

Chapitre 2
Génétique des populations
(Partie 1)

1. Définition et objectifs

La génétique des populations est une branche de la génétique née en 1920 et élaborée dans les années 30 par des mathématiciens et des biologistes (Wright, Fisher...) qui ont appliqué des bases mathématiques à l'hérédité mendélienne.

Contrairement à la génétique mendélienne qui étudie les mécanismes de transmission des caractères entre parents et descendants connus; la génétique des populations étudie le devenir des gènes au sein d'une population définie issue de croisements non contrôlés entre de nombreux parents.

L'étude du devenir d'un gène au sein d'une population, a pour but de connaître la fréquence de ses différents allèles; les mécanismes de leur transmission d'une génération à une autre; ainsi que les facteurs qui modifient ou limitent cette transmission au fil des générations.

Ainsi, l'étude des bases de cette science est nécessaire à tout améliorateur qui cherche à modifier génétiquement un caractère pour augmenter le potentiel génétique de ses animaux.

2 Notion de population

L'espèce est le groupe le plus général au sein duquel les individus sont susceptibles d'échanger des gènes entre eux (on dit qu'ils sont interféconds).

Une espèce est protégée par des barrières interspécifiques qui préviennent et empêchent les échanges de gènes entre individus appartenant à des espèces différentes.

L'espèce se décompose en sous unités appelées populations. On appelle population mendélienne le groupe d'individus appartenant à une même espèce ayant une reproduction sexuée et résidant dans des limites géographiques définies dans lesquels les croisements ont lieu.

Le stock génétique total de cette population est son pool de gène c'est-à-dire tous les gènes contenus dans chacun des génotypes individuels.

3- Constitution génétique d'une population

Pour définir la constitution génétique d'une population plusieurs sont utilisées :

- La constitution phénotypique : elle décrit les différents caractères et leurs fréquences dans la population
- La constitution génotypique : elle précise pour un caractère donné, les différents génotypes et leurs fréquences dans la population
- La constitution allélique : Elle fournit pour un gène considéré les différents allèles et leurs fréquences dans la population.

3-1- Calcul des fréquences génotypiques

3-1-1- Cas d'un gène autosomal

La fréquence d'un génotype dans la population est égale :

$$\frac{\text{Nombre d'individus porteurs de ce génotype}}{\text{Nombre total}}$$

Soit un gène A à deux formes alléliques A et a ; ce gène fournit trois génotypes :

AA , Aa et aa.

Dans un système de codominance, les trois génotypes correspondent à trois phénotypes différents, ainsi, si nous avons une population qui renferme **N** individus, on peut dénombrer les différents génotypes rencontrés, avec la répartition suivante :

X individus AA
Y individus Aa
Z individus aa
Avec : X+Y+Z= N

Les fréquences des trois génotypes seront les suivants :

$F(AA) = X/N$
 $F(Aa) = Y/N$
 $F(aa) = Z/N$

Avec $F(AA)+F(Aa)+F(aa)=1$

3-2- Calcul des fréquences alléliques

La fréquence d'un allèle dans la population est égale :

Nombre d'allèles de ce type / Nombre total d'allèles

$$F(A) = 2X + Y/2N = 2X/2N + Y/2N = X/N + Y/2N = F(AA) + \frac{1}{2} F(Aa) = p$$

$$F(a) = 2Z + Y/2N = 2Z/2N + Y/2N = Z/N + Y/2N = F(aa) + \frac{1}{2} F(Aa) = q$$

Ou bien, on peut déduire q, si on a p, sachant $p+q=1$ et donc $q=1-p$

Application : Les groupes sanguins MN chez l'homme

Les effectifs des groupes sanguins M et N chez les Américains de race blanche (d'après Stern, 1960), sont :

	Groupes			
	MM	MN	NN	
Effectifs :	1787	3039	1303	Total, N = 6129

a- Calculer les fréquences génotypiques MM, MN et NN.

