

# Lotka

## Tradução Técnica

### ■ **Excerto 7: Equilíbrio de Hardy-Weinberg**

### ■ **Excerpt 7: Hardy-Weinberg equilibrium**

### ◆ **Lucas Monteiro Nogueira**

Nº	Título	Idioma	Referência
1	<b>Équilibre de Hardy-Weinberg et l'urne gamétique</b> (Equilíbrio de Hardy-Weinberg e a urna gamética)	Francês → Português	Serre (2006)
2	<b>Mating frequencies and another proof of the Hardy-Weinberg law</b> (Frequências de cruzamento e outra prova da lei de Hardy-Weinberg)	Inglês → Português	Falconer e Mackay (1989)
3	<b>More than one locus: Gametic phase disequilibrium</b> (Mais de um locus: Desequilíbrio de fase gamética)	Inglês → Português	Falconer e Mackay (1989)

### ◆ **Referências**

1. Serre, J.-L. (2006). *Génétique des Populations: Cours et Exercices Corrigés*. Dunod.

2, 3. Falconer, D.S. e Mackay, T.F.C. (1989). *Introduction to Quantitative Genetics*. 3ª edição. Longman Scientific & Technical.

### ◆ **Índice**

#### 1. **Équilibre de Hardy-Weinberg et l'urne gamétique**

##### 1.1. **L'urne gamétique**

##### 1.2. **La relation de Hardy-Weinberg**

#### 2. **Mating frequencies and another proof of the Hardy-Weinberg law**

#### 3. **More than one locus: Gametic phase disequilibrium**

◆ 1. Équilibre de Hardy-Weinberg et l'urne gamétique

↳ Referência 1 (Serre, 2006), pp. 39, 42 – 43. Com modificações.

1.1. L'urne gamétique

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:

<b>Trois génotypes possibles</b>	<b>A1/A1</b>	<b>A1/A2</b>	<b>A2/A2</b>
<b>Fréquences génotypiques</b>	$D$	$H$	$R$

Ces fréquences  $D$ ,  $H$  et  $R$  sont quelconques et seront considérées, en première analyse, égales dans les deux sexes. On peut déduire les fréquences des allèles  $A1$  et  $A2$ , respectivement nommées  $p$  et  $q$ , soit:

$$p = D + \frac{H}{2} \quad (1.1)$$

$$q = R + \frac{H}{2} \quad (1.2)$$

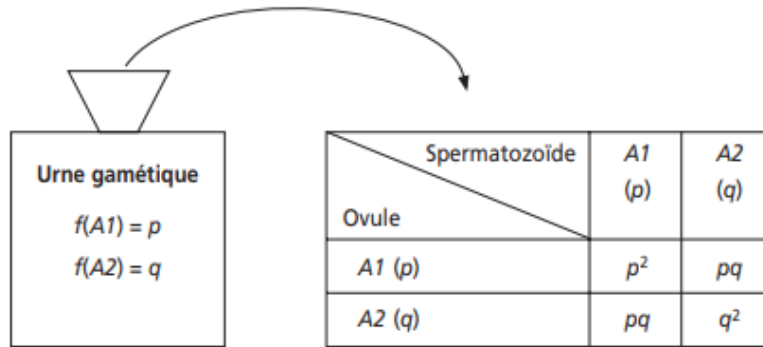
Quelle sera la constitution génétique (fréquences génotypiques et fréquences alléliques) à la génération suivante après un cycle vital? On peut répondre cette question au moyen du schéma de l'**urne gamétique**.

Si on considère que les couples se forment au hasard, c'est-à-dire qu'ils sont panmictiques (condition de **panmixie**), et que les gamètes eux-mêmes s'unissent au hasard (condition de **pangamie**) lors de la fécondation, on peut alors admettre que tout se passe comme si les couples mettaient en vrac leurs gamètes dans une urne gamétique de spermatozoïdes pour les mâles et d'ovules pour les femelles, et que tout descendant est issu de l'union de deux gamètes tirés au hasard dans chacune des deux urnes. Ce schéma est nommé schéma de l'urne gamétique (Figure 1). C'est d'ailleurs la réalité biologique pour les espèces végétales, et de nombreuses espèces animales aquatiques, qui émettent leurs gamètes dans le milieu environnant.

*Remarque: Sauf exception la panmixie s'accompagne toujours de la pangamie et on considère que la condition de panmixie inclue celle de la pangamie.*

Si les fréquences alléliques sont les mêmes dans les deux sexes, les urnes sont identiques. Compte tenu de la composition des urnes il sera alors possible de calculer la probabilité de tirage de chacun des trois génotypes possibles. D'où le schéma suivant:

**Figure 1.** Schéma de l'urne gamétique et conséquences mathématiques sur la diversité génétique.



La composition génétique de l'urne selon le schéma précédent suppose que chacun des adultes reproducteurs y a placé un nombre égal de gamètes, ce qui signifie qu'il y a *absence de sélection em matière de fertilité*.

Il suppose aussi que les homozygotes  $A1A1$  n'y déposent que des gamètes  $A1$ , ce qui signifie *l'absence de mutations*.

