II. L'évolution de la biodiversité génétique d'une population

Définitions :

- **Population** : groupe d'individus appartenant à une même espèce, vivant dans un même lieu et pouvant potentiellement se produire entre eux
- **Modèle de Hardy Weinberg** (= HW) : Modèle mathématique qui prédit la structure génétique d'une population sous certaines hypothèses (voir ci dessous).
- **Allèle :** Version d'un gène. Les allèles diffèrent par leur séquence nucléotidique.
- **Fréquence allélique:** Fréquence d'un allèle par rapport aux autres allèles du même gène dans une population.
- **Génotype** : Ensemble des allèles portés par un individu. On le limitera ici aux allèles du gène étudié. Si l'individu a 2 allèles identiques, on dit qu'il est **homozygote** pour ce gène. Si les allèles sont différents, on dit qu'il est **hétérozygote**
- **Fréquence génotypique** : occurrence d'un génotype donné (doublet d'allèle) par rapport au nombre total d'individus
- **Sélection naturelle** : force évolutive qui favorise la transmission de certains allèles dans un environnement donné : un allèle qui confère un avantage voit sa fréquence augmenter
- **Dérive génétique** : Variation aléatoire de la fréquence des allèles au cours des générations. Elle est plus importante dans les petites populations.

Les espèces changent au cours des temps parce que la composition génétique de l'espèce se modifie (composition génétique = nature et fréquence des allèles présents chez les individus de cette espèce). L'évolution ne peut s'étudier à l'échelle de l'individu, puisque la définition de l'espèce repose sur la capacité des individus à se reproduire entre eux. Il est donc nécessaire de considérer un groupe d'individus qui peuvent potentiellement se reproduire. On parle de **population**.

A. Le modèle théorique d'Hardy-Weinberg

1. Notions de fréquence allélique et de fréquence génotypique

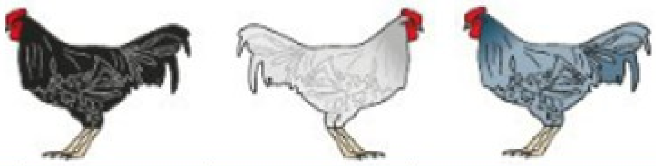
Dans une **population** on peut calculer la **fréquence d'un allèle**. On l'exprime comme une proportion de cet allèle dans la population par rapport aux autres allèles de ce gène. La somme des fréquences alléliques de tous les allèles d'un gène dans une population est donc par définition égale à 1.

On peut également calculer la **fréquence d'un génotype** qui correspond à la proportion de ce génotype par rapport aux autres génotypes possibles en considérant un gène. Dans une population, la somme des fréquences génotypiques de tous les génotypes possibles en considérant 1 gène est elle aussi égale à 1.

Attention : une fréquence est toujours comprise entre 0 et 1

Exemple 1 : La couleur des poulets andalous

La couleur des poulets andalous est déterminée par un gène qui existe sous 2 versions alléliques : l'allèle N qui détermine la couleur noire et l'allèle B qui détermine la couleur blanche. Les allèles N et B sont codominants : lorsqu'un poulet possède ces 2 allèles, son plumage est gris bleu.



Phénotype	[noir]	[blanc]	[Gris-bleu]
Génotype	(N//N)	(B//B)	(N//B)
Effectif dans la classe	15	10	8

Calcul des fréquences alléliques :

- fréquence de l'allèle N = *nombre d'allèle N / nombre total d'allèles* = *(2x15) + 8 / 2x 33* = *38 / 66* = *0.58*

- fréquence de l'allèle B = (2x10) + 8 / 2x 33 = 28 / 66 = 0.42

Fréquence de l'allèle N + fréquence de l'allèle B = 0.58 + 0.42 = 1

Calcul des fréquences génotypiques :

- fréquence du génotype (N//N) = *nombre d'individus ayant le génotypes (N//N) / nombre total d'individus dans la population* = 15 / 33 = 0.45

- fréquence du génotype (B//B) = 10 / 33 = 0.3

- fréquence du génotype (N//B) = 8 / 33 = 0.24

Somme des fréquences des 3 génotypes : 0.45 + 0.3 + 0.24 ≈ 1

=> ***La fréquence des allèles est différente de la fréquence des génotypes dans la population.***

