**La loi de Hardy-Weinberg et la nécessité d’un test statistique**

1. **La raison de ce document :**

On demande dans les programmes que les élèves soient capables de déterminer si l’équilibre de Hardy Weinberg est vérifié, en comparant des effectifs théoriques et observés de différents génotypes. Si les élèves sont capables de faire ces calculs, il est plus difficile de déterminer à partir de quand les écarts entre effectifs théoriques et observés justifient de rejeter l »hypothèse d’équilibre de Hardy Weinberg. Les tests statistiques nécessaires (Chi²) ne sont pas à leur programme. Après quelques rappels sur Hardy-Weinberg, je propose une feuille de calcul où les élèves peuvent rentrer leurs différents effectifs et le logiciel calcule le Chi² dans le cas simple d’un gène autosomique avec 2 allèles et un degré de liberté.

1. **Biblio :** Génétique des populations, J-L Serre, Dunod 2006.
2. **Place dans les programmes :**

**Programme  enseignement scientifique:**

|  |  |
| --- | --- |
| savoirs | Savoir-faire |
| Au cours de l’évolution biologique, la composition génétique des populations d’une espèce change de génération en génération. Le modèle mathématique de Hardy-Weinberg utilise la théorie des probabilités pour décrire le phénomène aléatoire de transmission des allèles dans une population. En assimilant les probabilités à des fréquences pour des effectifs de grande taille (loi des grands nombres), le modèle prédit que la structure génétique d’une population de grand effectif est stable d’une génération à l’autre sous certaines conditions (absence de migration, de mutation et de sélection). Cette stabilité théorique est connue sous le nom d’équilibre de Hardy-Weinberg. Les écarts entre les fréquences observées sur une population naturelle et les résultats du modèle s’expliquent notamment par les effets de forces évolutives (mutation, sélection, dérive, etc.).  | Pour la transmission de deux allèles dans le cadre du modèle de Hardy-Weinberg, établir les relations entre les probabilités des génotypes d’une génération et celles de la génération précédente. Produire une démonstration mathématique ou un calcul sur tableur ou un programme en Python pour prouver ou constater que les probabilités des génotypes sont constantes à partir de la seconde génération (modèle de Hardy-Weinberg). Utiliser des logiciels de simulation basés sur ce modèle mathématique. Analyser une situation d’évolution biologique expliquant un écart par rapport au modèle de Hardy-Weinberg. |

**Programme de spécialité SVT**

|  |  |
| --- | --- |
| connaissances | Capacités, attitudes |
| L’inéluctable évolution des génomes au sein des populations Dans les populations eucaryotes à reproduction sexuée, le modèle théorique de Hardy-Weinberg prévoit la stabilité des fréquences relatives des allèles dans une population. Mais, dans les populations réelles, différents facteurs empêchent d’atteindre cet équilibre théorique : l’existence de mutations, le caractère favorable ou défavorable de celles-ci, la taille limitée d’une population (effets de la dérive génétique), les migrations et les préférences sexuelles. | Comprendre et identifier les facteurs éloignant de l’équilibre théorique de Hardy-Weinberg, notamment l’appariement non aléatoire, la sélection, la population finie (dérive) |

Structure génétique d’une population : liens entre les fréquences alléliques et les fréquences génotypiques dans une population

Pour étudier la génétique d‘une population, il faut pouvoir déterminer et différencier les fréquences phénotypiques, les fréquences génotypiques et les fréquences alléliques (celles qui nous intéressent le plus mais ne sont pas toujours facilement accessibles.)

1. **Démonstrations de la loi de Hardy Weinberg :**

Dans le modèle de Hardy Weinberg, on choisit une population à sexes séparés, et à générations séparées.

Il y a deux façons de démontrer cette loi :

1. Par formation des couples possibles, en supposant la panmixie, l’absence de mutation et de sélection des gamètes ou des zygotes, ou des embryons.

Soit une population avec un gène à deux allèles, A1 et A2, de fréquences respectives p et q.

On a p+q=1

|  |  |  |  |
| --- | --- | --- | --- |
| Génotypes possibles | A1/A1 | A1/A2 | A2/A2 |
| Fréquences génotypiques | D | H | R |

p = D+H/2 et q = H/2+R

Si les couples se rencontrent au hasard, il y a 6 couples possibles et on peut prévoir les fréquences des génotypes de leurs descendants :

|  |  |  |
| --- | --- | --- |
| Couple possible | Fréquence du couple | Fréquence des descendants |
| A1/A1 | A1/A2 | A2/A2 |
| A1/A1\* A1/A1 | D² | 1 | 0 | 0 |
| A1/A1\* A1/A2 | 2\*D\*H | 0,5 | 0,5 | 0 |
| A1/A1\* A2/A2 | 2\*D\*R | 0 | 1 | 0 |
| A1/A2\* A1/A2 | H² | 0,25 | 0,5 | 0,25 |
| A1/A2\* A2/A2 | 2\*H\*R | 0 | 0,5 | 0,5 |
| A2/A2\* A2/A2 | R² | 0 | 0 | 1 |
| Totaux |  | D²+D\*H+(1/4)\*H² | D\*H+2\*D\*R+(1/2)\*H²+H\*R | (1/4)\*H²+ H\*R+R² |

On peut calculer la fréquence des différents génotypes dans la génération obtenue :

f(A1/A1) = D² + D\*H +(1/2)\*(1/4)\*H²

(D+H/2)² = D² + D\*H + (1/4)\*H² et D+H/2 = p

Donc f(A1/A1) = p²

De même, f(A2/A2) = q²

f(A1/A2) = D\*H+2\*D\*R+(1/2)\*H²+H\*R

or p\*q= (D+H/2)\*( H/2+R) = (1/2)\*D\*H +D\*R+(1/4)\*H²+(1/2)\*R\*H

donc f(A1/A2) = 2\*p\*q

Calculons la fréquence de l’allèle p dans la génération obtenue :

f(A1) = p²+pq = p(p+q)=p

Donc la fréquence de l’allèle A1 est inchangée, ainsi que celle de l’allèle A2.