Une fois obtenue par le dépôt des gamètes, la composition de l'urne est stable, ce qui signifie *l'absence de sélection gamétique*.

Le tirage de deux gamètes aboutit alors et seulement alors à la formation d'un génotype:

$A1/A1$ avec la probabilité $p^2$
ou $A1/A2$ avec la probabilité $2pq$
ou $A2/A2$ avec la probabilité $q^2$

## 1. Equilíbrio de Hardy-Weinberg e a urna gamética

### 1.1. A urna gamética

Consideremos uma população constituída de adultos reprodutores de uma geração  $i$ . No caso de um gene descrito por dois alelos ( $A1$  e  $A2$ ), a composição genética da população é constituída de:

<b>Três genótipos possíveis</b>	<b>A1/A1</b>	<b>A1/A2</b>	<b>A2/A2</b>
<b>Frequências genotípicas</b>	$D$	$H$	$R$

Tais frequências  $D$ ,  $H$  e  $R$  são arbitrárias, e, em primeira análise, serão consideradas iguais nos dois sexos. Podemos deduzir as frequências dos alelos  $A1$  e  $A2$ , respectivamente denotadas por  $p$  e  $q$ , como:

$$p = D + \frac{H}{2} \quad (1.1)$$

$$q = R + \frac{H}{2} \quad (1.2)$$

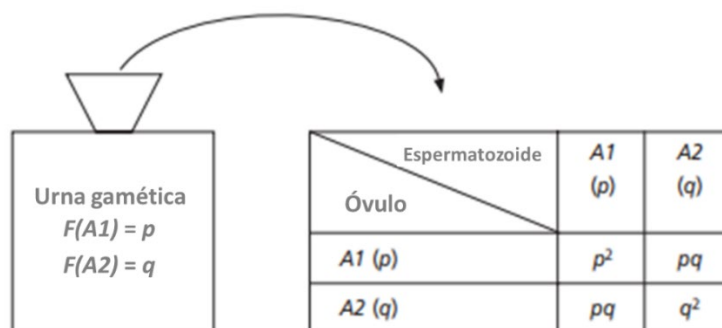
Qual será a composição genética (frequências genotípicas e frequências alélicas) da geração seguinte após um ciclo de vida? Podemos responder a essa pergunta por meio de uma representação esquemática da **urna gamética**.

Se considerarmos que os cruzamentos se formam ao acaso, ou seja, são *panmíticos* (condição de **panmixia**), e que os gametas unem-se eles próprios ao acaso (condição de **pangamia**) quando da fecundação, podemos imaginar que tudo se passa como se os casais inserissem seus gametas em urnas gaméticas de espermatozoides, no caso dos machos, e óvulos, no caso das fêmeas, e que todos os descendentes são formados a partir da união de dois gametas extraídos ao acaso de cada uma das duas urnas. Tal arranjo é chamado de *urna gamética* (Figura 1). De fato, esta é a realidade para espécies vegetais e um número substancial de animais aquáticos que liberam seus gametas no ambiente circundante.

*Nota: Salvo em algumas exceções, a hipótese de panmixia é sempre acompanhada de pangamia; consideramos, pois, que a condição de panmixia inclui a condição de pangamia.*

Se as frequências alélicas são as mesmas nos dois sexos, as urnas são idênticas. Tendo em conta a composição das urnas, será possível calcular a probabilidade de obter-se cada um dos três genótipos possíveis. Temos o seguinte esquema:

**Figura 1.** Esquema da urna gamética e consequências matemáticas para a diversidade genética.



A composição genética da urna supõe que cada um dos reprodutores adultos contribuiu com o mesmo número de gametas, o que significa que temos *ausência de seleção em termos de fertilidade*.

Supõe-se também que homozigotos A1A1 somente depositam gametas A1, o que implica *ausência de mutações*.

A composição da urna permanece estável após a deposição dos gametas, o que implica *ausência de seleção gamética*.

A extração de dois gametas resulta então, e apenas então, na formação de um genótipo:

A1/A1 com probabilidade $p^2$
ou A1/A2 com probabilidade $2pq$
ou A2/A2 com probabilidade $q^2$

## 1.2. 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.

<b>Les trois genotypes sont</b>	A1/A1	A1/A2	A2/A2
<b>Leurs fréquences sont égales à</b>	$p^2$	$2pq$	$q^2$

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.

*Équilibre de Hardy-Weinberg.* En effet, selon la formule de calcul des fréquences alléliques à partir des fréquences génotypiques, on peut obtenir :

$$f(A1) = p^2 + 2pq/2 = p^2 + pq = p(p + q) = p \quad (2.1)$$

$$f(A2) = q^2 + 2pq/2 = q^2 + pq = q(p + q) = q \quad (2.2)$$

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

## 1.2. A relação de Hardy-Weinberg

As novas frequências genóticas correspondem ao quadrado das frequências alélicas para os homocigotos, ou ao produto duplo das frequências alélicas para o heterocigoto.

