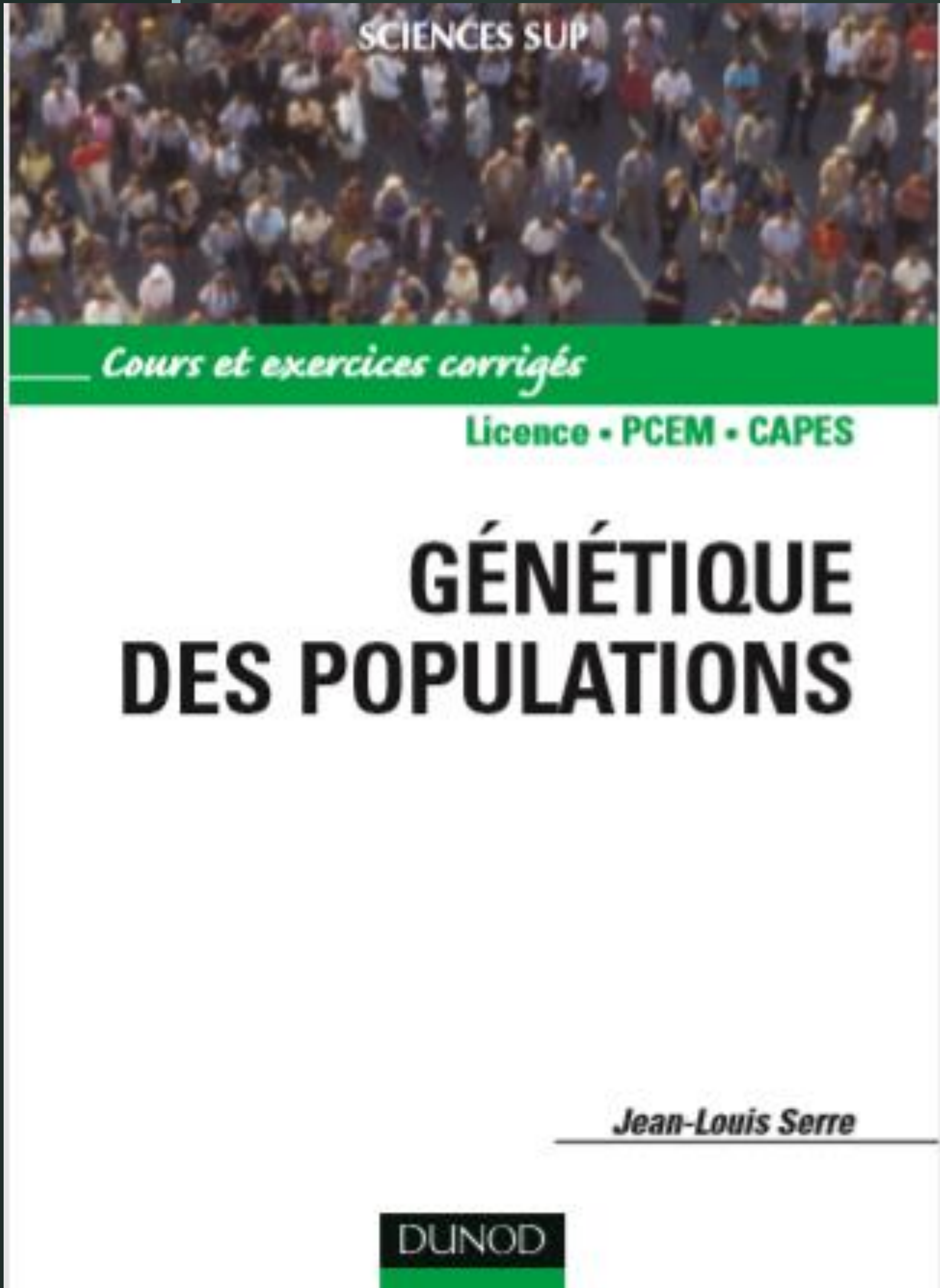


Les étudiants peuvent se servir de ces parties tirées de cette référence pour maitriser des notions fondamentales pour les Tds de la génétique des population

1.3 MESURE DE LA DIVERSITÉ GÉNÉTIQUE ET COMPOSITION GÉNÉTIQUE D'UNE POPULATION

1.3.1 La population

L'espèce est par définition un groupe génétiquement fermé au sein duquel les organismes sont susceptibles, par l'alternance méiose-fécondation de séparer ou de réunir les divers allèles de chacun des gènes et de concevoir des combinaisons génétiques nouvelles par la recombinaison génétique.

