Linkage disequilibrium due to random genetic drift*

By TOMOKO OHTA† AND MOTOO KIMURA

National Institute of Genetics, Mishima, Japan

1. INTRODUCTION

Linkage disequilibrium between two segregating loci may be built up in a population by two causes. One is epistatic interaction in fitness between the two loci and the other is random drift due to small population size. The former was studied by Kimura (1956), Lewontin & Kojima (1960), and many more recent investigators. They showed that at equilibrium permanent linkage disequilibrium may be maintained in an infinitely large population due to epistatic interaction in fitness. On the other hand, the latter, i.e. linkage disequilibrium due to random drift has not been investigated until quite recently. Hill & Robertson (1968) have shown its importance in small populations by demonstrating that the variance of disequilibrium coefficient D may become fairly large although its mean is zero.

In this report, some extension of their results will be presented with the use of an entirely different method of approach; Hill & Robertson used a moment generating matrix which was derived by taking expectations over the multinomial distribution of gametic frequencies. They obtained an exact analytical solution for the case of no recombination and no selection. Also, they made numerical studies which include some recombination and selection.

Here, we will make use of the method of the Kolmogorov backward equation as developed by Kimura (1957, 1962, 1964) and we will derive the analytical solution of the variance of linkage disequilibrium for the case of general value of recombination fraction when there is no selection.

2. BASIC THEORY AND RESULTS

Let us assume two loci in which pairs of alleles **A**, **a** and **B**, **b** are segregating and let the frequencies of the four types of gametes **AB**, **Ab**, **aB** and **ab** be respectively g_1, g_2, g_3 and g_4 $(g_1 + g_2 + g_3 + g_4 = 1)$. We will denote by N_e the variance effective number of the population as defined by Kimura & Crow (1963). Now, let

$$\phi \, (g_1, g_2, g_3, X_1, X_2, X_3; t)$$

be the probability density that the frequencies of **AB**, **Ab** and **aB** become respectively X_1, X_2 and X_3 at time t (measured with one generation as unit), given that

---

their initial frequencies are respectively g_1, g_2, g_3. Note here that we denote the relative frequencies of the four gamete types at the t-th generation by letter X with subscript 1, 2, 3 or 4 ($X_1 + X_2 + X_3 + X_4 = 1$).

Unless N_e is extremely small, ϕ satisfies the following Kolmogorov backward equation (1). An equivalent formula was given by Hill & Robertson (1966) in the form of the forward equation and by Ohta (1968) as the backward equation at steady state.

$$
\begin{aligned}
\frac{\partial \phi}{\partial T} ={}& \tfrac{1}{4}g_1(1-g_1)\frac{\partial^2 \phi}{\partial g_1^2} + \tfrac{1}{4}g_2(1-g_2)\frac{\partial^2 \phi}{\partial g_2^2} + \tfrac{1}{4}g_3(1-g_3)\frac{\partial^2 \phi}{\partial g_3^2} - \tfrac{1}{2}g_1 g_2 \frac{\partial^2 \phi}{\partial g_1 \partial g_2} - \tfrac{1}{2}g_1 g_3 \frac{\partial^2 \phi}{\partial g_1 \partial g_3} \\
& - \tfrac{1}{2}g_2 g_3 \frac{\partial^2 \phi}{\partial g_2 \partial g_3} + \tfrac{1}{2}N_e s_1 \left\{ g_1(1-g_1-g_2)\frac{\partial \phi}{\partial g_1} + g_2(1-g_1-g_2)\frac{\partial \phi}{\partial g_2} - g_3(g_1+g_2)\frac{\partial \phi}{\partial g_3} \right\} \\
& + \tfrac{1}{2}N_e s_2 \left\{ g_1(1-g_1-g_3)\frac{\partial \phi}{\partial g_1} - g_2(g_1+g_3)\frac{\partial \phi}{\partial g_2} + g_3(1-g_1-g_3)\frac{\partial \phi}{\partial g_3} \right\} \\
& - N_e c D \left\{ \frac{\partial \phi}{\partial g_1} - \frac{\partial \phi}{\partial g_2} - \frac{\partial \phi}{\partial g_3} \right\}.
\end{aligned}
\tag{1}
$$

In the above equation, $T = t/N_e$ and c is the recombination fraction between the two loci (A and B). Also, s_1 and s_2 are the additive selection coefficients at these loci, each defined as the difference in fitness between the two homozygotes at A and B loci respectively.

