

Module

Génétique des populations

L3-BD

Dr MESSAOUD KHELIFI Mouna

(mkm_etudiants@hotmail.fr)

Sommaire

I. Principales notions de la génétique des populations	1	4.1.1. Codominance : $A = B$	9
II. Structure génétique d'une population		4.1.2. Dominance et récessivité : $A > B$	10
1. Composition génétique d'une population	2	4.2. Gènes liés au sexe	10
2. Estimation de la composition génétique d'une population	2	4.3. Système multi-allélique : Trois allèles	10
2.1. Fréquences phénotypiques	2	4.3.1. Codominance : $A = B = C$	11
2.2. Fréquences génotypiques et fréquences alléliques	2	4.3.2. Dominance partielle : $A = B > C$	11
2.2.1. Cas de gènes autosomiques	2	4.3.3. Dominance totale : $A > B > C$	11
2.2.2. Cas de gènes liés au sexe	5	5. Conformité au modèle de Hardy-Weinberg : Test d'équilibre	11
Exercices	7	6. Conclusion	13
III. Modèle de Hardy-Weinberg		Exercices	13
1. Introduction	8	Annexe 1 : Application de la loi de Hardy-Weinberg	15
2. Étude théorique d'une population idéale	8	Annexe 2 : Table du Chi-deux	16
2.1. Système de reproduction panmictique	8	IV. Influence du régime de reproduction : Consanguinité	
2.2. Effectif infini de la population	8	1. Introduction	17
2.3. Absence de chevauchement	8	2. Consanguinité	17
2.4. Absence de mutations	8	2.1. Causes de la consanguinité	17
2.5. Absence de sélection	9	2.2. Principaux paramètres mesurant la consanguinité	17
2.6. Absence de migration	9	2.2.1. Coefficient de parenté (Φ)	18
3. Application de la loi de Hardy-Weinberg	9	2.2.2. Coefficient de consanguinité (f)	18
4. Généralisation de la loi de Hardy-Weinberg	9	2.3. Mesure de la parenté et de la consanguinité	18
4.1. Gènes autosomiques	9	2.4. Coefficients des parentés les plus courantes	18
		2.4.1. Germains : Frère et Sœur (Parenté de 1 ^{er} degré) ...	19
		2.4.2. Demi-germains : Demi-frère et Sœur (Parenté de	
		2 ^{ème} degré)	19

2.4.3. Cousins germains (Parenté de 3 ^{ème} degré)	19
2.4.4. Double cousins germains (Parenté de 3 ^{ème} degré)	19
2.4.5. Tante et Neveu (Parenté de 2 ^{ème} degré)	20
2.5. Structure génétique des populations consanguines	20
2.5.1. Coefficient moyen de consanguinité (F)	20
2.5.2. Conséquence de la consanguinité sur les fréquences génomiques	20
2.5.3. Conséquence de la consanguinité sur les fréquences alléliques	21
Exercices	21
Annexe 1 : Mesure de la parenté et de la consanguinité	23
Annexe 2 : Conséquence de la consanguinité sur les fréquences génomiques	24

4.2. Coefficient de sélection	28
4.3. Evolution d'une population soumise à la sélection	28
4.4. Etude de certains cas particuliers	29
4.4.1. Sélection contre un allèle létal récessif (Modèle 1)	29
4.4.2. Sélection contre un allèle sublétalement récessif (Modèle 2)	29
4.4.3. Sélection contre un allèle dominant (Modèle 3)	30
4.4.4. Sélection contre les homozygotes (Modèle 4)	30
4.4.5. Sélection contre les hétérozygotes (Modèle 5)	31
4.5. Différentes formes de sélection naturelle	31
4.5.1. Sélection stabilisante	31
4.5.2. Sélection directionnelle	31
4.5.3. Sélection diversifiante	31
Exercices	32

V. Influence des pressions évolutives

1. Introduction	25
2. Mutation	25
2.1. Impact des taux de mutation sur les fréquences alléliques	25
2.2. Vitesse du processus	25
3. Migration	26
3.1. Impact de la migration sur les fréquences alléliques	26
4. Sélection	27
4.1. Valeur sélective "w"	27

Principales notions de la génétique des populations

- **Génétique des populations** : Etude de la distribution et des changements de fréquences des versions d'un gène (allèles) dans les populations d'êtres vivants, sous l'influence des pressions évolutives (mutation, migration, sélection et dérive génétique).
- **Gène** : Unité d'hérédité contrôlant la manifestation et la transmission d'un caractère héréditaire.
- **Allèles** : Différentes formes que peut prendre un même gène. Ils occupent la même position (locus) sur les chromosomes homologues.
- **Locus** : Emplacement précis et invariant d'un gène sur un chromosome. Son pluriel est « **Loci** ».
- **Génome** : Ensemble du matériel génétique d'un individu ou d'une espèce (séquences codantes et non-codantes).
- **Génotype** : Ensemble des caractères génétiques d'un individu dont l'expression différentielle est visible dans le phénotype. Le génotype est donc la composition allélique de tous les gènes d'un individu. Prenant l'exemple d'un gène représenté par deux allèles : A et B, le génotype d'un individu pour ce gène est alors soit homozygote (AA ou BB), soit hétérozygote (AB).
- **Phénotype** : Ensemble des caractères observables chez un individu, résultant de l'interaction entre son génotype et les effets de son environnement.
- **Homozygote** : Individu d'une lignée pure qui porte en double exemplaire la même forme allélique d'un gène correspondant à un caractère héréditaire.
- **Hétérozygote** : Individu hybride qui porte deux formes alléliques différentes d'un gène correspondant à un caractère héréditaire.
- **Population** : Ensemble des individus de la même espèce qui ont la possibilité d'interagir entre eux au moment de la reproduction. La notion de population fait donc appel à des critères d'ordre spatiaux temporels et génétiques et résulte du fait que les individus d'une même espèce n'ont pas tous la possibilité de se rencontrer et de se croiser à cause de l'éloignement géographique ou de l'hétérogénéité de l'habitat. Il peut s'agir aussi, chez l'homme, de barrières culturelles, sociales, religieuses ou éthiques, qui limitent plus ou moins les possibilités d'unions entre individus, même géographiquement proches.
- **Mutation** : Changement de l'information génétique codée par l'ADN. La génétique des populations s'intéresse principalement aux mutations qui se produisent dans la lignée germinale donc aux taux de mutations par génération et leurs effets sur les fréquences alléliques des gènes concernés.
- **Migration** : Passage d'individus d'une population à une autre, par exemple sous forme de graines ou de pollen chez les plantes. Elle peut modifier les fréquences alléliques lorsque celles de la population qui fournit les migrants diffèrent de celles de la population qui les reçoit.

Structure génétique d'une population

1. Composition génétique d'une population

Pour définir la composition génétique d'une population plusieurs paramètres sont étudiés : les phénotypes, les génotypes et les allèles. Ainsi pour une variable donnée (exemple : la couleur des yeux) on doit préciser les différents états possibles (bleu, vert, marron.....) et les probabilités avec lesquelles ces états se rencontrent dans la population étudiée (0,1 / 0,2 / 0,6 ...). Ce premier descripteur est la **composition phénotypique** de la population.

De manière similaire, la **composition génotypique** précise pour le caractère considéré les différents génotypes possibles et leurs probabilités. La relation génotype-phénotype nécessite une analyse génétique préalable pour établir les effets éventuels de codominance ou de dominance et de récessivité.

Enfin, la **composition allélique** fournit, pour un locus considéré, les allèles présents dans la population et leurs probabilités.

Afin de déterminer ces différentes compositions, les techniques classiques de l'estimation statistique sont utilisées. On obtient ainsi, à partir d'un échantillon représentatif d'une population, les **fréquences** relatives des différents phénotypes, génotypes et allèles ; avec des valeurs comprises entre 0 et 1.

2. Estimation de la composition génétique d'une population

2.1. Fréquences phénotypiques

Les phénotypes sont par définition directement accessibles à l'observation, les **fréquences phénotypiques** peuvent être alors

calculées à partir d'un échantillon d'individus tirés aléatoirement dans la population. Par exemple, dans un échantillon de N individus dont N_G ont le corps gris et N_B ont le corps blanc, les fréquences phénotypiques de la population pour le caractère couleur du corps sont les suivantes :

$$\text{Fréquence du phénotype gris} = f_{q [G]} = N_G / N$$

$$\text{Fréquence du phénotype blanc} = f_{q [B]} = N_B / N$$

$$\text{Avec : } f_{q [G]} + f_{q [B]} = 1$$

AN : Soit $N = 60$, $N_G = 22$, $N_B = 38$

$$f_{q [G]} = N_G / N = 22 / 60 \rightarrow f_{q [G]} = 0,366$$

$$f_{q [B]} = N_B / N = 38 / 60 \rightarrow f_{q [B]} = 0,633$$

2.2. Fréquences génotypiques et fréquences alléliques

La détermination des fréquences génotypiques et alléliques est beaucoup plus complexe et dépend du caractère considéré (autosomique ou lié au sexe) et de la force des différentes formes alléliques du gène qui contrôle ce caractère (codominance, dominance et récessivité).

2.2.1. Cas de gènes autosomiques

- **Codominance**

La **codominance** désigne une relation équilibrée entre deux allèles qui leur permet de s'exprimer simultanément pour générer un phénotype intermédiaire (individus hétérozygotes) par rapport aux

phénotypes obtenus avec des allèles identiques (individus homozygotes).

Par exemple, le caractère groupe sanguin MN chez l'homme est gouverné par un gène à deux allèles M et N autosomiques et codominants. Trois phénotypes sont possibles M, MN et N, et correspondent aux génotypes MM, MN et NN, respectivement. Ainsi, chaque phénotype est associé uniquement à un seul génotype et réciproquement. Les **fréquences génotypiques** sont donc égales aux fréquences phénotypiques.

Si on considère que N_1 , N_2 et N_3 sont les effectifs respectifs des phénotypes M, MN et N ; et que N_t est l'effectif total de l'échantillon étudié avec $N_t = N_1 + N_2 + N_3$; Les fréquences génotypiques sont définies par les rapports suivants :

$$\text{Fréquence du génotype MM} = \mathbf{fq (MM)} = N_1 / N_t$$

$$\text{Fréquence du génotype MN} = \mathbf{fq (MN)} = N_2 / N_t$$

$$\text{Fréquence du génotype NN} = \mathbf{fq (NN)} = N_3 / N_t$$

$$\mathbf{Avec : } \mathbf{fq (MM) + fq (MN) + fq (NN) = 1}$$

Deux méthodes de calcul des **fréquences alléliques** sont possibles, un simple comptage des allèles ou une formule probabiliste.

Selon **la méthode de comptage**, la fréquence d'un allèle correspond au rapport du nombre de cet allèle au nombre total d'allèles occupants un locus donné, soit $2N$ pour un échantillon de N individus diploïdes. Dans le cas du caractère groupe sanguin MN chez l'homme :

N_1 : Nombre d'individu de phénotype M porteurs de deux allèles M.

N_2 : Nombre d'individu de phénotype MN porteurs d'un allèle M et d'un allèle N.

N_3 : Nombre d'individu de phénotype N porteurs de deux allèles N.