Cependant tous les individus d'une même espèce, s'ils sont potentiellement susceptibles de réaliser ce brassage peuvent en être pratiquement empêchés quand des barrières limitent les possibilités de croisements entre certains individus.

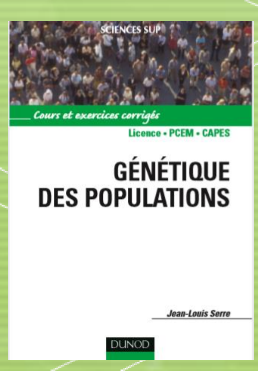
Il peut s'agir d'un isolement géographique, lié à l'existence d'une barrière naturelle comme un océan pratiquement infranchissable ou une chaîne montagneuse plus facilement franchissable, ou plus simplement par la distance qui limite la probabilité de croisements entre individus très éloignés.

Il peut s'agir d'un isolement écologique. Par exemple des plantes d'une même espèce occupant un même territoire semblent former une même population. Mais si une disparité dans la composition du sol décale la floraison entre les plantes du sol A et celles du sol B, les échanges génétiques entre les plantes des sols A et B seront limités. Il sera alors nécessaire de définir deux populations A et B. La définition et l'analyse de populations naturelles supposent donc une bonne connaissance de leur biologie et de leur biotope.

Il peut s'agir enfin, chez l'homme, de barrières culturelles, sociales ou ethniques, qui limitent plus ou moins les possibilités d'unions entre individus, même géographiquement proches.

Une espèce peut donc être subdivisée en sous-groupes au sein desquels la possibilité d'échanges génétiques entre individus est effective ; ces sous-groupes sont appelés populations et l'ensemble des allèles qu'ils partagent, pour chacun des gènes de l'espèce, en constitue le patrimoine génétique (pool allélique).

La mesure de la diversité génétique à l'intérieur des populations mais aussi entre les populations, l'origine et le devenir de ces diversités intra- et inter-populationnelles sont un enjeu important de la génétique des populations, notamment chez l'homme, en raison des polémiques qui ont accompagné la définition et l'usage du concept de race, ou les débats sur l'émergence de l'homme moderne à partir de l'*Homo erectus* (voir plus loin).



1.3.2 Variables d'état de la diversité et composition génétique d'une population

Il est possible de définir une chaîne de causalité liant la variabilité phénotypique des caractères et la diversité génétique sous jacente qui en est la cause (figure 1.6, à gauche). À chacun des niveaux hiérarchiques de la diversité, on peut associer des variables d'état qui mesurent la diversité génétique à ce niveau, les fréquences alléliques, les fréquences génotypiques, les fréquences phénotypiques (figure 1.6, à droite).

Les fréquences phénotypiques sont toujours accessibles directement par le dénombrement des phénotypes présents dans un échantillon, et la question se pose de savoir si il existe des relations mathématiques simples permettant, si on connaît la diversité à un niveau hiérarchique, d'en déduire la diversité à un autre niveau (figure 1.6, flèches doubles à droite). Si de telles relations sont disponibles alors la connaissance de la diversité en un point quelconque des niveaux hiérarchiques permettrait d'avoir une connaissance exhaustive de la diversité génétique de la population en tout autre point.

Or c'est bien le premier but de la génétique des populations que de savoir mesurer la diversité pour définir la composition génétique d'une population ou d'une espèce, au niveau des allèles et des génotypes.