In order to treat equation (1), we will transform the independent random variables from g_1, g_2 and g_3 into p, q and D such that

$$
\begin{aligned}
p &= g_1 + g_2, \\
q &= g_1 + g_3, \\
D &= g_1 g_4 - g_2 g_3 = g_1 - g_1^2 - g_1 g_2 - g_1 g_3 - g_2 g_3,
\end{aligned}
$$

where p is the frequency of gene $\mathbf{A}$, q is the frequency of gene $\mathbf{B}$ and D is the measure of linkage disequilibrium, all defined at $t = 0$. Then, equation (1) becomes as follows:

$$
\begin{aligned}
\frac{\partial \phi}{\partial T} ={}& \tfrac{1}{4}p(1-p)\frac{\partial^2 \phi}{\partial p^2} + \tfrac{1}{4}q(1-q)\frac{\partial^2 \phi}{\partial q^2} + \tfrac{1}{2}D\frac{\partial^2 \phi}{\partial q \partial p} + \tfrac{1}{2}D(1-2p)\frac{\partial^2 \phi}{\partial p \partial D} + \tfrac{1}{2}D(1-2q)\frac{\partial^2 \phi}{\partial q \partial D} \\
& + \tfrac{1}{4}\{ pq(1-p)(1-q) + D(1-2p)(1-2q) - D^2 \}\frac{\partial^2 \phi}{\partial D^2} \\
& + \tfrac{1}{2}\{ N_e s_1 p(1-p) + N_e s_2 D \}\frac{\partial \phi}{\partial p} + \tfrac{1}{2}\{ N_e s_2 q(1-q) + N_e s_1 D \}\frac{\partial \phi}{\partial q} \\
& + \tfrac{1}{2}\{ N_e s_1 D(1-2p) + N_e s_2 D(1-2q) - D(1+2N_e c) \}\frac{\partial \phi}{\partial D}.
\end{aligned}
\tag{2}
$$

In the above equation $\phi \equiv \phi(p, q, D, x, y, z; T)$ stands for the probability density that the gene frequencies and the disequilibrium coefficient become x, y and z at time T given that they are respectively p, q and D at $T = 0$. In the following, we will use letter z to represent the linkage disequilibrium at time T so that

$$
z = X_1 X_4 - X_2 X_3 = X_1(1 - X_1 - X_2 - X_3) - X_2 X_3.
$$

Note also that

$$X_1 = xy + z, \qquad X_3 = (1-x)y - z,$$
$$X_2 = x(1-y) - z, \quad X_4 = (1-x)(1-y) + z.$$

Now, let us define three functions as follows:

$$H = \iiint_O xy(1-x)(1-y)\,\phi\,dX_1\,dX_2\,dX_3, \tag{3}$$

$$I = \iiint_O z(1-2x)(1-2y)\,\phi\,dX_1\,dX_2\,dX_3, \tag{4}$$

$$J = \iiint_O z^2\phi\,dX_1\,dX_2\,dX_3. \tag{5}$$

Here the triple integrals are over the open interval O for which

$$0 < X_1 < X_1 + X_2 < X_1 + X_2 + X_3 < 1.$$

Notice here that H, I and J represent the expectations of

$$xy(1-x)(1-y), \quad z(1-2x)(1-2y) \quad \text{and} \quad z^2$$

respectively at time T. These three functions satisfy the same type of equation as (2), as may be seen, for example, by multiplying each of the terms of equation (2) by z^2 followed by integrating them over O thus changing ϕ into J.

