THE SAMPLING VARIANCE OF THE LINKAGE DISEQUILIBRIUM PARAMETER IN MULTI-ALLELE LOCI

NORIKAZU YASUDA

Division of Genetics, National Institute of Radiological Sciences, Chiba 280, Japan

Received 28.xii.77

SUMMARY

The linkage disequilibrium parameter Δ between two loci has been evaluated for multiple alleles. It shows that when there is random mating and chromosome pairs satisfy Hardy-Weinberg frequencies the disequilibrium between two multi-allele loci is due to unequal frequencies between coupling and repulsion phases in the double heterozygote.

The maximum likelihood estimators of Δ and the approximate variance of Δ are given for a large size of sample. A numerical example is also given.

1. INTRODUCTION

RECENTLY, considerable attention has been directed to associations between the frequencies of alleles at two polymorphic loci. Population geneticists have become interested in the adaptive significance of such associations in an attempt to understand the great number of enzyme polymorphisms found in natural populations (Prakash and Lewontin, 1968, 1971; Kojima, Gillespie and Tobari, 1970; Mukai, Mettler and Chigusa, 1971; Zouros and Krimbas, 1972; Charlesworth and Charlesworth, 1973; Franklin, 1973). In medical genetics, searches for associations between a variety of diseases and the HLA locus have been undertaken in the hope that these may yield pertinent clues as to the genetics and the pathogenesis of the disease in question (*e.g.* Svejgaard *et al.*, 1975). Thus it seems worth while to investigate some of the statistical problems of estimation as well as testing of linkage disequilibrium, Δ , one of the measures of such associations. Hill (1974) has showed statistical tests for $\Delta = 0$ and the estimation procedure in the case of two alleles.

In this communication, our attention is primarily directed to the evaluation of the sampling variance, which is necessary to make a significant test of the linkage disequilibrium parameter when it is not zero.

2. ANALYSIS

(i) Parameters

Consider two polymorphic loci, A and B , with two alleles A_1 and A_2 , and B_1 and B_2 , respectively. Let the frequencies of four haplotypes A_1B_1 , A_1B_2 , A_2B_1 and A_2B_2 respectively be r_{11} , r_{12} , r_{21} and r_{22} . Then, $x = r_{11} + r_{12}$ and $y = r_{11} + r_{21}$ are allelic frequencies of A_1 and B_1 , respectively. The linkage disequilibrium parameter may be defined as

$$\Delta = r_{11} - xy.$$

Another definition in terms of haplotype frequencies is

$$D = r_{11}r_{22} - r_{12}r_{21}. \quad (2)$$

When there is random mating and chromosome pairs satisfy Hardy-Weinberg frequencies, the second definition (2) indicates unequal frequencies between coupling and repulsion phases of the double heterozygote. It is easily verified that $\Delta = D$.

For multiple alleles at each locus, the separate linkage disequilibrium parameter for each haplotype can be defined as

$$\Delta_{ij} = r_{ij} - xy \quad (3)$$

where r_{ij} is the frequency of haplotype A_iB_j ($i = 1, 2, \dots, m_1; j = 1, 2, \dots, m_2$, in which m_1 and m_2 are the numbers of alleles at A and B locus, respectively). For $m_1 = m_2 = 2$, $\Delta_{11} = -\Delta_{12} = -\Delta_{21} = \Delta_{22} \equiv \Delta$, as it should be. Since the frequencies of A_1 and B_1 allele are respectively $x = r_{11} + r_{12} + \dots + r_{1m_2}$ and $y = r_{11} + r_{21} + \dots + r_{m_11}$, the Δ value will be

$$\begin{aligned} \Delta_{11} &= r_{11} - xy \\ &= r_{11} - [r_{11} + (r_{12} + \dots + r_{1m_2})][r_{11} + (r_{21} + \dots + r_{m_11})] \\ &= r_{11}[1 - r_{11} - (r_{12} + \dots + r_{1m_2}) - (r_{21} + \dots + r_{m_11}) \\ &\quad - (r_{12} + \dots + r_{1m_2})(r_{21} + \dots + r_{m_11})] \\ &= \sum_{s,t} D(11; st) \quad (s \neq 1, t \neq 1) \end{aligned}$$

where $D(11; st) = r_{11}r_{st} - r_{s1}r_{1t}$.

