Balancing Selection at Closely Linked, Overdominant Loci in a Finite Population

Montgomery Slatkin

Department of Integrative Biology, University of California, Berkeley, California 94720-3140

Manuscript received July 23, 1999
Accepted for publication November 19, 1999

ABSTRACT

High levels of allelic diversity and strong linkage disequilibrium are found in the major histocompatibility (MHC) system in humans and other vertebrates. This article proposes several descriptive statistics that quantify the extent and pattern of strong linkage disequilibrium between pairs of highly polymorphic loci. It also develops an approximate analytic theory incorporating the effects of balancing selection, mutation, recombination, and genetic drift at two closely linked loci and compares the theoretical predictions with published surveys of the MHC class II loci, DQA1 and DQB1, in humans and nonhuman primates. The descriptive statistics proposed include the fraction of complementary haplotypes (haplotypes with $D' = 1$), the fraction of excess haplotypes, and the numbers of alleles at each locus in complementary haplotypes with one or more alleles at the other locus. The model assumes the infinite alleles model of mutation and the symmetric overdominance model of selection. Analytic approximations in some cases are obtained in the strong selection, weak mutation (SSWM) limit introduced by J. Gillespie. The predictions of the approximate analysis are confirmed by simulation. Both the analytic theory and simulations show that relatively few haplotypes will be found when selection is strong and recombination is weak relative to genetic drift. The model can reproduce many of the observed patterns at DQA1 and DQB1 provided that the recombination rate is assumed to be very small.

BALANCING selection resulting from either overdominance in fitness or genetic self-incompatibility maintains many more alleles at a locus than would be expected at a neutral locus with the same mutation rate. Several theoretical studies provide approximations that allow prediction of the number of alleles maintained by strong balancing selection in finite populations (Kimura and Crow 1964; Takahata 1990; Sasaki 1992; Vekemans and Slatkin 1994; Slatkin and Muirhead 1999). Balancing selection affecting two closely linked loci can maintain substantial allelic diversity at each locus and also maintain strong linkage disequilibrium between them, but the problem is much harder to analyze. There are many theoretical studies of linkage and selection affecting diallelic loci in infinite populations (Lewontin and Kojima 1960; Karlin and Feldman 1970) and a few studies of linkage and selection when one of the loci has more than two alleles (Christiansen and Feldman 1975; Feldman *et al.* 1975). There are also several studies of linkage and mutation in finite populations (Hill and Robertson 1968; Griffiths 1981). But there are no analytic studies that model the combined effects of selection, mutation, and linkage in finite populations. In this article, I develop an approximate analytic theory for two closely linked overdominant loci, each of which has a potentially large number of alleles. The theory predicts the numbers of alleles maintained at each locus and the pattern of allelic association in haplotypes. The theory is based on Gillespie's (1984) strong-selection, weak-mutation (SSWM) approximation, which has been shown to be useful for modeling balancing selection, mutation, and genetic drift at a single locus (Sasaki 1989, 1992; Slatkin and Muirhead 1999).

Linkage disequilibrium in MHC: The present theoretical study was motivated in part by observations of linkage disequilibrium in the major histocompatibility (MHC) region in humans and other vertebrates. The MHC region contains numerous loci that are closely linked and highly polymorphic. Particularly good evidence of strong linkage disequilibrium is found in the class II loci, DQA1 and DQB1, in humans and some other primates. In humans, these two loci are ~20 kb apart on the short arm of chromosome 6. Although no recombinants between them have been reported, the recombination rate can be estimated roughly to be 0.0002 if 1 Mb is assumed to correspond to a rate of 0.01 (1 cM). Recombination rate is known to vary with location in the MHC region in humans (Cullen *et al.* 1997), so the recombination rate between DQA1 and DQB1 might well be <0.0002.

