DISEQUILIBRIUM, SELECTION, AND RECOMBINATION: LIMITS IN TWO-LOCUS, TWO-ALLELE MODELS

ALAN HASTINGS

Department of Mathematics, University of California, Davis, California 95616

ABSTRACT

All possible combinations of equilibria and fitnesses in two-locus, two-allele, deterministic, discrete-generation selection models are enumerated. This knowledge is used to obtain limits (which can be calculated to arbitrary precision) to the relationships among disequilibrium, selection and recombination for fixed values of allele frequencies. In all cases, the inequality $|rD| < s/10$ holds, where r is recombination and D is disequilibrium, and all selection coefficients lie between $1 - s$ and $1 + s$ times that of the double heterozygote. Linear programming techniques are used to calculate the minimum strength of selection needed to explain several observed nonzero values of D reported in the literature. One conclusion is that the failure to observe nonzero values of D is not surprising.

THE complete dynamic behavior of discrete generation, deterministic, one-locus, two-allele diploid selection models has been well known for over fifty years. Yet, knowledge of the behavior of two-locus, two-allele models is much more limited (KARLIN 1975; EWENS 1979). Even a knowledge of possible equilibrium-fitness combinations has heretofore been incomplete. However, in determining the effect of selection, two-locus models may be more important than one-locus models. When considering two loci, which adds the effect of recombination, measurements of linkage disequilibrium have been used in attempts to detect selection.

These attempts to use disequilibrium to infer selection have been based on results from an equilibrium analysis of special fitness models (see review in KARLIN 1975; EWENS 1979). As noted by LANGLEY (1977), efforts to detect linkage disequilibrium have not been very successful, although there have been some exceptions, such as in the study of *Drosophila montana* (ROBERTS and BAKER 1973; BAKER 1975). Two obvious questions are why have these efforts not been successful and what can one infer from a nonzero value of linkage disequilibrium?

More specific questions concerning the possible equilibrium behavior of *deterministic* two-locus, two-allele models with *general viabilities* can be formulated in the spirit of the general questions above. If the rate of recombination between two loci is known and a nonzero value of disequilibrium is observed, what is the minimum strength of selection? If the value of disequilibrium and strength of selection are known, what is the largest value of recombination possible? If recom-

bination and strength of selection are specified, what is the largest value of disequilibrium possible?

A complete answer to these questions would require a complete knowledge of all stable equilibria for the general two-locus, two-allele model. Although, using the techniques in this paper, such results may be possible, they would almost certainly be too unwieldy to be useful. Instead, in this paper, I provide a partial, but extremely useful, answer based on a complete knowledge of fitnesses corresponding to all equilibria of the general two-locus, two-allele model.

The approach is to reverse the usual procedure of specifying fitness values and solving for equilibria. Instead, I specify the chromosomal frequencies and then solve for the fitnesses. Then, the questions posed above can be answered by using linear programming (DANTZIG 1963).

Not only does this technique provide new results, but additionally it corresponds more closely to how population genetics models are used to interpret data. Fairly accurate observations of chromosomal frequencies are possible, and these are the inputs here. Judgments about selective strengths are possible, and these are outputs or can be used as constraints.

EQUILIBRIA OF THE TWO-LOCUS, TWO-ALLELE MODEL

The model is the standard two-locus, two-allele model (reviewed in KARLIN 1975). Let the alleles at locus A be A and a and at locus B, B and b. Then fitnesses can be expressed in terms of a symmetric $(w_{ij} = w_{ji})$ 4×4 matrix:

	AB	Ab	aB	ab
AB	w_{11}	w_{12}	w_{13}	w_{14}
Ab	w_{21}	w_{22}	w_{23}	w_{24}
aB	w_{31}	w_{32}	w_{33}	w_{34}
ab	w_{41}	w_{42}	w_{43}	w_{44}

$$(1)$$

Also, assume that there are no position effects, so that $w_{14} = w_{23}$. Let x_1, x_2, x_3, x_4 be the frequencies of chromosomes AB, Ab, aB, ab, respectively. Define the disequilibrium as

$$D = x_1 x_4 - x_2 x_3 \qquad (2)$$

Let r be the recombination rate between A and B. Finally, let

$$\varepsilon_i = -1, 1, 1, -1 \qquad \text{for } i = 1 \text{ to } 4. \qquad (3)$$