The definition of the separate linkage disequilibrium parameter $D(11; st)$ was first given by Kimura and Ohta (1971). For any haplotype, say, A_iB_j , we have

$$\Delta_{ij} = \sum_{s,t} D(ij; st) \quad (4)$$

in which

$$D(ij; st) = r_{ij}r_{st} - r_{sj}r_{it}. \quad (s \neq i, t \neq j). \quad (5)$$

This is a special case of Nagylaki's formula (Nagylaki, 1977). This shows that when mating is at random and chromosome pairs satisfy Hardy-Weinberg frequencies the Δ value, even in the case of a multiple allele system, measures in fact unequal frequencies between coupling and repulsion phases in the double heterozygote. Note that $D(ij; st) = D(st; ij)$ and $D(ij; st) = -D(it; sj)$. It can be shown that the number of independent D 's is $(m_1 - 1)(m_2 - 1)$. Obviously, $\Delta \equiv \Delta_{11} = D(11; 22) = -D(12; 21) = -D(21; 12) = D(22; 11) \equiv D$ for $m_1 = m_2 = 2$.

(ii) The sampling variance

The variance of the linkage disequilibrium parameter in a large size of sample can be calculated approximately in terms of haplotype frequencies and their variances and covariances. The approximate variance of D is then

$$V(D) = V(r_{ij}r_{st}) + V(r_{it}r_{sj}) - 2 \text{Cov}(r_{ij}r_{st}, r_{it}r_{sj}) \quad (6)$$

in which $V(\cdot)$ and $\text{Cov}(\cdot, \cdot)$ denote the variance and the covariance, respectively, and, for instance,

$$V(r_{ij}r_{st}) = r_{ij}^2 V(r_{st}) + r_{st}^2 V(r_{ij}) + 2r_{ij}r_{st} \text{Cov}(r_{ij}, r_{st}),$$

where we usually substitute estimates of haplotype frequencies for the parameter r 's.

From (4), we have

$$V(\Delta_{ij}) = \Sigma V(D) + 2\Sigma \text{Cov}(D, D')$$

where $\text{Cov}(D, D')$ denotes the covariance between two D 's and summations are taken for relevant haplotypes. The variance and covariance of haplotype frequencies can be obtained asymptotically by the method of maximum likelihood for any phenotype systems. It has been developed for this purpose a computer program ALLTYPE (Yasuda, 1969). The program will handle any factor union phenotype system (Cotterman, 1969) which will be constructed numerically by computer from a binary presentation of genes and phenotypes.

(iii) *The case of two alleles at each locus*

It is worth while to work out analytic formulae for the most simple case with two alleles at each locus. Suppose that n random zygotes sampled from a population, in which there is assumed to be random mating with each locus in Hardy-Weinberg equilibrium, are classified into nine possible phenotypes (both loci are codominant) according to their allelic constitution at two marker loci. Table 1 shows the observed numbers (n 's) and the

TABLE 1

Expected frequencies and observed numbers for a two-locus system each with two codominant alleles

Phenotype	Expected frequency	Observed number
A_1B_1	r_{11}^2	n_1
$A_1B_1B_2$	$2r_{11}r_{12}$	n_2
A_1B_2	r_{12}^2	n_3
$A_1A_2B_1$	$2r_{11}r_{12}$	n_4
$A_1A_2B_1B_2$	$2r_{11}r_{22} + 2r_{12}r_{21}$	n_5
$A_1A_2B_2$	$2r_{12}r_{22}$	n_6
A_2B_1	r_{21}^2	n_7
$A_2B_1B_2$	$2r_{21}r_{22}$	n_8
A_2B_2	r_{22}^2	n_9
Total	1	n

NOTE:—The symbol r_{ij} ($i, j = 1, 2$) is the frequency of haplotype A_iB_j .

expected frequencies of each phenotype. By chromosome counting (Webster, 1973), which supplies maximum likelihood estimates, we have

$$r_{11} = a_1 + bh, r_{12} = a_2 - bh, r_{21} = a_3 - bh \text{ and } r_{22} = a_4 + bh,$$

in which a 's and b are shown in terms of n 's such that

$$a_1 = (2n_1 + n_2 + n_4)/2n, a_2 = (n_2 + 2n_3 + n_6 + n_5)/2n,$$

$$a_3 = (n_4 + 2n_7 + n_8 + n_5)/2n, a_4 = (n_6 + n_8 + 2n_9)/2n \text{ and } b = n_5/2n.$$