<b>Os três genótipos são</b>	A1/A1	A1/A2	A2/A2
<b>As frequências dos genótipos são iguais a</b>	$p^2$	$2pq$	$q^2$

A relação assim estabelecida entre as frequências alélicas e as frequências genóticas é chamada relação de Hardy-Weinberg ou *relação panmítica*, dado que ela deriva diretamente da hipótese panmítica.

*Equilíbrio de Hardy-Weinberg.* Com efeito, de acordo com a fórmula de cálculo das frequências alélicas a partir das frequências genóticas, podemos obter:

$$f(A1) = p^2 + 2pq/2 = p^2 + pq = p(p + q) = p \quad (2.1)$$

$$f(A2) = q^2 + 2pq/2 = q^2 + pq = q(p + q) = q \quad (2.2)$$

Essa estabilidade da composição genética de uma população é chamada *equilíbrio de Hardy-Weinberg*.

▲ **Termos importantes**

- **Urne gamética** ("Urna gamética"): Conceito utilizado para representar a distribuição genotípica de uma população.
- **Panmixie** ("Panmixia"): Hipótese simplificadora segundo a qual cruzamentos ocorrem ao acaso.
- **Pangamie** ("Pangamia"): Hipótese simplificadora segundo a qual os gametas se unem ao acaso.

◆ **2. Mating frequencies and another proof of the Hardy-Weinberg law**

↳ *Referência 2/3 (Falconer e Mackay, 1989), pp. 13 – 14. Com modificações.*

Let us now look more closely into the breeding structure of a random-mating population, distinguishing the types of mating according to the genotypes of the pairs, and seeing what are the genotype frequencies among the progenies of the different types of mating. This provides a general method for relating genotype frequencies in successive generations. It also provides a proof of the Hardy-Weinberg law. The procedure is to obtain first the frequencies of all possible mating types according to the frequencies of the genotypes among the parents, and then to obtain the frequencies of genotypes according to the Mendelian ratios.

Consider a locus with two alleles, and let the frequencies of genes and genotypes in the parents be, as before:

	Genes		Genotypes		
	A <sub>1</sub>	A <sub>2</sub>	A <sub>1</sub> A <sub>1</sub>	A <sub>1</sub> A <sub>2</sub>	A <sub>2</sub> A <sub>2</sub>
Frequencies	<i>p</i>	<i>q</i>	<i>D</i>	<i>H</i>	<i>R</i>

There are altogether nine types of mating, and their frequencies when mating is random are found by multiplying together the marginal frequencies as shown in Table 1. Since the sex of the parent is irrelevant in this context, some of the types of mating are equivalent, and the number of different types reduces to six. By summation of the frequencies of equivalent types, we obtain the frequencies of mating types in the first two columns of Table 2. Now we have to consider the genotypes of offspring produced by each type of mating, and find the frequency of each genotype in the total progeny, assuming, of course, that all types of mating are equally fertile and all genotypes equally viable. This is done in the right-hand side of Table 2. Thus, for example, matings of the type A<sub>1</sub>A<sub>1</sub> × A<sub>1</sub>A<sub>1</sub> produce only A<sub>1</sub>A<sub>1</sub> offspring. So, of the total progeny, a proportion *D*<sup>2</sup> are A<sub>1</sub>A<sub>1</sub> genotypes derived from this type of mating. Similarly, one-quarter of the offspring of A<sub>1</sub>A<sub>2</sub> × A<sub>1</sub>A<sub>2</sub> matings are A<sub>1</sub>A<sub>1</sub>. So, this type of mating, which has a frequency of *H*<sup>2</sup>, contributes a proportion *H*<sup>2</sup>/4 of the total A<sub>1</sub>A<sub>1</sub> progeny. To find the frequency of each genotype in the total progeny we add the frequencies contributed by each type of mating. The sums, after

simplification, are given at the root of Table 2, and from the identities  $p = P + H/2$  and  $q = Q + H/2$ , they are seen to be equal to  $p^2$ ,  $2pq$ , and  $q^2$ . These are the Hardy-Weinberg equilibrium frequencies, and we have shown that they are attained by one generation of random mating, irrespective of the genotype frequencies among the parents.

**Table 1.**

			Genotype and freq. of female parent		
			$A_1A_1$	$A_1A_2$	$A_2A_2$
			$D$	$H$	$R$
Genotype and freq. of male parent	$A_1A_1$	$D$	$D^2$	$DH$	$DR$
	$A_1A_2$	$H$	$HD$	$H^2$	$HQ$
	$A_2A_2$	$R$	$PQ$	$HQ$	$Q^2$

**Table 2.**

Mating		Genotype and frequency of progeny		
Type	Frequency	$A_1A_1$	$A_1A_2$	$A_2A_2$
$A_1A_1 \times A_1A_1$	$P^2$	$P^2$	–	–
$A_1A_1 \times A_1A_2$	$2PH$	$PH$	–	–
$A_1A_1 \times A_2A_2$	$2PQ$	–	$2PQ$	–
$A_1A_2 \times A_1A_2$	$H^2$	$H^2/4$	$H^2/2$	$H^2/4$
$A_1A_2 \times A_2A_2$	$2HQ$	–	$HQ$	$HQ$
$A_2A_2 \times A_2A_2$	$Q^2$	–	–	$Q^2$
	<b>Sum</b>	$(P + H/2)^2$	$2(P + H/2)(Q + H/2)$	$(Q + H/2)^2$
	<b>=</b>	$p^2$	$2pq$	$q^2$