Le modèle de Hardy Weinberg postule qu’une population de très grande taille, panmictique, sans migration, sélection naturelle, mutations a des fréquences alléliques constantes au cours des générations.

L’équilibre de Hardy Weinberg est vérifié quand on a bien la relation mathématique attendue entre les fréquences génotypiques et les fréquences alléliques.

1. Par les probabilités directement (méthode plus mathématique, qui suppose le smêmes conditions.

Si la fréquence de l’allèle A1 est p et celle de l’allèle A2 est q.

La probabilité à la génération suivante (si panmixie…)d’obtenie un individu (A1/A1) est de p². celle d’obtenir un individu (A2/A2) est de q² et celle d’obtenir un individu (A1/A2) est de 2pq.

On peut ensuite tenir le même raisonnement que précédemment.

Cette méthode est plus rapide mais plus abstraite.

1. **Comment déterminer à partir d’effectifs observés si nous sommes dans le cas de l’équilibre de Hardy-Weinberg ?**

Il est difficile pour un élève de décider si les effectifs observés sont très éloignés ou pas des effectifs théoriques dans le cadre de la loi de Hardy Weinberg. Surtout que cela dépend de la taille de l’effectif, et du risque d’erreur que l’on est prêt à prendre. (NB. C’est la même chose pour déterminer à partir de résultats de F2 si deux gènes sont liés ou indépendants).

La solution à ce problème est d’utiliser des tests statistiques pas au programme de SVT (Vérifier en maths !).

Je vous propose donc un fichier Excel dans lequel **l’élève peut rentrer les effectifs observés, et théoriques des génotypes** dans le cas d’un gène avec deux allèles autosomiques, et d’un problème avec un degré de liberté (feuille calculseleves). Dans la feuille suivante (calculstableur), il n’a qu’à rentrer les effectifs observés et tout le reste est calculé automatiquement.



Tableau « vide »

**Exemple : Belin p 77 groupes sanguin MN**

On a :

|  |  |  |  |
| --- | --- | --- | --- |
| Génotypes | (M/M) | (M/N) | (N/N) |
| Nombre d’individus (effectifs observés) | 406 | 744 | 332 |

On peut alors calculer les fréquences alléliques : (on admet que si la population est très grande, les fréquences des allèles sont égales aux probabilités de « piocher » un allèle dans la population (« loi des grands nombres » citée dans le programme d’ES)

f(M) = (406+744/2)/(406+744+332) = 0,52

On en déduit f(N) = 1-f(M) = 0,48

On peut alors calculer les effectifs théoriques des différents génotypes si l’équilibre de Hardy-Weinberg est respecté.

f(M/M) = p² ; f(M/N) = 2pq et f(N/N) = q²

|  |  |  |  |
| --- | --- | --- | --- |
| Génotypes | (M/M) | (M/N) | (N/N) |
| Nombre d’individus observés | 406 | 744 | 332 |
| Effectifs théoriques | P²\*(406+744+332) =408 | 2pq/1482= 739 | q²\*(406+744+332)=334 |

Il faut alors décider si la distribution observée correspond bien à la distribution théorique si on respecte la loi de Hardy Weinberg. La feuille de calcul proposée est une aide à cette décision. (Ici, les valeurs sont très proches et ne prêtent pas trop à confusion).

Dans la 2ème feuille de calcul (calculstableur), il suffit de compléter les effectifs observés, dans la 1ère feuille de calcul (calculseleves), l’élève remplit les effectifs observés et théoriques :



On a la valeur du Chi² et la conclusion à tirer : il est inférieur à 3,84, donc on ne peut pas rejeter l’hypothèse que l’équilibre est bien atteint (sinon on a plus de 5% de chances de se tromper)

**Exemple Hachette p 63 groupes sanguin MN**



Encore une fois, le Chi2 est inférieur à la valeur de la table pour un risque de 5%, on ne peut pas rejeter l’hypothèse de l’équilibre.

1. **Attention aux degrés de libertés !**

Ne pas faire un problème « qui tourne en rond » !

Il n’est pas toujours simple d’obtenir les fréquences des génotypes, surtout quand un allèle est dominant. Par exemple pour le groupe sanguin rhésus et les allèles + (de fréquence p) et – (de fréquence q), d’après le phénotype on aura des individus [rhésus+] et des individus [rhésus-].

On peut alors être tentés de se dire que les individus [rhésus-] sont à la fréquence q², on peut donc calculer q, puis p, et obtenir les fréquences des différents génotypes. Mais en faisant cela, on a admis que la population était à l’équilibre de Hardy-Weinberg, et donc quand on calculera les effectifs théoriques, on trouvera forcément l’équilibre ! Il y a dans cet exemple 0 degré de liberté.

Il faut donc parfois trouver un autre moyen d’avoir les génotypes : par exemple si on peut faire une électrophorèse de protéines, un séquençage… pour distinguer les hétérozygotes des homozygotes.