dire : } Dq = -v q_e + u p_e = 0$$

$$Dq = -v q_e + u (1 - q_e) = 0$$

$$Dq = -v q_e + u - u q_e = 0$$

$$Dq = u - q_e (u + v) = 0$$

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

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.

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 voir indétectable à l'échelle de quelques dizaines voire quelques centaines de générations.

2. Les migrations

Dans la nature, les populations d'une même espèce ne sont pas génétiquement isolées. Il peut se produire à chaque génération des échanges d'adultes qui permettent des échanges de gènes, appelés flux géniques. Ces flux géniques sont généralement d'autant plus importants que les populations sont proches géographiquement.

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.

Le modèle de migration le plus simple, **appelé modèle insulaire**, 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).

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 :

$$m = n / (N + n)$$

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 Δp de l'allèle A entre ces 2 générations sera alors :

$$\Delta p = m (p_m - p_0) \text{ -----Explication}$$

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

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

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.

Remarque

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 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.

B. Sélection et adaptation

Les différents génotypes d'une population n'ont jamais les mêmes capacités reproductives ni les mêmes taux de survie.

Le phénotype global d'un individu et ses performances en termes de survie et de fertilité sera le résultat de sa constitution génétique à l'ensemble des loci, avec de possibles interactions entre gènes (épistasie, effets pléiotropes).

Ces différences de performance entre génotypes ont pour conséquence une variation des fréquences alléliques au cours du temps qui correspond au processus de sélection.

1. Valeur sélective et coefficient de sélection

A) Valeur sélective (fitness, ou valeur adaptative).

la valeur sélective (ou fitness) d'un génotype correspond au nombre de descendants viables et fertiles que produit en moyenne chaque individu de ce génotype à la génération suivante. La viabilité et la fertilité déterminent le nombre de descendants produits en moyenne par chaque génotype. La VS peut donc être donnée par la formule :

$$VS = \text{Survie} \times \text{Fécondité}$$

On distingue deux catégories de VS :

- **Fitness absolue**

On appelle fitness absolue, notée **W**, la valeur issue de la mesure de la probabilité de survie et de la fertilité de chaque catégorie génotypique et qui détermine directement leur nombre moyen de descendants. Pour un gène à deux allèles A et a, les valeurs de fitness seront notées de la façon suivante :

| Génotype | AA | Aa | aa |
|-----------------|-------------|-------------|-------------|
| Fitness absolue | W AA | W Aa | W aa |

- **Fitness relative**

La performance des génotypes est cependant toujours définie de façon relative = fitness relative notée **w** de façon à ce que la plus forte valeur de fitness soit égale à 1. La fitness relative d'une catégorie d'individus est le rapport entre sa fitness absolue **W** et la plus forte valeur de fitness absolue (**W**_{max}) observée dans la population.

$$\text{Fitness relative } w = \mathbf{W} / \mathbf{W}_{\max}$$

avec $0 < w < 1$

Dans le cas d'un gène à 2 allèles A et a :

$$w_{AA} = W_{AA} / W_{\max}$$

$$w_{Aa} = W_{Aa} / W_{\max}$$

$$w_{aa} = W_{aa} / W_{\max}$$

B) Coefficient de sélection

La différence entre **w** et **1** est appelée coefficient de sélection, noté **s**, qui est une mesure du taux de réduction de la fitness de chaque catégorie génotypique par rapport à la meilleure dans cette population.

$$w_{AA} = 1 - s_{AA}$$

$$w_{Aa} = 1 - s_{Aa}$$

$$w_{aa} = 1 - s_{aa}$$

Remarque : la fitness est très importante mais elle est difficile à mesurer car elle dépend de nombreuses caractéristiques individuelles et environnementales.