Then, the dynamics of the deterministic two-locus, two-allele model for discrete generations are given by (letting $'$ denote the next generation):

$$x'_i = \frac{x_i w_i + \varepsilon_i r D w_{14}}{\bar{w}} \qquad \text{for } i = 1 \text{ to } 4 \ , \qquad (4)$$

where

$$w_i = \sum_{j=1}^{4} w_{ij} x_j \qquad \text{for } i = 1 \text{ to } 4 \qquad (5)$$

and

$$\bar{w} = \sum_{i=1}^{4} w_i x_i \ . \tag{6}$$

To solve for equilibria of (4), set $x_i' = x_i$, yielding:

$$x_i \bar{w} = x_i w_i + \varepsilon_i r D w_{14} \quad \text{for } i = 1 \text{ to } 4. \tag{7}$$

Note the following about (7). First, if (7) holds for $i = 1$ to 3, then it holds for $i = 4$, since the x_i satisfy

$$\sum_{i=1}^{4} x_i = 1. \tag{8}$$

Second, note that if r and x_1, x_2. x_3 and x_4 are specified, (7) is a *linear system* of equations for the w_{ij}—a system of three equations with nine unknowns. Finally, if fitnesses are normalized so that w_{14} is one, and just x_1, x_2, x_3 and x_4 are specified, system (7) becomes a linear system of equations for r and the fitnesses.

Hence, given the chromosomal frequencies and recombination fraction, the possible fitnesses for which those frequencies are an equilibrium can be found, using a method for solving a linear system of equations, such as Gauss-Jordan elimination (DANTZIG 1963). There will be six free parameters. Similarly, if fitnesses are normalized, r and the fitnesses can be found once the chromosomal frequencies are specified. Hence, all possible combinations of equilibria and fitnesses for the two-locus two-allele model have been found.

However, the general solution is both unwieldy and difficult to interpret. Instead, in the following section, knowledge of the equilibrium-fitness combinations will be used to answer the questions posed above.

LINEAR PROGRAMMING AND EQUILIBRIA

Here, the methods used to answer the questions posed will be described. Using linear programming (DANTZIG 1963), one can maximize a quantity that is a linear function where the unknowns are subjected to inequalities that are linear in the unknowns. (The most practical method for doing this is the simplex method.) Hence, the goal is to express the questions in the introduction as linear programming problems, which will then be solved using standard (IMSL) computer routines.

The first problem is the following. *Given equilibrium chromosomal frequencies, with a nonzero value of* D *and limits to the strength of selection, what is the maximum value of* r *possible?* This can be phrased as a linear programming problem. Specify x_1, x_2, x_3, x_4. Normalize fitnesses so that $w_{14} = 1$. Assume that all other fitnesses satisfy

$$1 - s \le w_{ij} \le 1 + s \ ,$$

so that s is a measure of the strength of selection. (Note that three of the w_{ij} can be found linearly in terms of the other w_{ij} and r, as indicated in the previous section.) Maximize r for a fixed value of s. The reason I pose the problem as above, with s entering additively, is that then r_{max} (maximum value of r) becomes

a linear function of s (as is easily seen by writing out the linear programming problem); thus, it is necessary only to report r_{max}/s for different values of x_1, x_2, x_3, x_4. To find the minimum value of s for a given r, one uses $1/(r_{max}/s)$.

Since disequilibrium is of interest, it is more informative to report the results as a function of p_A, p_B and D, where p_A is the frequency of allele A, and p_B that of allele B. Here

$$p_A = x_1 + x_2$$

and

$$p_B = x_1 + x_3 \ .$$

Note that given the values of p_A, p_B and D, the frequencies x_i can be found.

