EXPECTED LINKAGE DISEQUILIBRIUM FOR A NEUTRAL LOCUS LINKED TO A CHROMOSOMAL ARRANGEMENT

CURTIS STROBECK

Department of Genetics, University of Alberta, Edmonton, Alberta, Canada

Manuscript received May 10, 1982
Revised copy accepted November 23, 1982

ABSTRACT

The expected value of the squared linkage disequilibrium is derived for a neutral locus associated with a chromosomal arrangement that is maintained in the population by strong balancing selection. For a given value of recombination, the expected squared linkage disequilibrium is shown to decrease as the intensity of selection maintaining the arrangement increases. The transient behavior of the expected square linkage disequilibrium is also derived. This theory applies to loci that are closely linked to inversions in Drosophila species and to loci closely linked to the differential segments of the translocation complexes in ring-forming species of Oenothera. In both cases the strong linkage disequilibria that have been observed in natural populations can be explained by random drift.

A large number of studies have been done to determine the extent to which significant linkage disequilibrium exists in natural populations (HEDRICK, JAIN and HOLDEN 1978). The studies in Drosophila species have generally found little or no linkage disequilibria between loci, but significant linkage disequilibria between inversions and loci closely linked to the inversions are consistently observed (PRAKASH and LEWONTIN 1968; KOJIMA, GILLESPIE and TOBARI 1970; LANGLEY, TOBARI and KOJIMA 1974; PRAKASH 1976; CHARLESWORTH, CHARLESWORTH and LOUKAS, 1977; CHARLESWORTH *et al.* 1979). These observed linkage disequilibria between inversions and closely linked loci have been explained as due to selection of loci associated with the inversions (PRAKASH and LEWONTIN 1968; LEWONTIN 1974), to association of neutral alleles with inversions when they were initially formed (NEI and LI 1975, 1980; ISHII and CHARLESWORTH 1977) or to random drift of neutral alleles at a locus associated with an inversion (NEI and LI 1975). It is this last explanation that is examined in this paper.

There are two aspects of the behavior of inversions in natural populations of Drosophila that need to be taken into account in any meaningful theory of random drift of neutral alleles at a locus that is closely linked to an inversion. First, since recombination between genes within an inversion is suppressed, inversions are inherited as a block. Second, inversion polymorphism is widespread and is maintained by strong balancing selection (DOBZHANSKY 1951).

This observed heterosis of inversion heterozygotes is thought to arise because sets of co-adapted genes are bound together by the inversions.

Strong linkage disequilibria have also been observed in the Oenothera species which are permanent translocation heterozygotes (LEVY and WINTERNHEIMER 1977; ELLSTRAND and LEVIN 1980). Permanent translocation heterozygosity in these species is maintained because there is alternate segregation of the chromosomes in the ring formed at meiosis and there exists a system of balanced lethals (CLELAND 1972). The balanced lethals are either zygotic, which results in 50% seed sterility, or gametophytic, in which the α complex is inherited through the egg and the β complex through the pollen. There is no recombination in the central segments of the chromosomes, which contain the balanced lethals, although recombination does occur in the paired regions at the ends of the chromosomes. This lack of recombination and alternate segregation of the chromosomes results in blocks of genes being inherited as a single unit (JENNER complexes) (CLELAND 1972). Therefore, the theory of random drift at a neutral locus closely linked to these complexes requires the same two assumptions as the theory of random drift at a neutral locus closely linked to an inversion, *i.e.*, the chromosomal arrangements are inherited as blocks and are maintained in the population by strong balancing selection.

Recurrence equations for the expected squared linkage disequilibrium of a neutral locus closely linked to a chromosomal arrangement are obtained using coefficients of identity. The equilibrium value assuming the infinite allele model (KIMURA and CROW 1964) and the transient behavior are obtained. It is shown that, as the intensity of selection maintaining the polymorphism increases, the expected squared linkage disequilibrium decreases.

