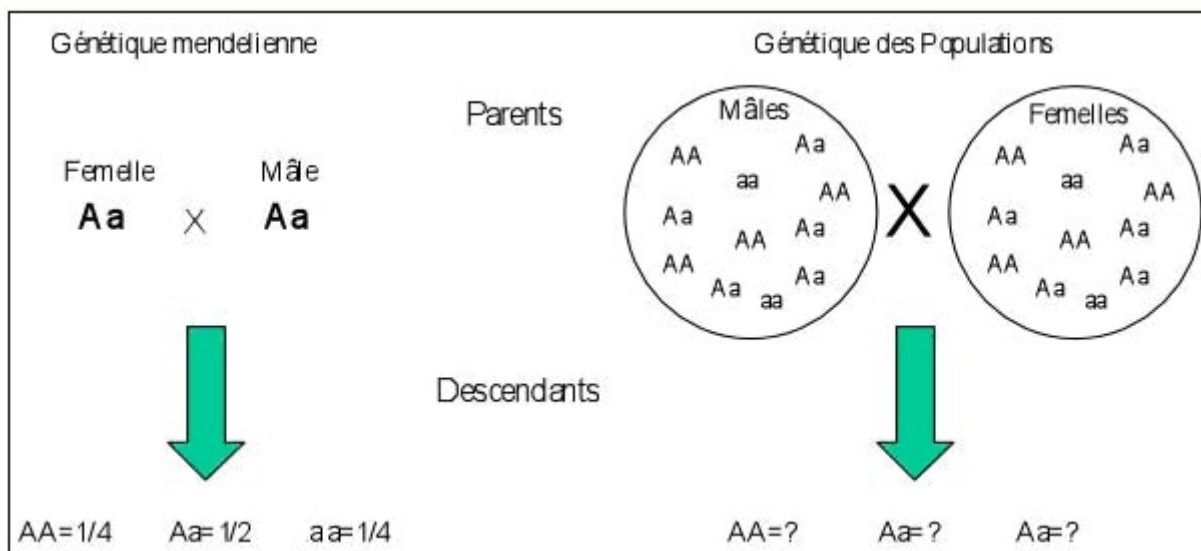


CHAPITRES VI : GENETIQUE DES POPULATIONS

INTRODUCTION-DEFINITIONS

La génétique formelle se base sur des croisements contrôlés par un expérimentateur.

La génétique des populations étudie les proportions des génotypes au sein d'un ensemble d'individus issus de **croisements non contrôlés** entre de nombreux parents. C'est donc une **application des principes de base de la génétique mendélienne à l'échelle des populations**.



L'isolat : est un groupe d'individus en isolement complet du reste du monde, du point de vue génétique.

La population:

- Groupe d'individus de la **même espèce** vivant dans une **zone géographique** déterminée.
- Ils ont en commun un stock de gènes appelé aussi **patrimoine génétique ou pool génétique**, qui est formé par tous les gènes contenus dans les génotypes individuels.
- A chaque génération, ce stock se répartira entre les gamètes qui donneront naissance à la génération suivante.

Comment étudier une population?

Pour simplifier l'étude d'une population, il faut prendre un **échantillon** où on va examiner le génotype ou le phénotype de chaque individu.

Si l'échantillon est **suffisamment grand et aléatoire**, la composition génétique de l'échantillon va représenter la population.

La réflexion va se faire en termes de **fréquences alléliques, fréquences génotypiques ou fréquences phénotypiques** qui caractériseront la population,

Fréquences alléliques, génotypiques ou phénotypiques

Soit un locus avec deux allèles (A , a), donc trois génotypes: AA , Aa , aa

N_{xx} = nombre des individus portant le génotype xx .

