**Chapitre 2 : Génétique des populations.**

La génétique des populations a pour objectif l’étude de la fréquence des gènes et des génotypes, et des facteurs susceptibles de modifier ces fréquences au cours des générations successives. Certains de ces facteurs comme la sélection, les mutations, la dérive génétique et les migrations peuvent changer la fréquence des gènes et des génotypes. La consanguinité (union entre sujets apparentés) peut modifier la fréquence des génotypes sans influencer la fréquence des gènes.

Les définitions suivant sont très important dans notre étude :

**Espèce :** Ensemble d’individus ayant des caractéristiques morphologiques, physiologique et chromosomique semblables et qui pouvez se reproduire entre eux, et le produit de leur croisement est fertile.

**Race :** Subdivision de l’espèce. Ensemble d’individus ayant des caractères génotypiques et phénotypiques communs les distinguant d’une autre race.

**Population :** Ensemble des individus de même espèce qui partagent le même espace géographique est qui se reproduit entre eux de façon effective.

1. **Fréquence génique et génotypique.**

**Structure génitique :** c’est à la fois la structure allélique et la structure génotypique.

**Structure allélique (génique) :** c’est la fréquence d’un gène allèle donné rapporté au nombre totale des gènes allèle dans un locus génétique dans une population.

**Structure génotypique :** c’est la fréquence d’un génotype donné rapporté au nombre totale des génotypes dans un locus génétique dans une population.

Pour parler de fréquences géniques (ou plutôt alléliques). Pour un gène autosomique, dans une population de N individus, il y a 2 N locus.

Si l’on considère un locus avec deux allèles **A** et **a**, **p** définit la proportion d’allèles **A** et **q** la proportion d’allèles **a**.

Deux méthodes sont utilisées pour l’estimation des fréquences allélique :

1. Sur la base des nombres des individus

Nous considérerons, que nous travaillons sur un locus codominant (on distingue les hétérozygotes des homozygotes) avec deux allèles chez un organisme diploïde. Admettons que nous ayons N individus dans une population. Parmi ces N individus, N1 se sont révélés être de génotype **AA**, N2 **Aa** et N3 **aa**. Notez que N = N1 + N2 + N3.

Soit **p** et **q**, les fréquences des allèles **A** et **a** respectivement dans l’échantillon de N ndividus. Il y a 2N allèles en tout puisque nous sommes chez des diploïdes.

 Il y a 2N1 et N2 allèles **A** chez les **AA** et **Aa** respectivement et 2N3 et N2 allèles **a** chez les **aa** et **Aa** respectivement. La fréquence des allèles **A** et **a** dans l’échantillon est donc :

p(A) = (2N1+N2)/2N

q(a) = (2N3+N2)/2N

1. Sur la base des fréquences génotypiques

L’estimation de la fréquence des gènes à partir des génotypes n’est possible que si tous les génotypes sont identifiables : les deux allèles sont codominants.

La meilleure estimation des la fréquence de ces allèles est :

 p (A)= f(AA) + 1/2 f(Aa)

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

si on utilisant les samboles suivant

D = f(AA)

H = f(Aa)

R = f(aa)

Donc on a

p (A)= D + ½ H

 q (a) = R + ½ H

1. **Equilibre de Hardy Weinberg.**

La loi de Hardy-Weinberg décrit les relations entre les fréquences génotypiques et les fréquences alléliques.

Proposée en 1908 indépendamment par le mathématicien anglais Hardy et le médecin allemand Weinberg, la loi de Hardy-Weinberg



Les conditions de la loi de Hardy-Weinberg sont comme suit:

* Population de très grande taille (∞).
* Panmixie pour un locus considéré (lors de la reproduction, les croisements s’effectuent au hasard pour les génotypes considérés).
* Absence de Migration entre population négligeable.
* AbsencedeMutation constante ou négligeable.
* Absence de sélection.

**Panmixie :** Tousles individus de la population ont la même chance de survie et de reproduction et la reproduction s’effectuent au hasard.