$$F(MM) = 1787/6129 = 0.3$$

$$F(MN) = 3039/6129 = 0.5$$

$$F(NN) = 1303/6129 = 0.2$$

b- Calculer les fréquences des deux allèles M et N.

$$\text{Fréquence de M} = (1787 \times 2 + 3039) / (2 \times 6129) = 0,55$$

$$\text{Fréquence de N} = (1303 \times 2 + 3039) / (2 \times 6129) = 0,45$$

OU

$$F(M) = F(MM) + 1/2 F(MN) = 0.3 + 1/2 (0.5) = 0.55$$

$$F(N) = F(NN) + 1/2 F(MN) = 0.2 + 1/2 (0.5) = 0.45$$

3-1-1- Cas d'un gène lié au sexe

Constitution génétique de la femelle est différente de celle du male :

La femelle est XX et le male est XY

Ainsi pour un locus à deux allèles A/a, la femelle a trois génotypes

- $X^A X^A$, $X^A X^a$ et $X^a X^a$ et le male seulement deux génotypes $X^A Y$ et $X^a Y$

Les deux sexes ont donc des constitutions génétiques différentes et s'ils sont en fréquences égales, le sexe homogamétique détient pour les gènes concernés 2/3 du pool génétique de la population, et le sexe hétérogamétique 1/3 seulement. Le calcul des fréquences alléliques se fait séparément entre la femelle et le male.

Explication :

Soit un locus à deux allèles A/a situé sur le chromosome X. dans une population où l'on dénombre Nf individus de sexe homogamétique et Nm individus de sexe hétérogamétique (avec Nf = Nm). Le nombre total d'allèles est 2Nf + Nm.

Prenant :

- f1, f2 et f3 comme effectifs respectifs de $X^A X^A$, $X^A X^a$ et $X^a X^a$ chez les femelles
- m1 et m2 comme effectifs respectifs de $X^A Y$ et $X^a Y$ chez les males.

En considérant chaque groupe sexuel comme un échantillon indépendant, on peut estimer les fréquences alléliques chez les femelles et chez les males et dans la population toute entière.

- femelle: $F(A) = 2f1 + f2/2Nf$
- Male : $F(A) = m1/Nm$

Dans la population toute entière : $F(A) = 2f1 + f2 + m1/2Nf + Nm$

Soit en utilisant la formule : $F(A) = 2/3 F(A)f + 1/3 F(A)m$ ceci dans le cas où le nombre de femelle est égale au nombre de males.

Si le nombre de femelles n'est pas égale au nombre de males on utilise la formule suivante :

$$F(A) = 2F(A)f \cdot Nf + F(A)m \cdot Nm / 2Nf + Nm$$

Application : Le gène couleur du pelage chez le chat

Chez le chat, la coloration du pelage est déterminée par un gène lié au sexe à deux allèles codominants (O et G). Les males $X^O Y$ et les femelles $X^O X^O$ sont entièrement oranges. Les males $X^G Y$ et les femelles $X^G X^G$ sont entièrement gris. Les femelles hétérozygotes présentent un phénotype en mosaïque (juxtaposition de plages orange et grises). Cette particularité permet de déterminer sans

ambiguïté les proportions phénotypiques et génotypiques à ce locus. Une étude portant sur 500 chats (150 mâles et 350 femelles) a donné les résultats suivants :

M : 30 X_OY et 120 X_GY ; F : 60 X_OX_O, 160 X_OX_G, 130 X_GX_G

- 1- Quelles les fréquences de ces deux allèles chez les mâles et chez les femelles ?
- 2- Quelles les fréquences de ces deux allèles dans la population toute entière ?