Next, we seek the solution of equation (2) for H, I and J for the case of no selection, i.e. $s_1 = s_2 = 0$. Let

$$H = [A_H\,pq(1-p)(1-q) + B_H(1-2p)(1-2q)\,D + C_H\,D^2]\,e^{\lambda T}, \tag{6}$$

$$I = [A_I\,pq(1-p)(1-q) + B_I(1-2p)(1-2q)\,D + C_I\,D^2]\,e^{\lambda T}, \tag{7}$$

$$J = [A_J\,pq(1-p)(1-q) + B_J(1-2p)(1-2q)\,D + C_J\,D^2]\,e^{\lambda T}, \tag{8}$$

where, A's, B's, C's and λ are constants.

If we substitute H for ϕ in equation (2) in which we assume $s_1 = s_2 = 0$, right-hand side becomes

$$[pq(1-p)(1-q)\{-A_H + \tfrac{1}{2}C_H\} + D(1-2p)(1-2q)\{\tfrac{1}{2}A_H - B_H(\tfrac{5}{2} + R) + \tfrac{1}{2}C_H\} + D^2\{2B_H - C_H(\tfrac{3}{2} + 2R)\}]\,e^{\lambda T}, \tag{9}$$

where $R = N_e c$. The left-hand side of equation (2) becomes

$$\lambda[pq(1-p)(1-q)\,A_H + D(1-2p)(1-2q)\,B_H + D^2 C_H]\,e^{\lambda T}. \tag{10}$$

By equating the two expressions (9) and (10), we have

$$\left. \begin{array}{r} -A_H + \tfrac{1}{2}C_H = \lambda A_H, \\ \tfrac{1}{2}A_H - B_H(\tfrac{5}{2} + R) + \tfrac{1}{2}C_H = \lambda B_H, \\ 2B_H - C_H(\tfrac{3}{2} + 2R) = \lambda C_H. \end{array} \right\} \tag{11}$$

From the above relation (11), we obtain the following cubic equation (12) which gives the three eigenvalues $(\lambda_1, \lambda_2, \lambda_3)$ for H. I and J have the same eigenvalues.

$$\lambda^3 + (5 + 3R)\lambda^2 + \left(\frac{27 + 38R + 8R^2}{4}\right)\lambda + \frac{9 + 26R + 8R^2}{4} = 0. \tag{12}$$

In order to solve (12), we transform the equation into the following form:

$$Y^3 - \tfrac{1}{12}(19 + 6R + 12R^2)\,Y + \tfrac{1}{108}(28 + 63R - 90R^2) = 0, \tag{13}$$

where $Y = \lambda + \tfrac{1}{3}(5 + 3R)$. Then, the three solutions are,

$$
\left.
\begin{aligned}
Y_1 &= \tfrac{1}{3}\sqrt{(19 + 6R + 12R^2)}\cos\tfrac{1}{3}\theta, \\
Y_2 &= \tfrac{1}{3}\sqrt{(19 + 6R + 12R^2)}\cos\tfrac{1}{3}(\theta + 4\pi), \\
Y_3 &= \tfrac{1}{3}\sqrt{(19 + 6R + 12R^2)}\cos\tfrac{1}{3}(\theta + 2\pi),
\end{aligned}
\right\}
\tag{14}
$$

where,

$$\theta = \cos^{-1}\left\{-\frac{28 + 63R - 90R^2}{(19 + 6R + 12R^2)^{\frac{3}{2}}}\right\}.$$

In terms of these solutions, the three eigenvalues $(0 > \lambda_1 > \lambda_2 > \lambda_3)$ are,

$$
\left.
\begin{aligned}
\lambda_1 &= Y_1 - R - \tfrac{5}{3}, \\
\lambda_2 &= Y_2 - R - \tfrac{5}{3}, \\
\lambda_3 &= Y_3 - R - \tfrac{5}{3}.
\end{aligned}
\right\}
\tag{15}
$$

The values of λ_1, λ_2 and λ_3 are all negative and their absolute values are tabulated in Table 1 for various values of $R = N_e c$ from 0 to 5. Figure 1 shows the relationship between λ's and R.