The second linear programming problem is more useful when dealing with specific data, knowing p_A, p_B, D and r. Here, one may not wish to normalize w_{14}, but to allow all fitnesses to satisfy

$$1 - s \leq w_{ij} \leq 1 + s \ ,$$

so that w_{14} can vary. Then, minimizing s, with r, p_A, p_B and D fixed, is a linear programming problem. The disadvantage of this slightly more general approach is that s_{min}/r is no longer a constant, so that it is not as useful for reporting general results. However, for small r (or s), s_{min}/r for this problem is very close to $1/(r_{max}/s)$ from the previous problem; therefore, the loss of generality in the previous problem is not very great.

<div align="center">RESULTS</div>

There are two sets of results reported here: (1) general results for the first problem of the previous section, and (2) specific results for the second problem, corresponding to experimental estimates of nonzero D reported in the literature.

For the general results, I used values of p_A from 0.5 through 0.9 and p_B from p_A through 0.9, in intervals of 0.1. (Given the symmetry of the problem, this accounts for all possible combinations of p_A and p_B, as explained below.) The values of D used were determined as follows. For given values of p_A and p_B, the maximum (D_{max}) and minimum (D_{min}) values of D are (LEWONTIN 1964):

$$D_{max} = \text{minimum of } p_A(1-p_B) \text{ and } p_B(1-p_A) \ ,$$
$$D_{min} = \text{maximum of } -p_Ap_B \text{ and } -(1-p_A)(1-p_B) \ .$$

Define

$$D_{int} = (D_{max} - D_{min})/10.$$

I have calculated values of r_{max}/s for D ranging from $D_{min} + D_{int}$ to $D_{max} - D_{int}$ in intervals of D_{int} (Figure 1). To illustrate the relationship for small D, the values of r_{max}/s for D ranging from $-D_{int}(0.8)$ to $(0.8)D_{int}$ (in intervals of $0.2D_{int}$) are in Figure 2. In all cases, $D = 0$ is omitted, since for that value r_{max}/s would be infinite. Also, $D = D_{max}$ (and $D = D_{min}$) are omitted, since for these values r_{max}/s would be zero. Since r_{max}/s is a monotone decreasing function of D for positive D (increasing for negative D), the values in the table can also be

Pa= 0.5   Pb= 0.5

D	-.200	-.150	-.100	-.050	0.050	0.100	0.150	0.200
Rmax/S	0.2925	0.5867	0.9450	1.6800	1.6800	0.9450	0.5867	0.2925

Pa= 0.5   Pb= 0.6

D	-.160	-.120	-.080	-.040	0.040	0.080	0.120	0.160
Rmax/S	0.2852	0.5927	1.0035	1.9240	1.9240	1.0035	0.5927	0.2853

Pa= 0.5   Pb= 0.7

D	-.120	-.090	-.060	-.030	0.030	0.060	0.090	0.120
Rmax/S	0.2790	0.6293	1.1760	2.5013	2.5013	1.1760	0.6293	0.2790

Pa= 0.5   Pb= 0.8

D	-.080	-.060	-.040	-.020	0.020	0.040	0.060	0.080
Rmax/S	0.2610	0.5973	1.1340	2.4960	2.4960	1.1340	0.5973	0.2610

Pa= 0.5   Pb= 0.9

D	-.040	-.030	-.020	-.010	0.010	0.020	0.030	0.040
Rmax/S	0.2430	0.5653	1.0920	2.4480	2.4480	1.0920	0.5653	0.2430

Pa= 0.6   Pb= 0.6

D	-.120	-.080	-.040	0.040	0.080	0.120	0.160	0.200
Rmax/S	0.3627	0.8187	1.7760	2.0000	1.0183	0.6440	0.4320	0.2232

Pa= 0.6   Pb= 0.7

D	-.090	-.060	-.030	0.030	0.060	0.090	0.120	0.150
Rmax/S	0.3720	0.9440	2.3520	2.5000	1.2720	0.7840	0.4720	0.2232