$N_{AA} + N_{Aa} + N_{aa} = N$ (nombre total des individus dans l'échantillon)

Fréquence génotypique : c'est la proportion des individus porteurs du génotype en question dans la population ou dans l'échantillon.

$$\begin{aligned} f_{(AA)} &= N_{AA} / N \\ f_{(Aa)} &= N_{Aa} / N \\ f_{(aa)} &= N_{aa} / N \\ f_{(AA)} + f_{(Aa)} + f_{(aa)} &= 1 \end{aligned}$$

Exemple :

Nous avons échantillonné 1000 individus. $N = 1000$

$$N_{AA} = 795$$

$$N_{Aa} = 190$$

$$N_{aa} = 15$$

Les fréquences génotypiques sont alors les suivantes :

$$f_{(AA)} = 795 / 1000 = 0,795$$

$$f_{(Aa)} = 190 / 1000 = 0,190$$

$$f_{(aa)} = 15 / 1000 = 0,015$$

$$f_{(AA)} + f_{(Aa)} + f_{(aa)} = 1$$

Fréquence allélique : c'est la proportion des allèles en question dans la population ou dans l'échantillon. $f_{(A)}$ et $f_{(a)}$

Soit N_x : nombre des allèles x dans l'échantillon; donc N_A et N_a

Chaque individu AA porte deux allèles A et chaque individu Aa n'en porte qu'un seul, donc :

$$N_A = 2 N_{AA} + N_{Aa}$$

Même logique pour l'allèle a .

$$N_a = 2 N_{aa} + N_{Aa}$$

Donc:

$$f_{(A)} = N_A / \text{nombre total des allèles} = N_A / 2N \text{ car on a 2 allèles par individu}$$

$$f_{(A)} = (2 N_{AA} + N_{Aa}) / 2N$$

$$f_{(a)} = N_a / 2N$$

$$f_{(a)} = (2 N_{aa} + N_{Aa}) / 2N$$

$$f_{(A)} + f_{(a)} = 1$$

Les fréquences alléliques peuvent également être calculées à partir des fréquences génotypiques :

$$f_{(A)} = f_{(AA)} + 1/2 f_{(Aa)}$$

par ce que tous les allèles portés par les individus AA sont A et que seulement la moitié des allèles sont A chez les hétérozygotes.

$$f_{(a)} = f_{(aa)} + 1/2 f_{(Aa)}$$

Comme on a seulement deux allèles: $f_{(A)} + f_{(a)} = 1$

Dans l'exemple précédent:

$$f_{(A)} = (2 \times 795 + 190) / 2 \times 1000 = 0,89$$

$$f_{(a)} = (2 \times 15 + 190) / 2 \times 1000 = 0,11$$

OU

$$f_{(A)} = 0,795 + (0,5 \times 0,190) = 0,89$$

$$f_{(a)} = 0,015 + (0,5 \times 0,190) = 0,11$$

Théoriquement, il n'est pas nécessaire de calculer les 2 fréquences puisque $f_{(A)} + f_{(a)} = 1$, mais calculez les 2 est un bon moyen de vérification.

Il faut toujours vérifier que la somme fasse bien 1 aussi.

$$f_{(A)} + f_{(a)} = 0,89 + 0,11 = 1$$

II- LOI ET EQUILIBRE DE HARDY-WEINBERG.

La transmission des différents allèles et leurs fréquences **semble très difficile à prévoir en réalité, à cause de plusieurs facteurs** comme: mutations, migrations, différence de survie ou de fécondité entre individus, transmission simultanée de très nombreux gènes polymorphes qui peuvent interagir entre eux...

Pour contourner ces difficultés, on doit étudier la transmission des caractères dans un cas simple appelé **population théorique idéale**, qui se définit par les caractéristiques suivantes :

1. **Les croisements sont entièrement aléatoires (reproduction panmictique): condition la plus importante.** Elle suppose que les individus ne choisissent pas leur partenaire sexuel ni en fonction de leur génotype, ni en fonction de leur phénotype et que la rencontre des gamètes se fait au hasard.
2. **Population vaste:** d'effectif très important (théoriquement infini).
3. **Absence de migration (flux migratoires):** population **close génétiquement**
4. **Absence de sélection:** Tous les individus, quel que soit leur génotype, ont la même capacité à se reproduire (fertilité) et à engendrer une descendance viable.
5. **Absence de mutation**

Dans une population idéale, on peut facilement calculer les fréquences génotypiques et alléliques de génération en génération.