Donc : Le nombre de l'allèle M dans la population est de $2N_1 + N_2$

Le nombre de l'allèle N dans la population est de $2N_3 + N_2$

Les fréquences des allèles M et N sont par conséquent les suivantes :

$$\text{Fréquence de l'allèle M} = \mathbf{fq (M)} = \mathbf{p} = (2N_1 + N_2) / 2N_t$$

$$\text{Fréquence de l'allèle N} = \mathbf{fq (N)} = \mathbf{q} = (2N_3 + N_2) / 2N_t$$

$$\mathbf{Avec : } \mathbf{p + q = 1}$$

Selon **la méthode probabiliste**, on peut calculer les fréquences alléliques à partir des fréquences génotypiques en utilisant les rapports suivants :

$$\text{Fréquence de l'allèle M} = \mathbf{fq (M)} = \mathbf{p} = \mathbf{fq (MM) + \frac{1}{2} fq (MN)}$$

$$\text{Fréquence de l'allèle N} = \mathbf{fq (N)} = \mathbf{q} = \mathbf{fq (NN) + \frac{1}{2} fq (MN)}$$

$$\mathbf{Avec : } \mathbf{p + q = 1}$$

AN : L'étude d'un échantillon de 1000 individus a permis d'estimer la composition génétique d'une population pour le caractère groupe sanguin MN chez l'homme :

Phénotypes	M	MN	N
Effectifs ($N_t = 1000$)	$N_1 = 350$	$N_2 = 500$	$N_3 = 150$
Fréquences phénotypiques	$f_{q [M]} = 350 / 1000$ $f_{q [M]} = 0,35$	$f_{q [MN]} = 500 / 1000$ $f_{q [MN]} = 0,5$	$f_{q [N]} = 150 / 1000$ $f_{q [N]} = 0,15$
Génotypes	MM	MN	NN
Fréquences génotypiques	$f_{q (MM)} = 350 / 1000$ $f_{q (MM)} = 0,35$	$f_{q (MN)} = 500 / 1000$ $f_{q (MN)} = 0,5$	$f_{q (NN)} = 150 / 1000$ $f_{q (NN)} = 0,15$
Allèles	M	M et N	N
Fréquences alléliques : Méthode de comptage	$p = (2 \times 350 + 500) / 2000$ $p = 0,6$	/	$q = (2 \times 150 + 500) / 2000$ $q = 0,4$
Fréquences alléliques : Méthode probabiliste	$p = 0,35 + 0,5 / 2$ $p = 0,6$	/	$q = 0,15 + 0,5 / 2$ $q = 0,4$

- **Dominance et récessivité**

L'allèle **récessif** d'un gène désigne un allèle qui ne s'exprime pas dans un génome qui contient un allèle **dominant** du même gène. C'est le cas de l'allèle récessif qui détermine le groupe sanguin O face aux allèles dominants qui définissent les groupes sanguins A et B.

Dans ce cas, seules les fréquences phénotypiques peuvent être estimées ainsi que les fréquences du génotype hétérozygote AB et du génotype homozygote récessif OO. En effet, il est impossible d'en déduire les fréquences génotypiques pour les phénotypes dominants présentant plusieurs possibilités génotypiques (AA et AO pour le groupe sanguin A ; BB et BO pour le groupe sanguin B). Les fréquences des allèles A et B ne peuvent également pas être calculées.

AN : L'étude d'un échantillon de 500 individus a permis d'estimer la composition génétique d'une population pour le caractère groupe sanguin ABO chez l'homme :

Phénotypes	A	B	AB	O
Effectifs ($N_t = 500$)	$N_1 = 213$	$N_2 = 56$	$N_3 = 15$	$N_4 = 216$
Fréquences phénotypiques	$f_{q [A]} = 213 / 500$ $f_{q [A]} = 0,426$	$f_{q [B]} = 56 / 500$ $f_{q [B]} = 0,112$	$f_{q [AB]} = 15 / 500$ $f_{q [AB]} = 0,03$	$f_{q [O]} = 216 / 500$ $f_{q [O]} = 0,432$
Génotypes	AA et AO	BB et BO	AB	OO
Fréquences Génotypiques	?	?	$f_{q (AB)} = f_{q [AB]}$ $f_{q (AB)} = 0,03$	$f_{q (OO)} = f_{q [O]}$ $f_{q (OO)} = 0,432$
Allèles	A et O	B et O	A et B	O
Fréquences alléliques	?	?	?	?

La connaissance des fréquences génotypiques et alléliques nécessite une information supplémentaire. Cette information peut provenir soit d'une expérimentation permettant de distinguer les homozygotes des hétérozygotes, soit de la connaissance de l'ascendance ou de la descendance des individus étudiés.

2.2.2. Cas de gènes liés au sexe

Lorsqu'un gène est situé sur un hétérochromosome, le sexe homogamétique possède deux locus et le sexe hétérogamétique un seul, c'est le cas des espèces dont les femelles sont XX et les mâles sont XY. Les fréquences alléliques peuvent être alors calculées indépendamment dans chacun des deux sexes.

Prenons l'exemple de la drosophile où le gène déterminant la forme de l'œil est situé sur le chromosome X. Ce gène possède deux allèles **codominants**, un allèle O donnant à l'œil une forme ovoïde et un allèle B donnant à l'œil une forme de barre. Chez les individus hétérozygotes les yeux sont réniformes. Dans ce système, on distingue cinq génotypes :

	Phénotypes	Ovoïde	Réniforme	Barre
Génotypes	♀	X_OX_O	X_OX_B	X_BX_B
	♂	X_OY	/	X_BY

Si dans une population on observe $N_{\text{♂}}$ mâles parmi lesquels N_O ont des yeux ovoïdes et N_B ont des yeux barres. Parmi les $N_{\text{♀}}$ femelles, N_1 ont des yeux ovoïdes, N_2 ont des yeux réniformes et N_3 ont des yeux barres. En considérant chaque groupe sexuel comme un échantillon indépendant, on peut estimer les fréquences des allèles O et B pour chacun des deux sexes et pour l'ensemble de la population en utilisant les rapports suivants :

• Chez les mâles

	Méthode de comptage	Méthode probabiliste
Allèle O	$p_{\text{♂}} = N_O / N_{\text{♂}}$	$p_{\text{♂}} = fq (X_OY)$
Allèle B	$q_{\text{♂}} = N_B / N_{\text{♂}}$	$q_{\text{♂}} = fq (X_BY)$

• Chez les femelles

	Méthode de comptage	Méthode probabiliste
Allèle O	$p_{\text{♀}} = (2N_1 + N_2) / 2N_{\text{♀}}$	$p_{\text{♀}} = fq (X_OX_O) + \frac{1}{2} fq (X_OX_B)$
Allèle B	$q_{\text{♀}} = (2N_3 + N_2) / 2N_{\text{♀}}$	$q_{\text{♀}} = fq (X_BX_B) + \frac{1}{2} fq (X_OX_B)$

• Dans la population générale

	Méthode de comptage	Méthode probabiliste
Allèle O	$p = (2N_1 + N_2 + N_O) / (2N_{\text{♀}} + N_{\text{♂}})$	$p = fq (X_OX_O) + \frac{1}{2} fq (X_OX_B) + fq (X_OY)$
Allèle B	$q = (2N_3 + N_2 + N_B) / (2N_{\text{♀}} + N_{\text{♂}})$	$q = fq (X_BX_B) + \frac{1}{2} fq (X_OX_B) + fq (X_BY)$

AN : Soit un échantillon de 1411 drosophiles (N_t), 716 femelles (N_f) et 695 mâles (N_m).

	Femelles			Mâles		Population générale				
Phéno	Ovoïde	Réniforme	Barre	Ovoïde	Barre	Ovoïde	Réniforme	Barre		
Eff	$N_1 = 282$	$N_2 = 361$	$N_3 = 73$	$N_O = 457$	$N_B = 238$	$N_1 + N_O = 282 + 457$ $N_1 + N_O = 739$		$N_2 = 361$	$N_3 + N_B = 73 + 238$ $N_3 + N_B = 311$	
fq phéno	$f_{q [O]} = 282 / 716$ $f_{q [O]} = 0,393$	$f_{q [R]} = 361 / 716$ $f_{q [R]} = 0,504$	$f_{q [B]} = 73 / 716$ $f_{q [B]} = 0,101$	$f_{q [O]} = 457 / 695$ $f_{q [O]} = 0,657$	$f_{q [B]} = 238 / 695$ $f_{q [B]} = 0,342$	$f_{q [O]} = 739 / 1411$ $f_{q [O]} = 0,523$		$f_{q [R]} = 361 / 1411$ $f_{q [R]} = 0,255$	$f_{q [B]} = 311 / 1411$ $f_{q [B]} = 0,220$	
Géno	$X_O X_O$	$X_O X_B$	$X_B X_B$	$X_O Y$	$X_B Y$	$X_O X_O$	$X_O Y$	$X_O X_B$	$X_B X_B$	$X_B Y$
fq géno	$f_{q (X_O X_O)} = 282 / 716$ $f_{q (X_O X_O)} = 0,393$	$f_{q (X_O X_B)} = 361 / 716$ $f_{q (X_O X_B)} = 0,504$	$f_{q (X_B X_B)} = 73 / 716$ $f_{q (X_B X_B)} = 0,101$	$f_{q (X_O Y)} = 457 / 695$ $f_{q (X_O Y)} = 0,657$	$f_{q (X_B Y)} = 238 / 695$ $f_{q (X_B Y)} = 0,342$	$f_{q (X_O X_O)} = 282 / 1411$ $f_{q (X_O X_O)} = 0,199$	$f_{q (X_O Y)} = 457 / 1411$ $f_{q (X_O Y)} = 0,323$	$f_{q (X_O X_B)} = 361 / 1411$ $f_{q (X_O X_B)} = 0,255$	$f_{q (X_B X_B)} = 73 / 1411$ $f_{q (X_B X_B)} = 0,051$	$f_{q (X_B Y)} = 238 / 1411$ $f_{q (X_B Y)} = 0,168$
Allèles	O	O et B	B	O	B	O		O et B		B
fq all : MC	$p = (2 \times 282 + 361) / 1432$ $p = 0,645$	/	$q = (2 \times 73 + 361) / 1432$ $q = 0,354$	$p = 457 / 695$ $p = 0,657$	$q = 238 / 695$ $q = 0,342$	$p = (2 \times 282 + 361 + 457) / (2 \times 716 + 695) = 1382 / 2127$ $p = 0,649$		/	$q = (2 \times 73 + 361 + 238) / (2 \times 716 + 695) = 745 / 2127$ $q = 0,350$	
fq all : MP	$p = 0,393 + 0,504 / 2$ $p = 0,645$	/	$q = 0,101 + 0,504 / 2$ $q = 0,353$	$p = f_{q [O]} = f_{q (X_O Y)}$ $p = 0,657$	$q = f_{q [B]} = f_{q (X_B Y)}$ $q = 0,342$	$p = 0,199 + 0,323 + 0,255 / 2$ $p = 0,649$		/	$q = 0,051 + 0,168 + 0,255 / 2$ $q = 0,346$	

Dans le cas de **dominance** et de **récessivité**, les fréquences alléliques pourront être estimées seulement chez les mâles, des informations supplémentaires sont nécessaires pour les calculer chez les femelles.