## 2. Frequências de cruzamento e outra prova da lei de Hardy-Weinberg

Podemos agora estudar mais a fundo a estrutura de cruzamentos de uma população sob cruzamentos aleatórios, distinguindo os tipos de cruzamento de acordo com os genótipos dos pares e verificando quais são as frequências genotípicas entre as progenias dos diferentes tipos de cruzamento. Tal abordagem fornece um método geral para relacionar frequências genotípicas em gerações sucessivas, e, ademais, resulta em uma prova da lei de Hardy-Weinberg. O procedimento envolve, primeiramente, obter as frequências de todos os tipos de cruzamento de acordo com as frequências genotípicas entre os pais, e, em seguida, o cálculo das frequências dos genótipos seguindo suas respectivas razões mendelianas.

Considere um *locus* com dois alelos, e permita que as frequências dos genes e genótipos dos pais sejam, como vimos na seção 2.1:

	Genes		Genótipos		
	$A_1$	$A_2$	$A_1A_1$	$A_1A_2$	$A_2A_2$
<b>Frequências</b>	$p$	$q$	$D$	$H$	$R$

Há ao todo 9 tipos de cruzamento, e suas frequências quando dos cruzamentos são obtidas multiplicando as frequências marginais listadas na Tabela 1. Uma vez que o sexo do parente é irrelevante nesse contexto, alguns tipos de cruzamento são equivalentes, e o número de tipos é reduzido para 6. Somando as frequências de tipos equivalentes, obtemos as frequências dos tipos de cruzamentos listadas nas duas primeiras colunas da Tabela 2. Temos então de considerar os genótipos da prole obtida com cada tipo de cruzamento e encontrar a frequência de cada genótipo na progenia total, supondo, evidentemente, que todos os tipos de cruzamento são igualmente férteis e todos os genótipos são igualmente viáveis. Isso é realizado no lado direito da Tabela 2. Portanto, temos que, por exemplo, cruzamentos do tipo  $A_1A_1 \times A_1A_1$  produzem uma prole do tipo  $A_1A_1$  apenas. Dentre a progenia total, uma proporção  $D^2$  são genótipos  $A_1A_1$  derivados desse tipo de cruzamento. Então, esse tipo de cruzamento, que tem frequência  $H^2$ , contribui em proporção  $H^2/4$  para a progenia total  $A_1A_1$ . Somamos as frequências de cada tipo de cruzamento para obter a frequência de cada genótipo na progenia total. Tais somas são apresentadas ao fim da Tabela 2, e, sabendo que  $p = P + H/2$  e  $q = Q + H/2$ , resultam em  $p^2$ ,  $2pq$  e  $q^2$ . Essas são as frequências de um equilíbrio de Hardy-Weinberg, as quais, como acabamos de mostrar, são obtidas após uma geração de cruzamentos aleatórios, independentemente das frequências genotípicas dos pais.

**Tabela 1.**

			Genótipo e freq. da mãe		
			$A_1A_1$	$A_1A_2$	$A_2A_2$
			$D$	$H$	$R$
Genótipo e freq. do pai	$A_1A_1$	$D$	$D^2$	$DH$	$DR$
	$A_1A_2$	$H$	$HD$	$H^2$	$HQ$
	$A_2A_2$	$R$	$PQ$	$HQ$	$Q^2$

**Tabela 2.**

Cruzamento		Genótipo e frequência da progenia		
Tipo	Frequência	$A_1A_1$	$A_1A_2$	$A_2A_2$
$A_1A_1 \times A_1A_1$	$P^2$	$P^2$	–	–
$A_1A_1 \times A_1A_2$	$2PH$	$PH$	–	–
$A_1A_1 \times A_2A_2$	$2PQ$	–	$2PQ$	–
$A_1A_2 \times A_1A_2$	$H^2$	$H^2/4$	$H^2/2$	$H^2/4$
$A_1A_2 \times A_2A_2$	$2HQ$	–	$HQ$	$HQ$
$A_2A_2 \times A_2A_2$	$Q^2$	–	–	$Q^2$
	<b>Soma</b>	$(P + H/2)^2$	$2(P + H/2)(Q + H/2)$	$(Q + H/2)^2$
	<b>=</b>	$p^2$	$2pq$	$q^2$



### 3. More than one locus: Gametic phase equilibrium

↳ Referência 2/3 (Falconer e Mackay, 1989), pp. 18 – 20. Com modificações.