Dans une génération de départ (génération 0):

Par convention, on pose les fréquences alléliques :

$$f_{(A)} = p \text{ et}$$

$$f_{(a)} = q$$

$$p + q = 1 \dots\dots\dots(1)$$

On sait que la somme des fréquences génotypiques dans cette génération est égale à 1:

$$f_{(AA)} + f_{(Aa)} + f_{(aa)} = 1 \dots\dots\dots(2)$$

Dans la génération suivante (génération 1):

Comme les unions se font au hasard, on peut construire un tableau de croisement à partir des gamètes de la population (génération 0) :

Gamètes		Mâles	
		A (p)	a (q)
Femelles	A (p)	AA (p ²)	Aa (pq)
	a (q)	Aa (pq)	aa (q ²)

La probabilité qu'un gamète (mâle ou femelle) soit A ou a est égale à la fréquence allélique de la génération initiale (génération 0).

Dans les 4 cases représentant les génotypes, on trouve les fréquences génotypiques de la génération suivante (génération 1).

$$f'_{(AA)} = p^2$$

$$f'_{(Aa)} = 2pq$$

$$f'_{(aa)} = q^2$$

La somme des fréquences génotypiques dans la génération 1 est

$$f'_{(AA)} + f'_{(Aa)} + f'_{(aa)} = p^2 + 2pq + q^2 = (p+q)^2 = 1 \dots\dots\dots(3)$$

Les fréquences alléliques de la génération 1 peuvent être calculées à partir des fréquences génotypiques:

$$p' = f_{(A)} = f'_{(AA)} + 1/2 f'_{(Aa)}$$

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

En remplaçant par les valeurs de chacune des fréquences génotypiques on obtient:

$$p' = f'(A) = p^2 + 1/2 (2 pq) = p^2 + pq = p (p + q) = p$$

$$q' = f'(a) = q^2 + 1/2 (2 pq) = q^2 + pq = q (p + q) = q$$

$$\text{Donc } p' + q' = p + q = 1 \dots\dots\dots(4)$$

D'après (1), (2), (3) et (4) on peut conclure que:

Les fréquences alléliques et les fréquences génotypiques restent stables de génération en génération. On dit alors que la population est en équilibre (équilibre de Hardy et Weinberg). Il existe une relation simple entre les fréquences alléliques et les fréquences génotypiques, donnée par la loi ou le modèle de Hardy-Weinberg (pour un gène à deux allèles) :
« Dans une population idéale la fréquence des génotypes sera le développement de $(p+q)^2$, p et q étant les fréquences alléliques $(p+q)^2 = p^2 + 2pq + q^2 = 1$ »

Exemple où $p = 0,89$ et $q = 0,11$.

Si la population est en équilibre de HW, alors les fréquences génotypiques sont les suivantes :

$$f_{(AA)} = p^2 = (0,89)^2 = 0,7921$$

$$f_{(Aa)} = 2 pq = 2 \times 0,89 \times 0,11 = 0,1958$$

$$f_{(aa)} = q^2 = (0,11)^2 = 0,0121$$

III- APPLICATIONS DE LA LOI DE HARDY-WEINBERG.

Dans les populations idéales

1. En cas de dominance/récessivité.

En cas d'un gène biallélique (A et a), la connaissance des fréquences phénotypiques permet de déduire la fréquence des génotypes.

Exemple:

Dans une population on a:

Fréquence du phénotype a, qu'on peut symboliser par: $f_{[a]} = 0,0006$

Fréquence du phénotype A, qu'on peut symboliser par: $f_{[A]} = 1 - 0,0006 = 0,9994$

Quelles sont les fréquences alléliques et génotypiques?

On a trois génotypes possibles AA, Aa et aa.