At DQA1, 16 alleles are known, 8 of which are in appreciable frequency worldwide; at DQB1, 25 alleles are known, 15 of which are in appreciable frequency worldwide (Marsh 1998). The two loci code for poly-

Address for correspondence: Montgomery Slatkin, Department of Integrative Biology, University of California, 3060 Valley Life Sciences Bldg., Berkeley, CA 94720-3140. E-mail: slatkin@socrates.berkeley.edu

TABLE 1

DQA1-DQB1 haplotype numbers and D' values in CEPH families

DQB1 alleles	DQA1 alleles						Total DQB1
	0101	0102	0103	0201	0301	0501	
0201	0	0	0	21 (0.51)	0	23 (0.38)	44
0301	0	0	0	0	16 (0.16)	30 (0.55)	46
0302	0	0	0	0	39 (1.0)	1	40
0303	0	0	0	14 (1.0)	1	0	15
0501	24 (1.0)	0	0	0	0	0	24
0502	0	4 (1.0)	0	0	0	0	4
0503	12 (0.61)	0	0	0	0	6 (0.13)	18
0601	0	0	2 (1.0)	0	0	0	2
0602	0	41 (1.0)	0	0	0	0	41
0603	0	0	10 (1.0)	0	0	0	10
0604	0	9 (1.0)	0	0	0	0	9
Total DQA1	36	54	12	35	56	60	253

The table shows haplotype numbers and D' values for 253 independent chromosomes from Centre d'Etude du Polymorphism Humain (CEPH) families (Begovich *et al.* 1992, Table III). The D' values were computed after the two haplotypes found only in a single copy (0501-0302) and (0301-0303) were deleted. There are two multi-A alleles (0102 with $i = 3$ and 0103 with $i = 2$). There is a significant excess (at the 1% level, based on a χ^2 test) of all haplotypes found in more than one copy, except for 0501-0503.

peptides, DQα and DQβ, that form a heterodimer. Kwok *et al.* (1993) showed that alleles at each locus could be grouped and that haplotypes formed from alleles in incompatible groups did not produce stable cell surface heterodimers. These "prohibited" haplotypes had not been present in population surveys (Fernandez-Viña *et al.* 1991), but recently Grahovac *et al.* (1998) found several of the supposedly prohibited haplotypes in their study of Siberian populations.

In humans, relatively few haplotypes are present and there is significant excess of almost all of them. Linkage disequilibrium is not complete and some alleles at DQA1 are in strong disequilibrium with two or more alleles at DQB1 (Begovich *et al.* 1992). Some typical data are presented in Table 1 and will be discussed in more detail later. DQA1 and DQB1 in chimpanzees and gorillas show no significant linkage disequilibrium (Gyllensten and Erlich 1993). In the rhesus monkey (*Macaca mulata*) DQA1 and DQB1 are in very strong linkage disequilibrium, and the pattern is different from that found in humans (Sauermann 1998). In a sample of 134 individuals, Sauermann (1998, Table 4) found 14 different haplotypes present in more than one copy. Each of the 14 haplotypes contained alleles at both loci that were in no other haplotypes in the sample. Sauermann (1998) attributed the difference between the patterns in rhesus monkeys and in humans to stronger selection acting on these two loci in the rhesus monkey.

All of these studies are based on samples of individuals collected for other purposes, so the data cannot be regarded as representing random samples from well-defined populations. The studies of Gyllensten and Erlich (1993) and Sauerman (1998) were of individuals bred in captivity, so some of the chromosomes sampled were possibly identical by descent because of past inbreeding. The nonindependence of chromosomes would affect the formal statistical analysis of the data. In this article, the data are compared only qualitatively to theoretical predictions. They are used to illustrate the application of the theory rather than to draw strong conclusions about the operation of selection in the different species discussed. Although larger samples would be desirable, they are not currently available.

LINKAGE DISEQUILIBRIUM AT MULTIALLELIC LOCI

If there are only two alleles at each of two loci, the coefficient of linkage disequilibrium, D, defined by Lewontin and Kojima (1960), is sufficient to describe the degree of nonrandom association between those loci. All other measures of nonrandom association can be expressed as functions of D and the two allele frequencies. With more than two alleles per locus, no single measure of linkage disequilibrium provides a complete description; what measure or measures are best depends on the pattern in the data and the purpose of the analysis. If the goal is to detect nonrandom association and to test its significance, then a measure based on the χ^2 statistic, as proposed by Feldman *et al.* (1975) and others, or the probability value from Fisher's exact test, used by Peterson *et al.* (1995), will serve. If, instead, the goal is to focus on a particular allele at one locus and characterize the extent of disequilibrium with markers at another locus, then the δ statistic, suggested by Bengtsson and Thomson (1981), is more appropriate.