The probability, h , that an individual with phenotype $A_1A_2B_1B_2$ has genotype A_1B_1/A_2B_2 can be obtained from a cubic equation

$$f(H) \equiv H^3 + c_1H^2 + c_2H - c_3 = 0$$

where

$$H = bh \text{ and } c_1 = [a_1 + a_4 - a_2 - a_3 - b]/2,$$

$$c_2 = [a_1a_4 + a_2a_3 - b(a_1 + a_4)]/2 \text{ and } c_3 = a_1a_4b/2.$$

A trial estimate for H may be chosen c_3/c_2 . Although the analytic solution of H can be expressed in terms of trigonometric functions, iterative processes may be more practical; namely, $H_{i+1} = H_i - f(H_i)/f'(H_i)$, where $f'(H) = 3H^2 + 2c_1H + c_2$ and the subscript i designates the number of iterative cycle.

The following change of variables may be convenient for calculation of variances. That is,

$$m = r_{11} + r_{12},$$

$$s = r_{11}/(r_{11} + r_{12})$$

and

$$t = r_{21}/(r_{21} + r_{22}).$$

Applying the method of maximum likelihood, the approximate variances and covariances are

$$\left. \begin{aligned} V(m) &= m(1-m)/2n, \\ \text{Cov}(m, s) &= \text{Cov}(m, t) = 0, \\ V(s) &= s(1-s)[s+t-2st-ms(1-s)]/(2nm \cdot \text{Det}), \\ \text{Cov}(s, t) &= -s(1-s)t(1-t)/(2n \cdot \text{Det}) \\ \text{and} \\ V(t) &= t(1-t)[s+t-2st-(1-m)t(1-t)]/[2n(1-m) \cdot \text{Det}], \end{aligned} \right\} \quad (7)$$

where $\text{Det} = s+t-2st-(1-m)t(1-t)-ms(1-s)$. The variances of haplotype frequencies are

$$\left. \begin{aligned} V(r_{11}) &= s^2V(m) + m^2V(s), \\ V(r_{12}) &= (1-s)^2V(m) + m^2V(s), \\ V(r_{21}) &= t^2V(m) + (1-m)^2V(t) \\ \text{and} \\ V(r_{22}) &= (1-t)^2V(m) + (1-m)^2V(t). \end{aligned} \right\} \quad (8)$$

The formulae (8) could be shown in terms of haplotype frequency (Yasuda, 1968).

The variance of D , according to large sample theory, should be

$$V(D) = (1-2m)^2(s-t)^2V(m) + m^2(1-m)^2[V(s) - 2\text{Cov}(s, t) + V(t)]. \quad (9)$$

Note that $D = m(1-m)(s-t)$, and the formula (9) agrees with Hill (1974) in case of $D = 0$, which corresponds to $s = t$ in the present terminology.

If one of loci is dominant, but the other is codominant, the maximum likelihood solution for m , s and t then could be calculated as follows.

$$m = [2(n_1 + n_2) + (n_3 + n_4)]/2n$$

$$s = 1 - \sqrt{(n_2 + k)/(n_1 + n_2 + k)}$$

and

$$t = 1 - \sqrt{(n_6 + k)/(n_5 + n_6 + k)},$$

in which n 's are the observed numbers shown in table 2, and k satisfies a cubic equation

$$g(k) \equiv \left[1 - \frac{n_1}{n_1 + n_2 + k}\right] \left[1 - \frac{n_5}{n_5 + n_6 + k}\right] - \left[1 - \frac{n_3}{n_3 + n_4 + 2k}\right]^2 = 0.$$

TABLE 2

Expected frequencies and observed numbers for a two-locus system each with two alleles, showing codominance (A) and dominance (B)

Phenotype	Expected frequency	Observed number
A_1B_1	$r_{11}^2 + 2r_{11}r_{12}$	n_1
A_1B_2	r_{12}^2	n_2
$A_1A_2B_1$	$2(r_{11}r_{21} + r_{11}r_{22} + r_{12}r_{21})$	n_3
$A_1A_2B_2$	$2r_{12}r_{22}$	n_4
A_2B_1	$r_{21}^2 + 2r_{21}r_{22}$	n_5
A_2B_2	r_{22}^2	n_6
Total	1	n

NOTE:—The symbol r_{ij} ($i, j = 1, 2$) is the frequency of haplotype A_iB_j .