Exercices

Exercice 1

Une étude du polymorphisme de l'alcool déshydrogénase chez un petit mammifère a révélé la présence de deux formes alléliques S et F. La population étudiée présentait les fréquences génotypiques suivantes :

$$fq(SS) = 0,16 \quad fq(SF) = 0,48 \quad fq(FF) = 0,36$$

1. Calculer les fréquences des allèles S et F.

Exercice 2

Chez le chat, la coloration du pelage est déterminée par un gène lié au sexe à deux allèles codominants (O et G). Les mâles X_OY et les femelles X_OX_O sont entièrement oranges. Les mâles X_GY et les femelles X_GX_G sont entièrement gris. Les femelles hétérozygotes X_OX_G présentent un phénotype en mosaïque (juxtaposition de plages oranges et gris). Cette particularité permet de déterminer sans ambiguïté la composition phénotypique et génotypique des populations à ce locus. Une étude portant sur 500 chats (150 mâles et 350 femelles) a donné les résultats suivants :

M : 30 X_OY 120 X_GY **F** : 56 X_OX_O 168 X_OX_G 126 X_GX_G

1. Quelles sont les fréquences des allèles O et G chez les mâles et chez les femelles ?
2. Calculer les fréquences de ces allèles dans l'ensemble de la population.

Exercice 3

Une étude de la diversité génétique d'une population d'Hêtre d'une forêt, a fourni les résultats suivants concernant le gène qui code pour l'enzyme 6-PGDH (2 allèles, A et B) :

Génotypes	AA	AB	BB
Effectifs	$N_1 = 125$	$N_2 = 31$	$N_3 = 4$

1. Calculer les fréquences alléliques à ces deux locus en utilisant la méthode probabiliste.

Exercice 4

Généralement, Il n'est pas possible de calculer les fréquences génotypiques à partir des seules fréquences alléliques, à moins de faire des hypothèses supplémentaires. Cela est néanmoins possible, lorsqu'un des génotypes manque.

1. Exprimer les fréquences des génotypes en fonction des fréquences alléliques p (allèle A) et q (allèle B) dans les cas suivants :

	Génotypes		
I	AA	AB	BB
II	AA	AB	BB
III	AA	AB	BB

Modèle de Hardy-Weinberg

1. Introduction

Avant d'étudier ce qui se produit dans une population qui évolue, c'est-à-dire dans laquelle les fréquences alléliques changent significativement, nous devons connaître ce qu'est l'état d'une population qui n'évolue pas, donc dans laquelle les fréquences alléliques ne changent pas significativement, c'est-à-dire qu'elles sont en état d'**équilibre génétique**. Il est fort probable que pareille population n'existe pas et qu'il ne s'agit que d'un concept théorique. Ce concept est néanmoins d'une très grande importance puisque sur lui repose toute la génétique des populations.

C'est le mathématicien anglais **G. H. Hardy** et le médecin allemand **W. Weinberg** qui, tous deux en 1908, et indépendamment l'un de l'autre, ont démontré que les fréquences alléliques ne dépendent pas de dominance ou de récessivité et qu'elles peuvent rester constantes d'une génération à la suivante, **lorsque certaines conditions sont respectées**. La démonstration de Hardy et Weinberg est d'une telle importance qu'on l'a formulé sous forme de loi, désignée **loi de Hardy-Weinberg** ou **loi de l'équilibre génétique**.



Godfrey Harold Hardy
1877 – 1947



Wilhelm Weinberg
1862 – 1937

2. Étude théorique d'une population idéale

Dans leur modèle, Hardy et Weinberg ont étudié la transmission des caractères dans une **population théorique idéale** qui présente des caractéristiques permettant de simplifier les calculs et les raisonnements. Ces caractéristiques sont les suivantes :

2.1. Système de reproduction panmictique

Dans ce système la formation des couples reproducteurs se fait complètement au hasard, les couples sont dits **panmictiques**. Cette hypothèse suppose que les individus ne choisissent pas leurs partenaires sexuels ni en fonction de leurs génotypes, ni en fonction de leurs phénotypes = **panmixie** et que les gamètes s'associent au hasard par rapport aux gènes considérés = **pangamie**.

2.2. Effectif infini de la population

La population étudiée doit avoir une taille infinie, la fréquence d'un événement est donc égale à sa probabilité (loi des grands nombres). Par exemple, la fréquence des « piles » peut être égale à 0,7 sur dix tirages « pile ou face » mais ne peut pas sur 100 000 tirages s'écarter notablement de sa probabilité qui est égale à 0,5.

2.3. Absence de chevauchement

Aucun croisement entre individus de générations différentes.

2.4. Absence de mutations

Les individus formant les couples produisent des gamètes, la fréquence des gènes présents dans ces gamètes n'est pas modifiée par des mutations. Ainsi, un individu AB produira toujours 50% de gamètes A et 50% de gamètes B.

4.1.2. Dominance et récessivité : A > B

Après croisement des gamètes portant ces allèles on a :

Phénotypes	A		B
Génotypes	AA	AB	BB
fq génotypiques	$f_{q(AA)} = p^2$	$f_{q(AB)} = 2pq$	$f_{q(BB)} = q^2$
fq phénotypiques	$f_{q[A]} = f_{q(AA)} + f_{q(AB)}$ $f_{q[A]} = p^2 + 2pq$		$f_{q[B]} = f_{q(BB)} = q^2$
Effectifs ($N_t = N_1 + N_2$)	N_1		N_2

L'estimation directe des fréquences alléliques est impossible car on ne peut pas estimer indépendamment les trois fréquences génotypiques. C'est alors que le modèle de Hardy-Weinberg présente la grande utilité de fournir le moyen **indirect** d'une telle estimation. En effet, on a :

$$f_{q[B]} = q^2 \quad \text{d'où} \quad q = \sqrt{f_{q[B]}} = \sqrt{N_2 / N_t}$$

On en tire évidemment p à partir de la formule $p + q = 1$. Connaissant p et q, on peut en déduire, par le calcul, les valeurs des fréquences génotypiques.

4.2. Gènes liés au sexe

Dans le cas des espèces dont les femelles sont XX et les mâles sont XY, prenant l'exemple d'un gène porté par X et présentant deux allèles A et B de fréquences respectives p et q. Supposons que les fréquences des deux allèles sont les mêmes chez les deux sexes. La loi de Hardy-Weinberg s'applique uniquement au sexe homogamétique (XX) :

♀ \ ♂	$X_A (p)$	$X_B (q)$	Y
$X_A (p)$	$X_A X_A (p^2)$	$X_A X_B (pq)$	$X_A Y (p)$
$X_B (q)$	$X_A X_B (pq)$	$X_B X_B (q^2)$	$X_B Y (q)$
	Femelles		Mâles

Les fréquences génotypiques à l'équilibre sont donc les suivantes :

Femelles	AA (p^2)	AB ($2pq$)	BB (q^2)
Mâles	A (p)	/	B (q)

4.3. Système multi-allélique : Trois allèles

Dans une population soumise aux conditions du modèle de Hardy-Weinberg, considérons trois allèles autosomiques A, B, C de fréquences respectives p, q, r ; identiques chez les deux sexes. On obtient après croisement des gamètes portant ces allèles :

♂ \ ♀	A (p)	B (q)	C (r)
A (p)	AA (p^2)	AB (pq)	AC (pr)
B (q)	AB (pq)	BB (q^2)	BC (qr)
C (r)	AC (pr)	BC (qr)	CC (r^2)

Equilibre de Hardy-Weinberg

$$p^2 + 2pq + q^2 + 2pr + 2qr + r^2 = (p + q + r)^2 = 1$$

Pour calculer les fréquences, il faut tenir compte des différents rapports de force qui peuvent exister entre ces trois allèles.

4.3.1. Codominance : A = B = C

Phénotypes	A	AB	AC	B	BC	C
Génotypes	AA	AB	AC	BB	BC	CC
fq géno = fq phéno	p^2	$2pq$	$2pr$	q^2	$2qr$	r^2

Ces fréquences sont celles de l'état d'équilibre, elles ne changent pas au cours des générations. Par ailleurs, les fréquences alléliques peuvent être calculées selon la méthode de comptage ou la méthode probabiliste.

4.3.2. Dominance partielle : A = B > C

Phénotypes	A		AB	B		C
Génotypes	AA	AC	AB	BB	BC	CC
fq génotypiques	p^2	$2pr$	$2pq$	q^2	$2qr$	r^2
fq phénotypiques	$p^2 + 2pr$		$2pq$	$q^2 + 2qr$		r^2

L'allèle C ne peut s'exprimer qu'en absence des deux autres allèles, au niveau du génotype homozygote CC.

$$\text{On sait que : } fq_{[C]} = r^2 \quad \text{donc} \quad fq_{(C)} = r = \sqrt{fq_{[C]}}$$

$$\text{Par ailleurs : } fq_{[A]} + fq_{[C]} = p^2 + 2pr + r^2 = (p + r)^2$$

$$\text{Donc : } p + r = \sqrt{fq_{[A]} + fq_{[C]}} \quad \text{d'où} \quad fq_{(A)} = p = \sqrt{fq_{[A]} + fq_{[C]}} - r$$

Connaissant p et r, on peut en déduire la valeur de q à partir de la formule $p + q + r = 1$ donc : $fq_{(B)} = q = 1 - (p + r)$.

4.3.3. Dominance totale : A > B > C

Phénotypes	A			B		C
Génotypes	AA	AB	AC	BB	BC	CC
fq génotypiques	p^2	$2pq$	$2pr$	q^2	$2qr$	r^2
fq phénotypiques	$p^2 + 2pq + 2pr$			$q^2 + 2qr$		r^2

L'allèle C ne peut s'exprimer qu'en absence des deux autres allèles, au niveau du génotype homozygote CC.

$$\text{On sait que : } fq_{[C]} = r^2 \quad \text{donc} \quad fq_{(C)} = r = \sqrt{fq_{[C]}}$$

$$\text{Par ailleurs : } fq_{[B]} + fq_{[C]} = q^2 + 2qr + r^2 = (q + r)^2$$

$$\text{Donc : } q + r = \sqrt{fq_{[B]} + fq_{[C]}} \quad \text{d'où} \quad fq_{(B)} = q = \sqrt{fq_{[B]} + fq_{[C]}} - r$$

Connaissant q et r, on peut en déduire la valeur de p à partir de la formule $p + q + r = 1$ donc : $fq_{(A)} = p = 1 - (q + r)$.