THEORY

Let $\mathbf{C}_1$ and $\mathbf{C}_2$ be two chromosomal arrangements differing by an inversion that are maintained at equilibrium in the population by a heterozygotic advantage. Let N be the total population size and N_{11}, N_{12} and N_{22} be the numbers of $\mathbf{C}_1\mathbf{C}_1$, $\mathbf{C}_1\mathbf{C}_2$ and $\mathbf{C}_2\mathbf{C}_2$ individuals in the population *after selection*. The number of $\mathbf{C}_1$ and $\mathbf{C}_2$ chromosomes in the population *after selection* are

$$N_1 = 2N_{11} + N_{12} \quad \text{and} \quad N_2 = 2N_{22} + N_{12} \tag{1}$$

and the proportions of the $\mathbf{C}_1$ chromosomes in $\mathbf{C}_1\mathbf{C}_1$ homozygotes and $\mathbf{C}_1\mathbf{C}_2$ heterozygotes *after selection* are

$$P_1 = 2N_{11}/N_1 \quad \text{and} \quad Q_1 = N_{12}/N_1 \tag{2a}$$

while

$$P_2 = 2N_{22}/N_2 \quad \text{and} \quad Q_2 = N_{12}/N_2 \tag{2b}$$

are the proportions of the $\mathbf{C}_2$ chromosomes in $\mathbf{C}_2\mathbf{C}_2$ homozygotes and $\mathbf{C}_1\mathbf{C}_2$ heterozygotes, respectively. If $1 - s$, 1, and $1 - t$ are the fitness values of $\mathbf{C}_1\mathbf{C}_1$,

C_1C_2 and C_2C_2 individuals, respectively, then

$$N_{11} = NP^2(1-s)/\bar{W}, \qquad N_{12} = 2NPQ/\bar{W}, \qquad N_{22} = NQ^2(1-t)/\bar{W}$$

and

$$N_1 = 2NP \quad \text{and} \quad N_2 = 2NQ \tag{3}$$

where

$$P = \frac{t}{s+t} \quad \text{and} \quad Q = \frac{s}{s+t}$$

and

$$\bar{W} = (1-s)P^2 + 2PQ + (1-t)Q^2.$$

Therefore,

$$P_1 = (1-s)P/\bar{W} \quad \text{and} \quad Q_1 = Q/\bar{W} \tag{4a}$$

and

$$P_2 = (1-t)Q/\bar{W} \quad \text{and} \quad Q_2 = P/\bar{W}. \tag{4b}$$

Let **A** be a neutral locus located on the same chromosome and let r be the frequency of recombination in C_1C_2 individuals between the **A** locus and the inversion. It is assumed that there are an infinite number of possible alleles at the **A** locus and that the mutation rate to new alleles, which differ from all preexisting alleles, is μ, *i.e.*, the infinite alleles model of KIMURA and CROW (1964).

The N_1 type C_1 chromosomes in the next generation are drawn randomly, with replacement, from the gametes produced by the individuals in the present generation. Therefore, the probability that a C_1 chromosome in the next generation comes from a C_1C_1 individual is P_1 and from a C_1C_2 individual, Q_1. Similarly, the probability that a C_2 chromosome comes from a C_2C_2 individual is P_2 and from a C_1C_2 individual, Q_2. If a chromosome comes from a C_1C_2 individual, then it has a probability r of being a recombinant.

Three coefficients of identity by descent are required to describe the probabilities that the genes of the **A** locus on two randomly chosen chromosomes are identical. If a_i^1 (a_i^2) denotes the gene at the **A** locus on an arbitrary C_1 (C_2) chromosome, then the three coefficients of identity are

$$\Phi_{11} = \text{Prob}(a_i^1 \equiv a_j^1) \qquad i \neq j$$

$$\Phi_{12} = \text{Prob}(a_i^1 \equiv a_j^2)$$

$$\Phi_{22} = \text{Prob}(a_i^2 \equiv a_j^2) \qquad i \neq j$$

("$\equiv$" is read "is identical to"). The recursion equations for the expected value of these three identity coefficients by descent over replicate populations are