Table 1. *The absolute values of three eigenvalues $(\lambda_1, \lambda_2, \lambda_3)$*
for various values of $N_e c$ from $N_e c = 0$ to $N_e c = 5$

$R = N_e c$	$|\lambda_1|$	$|\lambda_2|$	$|\lambda_3|$
0·0	0·500	1·500	3·000
0·2	0·668	1·663	3·269
0·4	0·776	1·878	3·545
0·6	0·843	2·128	3·829
0·8	0·885	2·393	4·122
1·0	0·912	2·664	4·424
2·0	0·967	3·960	6·073
3·0	0·982	5·129	7·888
4·0	0·989	6·224	9·787
5·0	0·993	7·282	11·726

Now let us obtain the constants, A's, B's and C's, in the formulas (6), (7) and (8). They can be determined by the initial conditions as follows: from the relation (11), we have

$$
\left.
\begin{aligned}
A_{Hi} &= \frac{C_{Hi}}{2(1 + \lambda_i)} \\
B_{Hi} &= \left(\frac{3 + 4R + 2\lambda_i}{4}\right)C_{Hi}
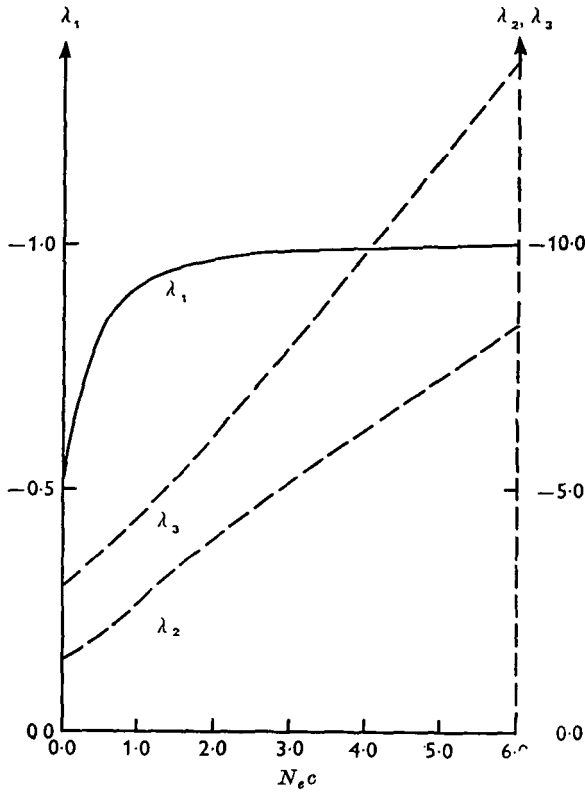\end{aligned}
\right\}
\quad (i = 1, 2, 3),
\tag{16}
$$

Fig. 1. The relationship between $N_e c$ and three eigenvalues (λ_1, λ_2, λ_3). The value of λ_1 (solid line) is given by the scale in the left ordinate, while the values of λ_2 and λ_3 (broken lines) are given by the scale in the right ordinate.

where the subscript i indicates that the subscripted quantity corresponds to eigenvalue λ_i. Therefore, the following relations hold for H, I and J.

$$\left. \begin{aligned} &\sum_{i=1}^{3} \frac{C_{Hi}}{2(1+\lambda_i)} = 1, \\ &\sum_{i=1}^{3} \left(\frac{3+4R+2\lambda_i}{4} \right) C_{Hi} = 0, \\ &\sum_{i=1}^{3} C_{Hi} = 0; \end{aligned} \right\} \tag{17}$$

$$\left. \begin{aligned} &\sum_{i=1}^{3} \frac{C_{Ii}}{2(1+\lambda_i)} = 0, \\ &\sum_{i=1}^{3} \left(\frac{3+4R+2\lambda_i}{4} \right) C_{Ii} = 1, \\ &\sum_{i=1}^{3} C_{Ii} = 0; \end{aligned} \right\} \tag{18}$$

4-2

$$\sum_{i=1}^{3} \frac{C_{Ji}}{2(1+\lambda_i)} = 0,$$

$$\sum_{i=1}^{3} \left(\frac{3+4R+2\lambda_i}{4}\right) C_{Ji} = 0,$$

$$\sum_{i=1}^{3} C_{Ji} = 1.$$

(19)