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

Une question centrale est de savoir si la loi de Hardy-Weinberg établie pour une population idéale s'applique également aux populations naturelles. L'application de cette loi dans les populations naturelles peut être vérifiée par un test statistique du Chi Deux (χ^2) qui permet de comparer les effectifs observés dans la population étudiée aux effectifs théoriques d'une population idéale. Ce test se résume en six étapes :

1. Emettre l'hypothèse que la population est à l'équilibre de Hardy-Weinberg.

- Calcul des fréquences alléliques.
- Calcul des effectifs théoriques des différents génotypes en utilisant la formule suivante :

Effectif théorique = Fréquence génotypique x Effectif total

- Calcul du « χ^2 calculé » en utilisant la formule suivante :

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

- Détermination du « χ^2 théorique » en utilisant les paramètres suivants (**Annexe 2**) :
 - Un risque α choisi par l'utilisateur qui est en général égal à 0,05.
 - Un nombre de degrés de liberté (**ddl**) égal à la différence entre le nombre de génotypes possibles et le nombre d'allèles du système génétique étudié.
- Comparaison entre le « χ^2 calculé » et le « χ^2 théorique », deux hypothèses (H_0) sont possibles :
 - χ^2 calculé < χ^2 théorique $\rightarrow H_0$ acceptée \rightarrow La population est à l'équilibre de Hardy-Weinberg.
 - χ^2 calculé > χ^2 théorique $\rightarrow H_0$ rejetée \rightarrow La population n'est pas à l'équilibre de Hardy-Weinberg.

AN : Chez l'homme, le groupe sanguin MN est déterminé par un gène à deux allèles codominants M et N. Une étude portant sur 730 australiens (N_t) a donné les résultats suivants :

Génotypes	MM	MN	NN
Effectifs observés	$N_1 = 22$	$N_2 = 216$	$N_3 = 492$

- H_0** : La population est à l'équilibre de Hardy-Weinberg.

2. Calcul des fréquences alléliques

$$fq_{(M)} = p = (2N_1 + N_2) / 2N_t = (2 \times 22 + 216) / 2 \times 730 \rightarrow p = 0,178$$

$$fq_{(N)} = q = (2N_3 + N_2) / 2N_t = (2 \times 492 + 216) / 2 \times 730 \rightarrow q = 0,822$$

3. Calcul des effectifs théoriques des différents génotypes

$$N_{1'} = p^2 \times N_t = (0,178)^2 \times 730 \rightarrow N_{1'} = 23,1$$

$$N_{2'} = 2pq \times N_t = (2 \times 0,178 \times 0,822) \times 730 \rightarrow N_{2'} = 213,6$$

$$N_{3'} = q^2 \times N_t = (0,822)^2 \times 730 \rightarrow N_{3'} = 493,2$$

4. Calcul du « χ^2 calculé »

$$\chi^2 \text{ cal} = (22 - 23,1)^2 / 23,1 + (216 - 213,6)^2 / 213,6 + (492 - 493,2)^2 / 493,2$$

$$\chi^2 \text{ cal} = 0,083$$

5. Détermination « χ^2 théorique »

$$\left. \begin{array}{l} \text{ddl} = 3 - 2 \rightarrow \text{ddl} = 1 \\ \alpha = 0,05 \end{array} \right\} \chi^2 \text{ théo} = 3,841$$

6. Comparaison entre le « χ^2 calculé » et le « χ^2 théorique »

$\chi^2 \text{ cal} < \chi^2 \text{ théo} \rightarrow H_0$ acceptée \rightarrow La population est à l'équilibre de Hardy-Weinberg.

6. Conclusion

Dans la plupart des cas, le modèle de Hardy-Weinberg constitue un bon descripteur de la structure génétique des populations naturelles car l'hypothèse de panmixie est souvent respectée alors que les effets des mutations, migration, sélection ne sont pas assez forts pour faire diverger les fréquences génotypiques des proportions du modèle de Hardy-Weinberg. Par ailleurs, Cet équilibre ne s'applique pas obligatoirement à tous les gènes d'une même population et peut ou non être rejeté en fonction du système génétique considéré.

Exercices

Pour tous les exercices on considère que $\alpha = 0,05$

Exercice 1

En Pologne, une étude portant sur un échantillon de 3100 personnes issues d'une population panmictique a révélé la présence de 1100 individus de groupe sanguin M, 1520 MN et 480 N ($M = N$). Pour le système Rhésus, 279 personnes étaient de groupe Rh- et 2821 de groupe Rh+ ($Rh+ > Rh-$).

- Calculer les fréquences des allèles M, N, Rh+ et Rh-.
- Quelle est la fréquence des hétérozygotes Rh+Rh- dans cette population ?

Exercice 2

L'analyse du polymorphisme de l'enzyme estérase 1 dans un échantillon de 300 personnes a révélé l'existence de trois niveaux de migration (A, B et C) à l'origine de six génotypes dont les effectifs sont les suivants :

Géno	AA	AB	AC	BB	BC	CC
Eff	$N_1 = 72$	$N_2 = 99$	$N_3 = 57$	$N_4 = 24$	$N_5 = 33$	$N_6 = 15$

- En admettant que $A = B = C$, la population est-elle à l'équilibre de Hardy-Weinberg ?

Exercice 3

Une étude du système du groupe sanguin ABO dans un échantillon de 5000 individus a permis d'identifier 2500 du groupe O, 1600 du groupe A, 700 du groupe B et 200 du groupe AB.

1. Calculer les fréquences des allèles A, B et O.

Exercice 4

Chez la drosophile, le gène responsable de la couleur des yeux est porté par le chromosome X. Ce gène est représenté par deux allèles, un allèle R dominant donnant à l'œil une couleur rouge et un allèle B récessif donnant à l'œil une couleur blanche. Dans une population panmictique du laboratoire, on a trouvé 340 mâles aux yeux rouges et 60 mâles aux yeux blancs.

1. Estimer la fréquence des allèles R et B.
2. Sachant que les fréquences alléliques sont les mêmes chez les mâles et chez les femelles, quel sera le pourcentage des femelles ayant les yeux blancs et celui des femelles ayant les yeux rouges dans cette population ?

Exercice 5

La coloration du pelage chez le lapin est déterminée par trois allèles : (S) sauvage, (H) himalayen et (A) albinos avec $S > H > A$ de fréquences respectives p, q et r. Si une population de lapins contenant des individus sauvages, himalayens et albinos est panmictique :

1. Quelles seront en fonction de p, q et r à la génération suivante, les fréquences attendues des différents génotypes et phénotypes ?

2. Dédurre des formules pour calculer les fréquences alléliques en fonction des différents phénotypes.

Un échantillon d'une population contient 168 lapins sauvages, 30 himalayens et 2 albinos.

3. Déterminer les fréquences alléliques p, q et r.
4. Calculer les proportions des différents génotypes parmi les lapins sauvages.

Annexe 1

- Application de la loi de Hardy-Weinberg

La structure génotypique de la nouvelle génération est :

Génotypes	AA	AB	BB
fq génotypiques	fq (AA) = p ²	fq (AB) = 2pq	fq (BB) = q ²

Les fréquences alléliques sont inchangées à la nouvelle génération. En effet, selon la formule de calcul des fréquences alléliques à partir des fréquences génotypiques on a :

$$fq(A) = p_1 = fq(AA) + \frac{1}{2} fq(AB) = p^2 + 2pq / 2 = p(p + q) = p$$

$$fq(B) = q_1 = fq(BB) + \frac{1}{2} fq(AB) = q^2 + 2pq / 2 = q(p + q) = q$$

Les fréquences génotypiques sont aussi constantes de génération en génération. En effet, si on étudie la deuxième génération on observera les croisements suivants :

Mâles \ Femelles	AA (p ²)	AB (2pq)	BB (q ²)
AA (p ²)	AA x AA (p ⁴)	AA x AB (2p ³ q)	AA x BB (p ² q ²)
AB (2pq)	AA x AB (2p ³ q)	AB x AB (4p ² q ²)	BB x AB (2pq ³)
BB (q ²)	AA x BB (p ² q ²)	BB x AB (2pq ³)	BB x BB (q ⁴)

Croisements	Fréquences des croisements	Fréquences génotypiques de la descendance		
		AA	AB	BB
AA X AA	p ⁴	p ⁴	/	/
AA X AB	2 (2p ³ q)	2p ³ q	2p ³ q	/
AA X BB	2 (p ² q ²)	/	2 (p ² q ²)	/
BB X BB	q ⁴	/	/	q ⁴
BB X AB	2 (2pq ³)	/	2pq ³	2pq ³
AB X AB	4p ² q ²	p ² q ²	2(p ² q ²)	p ² q ²
	α = 1	β = p²	Δ = 2pq	σ = q²

La somme de la colonne α est égale à 1

$$p^4 + 4p^3q + 6p^2q^2 + 4pq^3 + q^4 = 1$$

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

La somme de la colonne β est égale à p²

$$p^4 + 2p^3q + p^2q^2 = p^2 (p^2 + 2pq + q^2) = p^2$$

La somme de la colonne Δ est égale à 2pq

$$2p^3q + 4p^2q^2 + 2pq^3 = 2pq (p^2 + 2pq + q^2) = 2pq$$

La somme de la colonne σ est égale à q²

$$p^2q^2 + 2pq^3 + q^4 = q^2 (p^2 + 2pq + q^2) = q^2$$

Annexe 2

M.

TABLE DU CHI-DEUX :

$ddl \backslash \alpha$	0.90	0.80	0.70	0.50	0.30	0.20	0.10	0.05	0.02	0.01
1	0,0158	0,0642	0,148	0,455	1,074	1,642	2,706	3,841	5,412	6,635
2	0,211	0,446	0,713	1,386	2,408	3,219	4,605	5,991	7,824	9,210
3	0,584	1,005	1,424	2,366	3,665	4,642	6,251	7,815	9,837	11,341
4	1,064	1,649	2,195	3,357	4,878	5,989	7,779	9,488	11,668	13,277
5	1,610	2,343	3,000	4,351	6,064	7,289	9,236	11,070	13,388	15,086
6	2,204	3,070	3,828	5,348	7,231	8,558	10,645	12,592	15,033	16,812
7	2,833	3,822	4,671	6,346	8,383	9,803	12,017	14,067	16,622	18,475
8	3,490	4,594	5,527	7,344	9,524	11,030	13,362	15,507	18,168	20,090
9	4,168	5,380	6,393	8,343	10,656	12,242	14,684	16,919	19,679	21,666
10	4,865	6,179	7,267	9,342	11,781	13,442	15,987	18,307	21,161	23,209
11	5,578	6,989	8,148	10,341	12,899	14,631	17,275	19,675	22,618	24,725
12	6,304	7,807	9,034	11,340	14,011	15,812	18,549	21,026	24,054	26,217
13	7,042	8,634	9,926	12,340	15,119	16,985	19,812	22,362	25,472	27,688
14	7,790	9,467	10,821	13,339	16,222	18,151	21,064	23,685	26,873	29,141
15	8,547	10,307	11,721	14,339	17,322	19,311	22,307	24,996	28,259	30,578
16	9,312	11,152	12,624	15,338	18,418	20,465	23,542	26,296	29,633	32,000
17	10,085	12,002	13,531	16,338	19,511	21,615	24,769	27,587	30,995	33,409
18	10,865	12,857	14,440	17,338	20,601	22,760	25,989	28,869	32,346	34,805
19	11,651	13,716	15,352	18,338	21,689	23,900	27,204	30,144	33,687	36,191
20	12,443	14,578	16,266	19,337	22,775	25,038	28,412	31,410	35,020	37,566
21	13,240	15,445	17,182	20,337	23,858	26,171	29,615	32,671	36,343	38,932
22	14,041	16,314	18,101	21,337	24,939	27,301	30,813	33,924	37,659	40,289
23	14,848	17,187	19,021	22,337	26,018	28,429	32,007	35,172	38,968	41,638
24	15,659	18,062	19,943	23,337	27,096	29,553	33,196	36,415	40,270	42,980
25	16,473	18,940	20,867	24,337	28,172	30,675	34,382	37,652	41,566	44,314
26	17,292	19,820	21,792	25,336	29,246	31,795	35,563	38,885	42,856	45,642
27	18,114	20,703	22,719	26,336	30,319	32,912	36,741	40,113	44,140	46,963
28	18,939	21,588	23,647	27,336	31,391	34,027	37,916	41,337	45,419	48,278
29	19,768	22,475	24,577	28,336	32,461	35,139	39,087	42,557	46,693	49,588
30	20,599	23,364	25,508	29,336	33,530	36,250	40,256	43,773	47,962	50,892

Influence du régime de reproduction

« Écarts à la panmixie : Consanguinité »

1. Introduction

Chez de nombreuses espèces, les croisements ne se réalisent pas au hasard. L'hypothèse de panmixie n'étant pas respectée, la structure génétique des populations va s'écarter de celle d'une population théorique idéale. Ces croisements non aléatoires sont la conséquence d'un régime de reproduction appelé **consanguinité**.