Exemple 2 : On considère une population de 5 individus dont voici les génotypes et les phénotypes pour le gène du groupe sanguin :

Individu	Phénotype	Génotype
Individu A	[A]	(A//O)
Individu B	[O]	(O//O)
Individu C	[AB]	(A//B)
Individu D	[A]	(A//A)
Individu E	[AB]	(A//B)

Calcul des fréquences alléliques de cette population :

Fréquence de l'allèle A = 5/10 = 0.5

Fréquence de l'allèle B = 2/10 = 0.2

Fréquence de l'allèle O = 3/10 = 0.3

Calcul des fréquences génotypiques :

F(A//O) = 1/5 = 0.2

F(A//B) = 2/5 = 0.4

Calcul des fréquences phénotypiques :

f[A] = 2/5 = 0.4 f[0] = 1/5 = 0.2

f[B] = 0/5 = 0 f[AB] = 2/5 = 0.4

La fréquence des allèles est différente de la fréquence des génotypes elle-même différente de la fréquence des caractères phénotypiques dans la population.

Par exemple la fréquence du phénotype [A] est de 0.4

la fréquence du génotype (A//A) est 0.2

la fréquence de l'allèle A est de 0.5

2. Mise en évidence de l'équilibre de Hardy- Weinberg : la constance des fréquences alléliques au cours du temps sous certaines conditions.

On cherche à comprendre comment évoluent les fréquences alléliques et les fréquences génotypiques dans une population au cours du temps.

Le modèle de Hardy-Weinberg (=HW) suppose que dans une population « idéale » = dans laquelle les hypothèses ci-dessous sont respectées, **la fréquence de chaque allèle d'un gène et de chaque génotype est constante de génération en génération.**

Ce modèle est construit sur des hypothèses fortes :

- une population de grande taille
- une absence de mutations
- une reproduction aléatoire des individus
- une absence de migration et de sélection naturelle (les allèles ne modifient pas l'espérance de vie de l'individu).

Par exemple, si on suppose que la population française est à l'équilibre de HW pour les groupes sanguins (ce qui n'est pas tout à fait exact), si la fréquence de l'allèle A est aujourd'hui de 0,27 (=27%) elle restera de 0,27 de génération en génération. De la même façon si la fréquence du génotype (A//B) est aujourd'hui de 0,06 (=6%), elle restera de 0,06 de génération en génération.

Si on travaille avec la version la plus simple du modèle de HW, on considère un gène qui a uniquement deux allèles : l'allèle A, de fréquence p, l'allèle a, de fréquence q (avec p+q=1). Le modèle de HW prévoit donc que la fréquence p de l'allèle A gardera la même valeur chiffrée de génération en génération, tout comme la fréquence q de l'allèle a.

Il est en de même pour les fréquences génotypiques, que le modèle prévoit constantes de génération en génération.

Dans ce cas de figure simple (1 gène avec seulement 2 allèles), le modèle permet également d'estimer la fréquence de chaque génotype grâce à des formules (à connaître) :

Fréquences génotypiques :

f(AA) $= P^2$

f(Aa) $= 2pq$

f(aa) $= q^2$

Avec $p^2 + 2pq + q^2 = 1$

(/!\ A connaître)

3. La démonstration de Hardy : démonstration de la constance des fréquences alléliques au cours du temps (facultative)

Démontrons les formules ci-dessus (ce qui vous permettra de les mémoriser !) :
En considérant dans les cases grises la totalité des gamètes d'une génération (on raisonne ici à l'échelle de la population), on obtient le tableau suivant qui prévoit la fréquence des génotypes de la génération 2 :

		Gamète mâle (de la génération 1)		
		Allèle : **N** Fréquence : **P**	Allèle : **B** Fréquence : **q**	
Gamète Femelle (de la génération 1)	Allèle : **N** Fréquence : **P**	Génotype : **(N // N)** Fréquence du génotype : **P x P= P^2**	Génotype : **(N // B)** Fréquence du génotype : **p x q**	**Génération 2**
	Allèle : **B** Fréquence : **q**	Génotype :**(N // B)** Fréquence du génotype : **p x q**	Génotype : **(B // B)** Fréquence du génotype : **q x q = q^2**	

Fréquence des génotypes obtenus à la génération 2

On peut alors facilement démontrer que la fréquence de chaque allèle est restée constante, tout comme celle de chaque génotype.