From the above relations, (17), (18) and (19), we get

$$C_{H1} = \frac{2(1+\lambda_1)(1+\lambda_2)(1+\lambda_3)}{(\lambda_3-\lambda_1)(\lambda_2-\lambda_1)},$$

$$C_{H2} = \frac{2(1+\lambda_1)(1+\lambda_2)(1+\lambda_3)}{(\lambda_1-\lambda_2)(\lambda_3-\lambda_2)},$$

$$C_{H3} = \frac{2(1+\lambda_1)(1+\lambda_2)(1+\lambda_3)}{(\lambda_1-\lambda_3)(\lambda_2-\lambda_3)};$$

(20)

$$C_{I1} = \frac{2(1+\lambda_1)}{(\lambda_2-\lambda_1)(\lambda_3-\lambda_1)},$$

$$C_{I2} = \frac{2(1+\lambda_2)}{(\lambda_1-\lambda_2)(\lambda_3-\lambda_2)},$$

$$C_{I3} = \frac{2(1+\lambda_3)}{(\lambda_1-\lambda_3)(\lambda_2-\lambda_3)};$$

(21)

$$C_{J1} = -\frac{1+\lambda_1}{(\lambda_2-\lambda_1)(\lambda_3-\lambda_1)}(\lambda_2+\lambda_3+2R+\tfrac{5}{2}),$$

$$C_{J2} = -\frac{1+\lambda_2}{(\lambda_1-\lambda_2)(\lambda_3-\lambda_2)}(\lambda_1+\lambda_3+2R+\tfrac{5}{2}),$$

$$C_{J3} = -\frac{1+\lambda_3}{(\lambda_1-\lambda_3)(\lambda_2-\lambda_3)}(\lambda_1+\lambda_2+2R+\tfrac{5}{2}).$$

(22)

Finally, the general solutions for H, I and J are

$$H_t = \sum_{i=1}^{3} C_{Hi} \left[\frac{pq(1-p)(1-q)}{2(1+\lambda_i)} + \tfrac{1}{4}(3+4R+2\lambda_i)D(1-2p)(1-2q) + D^2\right] \exp\{\lambda_i t/N_e\},$$

(23)

$$I_t = \sum_{i=1}^{3} C_{Ii} \left[\frac{pq(1-p)(1-q)}{2(1+\lambda_i)} + \tfrac{1}{4}(3+4R+2\lambda_i)D(1-2p)(1-2q) + D^2\right] \exp\{\lambda_i t/N_e\},$$

(24)

$$J_t = \sum_{i=1}^{3} C_{Ji} \left[\frac{pq(1-p)(1-q)}{2(1+\lambda_i)} + \tfrac{1}{4}(3+4R+2\lambda_i)D(1-2p)(1-2q) + D^2\right] \exp\{\lambda_i t/N_e\},$$

(25)

where H_t, I_t and J_t represent the values of H, I and J at the t-th generation. In the

special case of $R = 0$ (no crossing over) and $D = 0$ (initial linkage equilibrium), H_t and J_t are reduced as follows:

$$H_t = \tfrac{1}{15}pq(1-p)(1-q)\,[6\exp\{-t/(2N_e)\} + 10\exp\{-3t/(2N_e)\} - \exp\{-3t/N_e\}], \quad (26)$$

$$J_t = \tfrac{1}{15}pq(1-p)(1-q)\,[6\exp\{-t/(2N_e)\} - 5\exp\{-3t/(2N_e)\} - \exp\{-3t/N_e\}]. \quad (27)$$

Thus, our results agree with those of Hill & Robertson (1968).

3. IMPLICATIONS OF THE RESULTS

The expected value of linkage disequilibrium (z in our terminology) has been worked out for two loci small populations for the case of no selection by Kimura (1963), Hill & Robertson (1966), and Karlin & McGregor (1968). If we denote the expected value of z at the t-th generation by D_t such that

$$D_t = \iiint_O z\phi\,dX_1\,dX_2\,dX_3,$$

then, with the continuous time model that we are using,

$$D_t = \exp\{-(2N_e c + 1)\,t/(2N_e)\}\,D_0, \quad (28)$$

where $D_0 = D$. However, the distribution of z is unknown. Even its variance has not been studied until very recently. As Hill & Robertson (1968) emphasize, the actually observed value of disequilibrium may be quite different from the expected value because of random drift. The mean of disequilibrium is permanently zero when there is no epistasis and if there is initial linkage equilibrium, but its variance may get quite large and may give large observed values of disequilibrium when the population size is small.