2. Consanguinité

La consanguinité est un régime de reproduction où les unions se font entre individus **apparentés**, c'est-à-dire ayant un ou plusieurs ancêtres communs. Un individu issu d'une union entre deux apparentés est dit **consanguin**. A un locus donné, il aura donc pu recevoir de ses deux parents des copies d'un même gène de l'ancêtre commun qui sont dits **identiques**. Noter que ce terme a, en génétique des populations, un sens très précis et restrictif : un homozygote AA n'a pas, en général, deux gènes identiques ; sauf s'il est consanguin.

2.1. Causes de la consanguinité

Les causes d'une fréquence élevée de croisements consanguins sont de trois grandes origines :

2.1.1. Système de reproduction : Autogamie

Il permet chez certains organismes hermaphrodites l'autofécondation. C'est un cas extrême de consanguinité puisque chaque individu se croise avec lui-même. Ce phénomène existe chez

de nombreuses espèces de plantes comme le maïs, le blé, le riz, la tomate, le coton..... Chez les animaux, l'autofécondation est plus rare, elle existe chez certains vers parasites, chez des mollusques et divers crustacés.

2.1.2. Choix d'un croisement avec un apparenté

Il peut s'observer dans certaines populations humaines caractérisées par des structures sociales ou des traditions qui favorisent l'union entre individus apparentés.

2.1.3. Petite taille des populations

Elle est à l'origine de croisements consanguins fréquents à cause du choix limité des conjoints. La probabilité de s'apparier avec un apparenté est donc importante même si les croisements se réalisent au hasard. Ce phénomène s'applique à toutes les espèces dont les effectifs sont constamment faibles (espèces rares) et également chez les espèces ayant des effectifs plus grands, mais dont les populations sont spatialement fragmentées ou avec de forte structure sociale, ce qui contraint les individus à se reproduire à l'intérieur d'un groupe d'effectif limité.

2.2. Principaux paramètres mesurant la consanguinité

Les notions de **parenté** et de **consanguinité** sont centrales pour comprendre le fonctionnement génétique et l'évolution des populations. Il faut bien distinguer la parenté et la consanguinité :

- Deux individus sont **apparentés** lorsqu'ils ont un ou plusieurs ancêtres communs.
- Un individu est **consanguin** lorsqu'il est issu d'un croisement entre deux individus apparentés.

2.2.1. Coefficient de parenté (Φ)

Le coefficient de parenté concerne deux individus, il correspond à la probabilité pour qu'un allèle pris au hasard à un locus donné chez un individu soit identique par descendance à un allèle pris au hasard au même locus chez l'autre individu.

2.2.2. Coefficient de consanguinité (f)

On appelle coefficient de consanguinité, la probabilité qu'un individu possède, à un locus donné, deux allèles identiques par descendance donc proviennent toujours de la copie sans mutation d'un allèle présent chez les ancêtres communs aux deux parents de l'individu consanguin.

Comme toute probabilité, le coefficient de consanguinité varie entre 0 et 1. Il est nul uniquement lorsque les parents de l'individu considéré ne possèdent aucun ancêtre commun. Sinon, la valeur du coefficient de consanguinité est d'autant plus élevée que le degré d'apparentement de ses parents est fort.

Remarque : D'après leurs définitions, le coefficient de consanguinité d'un individu « I » est égal au coefficient de parenté de ses deux parents « P et M ».

$$f_{(I)} = \Phi_{(P,M)}$$

2.3. Mesure de la parenté et de la consanguinité

Considérons un individu I et ses parents P et M apparentés par l'ancêtre commun A. La mesure de la consanguinité de l'individu I et de la parenté des individus P et M se fait grâce à la formule suivante (**Annexe 1**) :

$$f_{(I)} = \Phi_{(P,M)} = \sum (1/2)^{N+1} (1 + f)$$

$f_{(I)}$: Coefficient de consanguinité de l'individu I.

$\Phi_{(P,M)}$: Coefficient de parenté des individus P et M.

\sum : Somme des chaînes de parenté reliant les deux individus P et M en passant par l'ancêtre commun A.

N : Nombre de maillons sur une chaîne de parenté.

f : Coefficient de consanguinité de l'ancêtre commun A.

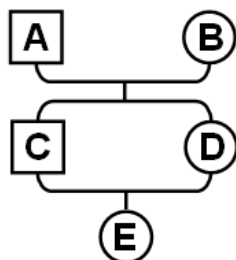
Remarques :

- Une chaîne de parenté ne peut pas passer deux fois par le même individu.
- Deux chaînes de parenté sont différentes dès qu'elles ont au moins un maillon différent.
- Si les parents de l'ancêtre commun ne sont pas apparentés, la valeur du coefficient de consanguinité de l'ancêtre commun est prise égale à 0.
- En absence d'informations concernant l'ascendance de l'ancêtre commun on considère qu'il n'est pas consanguin donc la valeur du coefficient de consanguinité de l'ancêtre commun est prise égale à 0.

2.4. Coefficients des parentés les plus courantes

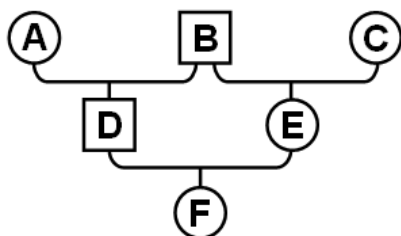
Les ancêtres communs n'ayant pas d'ascendance connue sont considérés comme non consanguins, la valeur de f est prise égale à 0 pour chacun d'eux.

2.4.1. Germain : Frère et Sœur (Parenté de 1^{er} degré)



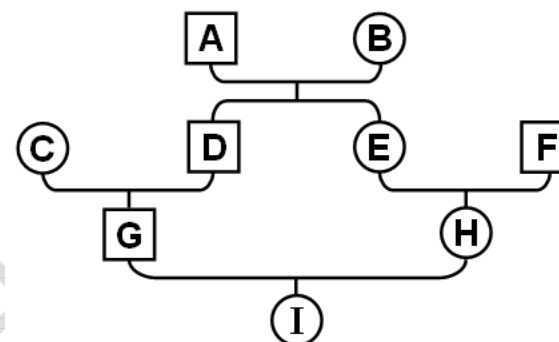
Σ	N	$f_{(E)} = \Phi_{(C,D)} = \sum (1/2)^{N+1}$
C - A - D C - B - D	2	$f_{(E)} = \Phi_{(C,D)} = 2 \times (1/2)^3 = 2 \times 1/8$ $f_{(E)} = \Phi_{(C,D)} = 1/4$

2.4.2. Demi-germain : Demi-frère et Sœur (Parenté de 2^{ème} degré)



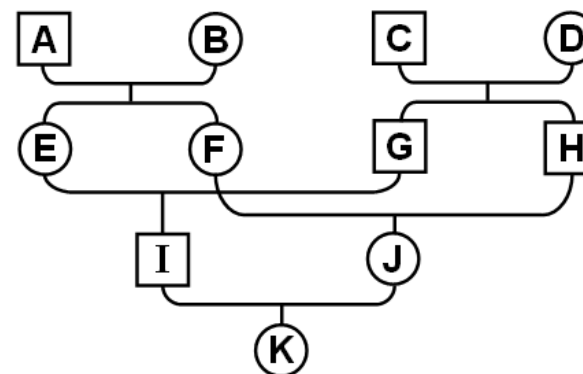
Σ	N	$f_{(F)} = \Phi_{(D,E)} = \sum (1/2)^{N+1}$
D - B - E	2	$f_{(F)} = \Phi_{(D,E)} = 1 \times (1/2)^3$ $f_{(F)} = \Phi_{(D,E)} = 1/8$

2.4.3. Cousins germain (Parenté de 3^{ème} degré)



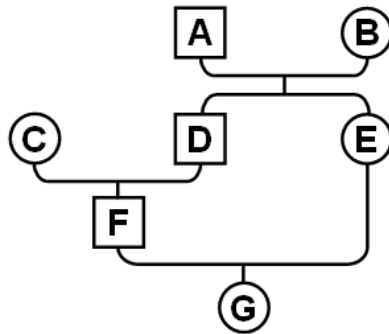
Σ	N	$f_{(I)} = \Phi_{(G,H)} = \sum (1/2)^{N+1}$
G - D - A - E - H G - D - B - E - H	4	$f_{(I)} = \Phi_{(G,H)} = 2 \times (1/2)^5 = 2 \times 1/32$ $f_{(I)} = \Phi_{(G,H)} = 1/16$

2.4.4. Double cousins germain (Parenté de 3^{ème} degré)



Σ	N	$f_{(K)} = \Phi_{(I,J)} = \sum (1/2)^{N+1}$
I - E - A - F - J I - E - B - F - J I - G - C - H - J I - G - D - H - J	4	$f_{(K)} = \Phi_{(I,J)} = 4 \times (1/2)^5 = 4 \times 1/32$ $f_{(K)} = \Phi_{(I,J)} = 1/8$

2.4.5. Tante et Neveu (Parenté de 2^{ème} degré)



Σ	N	$f_{(G)} = \Phi_{(F,E)} = \sum (1/2)^{N+1}$
F - D - A - E F - D - B - E	3	$f_{(G)} = \Phi_{(F,E)} = 2 \times (1/2)^4 = 2 \times 1/16$ $f_{(G)} = \Phi_{(F,E)} = 1/8$

2.5. Structure génétique des populations consanguines

2.5.1. Coefficient moyen de consanguinité (F)

Dans une population naturelle où existe des croisements entre apparentés, tous les individus n'ont pas le même coefficient de consanguinité. On caractérise alors la population par **un coefficient moyen de consanguinité F** qui est la moyenne des coefficients de consanguinité individuels.

$$F = \frac{(N_1 \times f_1) + (N_2 \times f_2) + \dots + (N_n \times f_n)}{N_t}$$

f_1, f_2, \dots, f_n : Coefficients de consanguinité relatifs à chaque type de consanguinité.

N_1, N_2, \dots, N_n : Effectifs relatifs à chaque type de consanguinité.

N_t : Effectif total de la population étudiée.