$$
\begin{aligned}
\Phi'_{11} = {} & (1-\mu)^2 \left(P_1^2 \left[\frac{1}{2N_{11}} + \left(1 - \frac{1}{2N_{11}}\right) \Phi_{11} \right] + 2P_1Q_1[(1-r)\Phi_{11} + r\Phi_{12}] \right. \\
& + Q_1^2 \left[\frac{1}{N_{12}} \{(1-r)^2 + 2r(1-r)\Phi_{12} + r^2\} + \left(1 - \frac{1}{N_{12}}\right) \{(1-r)^2\Phi_{11} \right. \\
& \left. \left. + 2r(1-r)\Phi_{12} + r^2\Phi_{22}\} \right] \right) \\
\Phi'_{12} = {} & (1-\mu)^2 \left(P_1P_2\Phi_{12} + P_1Q_2[(1-r)\Phi_{12} + r\Phi_{11}] + Q_1P_2[(1-r)\Phi_{12} + r\Phi_{22}] \right. \\
& + Q_1Q_2 \left[\frac{1}{N_{12}} \{(1-r)^2\Phi_{12} + 2r(1-r) + r^2\Phi_{12}\} \right. \\
& \left. \left. + \left(1 - \frac{1}{N_{12}}\right) \{(1-r)^2\Phi_{12} + r(1-r)\Phi_{11} + r(1-r)\Phi_{22} + r^2\Phi_{12}\} \right] \right) \\
\Phi'_{22} = {} & (1-\mu)^2 \left(P_2^2 \left[\frac{1}{2N_{22}} + \left(1 - \frac{1}{2N_{22}}\right) \Phi_{22} \right] + 2P_2Q_2[(1-r)\Phi_{22} + r\Phi_{12}] \right. \\
& + Q_2^2 \left[\frac{1}{N_{12}} \{(1-r)^2 + 2r(1-r)\Phi_{12} + r^2\} + \left(1 - \frac{1}{N_{12}}\right) \{(1-r)^2\Phi_{22} \right. \\
& \left. \left. + 2r(1-r)\Phi_{12} + r^2\Phi_{11}\} \right] \right)
\end{aligned}
\tag{5}
$$

If $N \gg 1$, $\mu \simeq 0\left(\frac{1}{N}\right)$, and $r \simeq 0\left(\frac{1}{N}\right)$, then these equations can be approximated by

$$
\begin{aligned}
\Phi'_{11} & \simeq \left(1 - \frac{1}{N_1} - 2\mu - 2Q_1r\right)\Phi_{11} + 2Q_1r\Phi_{12} + \frac{1}{N_1} \\
& = \left(1 - \frac{1}{2NP} - 2\mu - 2Q\tilde{r}\right)\Phi_{11} + 2Q\tilde{r}\Phi_{12} + \frac{1}{2NP} \\
\Phi'_{12} & \simeq Q_2r\Phi_{11} + (1 - 2\mu - Q_1r - Q_2r)\Phi_{12} + Q_1r\Phi_{22} \\
& = P\tilde{r}\Phi_{11} + (1 - 2\mu - \tilde{r})\Phi_{12} + Q\tilde{r}\Phi_{22} \\
\Phi'_{22} & \simeq 2Q_2r\Phi_{12} + \left(1 - \frac{1}{N_2} - 2\mu - 2Q_2r\right)\Phi_{22} + \frac{1}{N_2} \\
& = 2P\tilde{r}\Phi_{12} + \left(1 - \frac{1}{2NQ} - 2\mu - 2P\tilde{r}\right)\Phi_{22} + \frac{1}{2NQ}
\end{aligned}
\tag{6}
$$

where $\tilde{r} = r/\bar{W}$. These approximate equations are obtained by neglecting terms of $0\left(\frac{1}{N^2}\right)$ and less in equations (5) and using the relationships given in (1), (2), (3) and (4). Note that $\bar{W}$ enters into the recursion equations in only the form of

$\tilde{r} = r/\bar{W}$. Therefore, the intensity of selection maintaining the frequencies of $\mathbf{C}_1$ and $\mathbf{C}_2$ affects only the frequency of recombination in the population. The frequency of recombination in the population increases as the intensity of selection increases ($\bar{W}$ decreases) since the proportion of heterozygous $\mathbf{C}_1\mathbf{C}_2$ individuals in the population after selection increases.