General expressions for the expected values of $xy(1-x)(1-y)$, $z(1-2x)(1-2y)$ and z^2 are given by (23), (24) and (25) respectively. Numerical examples of the expected values of $z^2/[pq(1-p)(1-q)]$ and $z^2/[xy(1-x)(1-y)]$ in our terminology were given by Hill & Robertson (1968) and we only emphasize that for small values of $N_e c$ the effect of random drift is quite important for linkage disequilibrium. Among their two statistics, the latter, i.e. the expectation of $z^2/[xy(1-x)(1-y)]$, is quite interesting, although its general solution was not obtained here. However, J_t/H_t, which is the ratio of the expectations, seems to be not much different from the expectation of the ratio. Namely, the asymptotic value of J_t/H_t is equal to C_{J1}/C_{H1} or $-(\lambda_2 + \lambda_3 + 2R + 2\cdot 5)/[2(1+\lambda_2)(1+\lambda_3)]$ and this turns out to be roughly equal to $1/(4N_e c)$ when $N_e c$ is large, as can be inferred from Table 1.

Of the three eigenvalues given in Table 1, the one with the smallest absolute value (λ_1) is of special importance. It gives the asymptotic rate of decrease of H, I and J when t gets sufficiently large. That is,

$$H_t \sim C_{H1}\left[\frac{pq(1-p)(1-q)}{2(1+\lambda_1)} + \tfrac{1}{4}(3+4R+2\lambda_1)D(1-2p)(1-2q) + D^2\right]\exp\{\lambda_1 t/N_e\}$$
$$(t \to \infty), \quad (29)$$

$$I_t \sim C_{I1}\left[\frac{pq(1-p)(1-q)}{2(1+\lambda_1)} + \tfrac{1}{4}(3+4R+2\lambda_1)\,D(1-2p)(1-2q) + D^2\right]\exp\{\lambda_1 t/N_e\}$$
$$(t \to \infty), \quad (30)$$

$$J_t \sim C_{J1}\left[\frac{pq(1-p)(1-q)}{2(1+\lambda_1)} + \tfrac{1}{4}(3+4R+2\lambda_1)\,D(1-2p)(1-2q) + D^2\right]\exp\{\lambda_1 t/N_e\}$$
$$(t \to \infty). \quad (31)$$

It is well known that the heterozygosity at a single locus decreases exactly at the rate of $1/(2N_e)$ per generation. The above three formulas show that the simultaneous heterozygosity at the two loci decreases asymptotically at the rate of $|\lambda_1|/N_e$ per generation, because the expected frequency of double heterozygotes or the expected value of $2X_1X_4 + 2X_2X_3$ at the t-th generation is equal to

$$4H_t + 2I_t + 4J_t.$$

Furthermore, the expected frequencies of coupling and repulsion double heterozygotes may be obtained separately by noting that the expected value of

$$2X_1X_4 - 2X_2X_3$$

is equal to $2D_t$ by definition. Thus the expected frequencies of coupling (**AB/ab**) and repulsion (**Ab/aB**) double heterozygotes are

$$2H_t + I_t + 2J_t + D_t$$

and

$$2H_t + I_t + 2J_t - D_t$$

respectively.

When $N_e c = 0$, the rate of decrease of the frequency of double heterozygotes is $1/(2N_e)$. It gets larger with larger $N_e c$ as shown in Fig. 1 and finally it becomes $1/N_e$ when $N_e c = \infty$. The figure shows how two loci interact with each other by linkage.