With a trial estimate $-g(0)/g'(0)$, the solution k can be obtained by iterative processes. The variance and covariance for t and s become

$$\left. \begin{aligned} V(s) &= s(1-s)(2-s)(s+t-st)[(1-s)t(2-t) \\ &\quad + (1-m)\{2s(1-t)-t^2(1-s)\}]/(2nm \cdot \text{Det}'), \\ \text{Cov}(s, t) &= -s(1-s)(2-s)t(1-t)(2-t)(s+t-st)/(2n \cdot \text{Det}'), \\ \text{and} \\ V(t) &= t(1-t)(2-t)(s+t-st)[(1-t)s(2-s) \\ &\quad + m\{2t(1-s)-s^2(1-t)\}]/[2n(1-m) \cdot \text{Det}'], \end{aligned} \right\} \quad (10)$$

in which

$$\begin{aligned} \text{Det}' &= s(1-s)(2-s)t(1-t)(2-t) + (1-m)(1-t)s(2-s)[2s(1-t)-t^2(1-s)] \\ &\quad + m(1-s)t(2-t)[2t(1-s)-s^2(1-t)] + 2m(1-m)(s-t)^2(s+t-st). \end{aligned}$$

We also have $V(m) = m(1-m)/2n$ and $\text{Cov}(m, s) = \text{Cov}(m, t) = 0$. Thus the variances of D and r 's for the codominant-dominant system can be calculated by using (10) for the formulae (9) and (8), respectively.

For the case of both dominant loci, the relevant formulae have already been worked out (Bodmer and Bodmer, 1970; Mittal *et al.*, 1973).

In situations such as test crosses in *Drosophila* and pedigree studies in

man it may be possible to count directly each haplotype. In this case we have

$$V(D) = [r_{12}r_{21} + (r_{11} + r_{22})D - 4D^2]/n. \quad (11)$$

The expression (11) is identical to the formula of Hill (1974) although the latter was in terms of allelic frequencies. The variance of Δ_{ij} for multi-allelic loci becomes to

$$V(\Delta_{ij}) = \sum_{s,t} [r_{it}r_{sj} + (r_{ij} + r_{st})D(ij; st) - 4\{D(ij; st)\}^2]/n \\ + 8r_{ij}^2 \left(\sum_{s,t} r_{st}r_{it}r_{sj} \right) \left(\sum_{u,v} r_{uv}r_{iv}r_{uj} \right) / n^2 \quad (s \neq i, t \neq j, u \neq i, v \neq j). \quad (12)$$

(iv) *Numerical example*

Bender *et al.* (1977) have reported on a sample of South African Negroids ($n = 791$) being scored for each individual's *Bf* and *GLO* phenotypes. They found that no evidence for the linkage disequilibrium by a chi-square test with the 3×7 contingency table. They also have identified four alleles in the properdin factor *B* (*Bf*) locus, which is known in the major histocompatibility complex, and two alleles of the glyoxalase *I* (*GLO*) locus, located about ten centimorgans apart from *Bf* locus on chromosome 6 in man. As depicted in table 3, the phenotype data fit Hardy-Weinberg

TABLE 3

*Sample of South African Negroids classified according to their respective Bf and GLO phenotypes (expected values according to the Hardy-Weinberg law are given in parentheses)**