The equilibrium values for these approximate equations are

$$\begin{aligned} \hat{\Phi}_{11} &= (2\mu + \tilde{r})(1 + Q\Theta + PQR)/Z \\ \hat{\Phi} &= \tilde{r}(1 + 2PQ\Theta + PQR)/Z \\ \hat{\Phi}_{22} &= (2\mu + \tilde{r})(1 + P\Theta + PQR)/Z \end{aligned} \tag{7}$$

where

$$\begin{aligned} Z = {} & 2\mu(1 + P\Theta + PQR)(1 + Q\Theta + PQR) + P\tilde{r}(1 + Q\Theta + PQR)(1 + P\Theta) \\ & + Q\tilde{r}(1 + P\Theta + PQR)(1 + Q\Theta) \end{aligned}$$

and where $\Theta = 4N\mu$ and $R = 4N\tilde{r}$. If $P = Q = ½$ then the equilibrium values reduce to

$$\begin{aligned} \hat{\Phi}_{11} = \hat{\Phi}_{22} &= \frac{2\Theta + R}{2\Theta + R + \Theta^2 + \Theta R} \\ \hat{\Phi}_{12} &= \frac{R}{2\Theta + R + \Theta^2 + \Theta R}. \end{aligned} \tag{8}$$

It can be shown from the theory of perturbed matrices that these equilibrium values of the approximate equations (6) are the approximate equilibrium values of the exact equations (5). In fact, the absolute values of the differences between the exact and approximate equilibrium values are of order $1/N^2$ or less (theorem 5.9, NOBLE and DANIEL 1977).

The expected squared linkage disequilibrium and squared standard linkage disequilibrium can be obtained from these equilibrium values by a simple transformation. Let x_i (y_i) be the conditional frequency of the ith allele of the **A** locus, $\mathbf{A}_i$, given that it is associated with $\mathbf{C}_1$ ($\mathbf{C}_2$). Therefore, the frequency of $\mathbf{C}_1\mathbf{A}_i$ ($\mathbf{C}_2\mathbf{A}_i$) is Px_i (Qy_i) and the frequency of $\mathbf{A}_i$ in the population is

$$p_i = Px_i + Qy_i.$$

The linkage disequilibrium of the ith allele with $\mathbf{C}_1$ is defined by

$$D_i = Px_i - Pp_i = Px_i - P(Px_i + Qy_i) = PQ(x_i - y_i).$$

[D_i can also be defined by the relationship with $\mathbf{C}_2$, $D_i = -(Qy_i - Qp_i)$.]
If $N \gg 1$, then

$$\begin{aligned} \Phi_{11} &\simeq E\left(\sum_i x_i^2\right) \\ \Phi_{12} &\simeq E\left(\sum_i x_i y_i\right) \\ \Phi_{22} &\simeq E\left(\sum_i y_i^2\right) \end{aligned}$$

and

$$\begin{aligned} E\left(\sum D_i^2\right) &= E\left(\sum_i P^2Q^2[x_i - y_i]^2\right) = E\left(P^2Q^2\left[\sum_i x_i^2 - 2\sum_i x_iy_i + \sum_i y_i^2\right]\right) \\ &= P^2Q^2(\Phi_{11} - 2\Phi_{12} + \Phi_{22}) \\ E\left(\sum\sum_{i\neq j} p_ip_j\right) &= E\left(1 - \sum_i p_i^2\right) = 1 - E\left(\sum_i [Px_i + Qy_i]^2\right) \\ &= 1 - (P^2\Phi_{11} + 2PQ\Phi_{12} + Q^2\Phi_{22}) \end{aligned} \tag{9}$$

From (7)

$$E\left(\sum_i D_i^2\right) = 4P^2Q^2\mu(1 + 2N\mu + N\tilde{r})/Z \tag{10a}$$

and the expected squared standard linkage disequilibrium is

$$\begin{aligned} \sigma_d^2 &= \frac{E\left(\sum_i D_i^2\right)}{PQE\left(\sum\sum_{i\neq j} p_ip_j\right)} \\ &= \frac{PQ(1 + 2N\mu + N\tilde{r})}{PQ(1 + 2N\mu + N\tilde{r})(1 + \Theta + R) + (1 - 4PQ)(2N\mu + N\tilde{r})} \end{aligned} \tag{10b}$$