Pa= 0.6   Pb= 0.8

D	-.060	-.040	-.020	0.020	0.040	0.060	0.080	0.100
Rmax/S	0.3480	0.8960	2.2680	2.5000	1.2480	0.7560	0.4480	0.2088

Pa= 0.6   Pb= 0.9

D	-.030	-.020	-.010	0.010	0.020	0.030	0.040	0.050
Rmax/S	0.3240	0.8480	2.1840	2.5000	1.2240	0.7280	0.4240	0.1944

Pa= 0.7   Pb= 0.7

D	-.060	-.030	0.030	0.060	0.090	0.120	0.150	0.180
Rmax/S	0.5590	1.8880	2.5440	1.2500	0.7520	0.4935	0.3392	0.1770

Pa= 0.7   Pb= 0.8

D	-.040	-.020	0.020	0.040	0.060	0.080	0.100	0.120
Rmax/S	0.5220	1.7920	2.4960	1.2500	0.8320	0.5670	0.3584	0.1740

Pa= 0.7   Pb= 0.9

D	-.020	-.010	0.010	0.020	0.030	0.040	0.050	0.060
Rmax/S	0.4860	1.6960	2.4480	1.2500	0.8160	0.5460	0.3392	0.1620

Pa= 0.8   Pb= 0.8

D	-.020	0.020	0.040	0.060	0.080	0.100	0.120	0.140
Rmax/S	1.0440	2.2680	1.2480	0.8333	0.5760	0.4032	0.2773	0.1440

Pa= 0.8   Pb= 0.9

D	-.010	0.010	0.020	0.030	0.040	0.050	0.060	0.070
Rmax/S	0.9720	2.1840	1.2240	0.8333	0.6120	0.4368	0.2827	0.1389

Pa= 0.9   Pb= 0.9

D	0.010	0.020	0.030	0.040	0.050	0.060	0.070	0.080
Rmax/S	1.6960	1.0920	0.8160	0.6250	0.4704	0.3430	0.2331	0.1193

FIGURE 1.—The maximum values of r divided by s, r_{max}/s, where all fitnesses lie between $(1-s)w_{14}$ and $(1+s)w_{14}$, for values of D, p_A, p_B. (Read Pa as p_A and Pb as p_B.)

Pa= 0.5   Pb= 0.5

D	-.040	-.030	-.020	-.010	0.010	0.020	0.030	0.040
Rmax/S	2.01	2.55	3.60	6.74	6.74	3.60	2.55	2.01

Pa= 0.5   Pb= 0.6

D	-.032	-.024	-.016	-.008	0.008	0.016	0.024	0.032
Rmax/S	2.35	3.04	4.41	8.49	8.49	4.41	3.04	2.35

Pa= 0.5   Pb= 0.7

D	-.024	-.018	-.012	-.006	0.006	0.012	0.018	0.024
Rmax/S	3.12	4.15	6.18	12.24	12.24	6.18	4.15	3.12

Pa= 0.5   Pb= 0.8

D	-.016	-.012	-.008	-.004	0.004	0.008	0.012	0.016
Rmax/S	3.14	4.21	6.31	12.58	12.58	6.31	4.21	3.14

Pa= 0.5   Pb= 0.9

D	-.008	-.006	-.004	-.002	0.002	0.004	0.006	0.008
Rmax/S	3.09	4.16	6.26	12.53	12.53	6.26	4.16	3.09

Pa= 0.6   Pb= 0.6

D	-.032	-.024	-.016	-.008	0.008	0.016	0.024	0.032
Rmax/S	2.21	2.91	4.29	8.52	9.14	4.70	3.21	2.46

Pa= 0.6   Pb= 0.7

D	-.024	-.018	-.012	-.006	0.006	0.012	0.018	0.024
Rmax/S	3.01	4.10	6.24	12.63	12.76	6.38	4.24	3.16

Pa= 0.6   Pb= 0.8

D	-.016	-.012	-.008	-.004	0.004	0.008	0.012	0.016
Rmax/S	2.92	3.98	6.08	12.35	12.57	6.31	4.21	3.14