It should be noted here that this rate also gives the asymptotic rate of the loss of genetic heterogeneity in one of the loci when the two loci are segregating simultaneously. Now, let us compare the situation with the single locus case again. Kimura (1955) used letter Ω_t to stand for the probability that a particular locus is heterallelic and showed that

$$\Omega_t = \int_0^1 \phi(p, x; t)\, dx \sim 6p(1-p)\exp\{-t/(2N_e)\} \quad (t \to \infty). \quad (32)$$

This means that for large t the probability that two alleles coexist in the population decreases asymptotically at the rate of $1/(2N_e)$, although the heterozygosity decreases always at this rate. The rate $1/(2N_e)$ corresponds to the smallest eigenvalue of the Kolmogorov equation for the gene frequency distribution at a single locus. It is also called the rate of steady decay (cf. Kimura, 1955). For two segregating loci, the distribution function of gametic frequencies has not been obtained at present, but three eigenvalues of the corresponding Kolmogorov equation were obtained. Of these, the one with the smallest absolute value should give the asymptotic rate of the loss of one of the two pairs of alleles in the two loci. That is, the probability

that the four different gametic types coexist in the population decreases asymptotically with the rate $|\lambda_1|/N_e$ when t gets sufficiently large so that

$$\Omega_t \equiv \iiint_O \phi\, dX_1\, dX_2\, dX_3 \sim Q \exp\{\lambda_1 t/N_e\} \quad (t \to \infty), \tag{33}$$

where Q is a function of the initial gamete frequencies. Karlin & McGregor (1968) obtained the value of this rate by a different approach. Namely, they used the method of Markov chain with a model of discrete generations and listed the values for various combinations of $2N_e$ and c. When N_e is small, their results may be more exact than ours, but for $2N_e > 10$ the results given here may be sufficiently accurate, as the comparison of our results with theirs shows.

SUMMARY

The behaviour of linkage disequilibrium between two segregating loci in finite populations has been studied as a continuous stochastic process for different intensity of linkage, assuming no selection. By the method of the Kolmogorov backward equation, the expected values of the square of linkage disequilibrium z^2, and other two quantities, $xy(1-x)(1-y)$ and $z(1-2x)(1-2y)$, were obtained in terms of T, the time measured in N_e as unit, and R, the product of recombination fraction (c) and effective population number (N_e). The rate of decrease of the simultaneous heterozygosity at two loci and also the asymptotic rate of decrease of the probability for the coexistence of four gamete types within a population were determined. The eigenvalues λ_1, λ_2 and λ_3 related to the stochastic process are tabulated for various values of $R = N_e c$.

REFERENCES

HILL, W. G. & ROBERTSON, A. (1966). The effect of linkage on limits to artificial selection. *Genet. Res.* 8, 269–294.

HILL, W. G. & ROBERTSON, A. (1968). Linkage disequilibrium in finite populations. *Theor. Appl. Genet.* 38, 226–231.

KARLIN, S. & McGREGOR, J. (1968). Rates and probabilities of fixation for two locus random mating finite populations without selection. *Genetics* 58, 141–159.

KIMURA, M. (1955). Solution of a process of random genetic drift with a continuous model. *Proc. natn. Acad. Sci. U.S.A.* 41, 144–150.

KIMURA, M. (1956). A model of a genetic system which leads to closer linkage by natural selection. *Evolution* 10, 278–287.

KIMURA, M. (1957). Some problems of stochastic processes in genetics. *Ann. Math. Statist.* 28, 882–901.

KIMURA, M. (1962). On the probability of fixation of mutant genes in a population. *Genetics* 47, 713–719.

KIMURA, M. (1963). A probability method for treating inbreeding systems, especially with linked genes. *Biometrics* 19, 1–17.

KIMURA, M. (1964). Diffusion models in population genetics. *Jour. Applied Probability*, 1, 177–232.

KIMURA, M. & CROW, J. F. (1963). The measurement of effective population number. *Evolution* 17, 279–288.

LEWONTIN, R. C. & KOJIMA, K. (1960). The evolutionary dynamics of complex polymorphisms. *Evolution* 14, 458–472.

OHTA, T. (1968). Effect of initial linkage disequilibrium and epistasis on fixation probability in a small population, with two segregating loci. *Theor. Appl. Genet.* 38, 243-248.