[This definition of expected squared standard linkage disequilibrium is equivalent to that given by HILL and ROBERTSON (1968) if there are only two alleles at the **A** locus.] The expected squared standard linkage disequilibrium is shown in Figure 1 for $4N\mu = 1$ and $4Nr = 0$, 0.5, 2 and 8. If $P = Q = \frac{1}{2}$, then

$$E\left(\sum_i D_i^2\right) = \frac{N\mu}{2\Theta + R + \Theta^2 + \Theta R}$$

and

$$\sigma_d^2 = \frac{1}{1 + \Theta + R} \tag{11}$$

After considerable manipulation it can be shown that

$$\frac{\partial E\left(\sum_i D_i^2\right)}{\partial \tilde{r}} < 0 \quad \text{and} \quad \frac{\partial \sigma_d^2}{\partial \tilde{r}} < 0$$

and, therefore, both are strictly decreasing functions of $\tilde{r}$. Since $\tilde{r} = r/\bar{W}$ is an increasing function of the selection intensity maintaining the chromosomal arrangements at equilibrium, the expected squared linkage disequilibrium is a strictly decreasing function of the selection intensity.

The transient behavior of $E(\Sigma\ D_i^2)$ can also be studied using the approximate recursion equations for the coefficients of identity (6). These equations can be

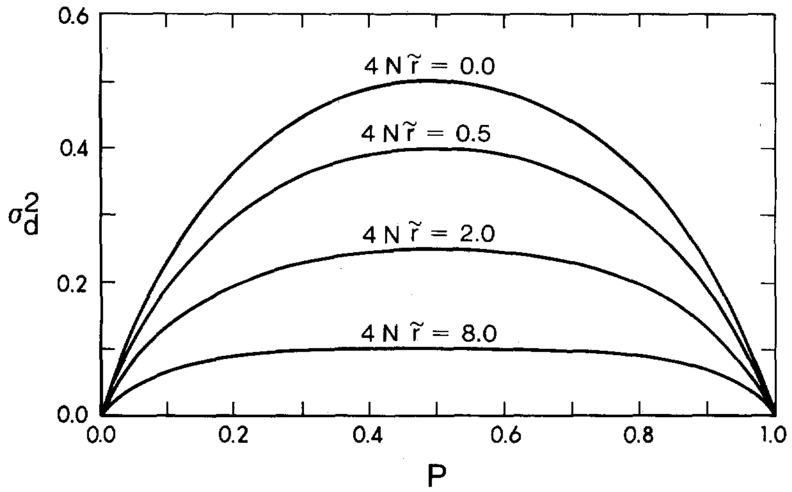

FIGURE 1.—The expected standard squared linkage disequilibrium, σ_d^2, for $\Theta = 4N\mu = 1.0$ and $R = 4N\tilde{r} = 0$, 0.5, 2 and 8.

written as

$$\mathbf{\Phi}' = \mathbf{J}\mathbf{\Phi} + \mathbf{b} \tag{12}$$

where

$$\mathbf{J} = \begin{pmatrix} 1 - \dfrac{1}{2NP} - 2\mu - 2Q\tilde{r} & 2Q\tilde{r} & 0 \\ P\tilde{r} & 1 - 2\mu - \tilde{r} & Q\tilde{r} \\ 0 & 2P\tilde{r} & 1 - \dfrac{1}{2NQ} - 2\mu - 2P\tilde{r} \end{pmatrix} \tag{13}$$

and $\mathbf{\Phi} = (\Phi_{11}, \Phi_{12}, \Phi_{22})^T$ and $\mathbf{b} = \left(\dfrac{1}{2NP}, 0, \dfrac{1}{2NQ}\right)^T$ are column vectors. At time t

$$\mathbf{\Phi}_t = \mathbf{J}^t\mathbf{\Phi}_o + (\mathbf{J}^{t-1} + \mathbf{J}^{t-2} + \cdots + \mathbf{J} + I)\mathbf{b}$$

where $\mathbf{\Phi}_t$ and $\mathbf{\Phi}_o$ are the column vectors of the coefficients of identity at time t and o, respectively.