AN : Une grande population où 16% des mariages sont réalisés entre cousins germains alors que les autres (84%) sont panmictiques présentera un coefficient moyen de consanguinité égal à :

$$F = (16/100) \times (1/16) + (84/100) \times 0 \rightarrow F = 1/100$$

2.5.2. Conséquence de la consanguinité sur les fréquences génotypiques

L'existence d'individus consanguins dans la population va modifier les fréquences des différents génotypes puisqu'elle va augmenter la fréquence des homozygotes dans la population et donc réduire la fréquence des hétérozygotes. En effet, considérons le cas simple d'un locus diallélique A et B de fréquence p et q dans une population où le taux moyen de consanguinité est F ; les fréquences des différents génotypes sont les suivantes (**Annexe 2**) :

$$f_{q(AA)} = p^2 + Fpq$$

$$f_{q(AB)} = (1 - F) 2pq \quad \text{Donc} \quad f_{q(AB)} = 2pq - 2Fpq$$

$$f_{q(BB)} = q^2 + Fpq$$

Pour un gène diallélique, la consanguinité se traduit par une réduction de la fréquence des hétérozygotes d'une quantité ($2Fpq$) et une augmentation de la fréquence des homozygotes d'une quantité ($2Fpq$) qui se répartit de moitié dans chaque catégorie d'homozygotes.

2.5.3. Conséquence de la consanguinité sur les fréquences alléliques

Si elle change les fréquences génotypiques, la consanguinité ne modifie pas les fréquences des allèles dans la population qui restent p et q dans le cas d'un gène diallélique. En effet :

$$fq(A) = fq(AA) + 1/2 fq(AB)$$

$$fq(A) = p^2 + Fpq + 1/2 \times (2pq - 2Fpq)$$

$$fq(A) = p^2 + Fpq + pq - Fpq$$

$$fq(A) = p^2 + pq$$

$$fq(A) = p(p+q) \xrightarrow{1}$$

$$fq(A) = p$$

De la même façon, on trouvera que la $fq(B) = q$

Exercices

Exercice 1

La composition génotypique d'une population consanguine est la suivante :

$$fq(AA) = 0,3 \quad fq(AB) = 0,2 \quad fq(BB) = 0,5$$

- Déterminer la valeur de F .

Exercice 2

- Calculer la valeur du coefficient de consanguinité des enfants issus des mariages suivants :

- Type 1** : Mariage entre un oncle et sa nièce.
- Type 2** : Mariage entre un frère et sa demi-sœur.
- Type 3** : Mariage entre cousins germains.

Une population comprend 5% de mariages de type 1, 10% de mariages de type 2, 30% de mariages de type 3 et 55% de mariages entre individus non apparentés (type 4).

- Calculer le coefficient moyen de consanguinité de cette population (on supposera que chaque type de mariage produit le même nombre de descendants).

Exercice 3

Soit un gène représenté par deux allèles A et B (avec $A > B$) de fréquences respectives p et q .

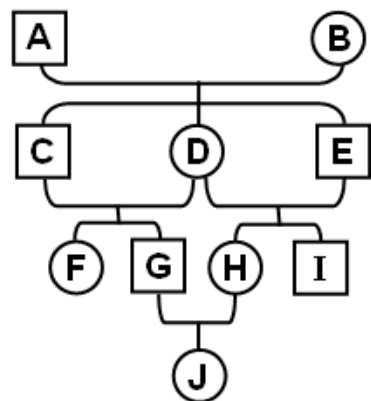
1. Quelles seront les fréquences des homozygotes pour l'allèle récessif de fréquence $q = 0,2$; dans le cas de panmixie et dans celui d'une consanguinité équivalente à celle entre double cousins germains qui concerne toute la population ?
2. Quelle conclusion peut-on tirer ?

Exercice 4

La couleur blanche de la laine chez le mouton est due à un gène dominant B, alors que la couleur noire est due à son allèle récessif b. Un échantillon de 900 moutons est composé de 891 blancs et 9 noirs.

1. Quelle est la fréquence des différents allèles ?

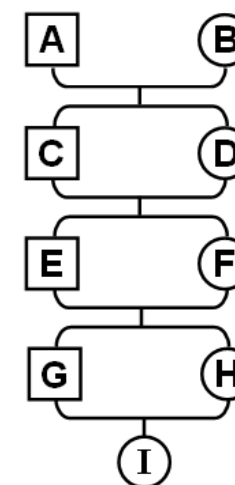
Soit le pédigrée suivant, représentant trois générations de croisements chez ces moutons :



2. Calculer le coefficient de parenté des individus D et E.
3. Calculer le coefficient de consanguinité de l'individu J.
4. Quelle est la probabilité pour qu'il soit blanc ? noir ?

Exercice 5

Soit le pédigrée d'un système d'accouplement frère-sœur de trois générations successives :

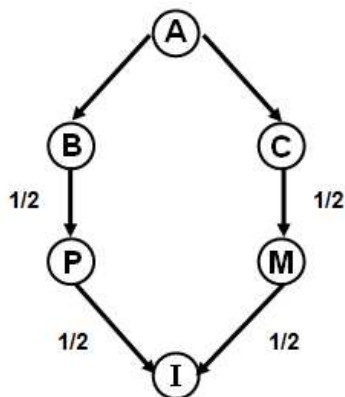


1. Calculer le coefficient de parenté des derniers ancêtres en prenant en considération que $f_{(A)} = 0,01$ et $f_{(B)} = 0,05$.

Annexe 1

- **Mesure de la parenté et de la consanguinité**

Considérons un individu I et ses parents P et M apparentés par l'ancêtre commun A.



A a transmis une copie de l'un de ses gènes à chacun de ses enfants B et C. A chaque génération ultérieure, ce gène a une probabilité 1/2 d'être transmis. La probabilité pour que le gène transmis par P à I provienne de A est finalement de $(1/2)^{N_P}$, N_P étant le nombre de maillons du père jusqu'à l'ancêtre commun. De même, la probabilité pour que le gène transmis par M à I provienne de A est de $(1/2)^{N_M}$, N_M étant le nombre de maillons de la mère jusqu'à l'ancêtre commun. Ces deux événements sont indépendants. La probabilité pour qu'un gène tiré au hasard chez P et un gène tiré au hasard chez M proviennent de A est donc $(1/2)^{N_P + N_M}$.

Cette probabilité assure l'**origine** commune des deux gènes, mais non leur **identité**, car chez A il y a deux gènes homologues. Les deux gènes transmis à B et C sont soit des copies d'un même gène de A

(probabilité 1/2), soit des copies de chacun des deux gènes homologues de A (probabilité 1/2) ; dans ce cas, ces deux gènes peuvent être quand même identiques, avec une probabilité f (coefficient de consanguinité de l'ancêtre commun). Puisque les deux événements sont indépendants, la probabilité que les deux tirages au sort donnent des gènes identiques est donc $(1/2 + 1/2 f_A)$.

Nous avons donc : $f_{(I)} = \Phi_{(P,M)} = (1/2)^{N_P + N_M} (1/2 + 1/2 f_A)$

Ce qui s'écrit : $f_{(I)} = \Phi_{(P,M)} = (1/2)^{N_P + N_M + 1} (1 + f_A)$

S'il existe plusieurs voies P-A-M par lesquelles P et M ont pu recevoir deux gènes identiques, ou plusieurs ancêtres en communs à P et M dont ils ont pu recevoir deux gènes identiques, la formule se généralise et on a alors :

$$f_{(I)} = \Phi_{(P,M)} = \sum (1/2)^{N+1} (1 + f)$$

$f_{(I)}$: Coefficient de consanguinité de l'individu I.

$\Phi_{(P,M)}$: Coefficient de parenté des individus P et M.

\sum : Somme des chaînes de parenté reliant les deux individus P et M en passant par l'ancêtre commun A.

N : Nombre de maillons sur une chaîne de parenté.

f : Coefficient de consanguinité de l'ancêtre commun A.

Annexe 2

- Conséquence de la consanguinité sur les fréquences génotypiques**

L'existence d'individus consanguins dans la population va modifier les fréquences des différents génotypes puisqu'elle va augmenter la fréquence des homozygotes dans la population et donc réduire la fréquence des hétérozygotes. En effet, considérons le cas simple d'un locus diallélique A et B de fréquence p et q dans une population où le taux moyen de consanguinité est F. Tirons un individu au hasard, les deux exemplaires du gène, présents chez cet individu, peuvent être **(voir le tableau)**.

Au total les fréquences des différents génotypes seront :

$$fq_{(AA)} = F p + (1 - F) p^2 = F p + p^2 - F p^2 = p^2 + F p (1 - p)$$

Avec $1 - p = q$ **Donc** $fq_{(AA)} = p^2 + F pq$

$$fq_{(AB)} = (1 - F) 2pq \quad \text{Donc} \quad fq_{(AB)} = 2pq - 2 Fpq$$

$$fq_{(BB)} = F q + (1 - F) q^2 = F q + q^2 - F q^2 = q^2 + F q (1 - q)$$

Avec $1 - q = p$ **Donc** $fq_{(BB)} = q^2 + F pq$

Pour un gène diallélique, la consanguinité se traduit par une réduction de la fréquence des hétérozygotes d'une quantité (2Fpq) et une augmentation de la fréquence des homozygotes d'une quantité (2Fpq) qui se répartit de moitié dans chaque catégorie d'homozygotes.

Probabilité Génotypes possibles	Identiques par ascendance		Pas identiques par ascendance		
	F		AA	AB	BB
Probabilité des différents génotypes dans la population	AA	1 ^{er} exemplaire A → probabilité p 2 ^{ème} exemplaire A → probabilité 1 (car identiques) Donc Probabilité AA → F p	1 ^{er} exemplaire A → probabilité p 2 ^{ème} exemplaire A → probabilité p Donc Probabilité AA → (1 - F) p ²	1 ^{er} exemplaire A → probabilité p 2 ^{ème} exemplaire B → probabilité q Ou 1 ^{er} exemplaire B → probabilité q 2 ^{ème} exemplaire A → probabilité p 2 ^{ème} exemplaire B → probabilité q Donc Probabilité AB → probabilité p Donc Probabilité AB → (1 - F) 2pq	1 ^{er} exemplaire B → probabilité q 2 ^{ème} exemplaire B → probabilité q Donc Probabilité BB → (1 - F) q ²
	BB	1 ^{er} exemplaire B → probabilité q 2 ^{ème} exemplaire B → probabilité 1 (car identiques) Donc Probabilité BB → F q	1 ^{er} exemplaire A → probabilité p 2 ^{ème} exemplaire A → probabilité p Donc Probabilité AA → (1 - F) p ²	1 ^{er} exemplaire A → probabilité p 2 ^{ème} exemplaire B → probabilité q Ou 1 ^{er} exemplaire B → probabilité q 2 ^{ème} exemplaire A → probabilité p 2 ^{ème} exemplaire B → probabilité q Donc Probabilité AB → probabilité p Donc Probabilité AB → (1 - F) 2pq	1 ^{er} exemplaire B → probabilité q 2 ^{ème} exemplaire B → probabilité q Donc Probabilité BB → (1 - F) q ²

Influence des pressions évolutives

1. Introduction

Les **pressions évolutives** sont des actions qui s'exercent sur la population conduisant à la modification de sa structure allélique. Il en existe quatre types, correspondant à l'abandon de diverses hypothèses de la loi de Hardy-Weinberg :

- La mutation.
- La migration.
- La sélection.
- La dérive génétique (liée à un effectif fini).