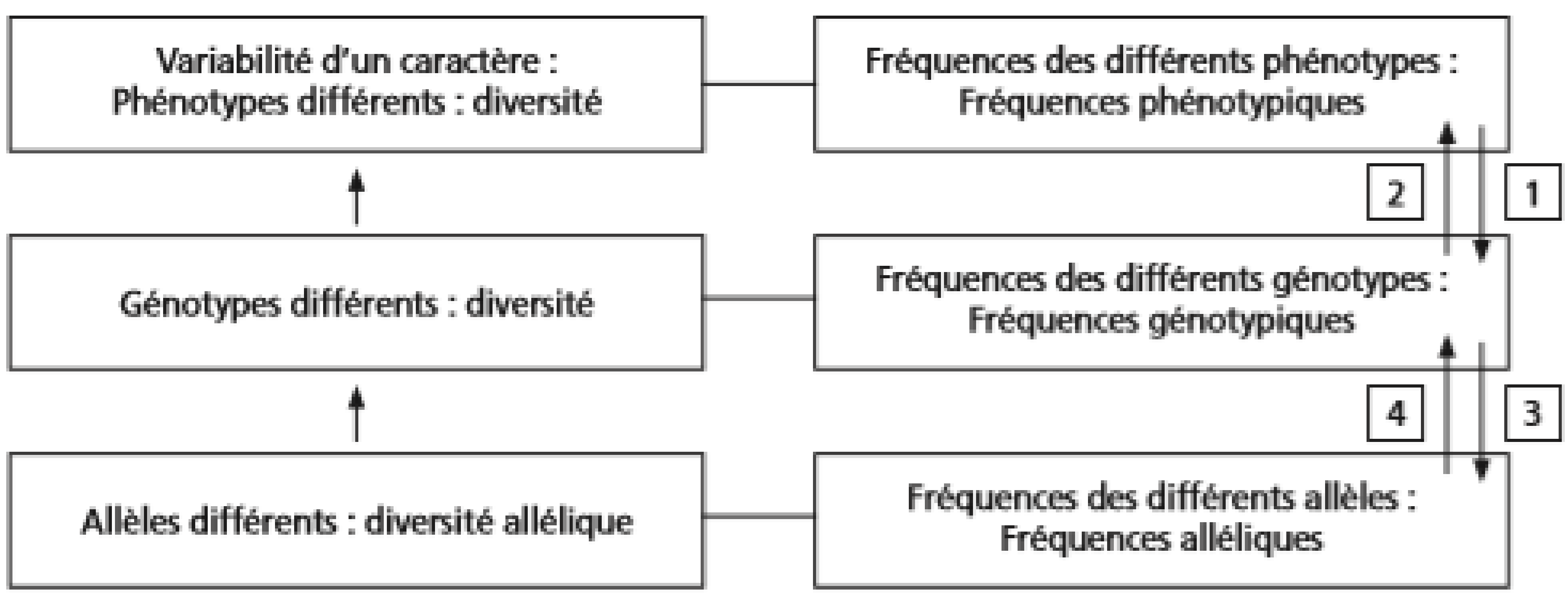
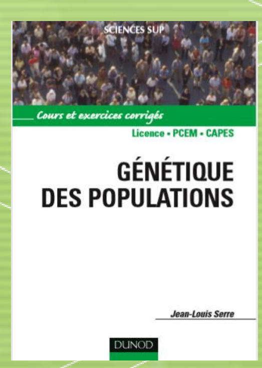


Figure 1.6 Niveaux hiérarchiques de la diversité génétique et variables d'états associés.

Dans un premier temps, on va définir la mesure de la diversité génétique relative à un gène (ou un marqueur) sachant qu'une fraction des problèmes de génétique des populations correspond à une telle situation. Ultérieurement, on généralisera nos résultats à l'étude simultanée de deux gènes ou marqueurs, situation très communément rencontrée en épidémiologie génétique, mais, comme cela a été dit, l'approche analytique devient impossible au-delà et il est nécessaire de recourir à des simulations informatiques.