The asymptotic rate of approach to the equilibrium values (7) is given by the largest eigenvalue of $\mathbf{J}$. The eigenvalues of $\mathbf{J}$ are

$$X_i = 1 - \frac{1}{2N} - 2\mu + \frac{X_i}{2N}$$

where the $X_i (i = 1 - 3)$ are the roots of the equation

$$\begin{aligned} X^3 &- \left[1 - \frac{P^2 + Q^2}{PQ} - 6N\tilde{r}\right]X^2 \\ &+ \left[1 + 8N^2\tilde{r}^2 - (1 - 2N\tilde{r})\frac{P^2 + Q^2}{PQ}\right]X - [1 + 2N\tilde{r}] = 0. \end{aligned} \tag{14}$$

The largest eigenvalue of J is always greater than $1 - \frac{1}{2N} - 2\mu$, i.e., (14) has a positive root since if $X = 0$ the lefthand side of (14) is less than zero.

If $P = Q = \frac{1}{2}$, i.e., $s = t$, then the eigenvalues of J are

$$\lambda_1 = 1 - \left(\frac{1 + 4N\mu + 2N\tilde{r} - \sqrt{1 + 4N^2\tilde{r}^2}}{2N} \right)$$

$$\lambda_2 = 1 - \left(\frac{1 + 4N\mu + 2N\tilde{r} + \sqrt{1 + 4N^2\tilde{r}^2}}{2N} \right)$$

$$\lambda_3 = 1 - \frac{1}{N} - 2\mu - \tilde{r}.$$

The largest eigenvalue is λ_1. The right eigenvectors are

$$\boldsymbol{\xi}_1 = \left(1, \frac{1 + \sqrt{1 + 4N^2\tilde{r}^2}}{2N\tilde{r}}, 1\right)^T$$

$$\boldsymbol{\xi}_2 = \left(1, \frac{1 - \sqrt{1 + 4N^2\tilde{r}^2}}{2N\tilde{r}}, 1\right)^T$$

$$\boldsymbol{\xi}_3 = (1, 0, -1)^T$$

and the left eigenvectors are

$$\boldsymbol{\eta}_1 = \left(\tfrac{1}{2}, \frac{1 + \sqrt{1 + 4N^2\tilde{r}^2}}{2N\tilde{r}}, \tfrac{1}{2}\right)\Big/K_1$$

$$\boldsymbol{\eta}_2 = \left(1/2, \frac{1 - \sqrt{1 + 4N^2\tilde{r}^2}}{2N\tilde{r}}, \tfrac{1}{2}\right)\Big/K_2$$

$$\boldsymbol{\eta}_3 = (1/2, 0, -\tfrac{1}{2})$$

where

$$K_1 = 2\left(1 + \frac{1 + \sqrt{1 + 4N^2\tilde{r}^2}}{4N^2\tilde{r}^2}\right)$$

and

$$K_2 = 2\left(1 + \frac{1 - \sqrt{1 + 4N^2\tilde{r}^2}}{4N^2\tilde{r}^2}\right)$$

are constants chosen so that $\boldsymbol{\eta}_i\boldsymbol{\xi}_i = 1$.

The solution of the recursion equations is

$$\begin{aligned} \boldsymbol{\phi}_t &= \sum_i c_i\boldsymbol{\xi}_i\lambda_i^t + \sum_i e_i\boldsymbol{\xi}_i(\lambda_i^{t-1} + \lambda_i^{t-2} + \lambda_i + 1) \\ &= \sum_i c_i\boldsymbol{\xi}_i\lambda_i^t + \sum_i e_i\boldsymbol{\xi}_i \frac{1}{1 - \lambda_i}(1 - \lambda_i^t) \end{aligned} \tag{15}$$

where

$$c_i = \boldsymbol{\eta}_i \boldsymbol{\phi}_o$$

and

$$e_i = \boldsymbol{\eta}_i \mathbf{b}$$

The vector of equilibrium values is

$$\hat{\Phi} = \sum_i e_i \xi_i \frac{1}{1 - \lambda_i}.$$

From (9), the expected squared linkage disequilibrium at time t is

$$E_t\left(\sum_i D_i^2\right) = \frac{1}{16}\left(\left[\frac{2N\tilde{r} - 1 - \sqrt{1 + 4N^2\tilde{r}^2}}{N\tilde{r}}\right]\left[c_1 - e_1\frac{1}{1 - \lambda_1}\right]\lambda_1^t + \left[\frac{2N\tilde{r} - 1 + \sqrt{1 + 4N^2\tilde{r}^2}}{N\tilde{r}}\right]\left[c_2 - e_2\frac{1}{1 - \lambda_2}\right]\lambda_2^t\right) + \frac{N\mu}{2\Theta + R + \Theta^2 + \Theta R}$$

Therefore, asymptotic rate to the equilibrium value of $E(\sum_i D_i^2)$ is also λ_1.