La loi Hardy-Weinberg se définit comme suit:

**Dans une grande population panmictique (Panmixie), on absence de migration, sélection et mutation les fréquences des gènes restent constantes d’une génération à l’autre.**

Si les hypothèses du modèle de Hardy Weinberg sont respectées on peut prédire exactement les fréquences génotypiques a partir des fréquences alléniques de la population :

|  |  |
| --- | --- |
| **Fréquence des allèles** | **Fréquence des génotypes** |
| Si on considère un **locus autosomique unique**pour **2 allèles**:• L’un dominant (**A)** de fréquence (**p)**• L’autre récessif (**a)** de fréquence (**q)**La somme des fréquences des 2 allèles dans la population est : **p+q= 1**.Cette formule est valable quelles que soient les fréquences **p** et **q** | Selon les lois de **l’hérédité monogénique** les individus formant la population se présentent avec **l’un des 3 génotypes** :• **(AA)** homozygote dominant• **(Aa)** hétérozygote• **(aa)** homozygote récessifAvec les fréquences respectives **(p2), (2 pq)** et **(q2).**La fréquence des 3 génotypes dans lapopulation suit **la loi binomiale** : **(p+q) 2 =1**La distribution des génotypes dans lapopulation est donnée par la formule :**p2 + 2 pq + q2 = 1** |

1. **Estimation des fréquences.**

Prenons l’exemple d’un locus qui peut être occupé par deux allèles **A** et **a**, tels que la proportion de gènes **A** est **p** et la proportion de gènes **a** est **q** :

 p+q =1

|  |  |
| --- | --- |
|   |  *Gamètes mâles          A (p)            a (q)*  |
|   *Gamètes  A (p)*  | *AA (p2)*  | *Aa (pq)*  |
|   Gamètes  a (q)  |       Aa (pq) |   aa (q2) |

Fréquence du génotype **AA : p2**
Fréquence du génotype **aa : q2**
Fréquence du génotype **Aa : 2pq**

f (A) = p2 + pq = p ( p+ q )= p
f (a) = q2 + pq = q ( p + q )= q

***En savoir plus***:
*Dans une population telle que définie précédemment, nous allons voir comment évolue la fréquence des gènes d’une génération à l’autre:*

Tableau 1

|  |  |  |  |
| --- | --- | --- | --- |
| *Accouplement possibles*  | *AA*  | *Aa*  |  *aa*  |
| *AA* |  *p4*  |  *2p3q*  |  *p2q2*  |
| *Aa* |  *2p3q* |  *4p2q2*  |  *2pq3*  |
| *Aa* |  *p2q2*  |  *2pq3*  |  *q4*  |

*Fréquence des accouplements aa x Aa = 2pq3 + 2pq3 = 4 pq3*

Tableau 2

|  |  |  |
| --- | --- | --- |
|   |   | *Génotypes des descendants* |
|   *Type d’accouplement* |    *fréquence* |      *AA* |      *Aa* | *aa* |
|   *AA X AA* |           *p4*  | *p4*  |  |  |
|   *AA X Aa* |       *4p3q* |  *2p3q* | *2p3q* |  |
|   *Aa X Aa* |      *4p2q2*  |  *p2q2*  | *2p2q2*  | *p2q2*  |
|   *aa X aa* |         *q4*  |  |  | *q4*  |
|   *aa X Aa* |      *4pq3*  |  | *2pq3* | *2pq3*  |
|   *AA x aa* |      *2p2q2*  |  | *2p2q2*  |  |

*Total :
AA : p2 (p4 + 2p3q + p2q2) = p2 (p2 + 2pq + q2) = p2
Aa : 2pq (p4 + 2p3q + p2q2) = 2pq(p2 + 2pq + q2) = 2pq
Aa : q2 (p4 + 2p3q + p2q2) = q2 (p2 + 2pq + q2) = q2

La proportion des génotypes reste donc inchangée à la deuxième génération, -c’est l’équilibre de Hardy-Weinberg.*

Soit:

* Les fréquences des génotypes sont appelées D pour (AA), H pour (Aa) , et R pour (aa) avec 0 =< [D,H,R] = < 1 et **D + H + R = 1**
* Les fréquences des allèles **p**, et **q** avec 0 =< [p,q] =< 1 et **p+q = 1**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Génotypes | AA |  Aa |  Aa |  |  |
| Effectifs | DN |  HN |  RN  |  (effectifs totaux N) |  |
| Fréquences génotypiques  | D |  H |  R |  avec (D+H+R) = 1  |  |

|  |  |  |  |
| --- | --- | --- | --- |
|  Fréquences allélique : | de  **A** = | **D + ½ H = p** |  |
|  | de   **a** = | **R + ½ H = q** | avec p+q=1 |

**NOTES**

Les fréquences génotypiques permettent, dans tous les cas, de calculer les fréquences alléliques.