The attainment of the [Hardy-Weinberg] equilibrium in genotype frequencies after one generation of random mating is true of all autosomal loci considered separately. But it is not true of the genotypes with respect to two or more loci considered jointly. To illustrate the point, suppose there were two populations, one consisting entirely of  $A_1A_1B_1B_1$  genotypes and the other entirely of  $A_2A_2B_2B_2$  genotypes. Suppose that these two populations were mixed, with equal numbers of each sex, and allowed to mate at random. With two alleles at each of two loci there are nine possible genotypes, but only three of these would appear in the first-generation progeny, the two original double homozygotes and the double heterozygote. There would be complete association between the traits determined by the two loci, and the two traits would appear to be determined by a single gene difference. With continued random mating the missing genotypes would appear in subsequent generations, but not immediately at their equilibrium frequencies, and the initial association between the traits would be progressively reduced. If the two loci were linked, the attainment of equilibrium frequencies would take longer because the appearance of the missing genotypes depends on recombination between the two loci. Disequilibrium with respect to two or more loci is called **gametic phase disequilibrium**, or **linkage disequilibrium**, irrespective of whether the loci are linked or not. Disequilibrium can arise from intermixture of populations with different gene frequencies, or from chance in small populations. Disequilibrium can also be produced, and maintained, by selection favoring one combination of alleles over another. The rate at which a random breeding population approaches equilibrium can be deduced as follows.

We first need a measure of the amount of disequilibrium. This is best expressed in terms of the frequencies of gametic types, rather than of zygotic genotypes. Consider two loci, each with two alleles, and gene frequencies as shown in Table 3.

**Table 3.**

<i>Genes</i>	$A_1$	$A_2$	$B_1$	$B_2$
<i>Gene frequencies</i>	$p_A$	$q_A$	$p_B$	$q_B$
<i>Gametic types</i>	$A_1B_1$	$A_1B_2$	$A_2B_1$	$A_2B_2$
<i>Frequencies (equilibrium)</i>	$p_A p_B$	$p_A q_B$	$q_A p_B$	$q_A q_B$
<i>Frequencies (actual)</i>	$r$	$s$	$t$	$u$
<i>Difference from equilibrium</i>	$+D$	$-D$	$-D$	$+D$

There are then four types of gamete. The population is in equilibrium if the gametes contain random combinations of the genes. The gametic frequencies at equilibrium therefore depend only on the gene frequencies, and are as shown in the table. Let the actual, nonequilibrium, frequencies be  $r$ ,  $s$ ,  $t$ , and  $u$ , as shown in the table. Each of these

differs from the equilibrium frequency by an amount  $D$ , two gametic types having a positive, and two a negative, deviation. The value of  $D$  for each gametic type is necessarily the same, except for the sign. The **amount of disequilibrium** is measured by  $D$ . The disequilibrium can be expressed by reference to genotypes by comparing the frequencies of coupling and of repulsion double heterozygotes. The genotype  $A_1B_1/A_2B_2$  can be called a **coupling heterozygote**, whether the two loci are linked or not. Its frequency is  $2ru$ . The **repulsion heterozygote** is  $A_1B_2/A_2B_1$  and its frequency is  $2st$ . If the population is in equilibrium, these two genotypes have equal frequencies. The relationship with  $D$  is

$$D = ru - st \quad (3)$$

Thus  $D$  is equal to half the difference in frequency between coupling and repulsion heterozygotes.

When a population in linkage disequilibrium mates at random, the amount of disequilibrium is progressively reduced with each succeeding generation. The rate at which this happens depends on the frequency of gametic types in two successive generations. This is perhaps easiest to visualize if the two loci are thought of as being linked on the same chromosome. The disequilibrium  $D$  in the progeny generation can be obtained from the frequency of any of the four gametic types, so let us consider only the  $A_1B_1$  type. This can appear in the progeny gametes in two ways. First, it can be produced as a non-recombinant from the genotype  $A_1B_1/A_xB_x$ , the subscript  $x$  meaning that either of the two alleles can be present. The frequency with which  $A_1B_1$  is produced in this way is  $r(1-c)$ ,  $r$  being the frequency of  $A_1B_1$  in the parental gametes and  $c$  the recombination frequency. Or, second, it can be produced as a recombinant from the genotype  $A_1B_x/A_xB_1$ . The frequency of the  $A_1B_x$  chromosome is  $p_A$  and that of the  $A_xB_1$  chromosome is  $p_B$ . So the frequency with which  $A_1B_1$  arises in this way is  $p_Ap_Bc$ . Therefore, the frequency of  $A_1B_1$  in the progeny gametes is