Si la population est en équilibre de HW, on a:

$$f_{(AA)} = p^2$$

$$f_{(Aa)} = 2 pq$$

$$f_{(aa)} = q^2$$

On sait que

$$f_{[A]} = f_{(AA)} + f_{(Aa)} = p^2 + 2 pq = 0,9994$$

$$f_{[a]} = f_{(aa)} = q^2 = 0,0006$$

On peut donc déduire les fréquences alléliques p et q puis les fréquences génotypiques:

$$q = \sqrt{q^2} = 0,0245$$

$$p = 1 - q = 0,9755$$

$$f_{(aa)} = q^2 = 0,0006$$

$$f_{(AA)} = p^2 = 0,9516$$

$$f_{(Aa)} = 2 pq = 0,0478$$

On vérifie par la somme des fréquences génotypiques qui est bien égale à 1

2. En cas de codominance

En cas d'un gène biallélique (A et B) avec codominance, les fréquences phénotypiques sont égales aux fréquences des génotypes.

$$f_{[B]} = f_{(BB)} = q^2$$

$$f_{[AB]} = f_{(AB)} = 2pq$$

$$f_{[A]} = f_{(AA)} = p^2$$

Si on connaît les valeurs des fréquences phénotypiques on peut donc déduire les fréquences alléliques p et q .

3. Cas des allèles liés au chromosome X

Soit un gène, porté par le chromosome X, avec deux allèles A et a et leurs fréquences respectives à la génération de départ p et q .

Quelles sont les fréquences génotypiques?

Considérons la descendance mâle et la descendance femelle séparément.

Descendance femelle

Gamètes mâles porteurs d'un X	Gamètes femelles		
		$X^A (p)$	$X^a (q)$
	$X^A (p)$	$X^A X^A (p^2)$	$X^A X^a (pq)$
	$X^a (q)$	$X^A X^a (pq)$	$X^a X^a (q^2)$

Tous les descendants sont des femelles, et on observe les proportions de HW.

Donc chez les femelles :

$$f(X^A X^A) = p^2$$

$$f(X^A X^a) = 2pq$$

$$f(X^a X^a) = q^2$$

Descendance mâle :

Gamètes mâles porteurs d'un Y	Gamètes femelles		
		$X^A (p)$	$X^a (q)$
	Y	$X^A Y (p)$	$X^a Y (q)$

Tous les descendants sont mâles et les proportions de HW ne sont pas respectées.

Donc chez les mâles :

$$f(X^A Y) = p$$

$$f(X^a Y) = q$$

La fréquence des phénotypes chez les mâles est une bonne estimation directe des fréquences alléliques dans la population.

On remarque qu'en cas de maladie RLX:

La fréquence du phénotype récessif chez les mâles est égale à la fréquence allélique de l'allèle récessif c'est-à-dire q .

La fréquence du phénotype récessif chez les femelles est donnée par q^2 ; il est donc plus rare chez les femelles que chez les mâles.

Dans les populations naturelles

La loi de Hardy-Weinberg ne s'applique qu'à des populations idéales. On peut vérifier l'application de la loi de Hardy-Weinberg dans les populations naturelles en cas de caractères codominants pour lesquels le calcul des fréquences alléliques est possible, et ceci grâce au **test de l'équilibre**.

Le principe du test est simple et peut être résumé en 3 étapes:

- 1. Echantillonnage:** On prend un échantillon d'une population, on compte les effectifs génotypiques réels (grâce à la codominance) et on calcule des fréquences alléliques réelles parmi les N individus échantillonnés, soit $p = f(A)$, et $q = f(a)$
- 2. On calcule les effectifs génotypiques attendus (N')** dans une population théorique idéale qui aurait le même effectif et les mêmes fréquences alléliques que la population étudiée soit:

$$f'_{(AA)} = p^2 \rightarrow N'_{AA} = p^2 N$$

$$f'_{(Aa)} = 2pq \rightarrow N'_{Aa} = 2pq N$$

$$f'_{(aa)} = q^2 \rightarrow N'_{aa} = q^2 N$$

3. Comparaison des effectifs observés et des effectifs attendus (comparaison des deux distributions) par un test statistique du Chi Deux (ou d'autres tests).