* Démonstration que **p = D + ½ H**, par dénombrement d'allèles :

Effectif population = N -> nombre d'allèles = 2N
**p** = nombre d’allèles **A** / nombre total des allèles = (2DN + HN) / 2N = D + H/2
de même pour a :
**q** = nombre d’allèles **a** / nombre total des allèles = (2RN + HN) / 2N = R + H/2

**II.1.1. EXERCICE**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| soit  | les phénotypes   | [A1] | [A1A2] | [A2] |
|  | les génotypes   | A1A1 | A1A2 | A2A2 |
|  | les effectifs   | 167 | 280 | 109 |    total N : 556 |  |

calculer les fréquences suivantes : F(P : phénotypes) , F(G : génotypes), F(A: allèles), F(gamètes) :
F(A) = F(gam), car il y a 1 allèle (de chaque gène) par gamète
Ici, de plus, F(p) = F(G), car allèles co-dominants

|  |  |  |  |
| --- | --- | --- | --- |
| F(P) = F(G) | 167/556 | 280/556 | 109/556 |
| Soit : | D=0.300 | H=0.504 | R=0.196    | vérifiez:Σ (D,H,R) =1 |

|  |  |
| --- | --- |
| F(A) = F(gam.)  | p = D+H/2 = (167+280/2)/ 556 soit 0.300+0.504/2 = 0.552 |
|    | q = R+H/2 = (109+280/2)/ 556 soit 0.196 + 0.504/2 = 0.448 |
|  |  |  |

 vérifiez: Σ (p,q)=1

**Maximum d’hétérozygotie**

 Si on donne toutes les valeurs possible de p (de p=0 jusqu’au p=1) on peut tracé les courbes suivants, présentant les variations des fréquences génotypiques en fonction des variations des fréquences allélique.

****

La figure montre la correspondance entre la fréquence allélique q de a et les fréquences génotypiques dans le cas de deux allèles en régime panmictique. La fréquence maximale des hétérozygotes H est alors atteinte lorsque

 p = q = 0.5 et H = 2pq = 0,50.

A l'inverse, lorsque l'un des allèles est rare (ex: q très petit), presque tous les sujets possédant cet allèle se trouvent sous la forme hétérozygote.

**Pour vérifier si la loi de Hardy -Weinberg s'applique à une population donnée pour un gène donné**

**On calcule les fréquences alléliques à partir des fréquences des génotypes**

Population N ---> nombre d'allèles = 2N

Nombre d’individus AA= N1
Nombre d’individus BB= N2
Nombre d’individus AB= N3

                                             2 N1+ N2
Fréquence de l’allèle A : p = -----------
                                              2N

                                             2 N3+ N2
Fréquence de l’allèle B : q = -----------
                                              2N

**On compare les nombres attendus aux nombres observés**

Nombres attendus = AA: p2 x N
                                     AB: q2 x N
                                     BB: 2pq x N

Nombres observés= fréquence des phénotypes

On fait un test de X2

***EXEMPLE***
*la fréquence des différents génotypes a été étudié dans trois gènes dans deux groupes de population: 95 Afro-Américains (AA), et 95 américains d’origine européenne (EA)*

Fréquence des différents génotypes

|  |  |  |  |
| --- | --- | --- | --- |
| *Polymorphisme*  |   *GJA-4 (C10009T)* |   *MMP-3 (–1171delA)* | *PAI-1 –668delG*   |
| *Genotype* | *C/C* | *C/T*  | *T/T*  | *-/-* | *-/A*  | *A/A* |   *-/-* | *-/G* |   *G/G* |
| *AA nombre           %* | *24(25.3)* | *52(54.9)* | *19(19.8  )* | *2(2.2)* | *19(19.8)* | *74(78)* | *7(7.6)* |   *35(37)* |   *53(55.4)* |
| *EA nombre           %* | *46(48.4)* | *42(44.1)* | *7(7.5)* | *30(32.3)* | *42(44.1)* | *23(23.6)* |   *37(38.7)*  |   *44(46.2)*  |   *14(15.1)*  |

*Pour le polymorphisme GJA-4 (C10009T)