2. Mutation

La **mutation** est le changement de l'information génétique codée par l'ADN. La génétique des populations s'intéresse principalement aux mutations qui se produisent dans la lignée germinale donc aux **taux de mutation** par génération et leurs effets sur les fréquences alléliques des gènes concernés.

2.1. Impact des taux de mutation sur les fréquences alléliques

Prenons l'exemple d'un gène représenté par deux allèles A et B de fréquences respectives p et q (avec $p + q = 1$). On considère que A mute en B avec un taux de mutation **u** et B mute en A avec un taux de mutation **v**. En effet, dans les conditions naturelles, la **mutation directe** ($A \rightarrow B$) est généralement accompagnée d'une **mutation reverse** ($B \rightarrow A$) pour produire un **nouvel état d'équilibre** caractérisé

par de nouvelles fréquences alléliques **p'** et **q'** (avec $p' + q' = 1$). Les changements dans les fréquences de A et de B sont représentés respectivement par **Δp** et **Δq** .

On a : $\Delta p = \text{Gain en A} - \text{Perte en A} = vq - up$

A l'équilibre : Gain en A = Perte en A

Donc : $\Delta p = 0 = vq - up$

$$vq = up$$

$$v(1-p) = up$$

$$v - vp = up$$

$$v = up + vp$$

$$v = p(u + v)$$

$$p' = \frac{v}{u + v}$$

On sait que : $p' + q' = 1$

$$\text{Donc : } q' = 1 - p' = 1 - \frac{v}{u + v} = \frac{u + v - v}{u + v}$$

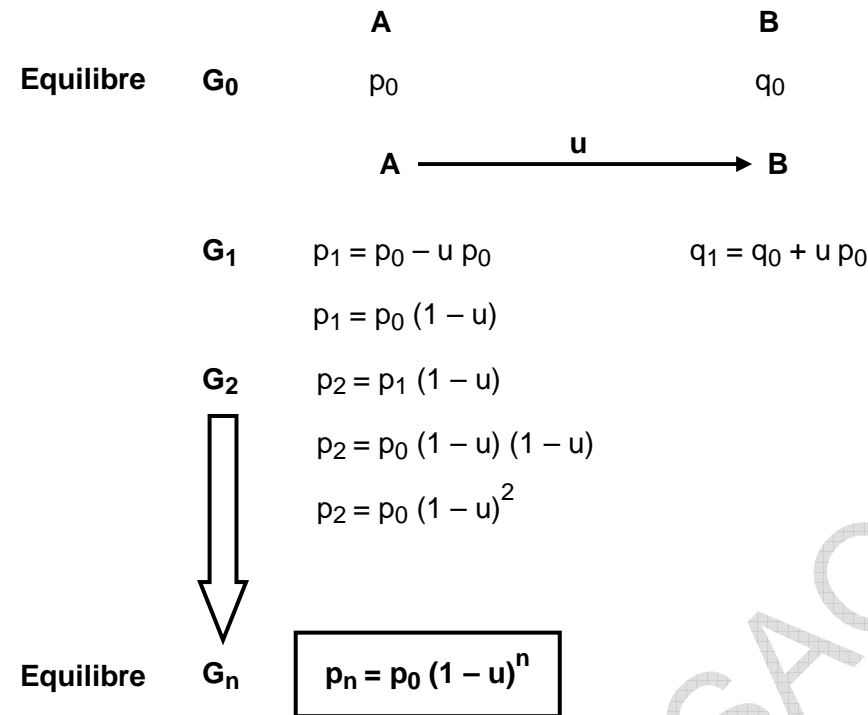
$$q' = \frac{u}{u + v}$$

Il est important de noter que les valeurs de p' et q' dépendent uniquement des taux de mutation et sont indépendantes des fréquences alléliques initiales p et q.

2.2. Vitesse du processus

Quand un facteur induit une variation des fréquences alléliques, il ne suffit pas de définir la valeur limite des fréquences alléliques à l'équilibre, il convient aussi d'estimer la vitesse avec laquelle la composition génétique de la population tend vers cet équilibre. Cette vitesse est estimée par le nombre des générations qui séparent les

deux états d'équilibre. Si le gène A, de fréquence p, mute en B à chaque génération à un taux de mutation u, alors que le taux de réversion est négligeable on a :



n : Nombre de générations qui séparent les deux états d'équilibre.

On a : $p_n / p_0 = (1 - u)^n$

$\text{Log} (p_n / p_0) = \text{Log} (1 - u)^n$

$\text{Log} p_n - \text{Log} p_0 = n \text{Log} (1 - u)$

$$n = \frac{\text{Log } p_n - \text{Log } p_0}{\text{Log} (1 - u)}$$

Les taux de mutations sont toujours des paramètres très faibles, de l'ordre de 10^{-5} à 10^{-10} , ce qui donne des vitesses d'évolution extrêmement lentes. Le rôle de la mutation est donc négligeable sur l'évolution de la fréquence des gènes, comparé à celui des autres pressions évolutives.

Il est néanmoins important de signaler qu'elle a un rôle primordial dans l'évolution puisqu'elle est la seule source de gènes nouveaux. Mais une fois qu'un nouveau gène est apparu par mutation, ce n'est plus elle qui va déterminer son devenir : si le nouvel allèle est défavorable, ou s'il est plus favorable que les anciens, c'est la sélection qui va déterminer l'évolution ultérieure de sa fréquence. Si le nouvel allèle ne présente ni avantage ni désavantage, c'est la dérive génétique qui interviendra.

3. Migration

La **migration** est le passage d'individus d'une population à une autre, par exemple sous forme de graines ou de pollen chez les plantes. Elle peut modifier les fréquences alléliques lorsque celles de la population qui fournit les migrants diffèrent de celles de la population qui les reçoit.

3.1. Impact de la migration sur les fréquences alléliques

La formulation mathématique de la migration et ses effets sur la structure génétique des populations est très complexe. On se contentera ici de l'étude d'un modèle simple appelé " **modèle de l'île** " où une population " **île** " reçoit à chaque génération des migrants d'une autre population " **continent** ", **sans qu'il ait des migrations dans le sens opposé.**

Représentons la fréquence du gène A par p_c dans la population continent qui fournit les migrants et par p_i dans la population île qui les reçoit. Soit m la fraction de migrants dans la population qui reçoit. La fréquence de A après une génération sera de :

$$p'_i = (1 - m) p_i + m p_c \quad \text{avec} \quad m = N_m / N_t = N_m / (N_m + N_i)$$

$(1 - m)$: Portion des non migrants.

$(1 - m) p_i$: Contribution faite par la population île.

$m p_c$: Contribution faite par les migrants de la population continent.

N_m : Effectif des migrants.

N_i : Effectif de la population île avant de recevoir les migrants.

N_t : Effectif de la population île après la réception des migrants.

Le changement dans la valeur de p est exprimé par Δp

Avec : $\Delta p = p_{i(1)} - p_i$

$$\Delta p = (1 - m) p_i + m p_c - p_i$$

$$\Delta p = p_i - m p_i + m p_c - p_i$$

$$\Delta p = m p_c - m p_i$$

$$\Delta p = m (p_c - p_i)$$

La fréquence de A dans la génération suivante peut donc être calculée suivant la formule :

$$p'_i = p_i + \Delta p$$

Une population qui reçoit, de manière répétée, des migrants d'une autre population sans que des migrations aient lieu en sens opposé, verra sa spécificité génétique disparaître progressivement. A terme, sa composition génétique sera identique à celle de la population d'origine des migrants. Les taux de migrations sont des paramètres beaucoup plus élevés que les taux de mutations, de l'ordre de 10^{-1} à 10^{-5} , ce qui donne des vitesses d'évolution beaucoup moins lentes.

4. Sélection

La **sélection naturelle** est le principal agent causal de l'évolution. Il s'agit d'un couple formé par un **complexe de facteurs du milieu** et une **population locale** d'une espèce. Le complexe de facteurs du milieu (le froid, la chaleur, la sécheresse, la compétition, les prédateurs....) favorise ou défavorise le bien-être et la reproduction chez les membres de la population. La population locale est formée de membres qui n'ont pas tous la même hérédité. Le génotype total de certains convient au complexe de facteurs qui constituent le milieu et fera qu'ils se porteront et se reproduiront bien dans ce milieu. Celui de certains autres, au contraire, ne convient pas à ces conditions de milieu et nuira à des degrés divers à leur survie ou à leur reproduction. Les premiers augmenteront en nombre alors que les seconds diminueront et pourront même disparaître totalement : c'est la **sélection naturelle**.

4.1. Valeur sélective "w"

En matière de sélection, chaque génotype est affecté d'un paramètre nommé **valeur sélective "w"**. L'idée est qu'un génotype est d'autant meilleur que les individus qui le possèdent, laissent plus de descendants. Dans une population, ce n'est pas donc le plus beau et le plus fort qui sera le mieux adapté mais le plus prolifique.

La valeur sélective d'un génotype peut être décomposée en deux facteurs :

- La **viabilité**, c'est-à-dire la probabilité pour un zygote ayant ce génotype de survivre jusqu'à l'âge reproducteur.
- La **fertilité**, c'est le nombre moyen de descendants par survivant.

Prenant le cas d'un gène diallélique (A et B), trois génotypes sont possibles : AA, AB et BB. Pour calculer la valeur sélective relative de chaque génotype, on procède tel qu'indiqué dans le tableau suivant :

Génotypes	N° initial d'individus	Viabilité	fertilité	N° moyen de descendants par génotype	Valeur sélective
AA	100	80	5	$0,8 \times 5 = 4,0$	$w_1 = 4/4 = 1$
AB	100	70	4	$0,7 \times 4 = 2,8$	$w_2 = 2,8/4 = 0,7$
BB	100	40	3	$0,4 \times 3 = 1,2$	$w_3 = 1,2/4 = 0,3$

Donc, on divise par lui-même le nombre moyen de descendants du génotype le plus prolifique, ce qui donne une valeur sélective optimum de "1" ; ensuite on divise par ce même nombre le nombre moyen de descendants de chaque autre génotype pour obtenir la valeur sélective de ces génotypes.

Selon les ouvrages, on trouve d'autres noms attribués à la valeur sélective ; par exemple : aptitude darwinienne, aptitude relative, valeur adaptative et fitness.

4.2. Coefficient de sélection

Le **coefficient de sélection "s"** est une mesure de la diminution de la valeur sélective d'un génotype par rapport à la valeur optimum "1".

$$s = 1 - w$$

donc

$$w = 1 - s$$

En d'autres termes, la valeur sélective exprime le degré d'avantage d'un génotype et le coefficient de sélection en exprime le degré de désavantage.

4.3. Evolution d'une population soumise à la sélection

Considérons une population possédant, pour le locus étudié, deux allèles A et B de fréquences respectives p et q (avec $p + q = 1$). Supposons que les conditions d'une population idéale sont respectées excepter pour la sélection. On attribue alors aux trois génotypes possibles AA, AB et BB ; les fréquences génotypiques p^2 , $2pq$ et q^2 ainsi que les valeurs sélectives w_1 , w_2 et w_3 . Etant donné les différences de viabilité et de fertilité entre ces trois génotypes, la contribution de chaque génotype à l'ensemble des gamètes de la génération suivante sera $w_1 p^2$, $w_2 2pq$ et $w_3 q^2$.