Pa= 0.6   Pb= 0.9

D	-.008	-.006	-.004	-.002	0.002	0.004	0.006	0.008
Rmax/S	2.82	3.86	5.92	12.06	12.37	6.23	4.17	3.13

Pa= 0.7   Pb= 0.7

D	-.024	-.018	-.012	-.006	0.006	0.012	0.018	0.024
Rmax/S	2.51	3.51	5.50	11.41	12.06	6.15	4.16	3.16

Pa= 0.7   Pb= 0.8

D	-.016	-.012	-.008	-.004	0.004	0.008	0.012	0.016
Rmax/S	2.39	3.36	5.27	10.97	11.66	5.97	4.06	3.09

Pa= 0.7   Pb= 0.9

D	-.008	-.006	-.004	-.002	0.002	0.004	0.006	0.008
Rmax/S	2.27	3.20	5.04	10.52	11.27	5.79	3.95	3.02

Pa= 0.8   Pb= 0.8

D	-.016	-.012	-.008	-.004	0.004	0.008	0.012	0.016
Rmax/S	1.52	2.30	3.82	8.32	9.54	5.03	3.51	2.74

Pa= 0.8   Pb= 0.9

D	-.008	-.006	-.004	-.002	0.002	0.004	0.006	0.008
Rmax/S	1.42	2.15	3.59	7.85	9.06	4.80	3.36	2.63

Pa= 0.9   Pb= 0.9

D	-.008	-.006	-.004	-.002	0.002	0.004	0.006	0.008
Rmax/S	0.27	0.70	1.53	3.99	5.68	3.23	2.39	1.96

FIGURE 2.—The maximum values of r divided by s, r_{max}/s, where all fitnesses lie between $(1-s)w_{14}$ and $(1+s)w_{14}$, for small values of D. Note that for $D=0$, r_{max}/s is infinite. (Read Pa as p_A and Pb as p_B.)

TABLE 1

s_{min} *(see text) for significant nonzero values of* D *reported in* ROBERT *and* BAKER *(1973) and* BAKER *(1975) for esterase loci in* D. montana

Pair of loci ("A" listed first)	Year	p_A	p_B	D	s_{min}
1 and 2	1970	0.274	0.637	− 0.109	0.003
1 and 2	1973	0.236	0.627	− 0.081	0.002
3 and 4	1970	0.568	0.346	− 0.163	0.006
3 and 4	1973	0.537	0.398	− 0.162	0.005

A value of 0.002 was used for r in all cases.

used to determine the maximum possible value of D, given p_A, p_B, r, and the maximum strength of selection.

To determine r_{max}/s for values of p_A, p_B and D not listed in Figures 1 and 2, make use of the symmetry of the model. The following sets of values for p_A, p_B and D all yield the same value of r_{max}/s.

$$\text{Set 1:} \quad p_A = a, \, p_B = b, D = d$$

$$\text{Set 2:} \quad p_A = b, \, p_B = a, D = d$$

$$\text{Set 3:} \quad p_A = a, \, p_B = 1 - b, D = -d$$

$$\text{Set 4:} \quad p_A = (1-a), \, p_B = b, D = -d$$

By combining the operations above, values of p_A, p_B and D not listed in the figures can be reduced to values listed in the figures.

The first specific case of observed nonzero D that I will use to determine the minimum strength of selection (under a selection hypothesis) is that of *Drosophila montana* (ROBERTS and BAKER 1973; BAKER 1975). Here, significant linkage disequilibrium was observed among four different closely linked esterase loci, at several different times. In particular, grouping active alleles as one "allele" and null alleles as the other, significant nonzero values of D were observed between esterase loci 1 and 2 and between 3 and 4. The loci are in the order 1–3–2–4. The total map distance between 1 and 4 is 0.37. To get an indication of the minimum strength of selection, s_{min} r was calculated for $r = .002$, as in the description of problem two above (see TABLE 1). The general picture that emerges is that selection coefficients ($2 \times s_{min}$) need be only 1% to explain these observed gamete frequencies.