Dans la population AA : nombres observés

Fréquence de l’allèle C : (24 X 2) + 52 = 100 = 0,526
                                         ----------------     -----
                                             95 X 2         190

Fréquence de l’allèle T : (19 X 2) + 52 = 90 = 0,474
                                         ----------------   ----
                                               95 X 2     190
Nombres attendus :
• génotype C/C : (0,526 ) 2 X 95    = 26,2
• génotype C/T : 2 X 0,526 X 0,474 X 95  = 47,3
• génotype T/T : ( 0,474 )2 X 95   = 21,3

x 2 : Σ(o-e)2 = (24-26.2)2 + (52-47.3)2 + (19-21.3)2 = 0.185+ 0.47 + 0.25 = 0.9
      --------        -----------      ------------         ----------
           e              26.2               47.3              21.3
    (p entre 0.3 et 0.5 : non significatif)*

**EXTENSION DE HW A D'AUTRES SITUATIONS GENIQUES**

**1. POUR UN GENE AUTOSOMAL, TRIALLELIQUE, CO-DOMINANT**

|  |  |  |  |
| --- | --- | --- | --- |
| 3 allèles  | A1, | A2, | A3 |
| de fréquences  | F(A1) = p,  | F(A2) = q,  | F(A3) = r |

|  |  |  |  |
| --- | --- | --- | --- |
|  | A1(p) | A2 (q) | A3 (r) |
| A1 (p) | p2 | pq | pr |
| A2 (q) | pq | q2 | qr |
| A3 (r) | pr | qr | r2 |

il y aura 6 génotypes

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | A1A1 | A1A2 | A1A3 | A2A2 | A2A3 | A3A3 |
| Fréquencesgénotypiques sous HW  | p2  | 2pq  | 2pr | q2 | 2qr | r2 |
|  |  |  |  |  |  |  |

**2. POUR UN GENE AUTOSOMAL, DIALLELIQUE, NON CO-DOMINANT**

A est dominant sur a, récessif; dans ce cas les génotypes (AA) et (Aa) ne pourront pas être distingués dans la population. Seuls les individus de phénotype [A] d'effectif N1 pourront être reconnus des individus de phénotypes [a] d'effectif N2.

Génotypes AA Aa aa

Phénotypes [A] [a]

Effectifs N1 N2 Total N

Fréquences

génotypiques 1-q2 q2

avec R=q2 = N2/N = N2 / N

Et la fréquence de l'allèle a = F(a) = √(q2)
Il s'agit d'une méthode couramment utilisée en génétique humaine pour calculer la fréquence des gènes récessifs rares.

Fréquences des homozygotes et des hétérozygotes pour des règnes récessifs rares chez l'homme

Exemple : Supposons une maladie récessive liée à des mutations homozygote d’un gène biallélique, **A** représentant l’allèle normal et **a** l’allèle muté.

Le phénotype des individus présentant les génotypes AA et Aa est identique.
Par contre, la proportion d’individus aa correspond à q2.
On peut donc en déduire q= √ q2 ; p= 1-q
La fréquence des hétérozygotes Aa correspond à 2pq et peut être calculée

2pq = 2 x √q2 x (1- √q2)

Si q est très petit, 1 - √q2 est très proche de 1, donc 2pq ~ √q2

*Exemple: la phénylcétonurie est une maladie récessive autosomique qui atteint un enfant sur 10 000.
q = √1 / 10 000 = 1 / 100
2 pq = 2 x 1/100 x (1 – 1/100)
1 – 1/100 = 99/100 : très peu différent de 1 ; donc 2pq ~ 2 X 1/100 ~ 1/50*

**Pour les maladies liées au chromosome X**

La situation est plus simple. Les hommes ne possédant qu’un seul chromosome X, la fréquence de l’allèle morbide est égale à la proportion de garçons qui sont atteints.

Maladie récessive liée au chromosome X

|  |
| --- |
|   Maladie récessive liée au chromosome X |
|   |  Garçon  |   Filles   |
|   Phénotype    |  Sain    |  Malade    |  Sain    |  Saine   |  (Malade)    |
|   Génotype |  XAY  |  Xa Y (I) |  XAXA |  XAXa (H)   |   XaXa |
|   Fréquence |  P |  Q |  p2  |  2pq |   q2  |

On connait I, l’incidence de la maladie dans la population masculine = q
La fréquence des filles hétérozygotes, H est donc égale à :
2 X q X (1 –q)
Si q est très petit : H= 2I (il y a deux fois plus de femmes hétérozygotes que de garçon atteints).