$$r' = r(1-c) + p_Ap_Bc \quad (4)$$

and the disequilibrium in the progeny generation is

$$D' = r' - p_Ap_B = r(1-c) - p_Ap_B(1-c)$$

$$\therefore D' = (r - p_Ap_B)(1-c)$$

$$\therefore D' = D(1-c) \quad (5)$$

If we take the process one generation further, we get

$$D'' = D'(1-c) = D(1-c)^2 \quad (6)$$

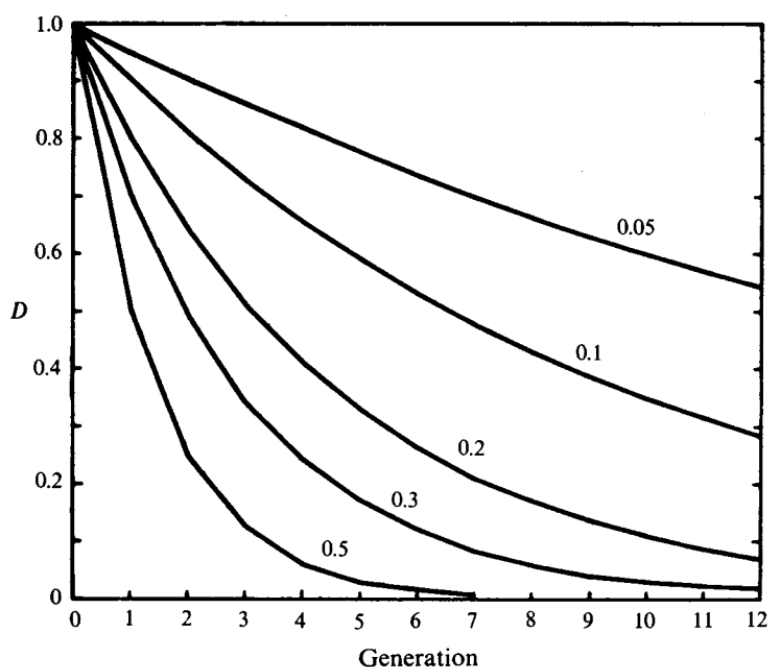
Thus, after any number  $t$  of generations, the disequilibrium is given by

$$D_t = D_0(1-c)^t \quad (7)$$

The loci do not have to be linked to be in disequilibrium. With unlinked loci  $c = 1/2$  and the amount of disequilibrium is halved by each generation of random mating. With linked loci the disequilibrium disappears more slowly. Figure 2 shows how the disequilibrium is reduced over 12 generations, with different degrees of linkage.

The approach to equilibrium given by the above equation applies equally to the disequilibrium of any number of loci considered jointly, provided  $(1 - c)$  is defined by the probability of a gamete passing through a generation without recombination between any of the loci. The larger the number of loci the smaller is the probability of no recombination; with two unlinked loci it is  $1/2$ , with three  $1/4$ , and with four  $1/8$ . Thus the multilocus disequilibrium decays faster than the 2-locus, which soon comes to dominate the total disequilibrium among a number of loci. A practical consequence of this is that when a number of loci are available for study, disequilibrium is more likely to be found with pairs of loci than with larger numbers considered jointly.

**Figure 2.** Approach to equilibrium under random mating of two loci, considered jointly. The graphs show the amount of disequilibrium,  $D$ , relative to the disequilibrium of generation 0. The five graphs refer to different degrees of linkage between the two loci, as indicated by the recombination frequency shown alongside each graph. The graph marked 0.5 refers to unlinked loci.



## 2. Mais de um locus: Desequilíbrio de fase gamética

A realização do equilíbrio de Hardy-Weinberg em termos de frequências genotípicas após uma geração de cruzamentos aleatórios aplica-se a todos os *loci* autossômicos considerados separadamente. No entanto, essa conclusão não é verdadeira para genótipos de dois ou mais *loci* tomados em conjunto. Para ilustrar essa observação, suponha a existência de duas populações, uma consistindo exclusivamente de indivíduos com

genótipo  $A_1A_1B_1B_1$  e a outra com genótipo  $A_2A_2B_2B_2$ . Suponha que as duas populações são misturadas com o mesmo número de indivíduos de cada sexo, e que estas podem cruzar aleatoriamente. Com dois alelos em cada um dos dois *loci* há nove genótipos possíveis, mas apenas três desses apareceriam na progenia de primeira geração, quais sejam, os dois homocigotos duplos originais e o heterocigoto duplo. Haveria associação completa entre as características determinadas pelos dois *loci*, e os dois fenótipos apareceriam como que determinados por uma única diferença genética. Com o progresso dos cruzamentos aleatórios, os genótipos ausentes apareceriam em gerações subsequentes, porém não em suas frequências de equilíbrio, e a associação inicial entre os fenótipos seria progressivamente reduzida. Ademais, se os dois *loci* fossem ligados, a realização das frequências de equilíbrio exigiria mais tempo porque o aparecimento dos genótipos ausentes dependeria da recombinação entre os dois *loci*. O desequilíbrio entre dois ou mais *loci* é denominado **desequilíbrio de fase gamética** ou **desequilíbrio de ligação**, e o termo se aplica independentemente de os *loci* serem ligados ou não. O referido desequilíbrio pode surgir com a mistura de populações dotadas de frequências genéticas distintas, ou ao acaso em pequenas populações. O desequilíbrio também pode ser produzido, ou mantido, por processos de seleção que favoreçam uma combinação de alelos em detrimento da outra. A taxa com a qual uma população sob cruzamentos aleatórios pode aproximar-se do equilíbrio pode ser deduzida da seguinte maneira. Primeiramente, precisamos de uma medida da quantidade de desequilíbrio. Para tanto, a melhor abordagem é exprimir a quantidade de desequilíbrio em termos das frequências de tipos gaméticos, ao invés dos genótipos zigóticos. Sejam dois *loci*, cada um com dois alelos e frequências genéticas distribuídas de acordo com a Tabela 3.