If $\tilde{r} \simeq 0\left(\frac{1}{N}\right)$ but N is small enough so that $u \ll \frac{1}{N}$, then the equilibrium value of $E(\sum_i D_i^2)$ is 0. However, the transient behavior in this case is still of interest and can be studied by setting $\mu = 0$. If

$$\boldsymbol{\Phi}_o = (\Phi_{11}(0), \Phi_{12}(0), \Phi_{22}(0))$$

then after considerable manipulation

$$E\left(\sum D_i^2\right) = \frac{1}{16}\left(\left(\frac{1 + 2N\tilde{r} - \sqrt{1 + 4N^2\tilde{r}^2}}{\sqrt{1 + 4N^2\tilde{r}^2}}\right)B_1\lambda_1^t - \left(\frac{1 + 2N\tilde{r} + \sqrt{1 + 4N^2\tilde{r}^2}}{\sqrt{1 + 4N^2\tilde{r}^2}}\right)B_2\lambda_2^t\right) \tag{17}$$

where

$$B_1 = \frac{1}{2}(1 - \Phi_{11}(0)) + \left(\frac{1 + \sqrt{4N^2\tilde{r}^2}}{2N\tilde{r}}\right)(1 - \Phi_{12}(0)) + \frac{1}{2}(1 - \Phi_{22}(0))$$

$$B_2 = \frac{1}{2}(1 - \Phi_{11}(0)) + \left(\frac{1 - \sqrt{4N^2\tilde{r}^2}}{2N\tilde{r}}\right)(1 - \Phi_{12}(0)) + \frac{1}{2}(1 - \Phi_{22}(0))$$

and since $\mu = 0$

$$\lambda_1 = 1 - \left(\frac{1 + 2N\tilde{r} - \sqrt{1 + 4N^2\tilde{r}^2}}{2N}\right)$$

$$\lambda_2 = 1 - \left(\frac{1 + 2N\tilde{r} + \sqrt{1 + 4N^2\tilde{r}^2}}{2N}\right).$$

NEI and LI (1980) derived (17) using diffusion theory. (They assumed that there were only two alleles at the **A** locus, and they calculated

$$E_t(d^2) = 16E_t(D^2) = 8E_t\left(\sum_i D_i^2\right).\Big)$$

It should be remembered that (17) is appropriate only to describe the behavior of $E(\Sigma_i D_i^2)$ if $\mu \ll \frac{1}{N}$ and is valid only if the time span is such that no new variation arises by mutation. If $\tilde{r} \gg \frac{1}{N}$ then the equilibrium value of $E(\Sigma_i D_i^2)$ is 0 and its transient behavior is essentially deterministic.

DISCUSSION

At equilibrium, the expected squared standard linkage disequilibrium is significant only if the rate of recombination is of the same order of magnitude as the mutation rate or less (see equation (10). The recombination rates between inversions and loci within the inversions due to double crossovers has been reviewed by ISHII and CHARLESWORTH (1977), and a rate of 10^{-4} per generation was assigned as typical. However, CHOVNICK (1973) examined recombination between alleles of the rosy locus in a paracentric inversion heterozygote in *Drosophila melanogaster*. He found a gene conversion rate of approximately 10^{-5} and the frequency of double crossovers to be much less. Moreover, for genes located near the breakpoints of an inversion, the frequency of double crossovers and gene conversions would be much lower due to lack of pairing. Therefore, depending on location of the gene within an inversion and the size of the inversion, the observed linkage disequilibrium could be accounted for by random drift.