<i>Bf</i>	<i>GLO</i>			
	1	1-2	2	Total
<i>F</i>	21 (23.4)	139 (131.0)	183 (183.6)	343 (337.9)
<i>F</i> - <i>F</i> 1	2 (3.5)	15 (15.8)	16 (16.7)	33 (36.0)
<i>F</i> - <i>S</i>	19 (19.3)	110 (112.5)	155 (163.6)	284 (295.4)
<i>F</i> - <i>S</i> 1	2 (1.8)	10 (10.3)	19 (14.7)	31 (26.8)
<i>F</i> 1	0 (0.1)	0 (0.5)	0 (0.4)	0 (0.1)
<i>F</i> 1- <i>S</i>	2 (1.5)	11 (6.8)	9 (7.4)	22 (15.7)
<i>F</i> 1- <i>S</i> 1	0 (0.1)	0 (0.6)	0 (0.7)	0 (1.4)
<i>S</i>	3 (4.0)	25 (24.1)	40 (36.5)	68 (64.6)
<i>S</i> - <i>S</i> 1	0 (0.7)	7 (4.4)	3 (6.6)	10 (11.7)
<i>S</i> 1	0 (0.0)	0 (0.2)	0 (0.3)	0 (0.5)
Total	49 (54.4)	317 (306.1)	425 (430.4)	791 = <i>n</i>

* After Bender *et al.*, 1977.

NOTE:—The chi-square values for the Hardy-Weinberg test are 14.04 (d.f. = 22) for *Bf*-*GLO*, 1.00 (d.f. = 1) for *GLO*, and 7.28 (d.f. = 6) for *Bf*.

frequencies either taking the two loci together ($\chi^2 = 14.04$, d.f. = 22), or *GLO* ($\chi^2 = 1.00$, d.f. = 1) or *Bf* ($\chi^2 = 7.28$, d.f. = 6) separately. The eight haplotype frequencies, and their variances and covariances, were calculated by the computer program ALLTYPE.

The eight values of Δ_{ij} , which were calculated from twenty-four separate linkage disequilibrium parameters $D(ij; st)$, together with the standard errors, are shown in table 4. None of them was significantly different from

TABLE 4

The linkage disequilibrium parameters for the Bf and GLO system in African Negroid ($n = 791$)

(Figures in parentheses are the standard errors)

Haplotype i, j	$\Delta_{ij} \times 10^3$	X^2	d.f.
$F, 1$	0.38 (0.20)	0.0025	1
$F, 2$	-0.38 (0.20)	0.0025	1
$F1, 1$	3.77 (2.60)	2.1025	1
$F1, 2$	-3.77 (2.60)	2.1025	1
$S, 1$	-3.96 (6.94)	0.3249	1
$S, 2$	3.96 (6.94)	0.3249	1
$S1, 1$	-0.19 (2.37)	0.0064	1
$S1, 2$	0.19 (2.37)	0.0064	1
Total	0.00	4.8726	

$$\begin{aligned} A \text{ combined } \chi^2 &= \frac{(4-1)(2-1)}{4.2} \cdot 4.8726 \\ &= 1.8272, \text{ d.f.} = 3 \end{aligned}$$

zero by the chi-square test $\chi_{ij}^2 = \Delta_{ij}^2 / V(\Delta_{ij})$ with one degree of freedom. A combined test for the two locus system may be performed by a statistic

$$\chi^2 = \frac{(m_1 - 1)(m_2 - 1)}{m_1 m_2} \sum \chi_{ij}^2$$

with $(m_1 - 1)(m_2 - 1)$ degrees of freedom. In the present example, $\chi^2 = 1.82$ with three degrees of freedom ($P > 0.5$). That is, the *Bf-GLO* system is in linkage equilibrium in African Negroids, corroborating the conclusion of Bender *et al.* (1977).

We lumped *F* and *F1*, and *S* and *S1*, respectively, in the *Bf* locus. Thus, two codominant systems with two alleles at each locus are constructed for purposes of illustration. We then have $n_1 = 23$, $n_2 = 154$, $n_3 = 199$, $n_4 = 23$, $n_5 = 131$, $n_6 = 183$, $n_7 = 3$, $n_8 = 32$, $n_9 = 43$ and $n = 791$. From these, we have $a_1 = 0.14096081$, $a_2 = 0.54740834$, $a_3 = 0.12136536$, $a_4 = 0.19026549$ and $b = 0.08280657$. A relation $a_1 + a_2 + a_3 + a_4 = 1$, may serve for a check on computation. Then, we obtain $c_1 = -0.21017699$, $c_2 = 0.03291848$ and $c_3 = 0.00111044$. A trial estimate of H is $c_3/c_2 = 0.03373303$ so that $f(H) = -0.00020077$ and $f'(H) = 0.02215243$. The improved value will be 0.04279614. Further iterative cycles gave $f(H) = -0.00000822$ and $f'(H) = 0.02042348$ so that $H = 0.04319862$ and $h = 0.52168107$. Thus, the maximum likelihood estimates of frequencies of haplotypes, $F-1$, $F-2$, $S-1$ and $S-2$, together with standard errors, are respectively $r_{11} = 0.18417 \pm 0.01130$, $r_{12} = 0.50420 \pm 0.01372$, $r_{21} = 0.07816 \pm 0.00876$ and $r_{22} = 0.23347 \pm 0.01191$. The transformed variables should be $m = 0.68837$, $s = 0.26755$ and $t = 0.25081$. The degree of linkage disequilibrium D then should be 0.00395 ± 0.00594 which is not significant ($\chi^2 = 0.44$, d.f. = 1) (table 5).