Similar calculations were performed for several cases of significant nonzero values of D observed in *Drosophila melanogaster* (LANGLEY, TOBARI and KOJIMA 1974) and the results are in Table 2. Here r was taken as one-half the map distance reported in (LANGLEY, TOBARI and KOJIMA 1974), taking into account the absence of recombination in males. Indicated selection coefficients ($2 \times s_{min}$) ranged from 0.04 to 0.18.

TABLE 2

s_{min} *(see text) for significant nonzero values of* D *reported in* LANGLEY, TOBAR *and* KOJIMA *(1974) for isozyme loci on chromosome 3 of* Drosophila melanogaster *in the Katsunuma population*

Pair of loci ("A" listed first)	p_A	p_B	D	r	s_{min}
Octanol dehydrogenase Aldehyde oxidase	0.842	0.784	0.043	0.037	0.026
Esterase C Leucine aminopeptidase	0.800	0.754	0.024	0.0465	0.092
Esterase 6 Phosphoglucomutase	0.170	0.667	0.023	0.033	0.021

Value of r is map distance divided by two.

DISCUSSION

Given particular equilibrium values of chromosomal frequencies at two loci, constraints have been found on the relationship between the maximum value of recombination and the minimum strength of selection. The limits imposed by these constraints can in fact be achieved (there exist equilibria with this much recombination or that little selection) if only existence of equilibria, not stability, is included. (The more complex role of stability will be pursued elsewhere.) It is important to note that if the true situation involves more than two loci, the results of this paper remain in force if fitnesses are replaced by induced fitnesses (what would be observed at two loci), as noted by EWENS and THOMPSON (1977).

Other results of a similar nature for general viabilities are those of NAGYLAKI (1976), KARLIN and CARMELLI (1975) and GINZBURG (personal communication). NAGYLAKI (1976), dealing with weak selection, deduced an asymptotic order of magnitude relationship between disequilibrium and selection coefficients, in the limit as the strength of selection goes to zero. KARLIN and CARMELLI (1975) reported numerical results for general viabilities indicating that large values of D and r were incompatible. Using a simple argument, GINZBURG demonstrated that $|rD| < 3s/4$. The distinguishing feature of the present paper is the precision of the results and that the limits can easily be calculated for any given equilibrium. Hence, general conclusions are possible.

First, consider the results presented in tabular form in Figures 1 and 2. These results provide limits to the relationship among three parameters in multilocus population genetics, recombination rate, r, disequilibrium, D, and selection, s. Note that in *all cases* the stringent inequality

$$|rD| \leq s/10, \tag{9}$$

holds. Note that, as D approaches 0 or D_{max} or D_{min}, the left-hand side of (9) approaches zero. Also, for values of p_A and p_B not near 0.5, the bound on D is stricter than that implied by (9), as demonstrated in Figures 1 and 2.

The general results provide a guide as to when the search for nonzero D (generated by deterministic equilibria) is likely to be successful. In particular, the results suggest that many of the negative results obtained in the search for disequilibrium (see review by LANGLEY 1977) are not surprising. Given the map distances and observed gene frequencies, selection coefficients would have to be extremely large to generate values of D large enough to be detected. Table IV of BROWN (1975) lists the sample sizes (number of random zygotes) needed to be 90% sure of rejecting the null hypothesis that $D = 0$, for given values of p_A, p_B and D. Combining the results of his table with the present paper, the difficulty of detecting deviations from linkage equilibrium is apparent. For example, let $p_A = 0.9$, $p_B = 0.5$, $s = .01$ (selection strength of about 2 percent) and let D be at the largest value possible as indicated by Figure 1. Then, if $r = .01$, more than 500 random zygotes are required, and if $r = .025$ more than 2300 random zygotes are required to be 90% sure of rejecting the null hypothesis that $D = 0$.

From a theoretical point of view, the results presented here suggest that population genetics theory should concentrate on those situations in which D is quite small, except for very small r. In both of these cases, ideas from one-locus theory may be useful in understanding equilibrium properties. Additionally, since these results imply that D is usually small, arguments for the evolution of recombination (MAYNARD SMITH 1978) should not be based on the presence of large disequilibrium.