**3. POUR UN GENE AUTOSOMAL, TRIALLELIQUE, NON CO-DOMINANT**

Exemple: les systèmes des groupes sanguins ABO. Bien que le système des groupes sanguins (ABO) chez l'homme soit souvent pris comme un exemple simple de polyallélie, il représente cependant un cas relativement complexe à cause de la codominance de A et B, de la présence d'un allèle nul 0 et de la dominance de A et B sur 0.
Si l'on désigne par

p la fréquence de l'allèle A
q la fréquence de l'allèle B
r la fréquence de l'allèle O

les diverses fréquences génotypiques et phénotypiques sont observées en appliquant la loi de Hardy-Weinberg.

|  |  |  |  |
| --- | --- | --- | --- |
| Phénotypes | Génotypes | Fréquences génotypiques | Fréquences phénotypiques |
| [A] | (AA) | p2  |  |
|  | (AO) | 2pr | p2+2pr |
| [B] | (BB) | q2 |  |
|  | (BO) | 2qr | q2+2qr |
| [AB] | (AB) | 2pq | 2pq |
| [O] | (OO) | r2 | r2  |

En utilisant les identités remarquables telles que:

p2 +2pr +r2 = (p+r)2
q2 +2qr +r2 = (q+r)2

avec

F[A] + F[O] = p2 +2pr +r2 = (p+r)2
F[B] + F[O] = q2 +2qr +r2 = (q+r)2
et F[O] = r2

**POUR UN GENE HETEROSOMAL (= gonosomique)**

**1. CHROMOSOME Y :** fréquence p et q chez le sujet XY; transmission aux descendants mâles

**2. CHROMOSOME X :**

Femelle XA XA p2

 XA Xa 2pq

 Xa Xa q2

Mâle XAY p

 XaY q

soit qm (Xa) chez l'homme, et qf (Xa) chez la femme, fréquences de l'allèle q :

* l'X des garçons (de génération n) est transmis par les mères (génération n-1)

donc qm (Xa) (n) = qf (Xa) (n-1)

* l'X portant l'allèle qf (Xa) des filles a:1/2 chance de venir du père 1/2 chance de venir de la mère

donc qf (Xa) (n) = (qm (Xa) (n-1) + qf (Xa) (n-1))/2

fréquence de l'allèle chez l'homme = fréquence chez la femme à la génération précédente
fréquence de l'allèle chez la femme = moyenne des fréquences des 2 sexes à la génération précédente

**calcul de la différence des fréquences allèliques entre les 2 sexes :**
soit d la différence entre qm (Xa) et qf (Xa)

qm (Xa) (n) - qf (Xa) (n) = ½ (qf (Xa) (n-1) - qm (Xa) (n-1))

= - ½ (qm (Xa) (n-1) - qf (Xa) (n-1))

Donc qm (Xa) (n) - qf (Xa) (n) = (- ½ )n (qm (Xa) (0) - qf (Xa) (0)) : tend vers zéro en 8 à 10 générations

**fréquence moyenne q :**

1/3 des X appartiennent à l'homme, 2/3 à la femme :

q*moy* = 1/3 qm (Xa) (n) + 2/3 qf (Xa) (n)

la fréquence moyenne est invariante
à l'équilibre, qm (Xa) = qf (Xa) = q*moy*

****

**Exercice** : Soit, en génération G0, 100% d'hommes normaux et 100% de femmes daltoniennes, calculer les fréquences du gène jusqu ‘ à G6 :

réponse

|  |  |
| --- | --- |
| G0: XNY | XDXD |
| G0 : qx(0) = 0.00 | qxx(0) = 1.00 |
| G1 : qx(1) = 1.00 | qxx(1) = 0.50 |
| G2 : qx(2) = 0.50 | qxx(2) = 0.75 |
| G3 : qx(3) = 0.75 | qxx(3) = 0.63 |
| G4 : qx(4) = 0.63 | qxx(4) = 0.69 |
| G5 : qx(5) = 0.69 | qxx(5) = 0.66 |
| G6 : qx(6) = 0.66 | qxx(6) = 0.60 |