1- Calcul des fréquences alléliques chez les femelles et chez les mâles :

- Chez les femelles :

$$F(O) f = 2 \times 60 + 160 / 2 \times 350 = 120 + 160 / 700 = 0.4$$

$$F(G) f = 2 \times 130 + 160 / 2 \times 350 = 420 / 700 = 0.6$$

- Chez les mâles

$$F(O) m = 30 / 150 = 0.2$$

$$F(G) m = 120 / 150 = 0.8$$

2- Calcul des fréquences alléliques dans la population toute entière

Le nombre de femelles étant différent de celui des mâles, donc on applique la formule suivante :

$$F(A) = \frac{2F(A)f \cdot N_f + F(A)m \cdot N_m}{2N_f + N_m}$$

$$F(O)p = \frac{2F(O)f \cdot N_f + F(O)m \cdot N_m}{2N_f + N_m} = \frac{2 \times 0.4 \times 350 + 0.2 \times 150}{700 + 150} = \frac{280 + 30}{850} = \frac{310}{850} = 0.36$$

$$F(G)p = \frac{2F(G)f \cdot N_f + F(G)m \cdot N_m}{2N_f + N_m} = \frac{2 \times 0.6 \times 350 + 0.8 \times 150}{700 + 150} = \frac{420 + 120}{850} = \frac{540}{850} = 0.64$$

4- Equilibre de HARDY-WEINBERG

4-1- Introduction

La connaissance des fréquences génotypiques et par voie de conséquence les fréquences alléliques nous permet de décrire la structure génétique d'une population. Mais, cette dernière n'est pas statique. En général, une population animale se reproduit pour donner une nouvelle génération. Il faut donc étudier le devenir de sa constitution génétique au cours des générations.

Cette étude est très difficile à réaliser compte tenu des nombreux facteurs qui peuvent modifier la fréquences des allèles (mutation, migration, différence de survie ou fécondité entre individus...).

4-2- Notion de population théorique idéale

Une première étape pour contourner les difficultés liées aux facteurs susceptibles de modifier la structure génétique de la population, est d'aborder la transmission des caractères dans un cas simple, appelé *population théorique idéale*, qui se définit de la manière suivante :

- Population d'effectif infiniment grand → pour minimiser les variations d'échantillonnage

- Population d'organismes diploïdes se reproduisant par le mode sexué et à méiose normale → un individu Aa produira toujours 50% des gamètes A et 50% des gamètes a.
- Population à générations non chevauchantes → aucun croisement entre individus de générations différentes.
- Absence de migration → Population close géographiquement
- Absence de sélection → tous les individus quelque soit leurs génotypes sont capables de se reproduire et engendrer une population viable.
- Absence de mutation → pas de passage d'un état allélique à un autre (ni A vers a, ni l'inverse).
- Population à reproduction panmictique (croisement au hasard entre individus) et pangamie (rencontre au hasard des gamètes).

4-3- Loi de Hardy-Weinberg

Dans une population théorique idéale, les fréquences des allèles et des génotypes au cours des générations suivent une loi normale, appelée loi de H-W qui constitue un modèle de référence en génétique des populations. Cette loi doit son nom à Hardy, mathématicien anglais et Weinberg, médecin allemand, qui l'ont établie indépendamment en 1908. La loi de HW stipule que les fréquences génotypiques et géniques ou alléliques (c'est-à-dire la structure génétique) reste stables de génération en génération.

On dit que la population est à l'équilibre

4-4- Démonstration de la loi

Soit une population théorique idéale renfermant un gène à deux allèles **A/a** dont les fréquences sont **p** et **q**. Ces fréquences sont également les fréquences de différentes catégories de gamètes produits par les individus de cette population (méiose normale).

Ainsi les males de cette génération **n** produiront **p** gamètes porteurs de **A** et **q** gamètes porteurs de **a**. de la même manière les femelles produiront **p** gamètes porteurs de **A** et **q** gamètes porteurs de **a**.