In the ring-forming species, *Oenothera biennis* (LEVY and WINTERNHEIMER 1977) and *O. laciniata* (ELLSTRAND and LEVIN 1980) strong linkage disequilibria between translocation complexes and gene loci have also been observed. Ring-forming species of Oenothera can either be cross pollinated or partially selfed, and permanent heterozygosity is maintained by a system of balanced lethals (*i.e.*, $P = Q = \frac{1}{2}$). It has been assumed throughout this paper that there was random mating. However, the equilibrium values for Φ_{11}, Φ_{12} and Φ_{22} in a partially selfing population are given by (8) if $1 - S \gg 1/N$ where S is the rate of selfing (STROBECK 1980). The observed outcrossing rates in partially selfed ring-forming species of Oenothera varies from 0.006 to 0.20 (CLELAND 1972). Therefore, the expected squared standard linkage disequilibrium for these species of Oenothera is given by (11) if the population size is much larger than ~200 (assuming an outcrossing rate of 0.005). Since in the ring forming of Oenothera recombination occurs only at the tips of the chromosomes and not in the central differentiated segments, random drift can account for the strong linkage disequilibria observed in those species between genes located within these central segments and the translocation complex.

This work was supported by NSERC grant A0502 to C. STROBECK.

LITERATURE CITED

CHARLESWORTH, B., D. CHARLESWORTH and M. LOUKAS, 1977 Linkage disequilibrium studies on *D. subobscura* populations. Drosophila Inform. Serv. **52:** 79.

CHARLESWORTH, B., D. CHARLESWORTH, M. LOUKAS and K. MORGAN, 1979 A study of linkage disequilibrium in British populations of *Drosophila subobscura*. Genetics **92:** 983–994.

CHOVNICK, A., 1973 Gene conversion and transfer of genetic information within the inverted region of inversion heterozygotes. Genetics **75:** 123–131.

CLELAND, R. E., 1972 *Oenothera: Cytogenetics and Evolution.* Academic Press, New York.

DOBZHANSKY, T., 1951 *Genetics and the Origin of Species.* Columbia University Press, New York.

ELLSTRAND, N. C. and D. A. LEVIN, 1980 Association of alleles with chromosomal complexes in the permanent translocation heterozygote, *Oenothera laciniata.* Heredity **44:** 169–176.

HEDRICK, P., S. JAIN and L. HOLDEN, 1978 Multilocus system in evolution. In: *Evolutionary Biology,* Vol. 11, Edited by M. K. HECHT, W. C. STEERE and B. WALLACE. Plenum, New York.

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

ISHII, K. and B. CHARLESWORTH, 1977 Associations between allozyme loci and gene arrangements due to hitch-hiking effects of new inversions. Genet. Res. **30:** 93–106.

KIMURA, M., and J. F. CROW, 1964 The number of alleles that can be maintained in a finite population. Genetics **49:** 725–738.

KOJIMA, K., J. H. GILLESPIE and Y. N. TOBARI, 1970 A profile of Drosophila species' enzymes assayed by electrophoresis. I. Number of alleles, heterozygosities and linkage disequilibrium in glucose-metabolizing systems and some other enzymes. Biochem. Genet. **4:** 627–637.

LANGLEY, C. H., Y. N. TOBARI and K. KOJIMA, 1974 *Linkage disequilibrium in natural populations of Drosophila melanogaster.* Genetics **78:** 921–936.

LEVY, M. and P. L. WINTERNHEIMER, 1977 Allozyme linkage disequilibria among chromosome complexes in the permanent translocation heterozygote, *Oenothera biennis.* Evolution **31:** 465–476.

LEWONTIN, R. C., 1974 *The Genetic Basis of Evolutionary Change.* Columbia University Press, New York.

NEI, M. and W. -H. LI, 1975 Probability of identical monomorphism in related species. Genet. Res. **26:** 31–43.

NEI, M. and W. -H. LI, 1980 Non-random association between electromorphs and inversion *chromosomes in finite populations. Genet.* Res. **35:** 65–83.

NOBLE, B. and J. W. DANIEL, 1977 Applied Linear Algebra. Prentice-Hall, Englewood Cliffs, New Jersey.

PRAKASH, S., 1976 Gene differences between third-chromosome inversions of *Drosophila pseudoobscura.* Genetics **84:** 787–790.

PRAKASH, S. and R. C. LEWONTIN, 1968 A molecular approach to the study of genic heterozygosity in natural populations. III. Direct evidence of coadaptation in gene arrangements of Drosophila. Proc. Natl. Acad. Sci. USA **59:** 398–405.

STROBECK, C., 1980 Heterozygosity of a neutral locus linked to a self-incompatibility locus or a balanced lethal. Evolution **34:** 779–788.

Corresponding Editor: B. S. WEIR