Finally, for fixed values of the limits to the selection coefficients, the decline in the maximum possible value of D (say for D positive) as recombination, r, is increased does not fit any simple function. In particular, $D_{max}(r)$ is not an exponential function of r.

I will now turn to the interpretation of experimental estimates of disequilibrium (D), discussing first general questions and then the results reported in the previous section. Using the results of this paper, the minimum strength of selection necessary to maintain an observed disequilibrium D (at a deterministic equilibrium) can be calculated. There are three general categories into which s_{min} might fall. First, s_{min} might be so small that selection would be impossible to detect directly. Another possibility would be for s_{min} to be so large as to stretch the credulity of a selective explanation for the observed D. Finally, s_{min} might fall between the two cases above, suggesting that more direct measures of selection might prove successful.

All three general classes are represented in the cases reported in the previous section of s_{min} calculated from observed nonzero D. In particular the best documented case, that of D. montana, falls on the borderline of possibly detectable differences. However, since the s_{min} values on the order of one percent represent lower bounds, a search for the selective differences might just prove successful. However, in case of disequilibrium between esterase C and leucine aminopeptidase in Drosophila melanogaster (Table 2), where selective differences of at least 18% are indicated, an explanation for this D value based on selection seems improbable.

In applying the results of this paper directly to data, several cautions are in order. First, the census time assumed in the model is just after mating. However, if selective differences are small, census time should not greatly affect estimates of D. Second, the results of the present paper are for populations in equilibrium. However, if populations are near equilibrium, the results of this paper almost certainly would apply.

Although the results of this paper do not apply directly to more loci, more alleles and the effect of mutation, the techniques used here can be applied. Research along these lines is currently in progress. The other avenue of research suggested by the results of this paper is to use the complete knowledge of all equilibrium-fitness combinations for general two-locus, two-allele models obtained here to examine these models in more depth.

LITERATURE CITED

Baker, W. K., 1975   Linkage disequilibrium over space and time in natural populations of *Drosophila montana*. Proc. Natl. Acad. Sci. U.S. **72**: 4095–4099.

Brown, A. H. D., 1975   Sample sizes required to detect linkage disequilibrium between two or three loci. Theoret. Pop. Biol. **8**: 184–201.

Dantzig, G. B., 1963   *Linear Programming and Extensions*. Princeton University Press, Princeton, New Jersey.

Ewens, W. J., 1979   *Mathematical Population Genetics*. Springer-Verlag, Berlin.

Ewens, W. J. and G. Thomson, 1977   Properties of equilibria in multi-locus genetic systems. Genetics **87**: 807–819.

Karlin, S., 1975   General two-locus selection models: some objectives, results and interpretations. Theoret. Pop. Biol. **7**: 364–398.

Karlin, S. and D. Carmelli, 1975   Numerical studies on two-loci selection models with general viabilities. Theoret. Pop. Biol. **7**: 399–421.

Langley, C., 1977   Nonrandom associations between allozymes in natural populations of *Drosophila melanogaster*. pp. 265–274. In: *Measuring Selection in Natural Populations*. Edited by F. B. Christiansen and T. M. Fenchel. Springer-Verlag, New York.

Langley, C. H., Y. N. Tobari and K. I. Kojima, 1974   Linkage disequilibrium in natural populations of *Drosophila melanogaster*. Genetics **78**: 921–936.

Lewontin, R. C., 1964   The Interaction of selection and linkage. I. General considerations; heterotic models. Genetics **49**: 49–67.

Maynard Smith, J., 1978   *The Evolution of Sex*. Cambridge University Press, Cambridge.

Nagylaki, T., 1976   The evolution of one- and two-locus systems. Genetics **83**: 583–600.

Roberts, R. M. and W. K. Baker, 1973   Frequency distribution and linkage disequilibrium of active and null esterase isozymes in natural populations of *Drosophila montana*. Am. Naturalist **107**: 709–726.

Corresponding editor: W. J. Ewens