- Fréquence de l'allèle N (à la génération 2) :

$$\frac{(2x\, p^2 \times Nt) + (pq \times Nt) + (pq \times Nt)}{2 \times (p^2 + 2pq + q^2) \times Nt} = \frac{2 \times Nt \times (p^2 + pq)}{2 \times 1 \times Nt} = \frac{2xpxNt \times (p+q)}{2 \times Nt} = p \times 1 = p$$

- Fréquence de l'allèle B (à la génération 2) :

$$\frac{(2 \times q^2 \times Nt) + (pq \times Nt) + (pq \times Nt)}{2 \times (p^2 + 2pq + q^2) \times Nt} = \frac{2 \times Nt \times (q^2 + pq)}{2 \times 1 \times Nt} = \frac{2 \times q \times Nt \times (p+q)}{2 \times Nt} = q \times 1 = q$$

Basé sur des connaissances en probabilité simples, ce modèle prévoit un maintien des **fréquences alléliques** et **génotypiques** au cours du temps. On dit alors que la population est à l'équilibre d'Hardy-Weinberg.

4. Vérifier si une population est à l'équilibre de HW (indispensable !) /!\

En exercice, vous devez savoir déterminer si une population est à l'équilibre de HW.

Voir fiche méthode et énoncés des exercices sur les pages suivantes

B. Les écarts observés au modèle d'Hardy-Weinberg

La plupart du temps, les populations ne sont pas à l'équilibre de HW. Cet **écart** est observé à partir du moment où l'une des hypothèses de départ du modèle n'est plus respectée.

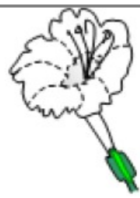
- Si la population est petite, alors la **dérive génétique** devient importante
- L'apparition de nouvelles mutations modifie la fréquence des allèles et donc leur probabilité de transmission, il en est de mêmes des flux migratoires
- Si certains allèles permettent d'être plus choisis lors de la reproduction (**sélection sexuelle**) leur probabilité de transmission augmente, il en est de même pour les allèles qui augmentent l'espérance de vie de l'individu les portant (**sélection naturelle**).

Application du cours – Hardy-Weinberg (remplir sur la fiche méthodologique)

Exploitez les données de ces deux exemples afin de dire si les populations sont à l'équilibre de HW.
Dans le cas contraire, proposez une explication à l'écart observé.

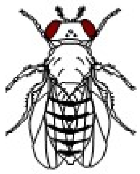
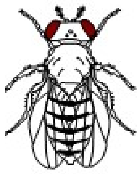
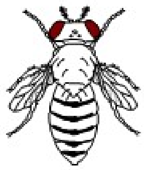
Exemple 1 :

On considère des fleurs « Belle de nuit » dont la couleur est déterminée par un seul couple d'allèle, avec deux allèles A et B codominants. On échantillonne toutes les fleurs présentes dans une prairie, et on note leur couleur.

Phénotype	 [rouge]	 [rose]	 [blanc]
Genotype des individus	AA	AB	BB
Effectif	166	187	47

Exemple 2 :

On considère à présent une population de drosophiles (mouches du vinaigre) élevée en laboratoire, et repiquée dans un nouveau milieu à chaque génération. On s'interesse à un gène qui contrôle la forme des ailes, il existe dans la population 2 allèles : l'allèle ***vg+***, dominant responsable d'une aile sauvage, ainsi que l'allèle ***vg-***, recessif, provoquant une aile raccourcie. La nourriture se situe en haut de la cage, il faut voler pour y parvenir. Les individus sont dénombrés et génotypés après 20 générations.

Phénotype	 [ailes sauvages]	 [Ailes sauvage]	 [Ailes vestigiales]
Genotype des individus	(vg+//vg+)	(vg+//vg-)	(vg-//vg-)
Effectif	200	220	5

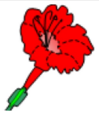
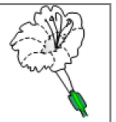
Méthodologie : comment vérifier si une population est à l'équilibre de Hardy-Weinberg ?

Pour un gène avec 2 allèles A et a, de fréquence allelique respectives *p* et *q*.