**Tabela 3.**

<i>Genes</i>	$A_1$	$A_2$	$B_1$	$B_2$
<i>Frequência do gene</i>	$p_A$	$q_A$	$p_B$	$q_B$
<i>Tipo gamético</i>	$A_1B_1$	$A_1B_2$	$A_2B_1$	$A_2B_2$
<i>Frequência (equilíbrio)</i>	$p_Ap_B$	$p_Aq_B$	$q_Ap_B$	$q_Aq_B$
<i>Frequência (atual)</i>	$r$	$s$	$t$	$u$
<i>Diferença em relação ao equilíbrio</i>	$+D$	$-D$	$-D$	$+D$

Há quatro tipos de gameta. Uma população estará em equilíbrio se os gametas contiverem combinações aleatórias dos genes. Portanto, as frequências gaméticas no equilíbrio dependem apenas das frequências de gene e estão listadas na tabela anterior. Como mostra a tabela, sejam  $r$ ,  $s$ ,  $t$  e  $u$  as frequências reais, as quais estão ainda fora de equilíbrio. Cada uma de tais frequências diverge da frequência de equilíbrio por uma quantidade  $D$ ; dois tipos de gameta apresentam desvio positivo,  $+D$ , e outros dois apresentam desvio negativo,  $-D$ . O valor de  $D$  para cada tipo gamético é necessariamente o mesmo, exceto

pelo sinal. Diz-se que a **quantidade de desequilíbrio** é medida por  $D$ . O desequilíbrio pode ser expresso com recurso aos genótipos pertinentes, comparando as frequências de heterozigotos duplos de associação e repulsão. O genótipo  $A_1B_1/A_2B_2$  pode ser denominado **heterozigoto de associação**, sejam os dois *loci* ligados ou não. Sua frequência é  $2ru$ . O **heterozigoto de repulsão** é  $A_1B_2/A_2B_1$  e sua frequência é  $2st$ . Se a população estiver em equilíbrio, os dois referidos genótipos terão frequências iguais. A relação de tais frequências com o parâmetro de quantidade de desequilíbrio  $D$  é

$$D = ru - st \quad (3)$$

Portanto,  $D$  é igual a metade da diferença entre a frequência entre heterozigotos de associação e heterozigotos de repulsão.

Quando uma população em desequilíbrio de ligação cruza aleatoriamente, a quantidade de desequilíbrio é reduzida com cada geração subsequente. A taxa com a qual isso ocorre depende da frequência de tipos gaméticos em duas gerações sucessivas. Isso é particularmente fácil de visualizar quando supomos que os dois *loci* estão ligados no mesmo cromossomo. O desequilíbrio  $D$  na geração de progenia pode ser obtido a partir da frequência de qualquer um dos quatro tipos de gameta, portanto consideramos apenas o tipo  $A_1B_1$ . Este pode aparecer nos gametas da progenia em duas formas. Primeiramente, este pode ser produzido em forma não-recombinante com genótipo  $A_1B_1/A_xB_x$ , onde o subscrito  $x$  significa que qualquer um dos dois alelos pode estar presente. A frequência com a qual  $A_1B_1$  é produzido dessa forma é  $(1 - c)r$ , onde  $r$  é a frequência de  $A_1B_1$  nos gametas parentais e  $c$  é a frequência de recombinação. Ou, em uma segunda hipótese, o gameta  $A_1B_1$  pode ser produzido como um recombinante oriundo do genótipo  $A_1B_x/A_xB_1$ . A frequência do cromossomo  $A_1B_x$  é  $p_A$  e aquela do cromossomo  $A_xB_1$  é  $p_B$ . Portanto, a frequência com a qual  $A_1B_1$  ocorre dessa maneira é  $p_A p_B c$ . Segue que a frequência de  $A_1B_1$  na progenia de gametas é

$$r' = r(1 - c) + p_A p_B c \quad (4)$$

e o desequilíbrio na geração de progenia é

$$D' = r' - p_A p_B = r(1 - c) - p_A p_B (1 - c)$$

$$\therefore D' = (r - p_A p_B)(1 - c)$$

$$\therefore D' = D(1 - c) \quad (5)$$

Se conduzirmos o processo por mais uma geração, obtemos

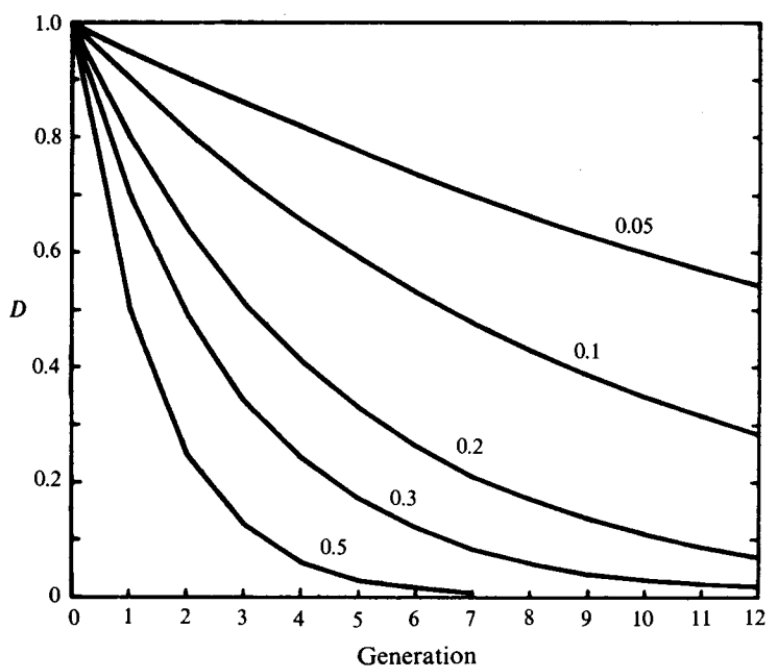
$$D'' = D'(1 - c) = D(1 - c)^2 \quad (6)$$