Assuming that either the *F-F* and *F-S* phenotypes in the *Bf* locus or the 1-1 and 1-2 phenotypes in the *GLO* locus could not be distinguished in the data, we would have two dominant-codominant and one dominant-dominant systems. As an example, let allele 2 in the *GLO* locus be recessive. Then, we have $n_1 = 177$, $n_2 = 199$, $n_3 = 154$, $n_4 = 183$, $n_5 = 35$ and $n_6 = 43$. A trial estimate of k will be $-g(0)/g'(0) = -(-0.00310904)/$

TABLE 5

A comparison of the linkage disequilibrium parameter (D) when the Bf and GLO loci are assumed to show either codominance or dominance with two alleles

<i>Bf</i>	<i>GLO</i>	$D(\sigma_D) \times 10^3$	$\chi^2 = D^2/V(D)$
Codominant	Codominant	3.95 (5.94)	0.44
Codominant	Dominant	3.52 (6.03)	0.34
Dominant	Codominant	6.04 (7.01)	0.74
Dominant	Dominant	2.98 (7.28)	0.17

0.00668025 = 0.46540773. One cycle of iteration gave $k = 0.46616764$. The second iteration resulted in $k = 0.46615263$, of which five decimal places are accurate. We have $s = 0.27210$ and $t = 0.25572$. Since $m = [2 \times (177 + 199) + (154 + 183)] / (2 \times 791) = 0.68837$, the frequencies of haplotype $F-1$, $F-2$, $S-1$ and $S-2$ became $r_{11} = 0.18731 \pm 0.01194$, $r_{12} = 0.50106 \pm 0.01430$, $r_{21} = 0.07969 \pm 0.00919$ and $r_{22} = 0.23194 \pm 0.01227$, respectively. Finally, we have $D = 0.00352 \pm 0.00603$, which is obviously not significant ($\chi^2 = 0.34$, d.f. = 1). In table 4, the degree of linkage disequilibrium D , together with the standard error, is summarised for the sake of comparison of the different models.

3. DISCUSSION AND CONCLUSIONS

The aims of this investigation were to show the degree of linkage disequilibrium for two loci, each with multiple alleles and with unequal frequencies of coupling and repulsion phases in double heterozygotes, with mating at random and chromosome pairs satisfying Hardy-Weinberg frequencies. It was also the aim to develop a test for the null hypothesis that $\Delta_{ij} = 0$, and for estimation of the parameters. In some current work significance is tested by a contingency chi-square without estimation of the degree of linkage disequilibrium (Sinnock and Sing, 1972; Mukai *et al.*, 1974; Bender *et al.*, 1977). Although the numerical example shown here did not reveal any disequilibrium, the procedure developed in this communication allows not only for estimation, but also for testing the degree of linkage disequilibrium.

We are not going to discuss the efficiency of the sampling variance of D . A discussion has been provided by Hill (1974).

Finally, it may be worth while to mention that the analytic formulae developed could be applied to blood group systems such as MNSs and Rh in certain circumstances.

Acknowledgments.—We are grateful to Mr Kenjiro Fukuhisa whose generosity in running programs on an electronic computer, TOSBAC 3400, installed in the National Institute of Radiological Sciences, made this study possible. The work was supported in part by a fund for promoting multiministerial Projects from Science and Technology Agency, Japan.