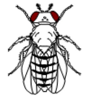
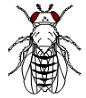
1. On mesure les effectifs des différents **génotypes** (en général, cela est donné).

Génotypes	AA	Aa	aa	Total
Effectifs	n_{AA}	n_{Aa}	n_{aa}	N

Ex 1 : Les belles de nuits

Phénotype	 [rouge]	 [rose]	 [blanc]
Genotype des individus	AA	AB	BB
Effectif	166	187	47

Ex 2 : Les drosophiles

Phénotype	 [ailes sauvages]	 [Ailes sauvage]	 [Ailes vestigiales]
Genotype des individus	(vg+//vg+)	(vg+//vg-)	(vg-//vg-)
Effectif	200	220	5

2. On calcule les fréquences alléliques à partir de ce tabeau. Comme chaque individu est diploïde, le nombre total d'allèles équivaut à 2 fois le nombre d'individus.

$$p = \frac{(2n_{AA} + n_{Aa})}{2N} \quad \text{et} \quad q = 1 - p = \frac{(2n_{aa} + n_{Aa})}{2N}$$

Ex 1 :

p= $\frac{(2x166) + 187}{2*400}$ = 0.65

q= 1-p = 1 – 0.65 = 0.35

Ex 2 :

p= $\frac{(2x200) +220}{2*425}$ = 0.73

q= 1 - 0.73 =0.27

3. En utilisant ces fréquences allélique calculées, on calcule les effectifs attendus si la population était l'équilibre de HW. Pour cela, on multiplie la fréquence de chaque génotype par l'effectif total.

Génotypes	AA	Aa	aa
Fréquence attendue si équilibre de HW	p^2	$2pq$	q^2
Effectifs attendus si équilibre de HW	p^2N	$2pqN$	q^2N

Ex 1 :

Génotypes	AA	AB	BB
Fréquence attendue si équilibre de HW	0.65^2= 0.42	2x0.65x0.35 =0.46	0.35^2 =0.12
Effectifs attendus si équilibre de HW	0.42x400 = 168	0.46x400 = 184	0.12x400 =48

Ex 2 :

Génotypes	(vg+ //vg+)	(vg+ //vg-)	(vg- //vg-)
Fréquence attendue si équilibre de HW	0.73^2 = 0.53	2x0.73x0.27 = 0.39	0.27^2 =0.07
Effectifs attendus si équilibre de HW	0.53x425 = 225	0.39x425 = 166	0.07x425 = 30

4. On compare ces effectifs attendus avec les effectifs observés (ceux du 1). Deux cas de figures se présentent :

-> **les données observées ≈ données attendues**
-> la population est à l'équilibre de HW
-> toutes les hypothèse initiales (voir cours) sont vérifiées

- >**les données observées ≠ données attendues**
-> la population n'est pas à l'équilibre de HW
- >Au moins une hypothèse n'est pas vérifiée
(en général l'information est dans un document !)

Ex 1 — Conclusion : **Les effectifs attendus pour une population à l'équilibre sont proches des effectifs observés => La population est donc à l'équilibre de Hardy-Weinberg.**
Les hypothèses initiales sont donc vérifiées.

Ex 2 — Conclusion : **Les effectifs attendus sont significativement différents des effectifs observés.**
La population n'est pas donc à l'équilibre de Hardy-Weinberg.
Au moins une des hypothèses initiales n'est pas vérifiée.
- ailes vestigiales : plus de difficultés à se nourrir car le vol nécessaire pour accéder aux ressources.
- Espérance de vie est réduite, moins de vitalité :
> ils se reproduisent moins et ne transmettent pas leur allèle à la descendance
> la fréquence de *vg-* n'est pas stable, elle décroit au cours des générationssous l'effet de la sélection naturelle

Remarque : Il n'est pas possible d'obtenir parfaitement les mêmes valeurs. En recherche, on réalise un test statistique afin de dire si les différences sont significatives ou non. Vous n'aurez que des résultats sans ambiguïté.

Variante : Parfois il vous sera donné directement la fréquence des génotypes dans la population, sans vous donner l'effectif total.

Exemple *: la fréquence des malades atteints par la mucosvisidose (a//a) est de 1/3500*

On en déduit alors que f(aa) = q^2=1/3500

$q = \sqrt{1/3500}$ or $p+q = 1$

$p = 1 - \sqrt{1/3500}$