Consequentemente, após qualquer número  $t$  de gerações, o desequilíbrio é dado por

$$D_t = D_0(1 - c)^t \quad (7)$$

Os *loci* não precisam estar ligados para estar em desequilíbrio. Para *loci* não ligados,  $c = 1/2$  e a quantidade de desequilíbrio será reduzida pela metade em cada geração sucessiva de cruzamentos aleatórios. O desequilíbrio é reduzido mais lentamente com *loci* ligados. A Figura 2 mostra como o desequilíbrio é reduzido ao longo de 12 gerações para diferentes níveis de *linkage*.

**Figura 2.** Convergência ao equilíbrio para cruzamentos aleatórios de dois loci considerados em conjunto. Os gráficos mostram a quantidade de desequilíbrio,  $D$ , relativa ao desequilíbrio da geração 0. As cinco curvas referem-se a diferentes níveis de ligação entre os dois loci, como indica a frequência de recombinação ao lado de cada curva. A curva com indicação '0.5' refere-se a loci não ligados.



A abordagem de equilíbrio descrita pela equação anterior aplica-se igualmente ao desequilíbrio de qualquer número de *loci* considerados em conjunto, desde que  $(1 - c)$  seja definido pela probabilidade de um gameta passar por uma geração sem recombinação entre qualquer um dos *loci*. Quanto maior for o número de *loci*, menor será a probabilidade de não haver recombinação; com dois *loci* não ligados esta é  $1/2$ , com três é  $1/4$ , e com quatro é  $1/8$ . Portanto, o desequilíbrio 'multi-locus' decai mais rapidamente que o desequilíbrio 'duplo-locus', o qual logo passa a dominar o desequilíbrio total entre um número expressivo de *loci*. Uma consequência prática disso é que, quando há uma variedade de *loci* disponíveis para estudo, é mais fácil encontrar desequilíbrio entre pares de *loci* do que em números maiores considerados em conjunto.

▲ **Termos importantes**

■ **Gametic phase disequilibrium/linkage disequilibrium** (“Desequilíbrio de fase gamética/desequilíbrio de ligação”): Associação não-aleatória de alelos de diferentes *loci* em uma dada população. Diz-se que um desequilíbrio de ligação ocorrerá se a frequência dos alelos de dois (ou algum outro inteiro  $> 2$ ) *loci* sob estudo for diferente da frequência esperada de uma associação aleatória de tais alelos, sugerindo assim uma associação preferencial entre os alelos. Apesar de o termo ‘desequilíbrio de ligação’ sugerir a existência de *linkage* para a ocorrência de tais desvios, um GPD/LD não exige que os *loci* cromossômicos pertinentes estejam ligados.

■ **Amount of disequilibrium** (“Quantidade de desequilíbrio”): Parâmetro que fornece uma indicação quantitativa do desequilíbrio de ligação existente entre dois alelos. Este é por vezes denominado *coefficient of linkage disequilibrium* (“coeficiente de desequilíbrio de ligação”), embora Falconer e Mackay (1989) refiram-se a ele simplesmente como *amount of disequilibrium*.

■ **Coupling heterozygote/repulsion heterozygote** (“Heterozigoto de associação/heterozigoto de repulsão”): Duplo-heterozigotos utilizados como referências na análise da quantidade de desequilíbrio em uma situação de desequilíbrio de ligação.



A **Lotka** oferece uma variedade de serviços de tradução, revisão e composição.

[www.lotkatranslation.com](http://www.lotkatranslation.com)



#### **Tradução técnica**

Traduzimos artigos, manuscritos e outros gêneros textuais em 8 idiomas.



#### **Revisão técnica**

Melhoramos a qualidade gramatical e estilística do seu documento.



#### **Jargão e terminologia**

Dominamos as terminologias e jargões de diversas áreas de engenharia e ciências naturais.



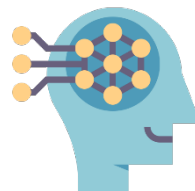
#### **Edição de referências bibliográficas**

Preparamos listas de referências em 4 padrões diferentes (ABNT, Harvard, APA, MLA).



#### **Edição de expressões matemáticas**

Preparamos as equações, reações químicas, tabelas e outros elementos especiais do seu documento.



#### **Revisão de documentos gerados por inteligência artificial**

Tecnologias de inteligência artificial generativa são ainda incipientes e propensas a erro. A Lotka pode melhorar a qualidade técnica de textos gerados ou traduzidos por IA.