4. REFERENCES

- BENDER, K., MAUFF, G., AND HITZEROTH, H. W. 1977. No evidence for linkage disequilibrium between *Bf* and *GLO* in African negroids. *Hum. Genet.*, 38, 227-230.
 BODMER, J. G., AND BODMER, W. F. 1970. Studies on African pygmies. VI. A comparative study of the HL-A polymorphism in the Babinga pygmies and other African and Caucasian populations. *Amer. J. Hum. Genet.*, 22, 396-411.

- CHARLESWORTH, B., AND CHARLESWORTH, D. 1973. A study of linkage disequilibrium in populations of *Drosophila melanogaster*. *Genetics*, 73, 351-359.
- COTTERMAN, C. W. 1969. Factor-union phenotype systems. In *Computer Application in Genetics*, ed. N. E. Morton, pp. 1-19. Univ. Hawaii Press.
- FRANKLIN, I. R. 1973. Selection, migration and genetic drift in natural populations of *D. melanogaster*. *Genetics*, 74, s84.
- HILL, W. G. 1974. Estimation of linkage disequilibrium in randomly mating populations. *Heredity*, 33, 229-239.
- KIMURA, M., AND OHTA, T. 1971. *Theoretical Aspects of Population Genetics*, pp. 92-93. Princeton Univ. Press.
- KOJIMA, K., GILLESPIE, J., AND TOBARI, Y. N. 1970. A profile of *Drosophila* species' enzymes estimated by electrophoresis. I. Number of alleles, heterozygosities and linkage disequilibrium in glucose-metabolising systems and some other enzymes. *Biochem. Genet.* 4, 626-637.
- MITTAL, K. K., HASEGAWA, T., TING, A., MICKEY, M. R., AND TERASAKI, P. I. 1973. Genetic variation in the HL-A system between Ainu, Japanese and Caucasians. In *Histo-, compatibility Testing 1972*, eds. J. Dausset and J. Colombani, pp. 187-195. Munksgaard, Copenhagen.
- MUKAI, T., METTLER, L. E., AND CHIGUSA, S. I. 1971. Linkage disequilibrium in a local population of *Drosophila melanogaster*. *Proc. Nat. Acad. Sci. U.S.A.*, 69, 2474-2478.
- MUKAI, T., WATANABE, T. K., AND YAMAGUCHI, O. 1974. The genetic structure of natural populations of *Drosophila melanogaster*. XII. Linkage disequilibrium in a large local population. *Genetics*, 77, 771-793.
- NAGYLAKI, T. 1977. The evolution of one- and two-locus system, II. *Genetics*, 35, 347-354.
- PRAKASH, S., AND LEWONTIN, R. C. 1968. A molecular approach to the study of genetic heterozygosity in natural populations. III. Direct evidence of coadaptation in gene arrangements of *Drosophila*. *Proc. Nat. Acad. Sci. U.S.A.*, 59, 398-405.
- PRAKASH, S., AND LEWONTIN, R. C. 1971. A molecular approach to the study of genetic heterozygosity in natural populations. V. Further direct evidence of coadaptation in inversion of *Drosophila*. *Genetics*, 69, 405-408.
- SINNOCK, P., AND SING, G. E. 1972. Analysis of multi-locus genetic systems in Tecumseh, Michigan. II. Consideration of the correlation between non-alleles in gametes. *Amer. J. Hum. Genet.*, 24, 393-415.
- SVEJGAARD, A., PLATZ, P., RYDER, L. P., NIELSEN, L. S., AND THOMSEN, M. 1975. HL-A and disease associations—A survey. *Transplant. Rev.*, 22, 3-43.
- WEBSTER, T. P. 1973. Adaptive linkage disequilibrium between two esterase loci of a salamander. *Proc. Nat. Acad. Sci. U.S.A.*, 70, 1156-1160.
- YASUDA, N. 1968. A gene counting method of maximum likelihood for estimating gene frequency. In *Annual Report of National Institute of Genetics. Misima*, 18, 73-75.
- YASUDA, N. 1969. Estimation of inbreeding coefficient from mating type frequency and gene frequency. In *Computer Application in Genetics*, ed. N. E. Morton, pp. 87-102. Univ. Hawaii Press.
- ZOUROS, E., AND KRIMBAS, C. B. 1972. Linkage disequilibrium in natural populations of *Drosophila subobscura* maintained by selection. *Genetics*, 71, s71.