Après un cycle de sélection, les nouvelles fréquences alléliques p' et q' s'exprimeront de la façon suivante :

$$p' = \frac{w_1 p^2 + \frac{1}{2} w_2 2pq}{W}$$

et

$$q' = \frac{w_3 q^2 + \frac{1}{2} w_2 2pq}{W}$$



$$p' = \frac{w_1 p^2 + w_2 pq}{W}$$

$$q' = \frac{w_3 q^2 + w_2 pq}{W}$$

W représente la **valeur sélective moyenne de la population**

$$W = w_1 p^2 + w_2 2pq + w_3 q^2$$

On peut alors calculer la variation des fréquences alléliques d'une génération à la suivante sous l'effet de la sélection :

$$\Delta p = p' - p \quad \text{et} \quad \Delta q = q' - q$$

4.4. Etude de certains cas particuliers

Prenons l'exemple d'un gène représenté par deux allèles A et B de fréquences respectives p et q (avec p + q = 1).

4.4.1. Sélection contre un allèle létal récessif (Modèle 1)

Génotype	AA	AB	BB	Total
fq géno avant la sélection	p ²	2pq	q ²	1
w	1	1	0	/
W	p ² + 2pq + 0 = p(p + 2q) = p[(p + q) + q]			W = p(1 + q)
fq géno après la sélection	p ² / W	2pq / W	0	1

$$q' = \frac{w_3 q^2 + w_2 pq}{W} = \frac{\cancel{p}q}{\cancel{p}(1 + q)} \rightarrow q' = \frac{q}{1 + q}$$

$$\Delta q = q' - q = \frac{q}{1 + q} - q = \frac{q - q(1 + q)}{1 + q} \rightarrow \Delta q = -\frac{q^2}{1 + q}$$

4.4.2. Sélection contre un allèle sublétal récessif (Modèle 2)

Dans ce cas de figure, les homozygotes récessifs produisent des descendants mais moins que les autres génotypes.

Génotype	AA	AB	BB	Total
fq géno avant la sélection	p ²	2pq	q ²	1
w	1	1	1 - s	/
W	p ² + 2pq + (1 - s)q ²			W = 1 - sq ²
fq géno après la sélection	p ² / W	2pq / W	(1 - s)q ² / W	1

$$q' = \frac{(1 - s)q^2 + pq}{1 - sq^2}$$

et

$$\Delta q = \frac{-sq^2(1 - q)}{1 - sq^2}$$

4.4.3. Sélection contre un allèle dominant (Modèle 3)

Lorsque A est complètement dominant et que la sélection élimine tous les dominants (gène létal dominant ; $s = 1$), la fréquence de cet allèle est réduite à 0 en une génération.

Lorsque $s < 1$ (gène sublétal dominant) la sélection agit comme suit :

Génotype	AA	AB	BB	Total
fq géno avant la sélection	p^2	$2pq$	q^2	1
w	$1 - s$	$1 - s$	1	/
W	$(1 - s) p^2 + (1 - s) 2pq + q^2$			$W = 1 - sp(2 - p)$
fq géno après la sélection	$(1 - s) p^2 / W$	$(1 - s) 2pq / W$	q^2 / W	1

$$p' = \frac{p - sp}{1 - sp(2 - p)}$$

et

$$\Delta p = \frac{-sp(1 - p)^2}{1 - sp(2 - p)}$$

4.4.4. Sélection contre les homozygotes (Modèle 4)

Cette situation se présente lorsque la coexistence de deux allèles chez l'hétérozygote produit l'hétérosis qui désigne l'augmentation des capacités et/ou la vigueur d'un hybride par rapport aux lignées pures. Les hétérozygotes sont alors adaptativement supérieurs aux homozygotes et la sélection conduit à un état d'équilibre qui assure la conservation des deux allèles dans la population contrairement aux systèmes précédents où l'allèle défavorisé est éliminé avec le temps.

Génotype	AA	AB	BB	Total
fq géno avant la sélection	p^2	$2pq$	q^2	1
w	$1 - s_A$	1	$1 - s_B$	/
W	$(1 - s_A) p^2 + 2pq + (1 - s_B) q^2$			$W = 1 - s_A p^2 - s_B q^2$
fq géno après la sélection	$(1 - s_A) p^2 / W$	$2pq / W$	$(1 - s_B) q^2 / W$	1

$$q' = \frac{q(1 - s_B q)}{1 - s_A p^2 - s_B q^2}$$

et

$$\Delta q = \frac{pq(s_A p - s_B q)}{1 - s_A p^2 - s_B q^2}$$

On distingue les trois cas suivants :

	Δq	$f_{q(B)} = q'$
$s_A p > s_B q$	$\Delta q > 0$ (valeur positive)	\nearrow
$s_A p < s_B q$	$\Delta q < 0$ (valeur négative)	\searrow
$s_A p = s_B q$	$\Delta q = 0$	Point d'équilibre

Les fréquences de A et B au point d'équilibre (\bar{p} et \bar{q} , respectivement) dépendent uniquement des coefficients de sélection :

$$\left. \begin{aligned} \text{A l'équilibre : } s_A \bar{p} &= s_B \bar{q} \\ s_A \bar{p} - s_B \bar{q} &= 0 \\ s_A (1 - \bar{q}) - s_B \bar{q} &= 0 \\ s_A - s_A \bar{q} - s_B \bar{q} &= 0 \\ s_A &= s_A \bar{q} + s_B \bar{q} \\ s_A &= \bar{q} (s_A + s_B) \end{aligned} \right\}$$

$$\bar{q} = \frac{s_A}{s_A + s_B}$$

La valeur de \bar{p} peut être obtenue de la même façon :

$$\bar{p} = \frac{s_B}{s_A + s_B}$$

4.4.5. Sélection contre les hétérozygotes (Modèle 5)

Dans ce cas, comme précédemment, il y a production d'un équilibre des fréquences. Néanmoins, cet équilibre est instable parce que tout écart à partir du point d'équilibre déclenche un processus de diminution progressive qui conduit à l'élimination de l'allèle défavorisé au moment de l'écart initial et à la fixation de l'autre. Dans le modèle qui suit, on suppose que les deux homozygotes ont la même valeur sélective. Dans ce cas, l'équilibre sera atteint lorsque les deux allèles auront la même fréquence ($f_q = 0,5$).

Génotype	AA	AB	BB	Total
fq géno avant la sélection	p^2	$2pq$	q^2	1
w	1	$1 - s$	1	/
W	$p^2 + (1 - s) 2pq + q^2$			$W = 1 - s 2pq$
fq géno après la sélection	p^2 / W	$(1 - s) 2pq / W$	q^2 / W	1

$$q' = \frac{q - s pq}{1 - s 2pq}$$

et

$$\Delta q = \frac{s 2pq (q - 1/2)}{1 - s 2pq}$$

4.5. Différentes formes de sélection naturelle

4.5.1. Sélection stabilisante

Elle agit lorsque les individus présentant un phénotype moyen possèdent la plus grande valeur sélective. Elle permet donc de maintenir la population stable génétiquement puisqu'elle favorise les hétérozygotes et élimine les homozygotes (**Modèle 4**).

4.5.2. Sélection directionnelle

Elle intervient quand le milieu change progressivement dans une direction particulière. Les individus à qui ce changement convient, et qui de ce point de vue ne correspondent pas à la moyenne de la population, produisent donc plus de descendants (**Modèle 1, 2 et 3**).

4.5.3. Sélection diversifiante

On parle de sélection diversifiante lorsque les individus présentant des phénotypes extrêmes possèdent les plus grandes valeurs sélectives. Elle est donc dirigée contre les hétérozygotes et en faveur des homozygotes (**Modèle 5**).

Exercices

Exercice 1

En prenant des taux de mutation de $u = 10^{-5}$ et $v = 10^{-4}$

1. Quelles seront les fréquences d'équilibre de deux allèles de fréquence p et q ?

Exercice 2

En prenant le cas d'un taux de mutation $u = 10^{-5}$ par gamète et par génération, quelle sera la fréquence d'équilibre (p) d'un allèle A :

1. Au terme de 1000 générations pour $p_0 = 1$.
2. Au terme de 2000 générations pour $p_0 = 0,5$.
3. Au terme de 10000 générations pour $p_0 = 0,1$.

Exercice 3

Une population renferme 4% d'individus homozygotes pour l'allèle récessif B. Sachant que le taux de mutation directe $A \rightarrow B$ est $u = 2 \times 10^{-5}$ et celui de la mutation reverse $B \rightarrow A$ est $v = 1,2 \times 10^{-4}$.

1. Estimer l'évolution de la fréquence des deux allèles.

Exercice 4

Déterminer la fréquence d'un gène A dans une population après une génération de migration, sachant que :

- La fréquence des migrants est de 0,03.
- La fréquence de A chez les autochtones est de 0,4.
- La fréquence de A chez les migrants est de 0,8.

Exercice 5

La composition génotypique de deux populations est la suivante :

	AA	AB	BB
Population 1	0,49	0,42	0,09
Population 2	0,1	0,4	0,5

Au cours de l'histoire de ces deux populations, des migrants de la population « 2 » ont colonisé la population « 1 » et après plusieurs générations dans cette dernière les fréquences génotypiques sont devenues les suivantes :

	AA	AB	BB
Population 1'	0,25	0,5	0,25

En supposant que la population « 2 » est beaucoup plus grande que la population « 1 » et que l'apport s'est effectué à un taux constant à chaque génération :

1. Quelle est la proportion de gènes de la population « 1' » provenant de la population 2 à chaque fois ?

Exercice 6

Un allèle récessif de fréquence $q = 0,5$ a un coefficient de sélection de $s = 0,53$.

1. Calculer la fréquence de q à la génération suivante.
2. Estimer la variation de la fréquence de cet allèle après une génération de sélection.

Exercice 7

Sous l'influence d'un coefficient de sélection $s = 0,9$ contre les hétérozygotes :

1. Calculer le changement de la fréquence q d'une génération à la suivante d'un allèle récessif B à partir d'une fréquence initiale de $0,6$.
2. Calculer la fréquence de B à la génération suivante.

Exercice 8

Soit trois allèles A , B et C de fréquences respectives de $p = 0,2$, $q = 0,3$ et $r = 0,5$; les valeurs sélectives de chaque génotype sont les suivantes :

Génotypes	AA	AB	AC	BB	BC	CC
w	$w_1 = 0,8$	$w_2 = 1$	$w_3 = 1$	$w_4 = 0,7$	$w_5 = 1$	$w_6 = 0,9$

1. Quelles fréquences doit-on attendre pour A , B et C à la génération suivante ?

Exercice 9

Dans une population, un gène est représenté par deux allèles A et B de fréquences respectives de $p = 0,6$ et $q = 0,4$. Une sélection est exercée contre les homozygotes avec un coefficient de sélection de $s_A = 0,5$ contre le génotype AA et $s_B = 0,3$ contre le génotype BB .

1. Estimer la variation de la fréquence de l'allèle B après une génération de sélection.