

Le modèle de Hardy-Weinberg – Application à l'étude génétique des populations

[Voir le diaporama de cours →](#)

Le modèle de Hardy-Weinberg décrit, au cours des générations une situation d'équilibre génotypique dans une population. Il montre que, sous certaines hypothèses, dans une population à reproduction sexuée les fréquences alléliques restent constantes d'une génération à la suivante.

1 - Les hypothèses simultanées de construction du modèle

- **H¹₀ Pas de mutations** – il n'y a pas de nouveaux allèles qui apparaissent. On peut faire l'hypothèse de mutations sur les allèles observés mais alors les fréquences de mutation $A \rightarrow a$ et $a \rightarrow A$ sont égales.
- **H²₀ Pas de sélection naturelle**. Aucun génotype ne produit un phénotype qui soit avantage ou désavantage. Cette hypothèse recouvre celle d'absence de survie différentielle des zygotes.
- **H³₀ Pas de recouvrement des populations**, les individus se reproduisent dans leur cohorte (classe d'âge). Cette hypothèse a peu de poids dans le modèle.
- **H⁴₀ Pas de sélection sexuelle**, les accouplements se font au hasard (=panmixie)
- **H⁵₀ Pas de sélection gamétique** (=pangamie).
- **H⁶₀ Pas de migrations** (aucune arrivée ni départ d'individus)
- **H⁷₀ Effectif important** (donc pas de dérive génétique)

Remarque : Il peut sembler que ces hypothèses soient valides pour des populations d'oursins et pourtant...

2 – Établissement du modèle

Dans la suite on assimile probabilités et fréquences.

On se limite au cas d'un seul gène autosomique à 2 allèles codominants notés A et a.

A- Démonstration pas à pas

À la génération G: pour un couple d'allèles codominants

Tableau des fréquences génotypiques à la génération G

Génotype →	AA	Aa	aa	Remarques N = effectif total D+H+R=1 D, H, R entre 0 et 1
Effectif	D.N	H.N	R.N	
Fréquence	D	H	R	

Tableau des fréquences alléliques (fréquence des gamètes, haplotypes) à la génération G

Allèle →	A	a	Remarques 2.N = effectif total = nombre de gamètes On pose : $p = D + H/2$ $q = R + H/2$ $p+q=1$, p et q entre 0 et 1
Effectif	2.D.N + H.N	2.R.N + R.N	
Fréquence	$D + H/2 = p$	$R + H/2 = q$	

Échiquier de croisement (échiquier de Punnett), prévision des génotypes de G+1

Gamètes mâles → Gamètes femelles	A (p)	a (q)	Remarques Rencontre des gamètes au hasard Tous les zygotes sont viables Mêmes fréquences alléliques dans les 2 sexes car gène autosomique
A (p)	AA (p^2)	Aa (pq)	
a (q)	Aa (qp)	aa (q^2)	

Nouvelles fréquences génotypiques à la génération G+1 :

$$f(AA) = p^2, f(Aa) = 2pq, f(aa) = q^2$$

à condition que les hypothèses d'absence de sélection, de dérive, de migration et de mutation soient valides.

Nouvelles fréquences alléliques (gamétiques) à la génération G+1

$f(A) = p^2 + pq$ car un individu de génotype AA produit des gamètes A avec une fréquence de 1. Par ailleurs un individu Aa produit des gamètes A avec une fréquence de 1/2.

De même pour $f(a) = q^2 + qp$.

Ce qui donne avec $p + q = 1$

$$\begin{aligned} f(A) &= p^2 + pq = p(p + q) = p & \text{et} & & f(a) &= pq + q^2 = q(p + q) = q \\ & & \text{d'où} & & & & \\ f(A) &= p = D + H/2 & \text{et} & & f(a) &= q = R + H/2 \end{aligned}$$

On retrouve exactement les mêmes fréquences alléliques qu'à la génération précédente.

Les proportions des génotypes et des allèles restent donc bien constantes de génération en génération... en l'absence de forces évolutives !