Les résultats des croisements qui déterminent la composition génétique de la génération **n+1** peut être obtenue par un échiquier de croisement dont les différents génotypes obtenus résultent du développement du binôme :

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

La génération **n+1** sera donc constituée de :

- **p^2 zygotes AA ; $2pq$ zygotes Aa et q^2 zygotes a**

	Male		
Femelle		A p	a q
A p		AA p^2	Aa pq
a q		Aa pq	aa q^2

- Fréquences phénotypiques : p^2 AA ; $2pq$ Aa et q^2 pour aa
- Fréquences génotypiques : $[A] = p^2 + 2pq$ et $[a] = q^2$

Dans une population théorique idéale, ces fréquences seront également celles reproducteurs de la génération $n+1$, pour lesquels la fréquence de chacun des deux allèles sera égale à la fréquence génotypiques des homozygotes + la moitié de la fréquence d'hétérozygotes.

$$\text{Pour A} \rightarrow p_1 = p^2 + (1/2) 2pq = p^2 + pq = p(p+q) = p \quad p+q=1$$

$$\text{Pour a} \rightarrow q_1 = q^2 + (1/2) 2pq = q^2 + pq = q(p+q) = q$$

Les fréquences alléliques n'ont donc pas changé, ce qui donne à la génération suivante $n+2$ les mêmes fréquences génotypiques qu'à la génération précédente soit :

- p^2 AA ; $2pq$ Aa et q^2 aa.

Le système est donc stable aussi bien pour les fréquences génotypiques, que pour les fréquences alléliques. On dit qu'on est à l'équilibre de Hardy-Weinberg.

4-5- Application et utilisation du modèle de Hardy-Weinberg

4-5-1- Test d'équilibre ou test de conformité d'une population

Le test d'équilibre permet de savoir si une population donnée obéit à la loi de HW ou non. Le principe du test est simple et peut être résumé en 3 étapes :

- Echantillonnage d'une population \rightarrow dénombrement des effectifs génotypiques réels (possibles en cas de dominance) et calcul des fréquences alléliques réelles parmi les N individus échantillonnés, soit $p = f(A)$ et $q = f(a)$.
- Calcul des effectifs génotypiques attendus dans une population idéale qui aurait le même effectif et les mêmes fréquences alléliques que la population étudiée soit :
 - $AA = p^2 \times N$ $Aa = 2pq \times N$ $aa = q^2 \times N$
- Comparaison des effectifs observés et des effectifs théoriques ou attendus par un test statistique de X^2 (chi deux) :
 - $X^2 = \sum (\text{eff. Observés} - \text{effe. Théoriques})^2 / \text{effe. Théoriques}$

La valeur du X^2 calculée sera comparée à une valeur seuil lue dans la table, en fonction d'un nombre de degré de liberté (ddl) égale à la différence entre le nombre de génotype et le nombre d'allèles du système génétique étudié et une valeur α qui représente le risque d'erreur qui est généralement de 5%..

On conclue de la manière suivante :

- Si X^2 calculé est inférieur au X^2 table \rightarrow la population étudiée est à l'équilibre.
- Si X^2 calculé est supérieur au X^2 table \rightarrow la population étudiée n'est pas à l'équilibre.

Application :

Chez l'homme le groupe sanguin MN est déterminé par un gène à deux allèles co-dominants Met N. une étude portant sur 730 aborigènes australiens a donnée les résultats suivants :

22 MM 216 MN 492 NN

- Cette population est à l'équilibre ?

-

4-5-2- Estimation des fréquences alléliques

Lorsqu'un caractère est due à un gène à 2 allèles A et a où A est totalement dominante sur a. seuls deux phénotypes peuvent être distingués dans la population :

- Phénotype [A] correspond aux deux génotypes AA et Aa
- Phénotype [a] correspond au génotype aa

Contrairement au système codominant, il n'est pas possible de calculer les fréquences alléliques dans la population, car les proportions respectives des génotypes AA et Aa ne sont pas connues.

Le modèle de HW va permettre de donner une estimation de ces fréquences à partir de la fréquence du phénotype homozygote récessif qui correspond à q^2 si bien sur la population est à l'équilibre.