2.3 LE MODÈLE DE HARDY-WEINBERG

Le modèle théorique général de Hardy-Weinberg sera établi dans le cas le plus simple, le plus général et le plus utile d'un gène autosomique di-allélique, pour une population d'organismes à sexes séparés, présentant des générations séparées.

La généralisation du modèle à un cas multi-allélique, le cas particulier des gènes liés au sexe, le cas des générations chevauchantes ou le cas de deux gènes étudiés simultanément seront exposés au chapitre 3.

Considérons une population constituée des adultes reproducteurs de la génération i.

Dans le cas d'un gène di-allélique (allèles *A1* et *A2*), la composition génétique de la population est constituée des :

– **trois génotypes possibles :**	***A1/A1***	***A1/A2***	***A2/A2***
– **de fréquences génotypiques :**	***D***	***H***	***R***

Ces fréquences D, H et R sont quelconques et seront considérées, en première analyse, égales dans les deux sexes.

On sait (voir chapitre 1) en déduire les fréquences des allèles *A1* et *A2*, respectivement nommées p et q, soit :

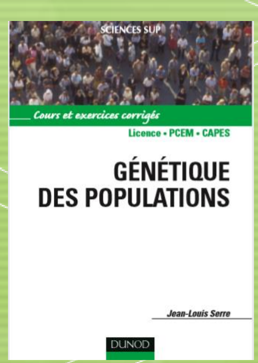
$$p = D + H/2$$

$$q = R + H/2$$

Quelle sera la constitution génétique (fréquences génotypiques et fréquences alléliques), à la génération suivante après un cycle vital ?

Il suffit de se reporter au cycle vital présenté dans le paragraphe précédent pour obtenir la solution en fonction des conditions que nous aurons dû définir et qui forment les « conditions de l'équilibre de Hardy-Weinberg ».

On peut établir ce modèle de deux manières différentes, soit par la formation des couples en suivant pas à pas le cycle vital, soit par le schéma de l'urne gamétique.



2.3.1 Établissement du modèle de Hardy-Weinberg par le cycle vital

a) *Formation des couples : condition de panmixie*

La première étape du cycle vital (**1** dans l'encadré ci-dessus) est la formation de couples reproducteurs. Des règles d'union peuvent exister : on fera l'hypothèse qu'ils se forment au hasard, les couples sont dits **panmictiques** (**condition de panmixie**).

Dans ce cas, on peut générer six types possibles de couples (tableau 2.1).

> ***Remarque :*** il convient de noter que la panmixie ne signifie pas que les six types de couples sont équiprobables ou équifréquents (1/6) mais que leurs probabilités respectives sont fonction de la fréquence, dans la population, de chacun des génotypes associés dans le couple. Pour prendre un exemple caricatural les couples noirs × noirs ne peuvent pas avoir la même probabilité ou la même fréquence dans un pays comme la Suède ou le Sénégal.

b) *Probabilité et fréquences des évènements : condition d'effectif infini de la population*

Dans une population naturelle concrète, ce qui nous importe, et ce qui compte, ce sont les fréquences des couples, les fréquences de leurs descendants, les fréquences des génotypes, les fréquences alléliques, et non les probabilités de ces évènements.

On sait que la fréquence d'un évènement est égale à sa probabilité si le nombre de tirages est très grand (loi des grands nombres) ; par exemple la fréquence des « piles » peut être égale à 0,7 sur dix tirages mais ne peut, sur 100 000 tirages, s'écarter notablement de sa probabilité égale à 0,5.

Afin de pouvoir considérer que les fréquences des couples ou des génotypes sont égales à leurs probabilités respectives, nous considérerons que la population est de taille infinie (concrètement suffisamment grande pour y appliquer la loi des grands nombres aux évènements étudiés). Le seuil grand/petit n'est pas définissable dans l'absolu et ne sera discuté qu'à la fin de l'ouvrage (voir chapitre 8).