Pour faire un test du Chi deux:

- on émet une hypothèse H_0 : égalité entre la distribution observée et la distribution théorique (la population suit la loi de Hardy-Weinberg)
- on calcul de la distance χ^2 permettant de tester l'hypothèse H_0 .

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

- La somme est effectuée sur tous les génotypes et la valeur χ^2 est comparée à une valeur seuil, lue dans une table du χ^2 (voir à la fin de ce document), en fonction de 2 paramètres :
 - o un risque α choisi par l'utilisateur qui est en général 5%,
 - o un degré de liberté (ddl) égal à la différence entre le nombre de génotypes et le nombre d'allèles du système génétique étudié.
- si χ^2 calculé < à χ^2 seuil, H_0 est acceptée et on conclut que la population suit la loi de Hardy-Weinberg donc est à l'équilibre.
- si χ^2 calculé > à χ^2 seuil, H_0 est rejetée et on conclut que la population ne suit pas la loi de Hardy-Weinberg avec un risque $\alpha = 5\%$ de se tromper.

Exemple

Chez l'homme, le groupe sanguin MN est déterminé par un gène à deux allèles codominants M et N, ce qui permet d'attribuer un génotype à chaque individu échantillonné, puis d'estimer les fréquences alléliques dans la population.

Une étude portant sur 730 aborigènes australiens a donné les résultats suivants :

22 MM, 216 MN, 492 NN

1- Calcul des effectifs génotypiques réels et des fréquences p et q des allèles M et N:

Comme il s'agit d'une codominance, les effectifs génotypiques = effectifs phénotypiques, donc:

$$N_{MM} = 22$$

$$N_{MN} = 216$$

$$N_{NN} = 492$$

Calcul des fréquences p et q des allèles M et N:

$$p = (22 + 1/2 \times 216) / 730 = 0,178 \text{ pour l'allèle M}$$

$$q = 492 + 1/2 \times 216 / 730 = 0,822 \text{ pour l'allèle N.}$$

2- Calcul des effectifs théoriques attendus (N') des différentes catégories génotypiques:

$$N'_{MM} = p^2 \times 730 = (0,178)^2 \times 730 = 23,1$$

$$N'_{MN} = 2pq \times 730 = (2 \times 0,178 \times 0,822) \times 730 = 213,6$$

$$N'_{NN} = q^2 \times 730 = (0,822)^2 \times 730 = 493,2$$

3- Comparaison des effectifs observés et des effectifs attendus par le test du Chi deux

$$\chi^2 = (22-23,1)^2/23,1 + (216-213,6)^2/213,6 + (492-493,2)^2/493 = 0,083$$

La valeur seuil pour 3-2=1 degré de liberté et un risque de 5% est 3,84.

La valeur de la statistique X^2 étant très inférieure à la valeur seuil, on conclut qu'il n'y a pas de différence significative entre la distribution observée et la distribution théorique. On admet donc que la population d'aborigènes australiens est à l'équilibre de Hardy-Weinberg.

Si une population est considérée à l'équilibre de Hardy-Weinberg après un test statistique, ceci n'implique pas que toutes les conditions d'application de cette loi soient respectées (effectif infini, absence de mutation, absence de sélection, etc...) mais l'hypothèse la plus importante qui doit être respectée est la **panmixie**.

Si une population n'est pas considérée à l'équilibre de Hardy-Weinberg après un test statistique, ceci implique la recherche de la cause de cette situation :

- Consanguinité (pas d'unions aléatoires) : Croisement entre individus apparentés ce qui favorise l'homozygotie.
- Sélection : c'est le fait qu'un allèle confère soit un avantage sélectif et se propage à travers le patrimoine génétique, soit il confère un désavantage sélectif et disparaît de lui.
- Migration : entrée ou sortie d'individus dans le groupe.
- Mutation : cause d'apparition de nouveaux allèles pour un gène donné, changeant ainsi la fréquence des allèles précédents.

TABLE DU CHI-DEUX : X^2

α	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