- $q^2 \rightarrow aa$ et donc $q = \sqrt{q^2}$; une fois que j'ai déterminé q, je peux calculer $p = 1 - q$
- Parmi les individus de phénotype [A], je peux déterminer les proportions génotypiques AA et Aa, en utilisant le modèle de HW :
 - La fréquence de AA = $p^2 / p^2 + 2pq$
 - La fréquence de Aa = $2pq / p^2 + 2pq$

5- Généralisation de la loi

5-1- Transmission d'un gène à plusieurs allèles (système multiallélique)

La loi de HW s'applique également à des gènes qui existent sous plus de 2 états alléliques. L'équilibre correspond alors à l'association aléatoire des différents allèles pour former les génotypes dont la fréquence reste stable de génération en génération.

Démonstration :

Dans une population diploïde, à reproduction sexuée, on considère 3 allèles à dominance hiérarchique $A > a' > a$ tel que leurs fréquences alléliques respectives sont : p, q et r.

Les croisements aléatoires donneront naissance à 3 génotypes correspondants à 6 génotypes différents. Ces derniers résultent du développement du trinôme $(p+q+r)^2$.

$$\begin{array}{ccccccc}
 (p + q + r)^2 = & p^2 & + & 2pq & + & 2pr & + & q^2 & + & 2qr & + & r^2 \\
 \text{Génotypes :} & AA & & Aa & & Aa & & a'a' & & a'a & & aa \\
 \text{Phénotypes :} & & & [A] & & & & [a'] & & & & [a]
 \end{array}$$

Les fréquences à la génération suivante sont :

- $P' = p^2 + pq + pr = p(p + q + r) = p$ sachant que $p + q + r = 1$
- $q' = q^2 + pq + qr = q(p + q + r) = q$
- $r' = r^2 + pr + qr = r(r + p + q) = r$

Les fréquences alléliques ne changent pas au cours des générations comme dans le cas du diallélisme.

- **Calcul des fréquences alléliques dans ce cas :**

On peut calculer les fréquences de l'allèle le plus dominant A en prenant en considération d'une part, le phénotype du à cet allèle et d'autre part les autres phénotypes. Ces derniers peuvent être considérés comme étant dus à un allèle hypothétique a^x récessif par rapport à

Soit p la fréquence de A et q^x la fréquence de a^x

$(q^x)^2$ est la fréquence des phénotypes autres que A (c'est-à-dire dans notre cas celui de a' et a)
 $\rightarrow q^x = \sqrt{(q^x)^2}$

Une fois que nous déterminons q^x nous pouvons calculer la fréquence p de A, soit :

$$- p + q^x = 1 \rightarrow p = 1 - q^x$$

je détermine la fréquence de r de a, sachant que $r^2 = \frac{[a]}{[A]+[a'] + [a]}$

et finalement je détermine $q = 1 - (p + r)$

Exemple dans le TD N°4 (exercice 3).

5-2- Transmission d'un gène lié au sexe

Dans une population à l'équilibre, les fréquences alléliques sont les mêmes dans les deux sexes et dans la population globale. Ainsi, on peut bien utiliser les résultats chez les femelles comme chez les males.

En pratique, pour obtenir la fréquence allélique de la population d'un locus A/a, il suffit de calculer les fréquences chez le sexe hétérogamétique du fait de la simplicité des calculs. En effet :

- la fréquence des individus $X^A Y$ correspond bien à la fréquence de A
- la fréquence des individus $X^a Y$ correspond bien à la fréquence de a.

Les fréquences génotypiques dans chacun des deux sexes pour une population à l'équilibre sont représentés dans le tableau suivant :

Sexe homogamétique X X	Sexe hétérogamétique X Y
- $X^A X^A$	- $X^A Y$
- $X^A X^a$	- $X^a Y$
- $X^a X^a$	