Cette condition d'effectif infini s'ajoute à la condition de panmixie. Le tableau de formation des couples et de leurs descendants se présente alors ainsi :

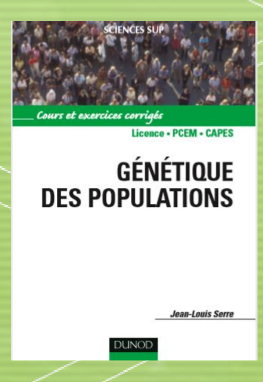
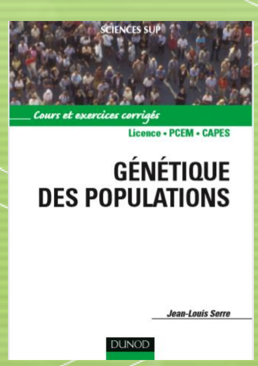


TABLEAU 2.1 FRÉQUENCES DES COUPLES PANMICTIQUES ET DE LEURS DESCENDANTS POUR UN GÈNE DI-ALLÉLIQUE.

Types de couples	Fréquences des couples	Fréquences des descendants *A1/A1*	Fréquences des descendants *A1/A2*	Fréquences des descendants *A2/A2*
A1/A1 x A1/A1	D^2	1	0	0
A1/A1 x A1/A2	$2DH$	1/2	1/2	0
A1/A1 x A2/A2	$2DR$	0	1	0
A1/A2 x A1/A2	H^2	1/4	1/2	1/4
A1/A2 x A2/A2	$2RH$	0	1/2	1/2
A2/A2 x A2/A2	R^2	0	0	1
TOTAL	1	$D^2 + DH + H^2/4$	$DH + 2DR + H^2/2 + RH$	$R^2 + RH + H^2/4$

Maintenant que les couples sont formés et leurs fréquences connues (sous les conditions de panmixie et d'effectif infini), il s'agit de réaliser les fécondations pour obtenir les adultes reproducteurs de la génération suivante de manière à avoir réalisé le cycle vital d'une génération.



2.3.3 Bilan du modèle de Hardy-Weinberg

a) *La relation de Hardy-Weinberg*

Les nouvelles fréquences génotypiques correspondent soit au carré des fréquences alléliques pour les homozygotes, soit au double produit des fréquences alléliques pour l'hétérozygote.

Les trois génotypes sont	***A1/A1***	***A1/A2***	***A2/A2***
Leurs fréquences sont égales à	p^2	$2pq$	q^2

La relation ainsi établie entre les fréquences alléliques et les fréquences génotypiques est appelée « relation de Hardy-Weinberg » ou « relation panmictique » car elle découle directement de l'hypothèse panmictique.

Remarque : cette relation mathématique permet, en supposant que l'hypothèse panmictique soit valide, de remonter aux fréquences génotypiques (donc phénotypiques) quand on ne connaît que les fréquences alléliques (voir figure 1.6 et paragraphe 2.3.5.c).

b) *L'équilibre de Hardy-Weinberg*

Les fréquences alléliques sont inchangées à la génération suivante.

En effet, selon la formule de calcul des fréquences alléliques à partir des fréquences génotypiques (voir chapitre précédent), on a :

$$f(A1) = p^2 + 2pq/2 = p^2 + pq = p\,(p + q) = \boldsymbol{p}$$
$$f(A2) = q^2 + 2pq/2 = q^2 + pq = q\,(p + q) = \boldsymbol{q}$$

Cette stabilité de la composition génétique de la population est appelée « équilibre de Hardy-Weinberg ».

c) *Les conditions de l'équilibre de Hardy-Weinberg*

Les conditions supposées réalisées dans le modèle de l'équilibre de Hardy-Weinberg peuvent se regrouper en trois grands groupes :

condition 1 : la population est panmictique ;
condition 2 : la population est de taille quasi infinie (loi des grands nombres applicable) ;
condition 3 : mutation, sélection, migration sont inexistantes (ou négligeables).