2. Cas simple de modèle de sélection : sélection contre l'homozygote récessif

Le modèle le plus simple de sélection est celui d'un gène à deux allèles A et a où l'allèle A est totalement dominant sur l'allèle a et où seul le génotype aa est désavantagé (on dit aussi contre-sélectionné). Les génotypes AA et Aa ont donc les mêmes performances car ils présentent le même phénotype [A]. Ceci se produit pour de nombreuses maladies génétiques récessives chez l'homme dont seul l'homozygote aa est atteint par des syndromes qui réduisent sa survie ou sa fertilité.

Dans ce cas de sélection contre l'homozygote récessif, on supposera que la différence de fitness est uniquement due à une plus faible survie du génotype aa qui subit un coefficient de sélection s. Les valeurs de fitness relatives des différents génotypes peuvent alors être notées :

| Génotypes | AA | Aa | aa |
|-----------|------------|------------|----------------|
| Fitness | $w_{AA}=1$ | $w_{Aa}=1$ | $w_{aa} = 1-s$ |

Si la population se reproduit au hasard (panmixie) et que les fréquences des allèles A et a sont respectivement p et q, les fréquences génotypiques avant l'action de la sélection (parmi les zygotes) suivent la loi de Hardy-Weinberg. Après l'action de la sélection, ces fréquences génotypiques seront modifiées car une partie des individus de génotype aa ne survivront pas jusqu'à la phase adulte. Les fréquences génotypiques chez les adultes ne

suiront donc plus la loi de Hardy-Weinberg et les fréquences alléliques vont varier entre le stade juvénile et le stade adulte. Ces variations peuvent être calculées de la façon suivante :

| | | | |
|---------------------------|------------|------------|--------------|
| Génotypes | AA | Aa | aa |
| Fréquence avant sélection | p^2 | $2pq$ | q^2 |
| Fitness associée | $w_{AA}=1$ | $w_{Aa}=1$ | $w_{aa}=1-s$ |
| après sélection | p^2 | $2pq$ | $q^2(1-s)$ |

Puisque seulement une proportion $(1-s) q^2$ du génotype aa survit, et donc qu'une proportion sq^2 meurt, la somme des génotypes restant après sélection est différente de 1 et vaut $1-sq^2$ ce qui correspond à la fitness moyenne de la population noté.

$$= p^2 + 2pq + q^2(1-s) = 1-sq^2$$

Les nouvelles fréquences génotypiques parmi les survivants seront après sélection :

| | | | |
|---------------------------|--------------|--------------|-------------------|
| Génotypes | AA | Aa | aa |
| Fréquence après sélection | $p^2/1-sq^2$ | $2pq/1-sq^2$ | $q^2(1-s)/1-sq^2$ |

la nouvelle fréquence allélique pour a devient :

$$f(a) q_1 = f(aa) + 1/2 f(Aa)$$

$$q_1 = (q^2(1-s) + pq)/(1-sq^2)$$

La variation de fréquence allélique Dq de l'allèle a entre deux générations successives est donc :

$$Dq = (-sq^2)/(1-sq^2)$$

3. Cas de sélection générale :

Le modèle de sélection général est celui d'un gène à deux allèles A et a où l'allèle A est totalement dominant sur l'allèle a. Dans ce cas de sélection, on supposera que la différence de valeur sélective est uniquement due à une plus faible survie des certains génotypes qui subissent un coefficient de sélection s .

Si la population se reproduit au hasard (panmixie) et que les fréquences des allèles A et a sont respectivement p et q, les fréquences génotypiques avant l'action de la sélection (parmi les zygotes) suivent la loi de Hardy-Weinberg. Après l'action de la sélection, ces fréquences génotypiques seront modifiées car une partie de certains génotypes ne survivront pas jusqu'à la phase adulte.