De plus comme $p + q = 1 \Rightarrow (p + q)^2 = 1^2 \Rightarrow$

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

B- Démonstration plus intuitive : « la soupe de gamètes dans le modèle *oursin, jm* »

Si $f_A = p$ et $f_a = q$ alors l'ovule est « tiré au sort » dans l'ensemble des gamètes produits par les femelles avec une probabilité de $(p + q) = 1$ sous les hypothèses HW.

De même pour le spermatozoïde $(p + q) = 1$ si femelles et mâles ont les mêmes fréquences génotypiques (gène autosomique, population en équilibre de HW). La probabilité de tirer au sort un couple (ovule, spermatozoïde) dans cette « soupe de gamètes » tous équiprobables est $(p + q) \times (p + q) = 1^2$ (événements simultanés indépendants, événement certain sous les hypothèses HW ; ni migration, ni mutations).

$$\text{D'où en développant } p^2 + 2pq + q^2 = 1$$

Identification des termes :

$p^2 = f_A \times f_A = f_{(AA)}$, $q^2 = f_a \times f_a = f_{(aa)}$, $2pq = (f_A \times f_a) + (f_a \times f_A) = f_{(Aa)}$ (probabilité du « et » et du « ou », év. Indépendants)

On peut retrouver ce résultat à l'aide d'un échiquier de croisement.

Si les hypothèses HW sont valides alors il est possible de prévoir les fréquences génotypiques à partir des fréquences alléliques.

La comparaison entre ces valeurs calculées et les fréquences génotypiques observées permet de déterminer si la population est en équilibre (de HW) ou bien si elle est soumise à un facteur évolutif.

Pratique du test : [Attention à la notation ; D et D_x désignent 2 grandeurs différentes]

- 1- Les effectifs observés donnent les fréquences génotypiques (observées)
- 2- Calculer les fréquences alléliques observées p et q
- 3- Si HW (hypothèses), alors $D=p^2$, $H=2pq$, $R=q^2$ sont les Fréquences Génotypiques Théoriques sous HW. Les Fréquences Génotypiques Calculées donnent les Effectifs Calculés (E_i);
- 4- Comparaison Effectifs Observés - Effectifs Calculés: $D_x^2 = \sum (O_i - E_i)^2 / E_i$
- 5- Si D_x^2 significatif par rapport à χ^2 (seuil, ddl): On rejette au moins l'une des hypothèses
 - Dérive, Consanguinité ?
 - Sélection ?
 - Mutations, migrations ?

Nb : Recalculer les fréquences génotypiques théoriques à partir des fréquences alléliques observées et utiliser ces valeurs comme un critère applicable à la population est sans biais.

En effet la connaissance des fréquences alléliques observées n'apporte aucune information sur les fréquences génotypiques réelles. Autrement dit, en passant des fréquences génotypiques observées aux fréquences alléliques il y a perte de l'information sur la répartition des allèles dans les différents génotypes.

Par exemple, avec $f_A = 0,5$ et $f_a = 0,5$, on peut avoir n'importe laquelle des distributions réelles suivantes ;

Effectifs	Eff _(AA)	Eff _(Aa)	Eff _(aa)
Distribution 1	500	0	500
Distribution 2	400	200	400
Distribution 3	50	900	50
Distribution HW	250	500	250

Bilan : Le modèle de HW

a) Sous les hypothèses de HW, le modèle prévoit que la reproduction sexuée ne modifie pas les fréquences alléliques d'une génération à la suivante.

b) En connaissant la composition génotypique de la population à la génération n, le modèle de Hardy-Weinberg permet de prévoir, sous certaines hypothèses la composition génotypique à la génération n+1.

c) En connaissant les fréquences alléliques à la génération n, ce modèle permet de calculer les fréquences génotypiques dans cette génération si les hypothèses sont valides.

C'est à dire si cette génération est issue de la précédente en l'absence de « forces évolutives ».

La composition génotypique de la population à la génération n permet toujours de calculer les fréquences alléliques à cette génération. Mais les fréquences alléliques ne permettent de calculer les fréquences génotypiques que sous les hypothèses de HW.

Ce modèle mathématique permet, par comparaison avec les résultats réels observés, de mettre en évidence, ou non, la présence de forces évolutives s'appliquant sur les phénotypes contrôlés par le gène étudié.