Les fréquences génotypiques chez les adultes ne suivront donc plus la loi de Hardy-Weinberg et les fréquences alléliques vont varier entre le stade juvénile et le stade adulte. Ces variations peuvent être calculées de la façon suivante :

| Génotypes | AA | Aa | aa |
|---------------------------|-------------|-------------|-------------|
| Fréquence avant sélection | p^2 | $2pq$ | q^2 |
| Fitness associée | S_1 | S_2 | S_3 |
| Fréquence après sélection | $p^2 S_1/T$ | $2pq S_2/T$ | $q^2 S_3/T$ |

$$\text{Avec } T = p^2 S_1 + 2pq S_2 + q^2 S_3$$

la nouvelle fréquence allélique pour a devient :

$$f(a) \rightarrow q_1 = f(aa) + 1/2 f(Aa)$$

La variation de fréquence allélique Dq de l'allèle a entre deux générations successives est donc :

$$Dq = \frac{pq (p(S_2 - S_1) + q (S_3 - S_2))}{p^2 S_1 + 2pq S_2 + q^2 S_3}$$

$$Dp = - Dq$$

Puisque s, p, et q sont des paramètres toujours positifs compris entre 0 et 1, les conclusions suivantes peuvent être données:

- Dq est toujours négatif car le numérateur (spq^2) et le dénominateur ($1-sq^2$) sont toujours positifs et compris entre 0 et 1. Ainsi, la fréquence de l'allèle a ne cesse de diminuer au cours des générations sous l'effet de la sélection.

- L'équilibre sera atteint ($Dq=0$) lorsque $q=0$, c'est-à-dire lorsque que l'allèle a sera complètement éliminé de la population (l'autre cas où $Dq=0$ correspond à l'absence de A dans la population $p=0$ et $q=1$)
 - plus la valeur de s est forte plus la valeur absolue Dq est élevée
 - plus la valeur de q est faible plus la valeur absolue Dq est faible ce qui implique que la diminution de la fréquence de l'allèle a par sélection est d'autant plus faible que la fréquence de a est faible.

Remarque : La fitness d'un individu dépend en effet :

1- de la composition génétique à l'ensemble des loci et pas seulement du génotype à un locus donné.

2- des conditions environnementales. Un génotype peut être avantagé dans certains environnements alors qu'il sera désavantagé dans d'autres conditions (interaction génotype-environnement).

3- de la composition génétique de la population à laquelle appartient cet individu. La fitness d'un individu pourra en effet être soit faible soit élevée en fonction de la présence des différents autres génotypes de la population.

La fitness est donc mesurée de façon indirecte à partir des variations de la fréquence des allèles observées entre générations successives.

A. Dérive génétique aléatoire

Les facteurs dispersifs provoquent la Dérive génétique aléatoire qui ont pour conséquence la dispersion, avec le temps, des fréquences géniques dans des populations d'effectif réduit (isolats restreints).

Plus la taille d'une population est petite, plus ces effets dispersifs sont importants. La dérive génétique et la sélection provoquent des variations de fréquence des allèles à l'intérieur d'une population.

Dans les populations réduites certains allèles peuvent perdre par hasard alors que d'autres sont retenus. Si un individu se trouve porteur d'un allèle, par ailleurs rare, et a de nombreux descendants, l'allèle, qu'il soit nuisible ou non, peut atteindre une fréquence importante à la suite d'une dérive génétique aléatoire.

Les effets des facteurs dispersifs diffèrent fondamentalement des facteurs systématiques car ils ne mènent pas nécessairement à l'établissement de caractères plus ou moins favorable à l'espèce mais qu'ils peuvent provoquer la fixation de caractères non désirables et ne peuvent donc être prévus

La dérive génétique est l'effet du hasard : si les gamètes de quelques individus seulement forment la génération suivante, alors les allèles de ces individus ne sont pas forcément représentatifs de la génération parentale. On peut prendre pour analogie un lancer de dés : si on lance un dé 6 fois, la probabilité pour obtenir un seul 1 est faible, alors que si on lance le dé 1000 fois, la proportion de 1 obtenus est beaucoup plus proche de 1/6. Ainsi la dérive génétique est d'autant plus marquée que la population est de petite taille.