In other situations, there is obvious strong linkage

disequilibrium involving most or all alleles at both loci and there is no reason to focus on any particular allele at either locus. As in the data presented in Table 1, very few of the many possible haplotypes are found and almost every haplotype shows significant excess or deficiency. Furthermore, many or most of the values of Lewontin's D' (1964) statistic (defined to be the ratio of the coefficient of linkage disequilibrium to its maximum possible value) for haplotypes present in more than one copy are 1 or nearly 1. Very strong linkage disequilibrium of this kind can be generated by strong overdominant selection affecting both loci, as shown in the next several sections. To compare theory and data, some descriptive statistics are necessary to quantify this pattern.

Fraction of complementary haplotypes: I adapt the terminology of Feldman *et al.* (1975) and call haplotypes, which are present in more than a specified number of copies and for which $D' = 1$, *complementary* haplotypes. Allowing for a threshold is necessary because haplotypes found in only one or a few copies may be artifacts or aberrant for some reason. For the data in Table 1, I use a threshold value of two copies, which means that two haplotypes, 0501-0302 and 0301-0303, are ignored. One descriptive statistic is the fraction of haplotypes that are complementary. For the data in Table 1, there are 14 haplotypes found in more than one copy, and 8 of them have $D' = 1$, so that fraction, which is denoted by CF (for complementary fraction), is 0.571.

Fraction of extra haplotypes: Another descriptive statistic is the number of haplotypes relative to the total possible number. The minimum possible number of haplotypes is the larger of k_A and k_B, and is 11 for the data in Table 1. The maximum number is $k_A k_B$, which is 66 for the data in Table 1. An index describing the proportion of haplotypes greater than the minimum, which I call the fraction extra (FE) haplotypes, is $FE = (k_H - k_M)/(k_A k_B - k_M)$, where k_H is the number of haplotypes and k_M is the larger of k_A and k_B. This fraction is 0 if the minimum number is found and 1 if all possible haplotypes are found. For the data in Table 1, FE = 0.05.

Multi-A, multi-B, and unique haplotypes: It is useful to characterize the complementary haplotypes in which each allele is found. If an allele at A is in a complementary haplotype with only one allele at B, that haplotype will be called a *unique* haplotype, and the number of unique haplotypes will be denoted by *K*. If an allele at A is in complementary haplotypes with *i* alleles at B, that will be defined as one class *i* multi-A haplotype, indicating that one allele at A is associated with multiple alleles at B. The number of multi-A haplotypes will be denoted by *I*, and the fraction of multi-A haplotypes that are class *i* will be denoted by ϕ_i. If an allele at B is in complementary haplotypes with *j* alleles at A, that will be defined to be a class *j* multi-B haplotype, indicating that one allele at B is associated with multiple alleles at A. The number of multi-B haplotypes will be denoted by *J*, and the fraction that are class *j* will be denoted by ρ_j. An example of the use of this notation is given in Table 2.

TABLE 2

State space and notation for the model with complete linkage: haplotypes found in a hypothetical data set with 13 alleles at the A locus and 11 alleles at the B locus

A_1B_1, A_1B_2, A_1B_3 ($i = 3$)
A_2B_4, A_2B_5 ($i = 2$)
A_3B_6, A_4B_7, A_5B_8 ($K = 3$)
A_6B_9, A_7B_9, A_8B_9, $A_{10}B_9$ ($j = 4$)
$A_{11}B_{10}$, $A_{12}B_{10}$ ($j = 2$)
$A_{13}B_{11}$, $A_{14}B_{11}$ ($j = 2$)

In this case, $I = 2$, $J = 3$, $K = 3$. At the symmetric, deterministic equilibrium described in the text, the frequencies of gametes in each row containing more than one are equal. The frequency of those in the first row would be x_3, those in the second row would be x_2, those in the third row would be z, those in the fourth row would be y_4, those in the fifth and sixth rows would be y_2, and $\phi_2 = \phi_3 = 1/2$, $\rho_2 = 2/3$, and $\rho_4 = 1/3$.

In Table 1, $J = K = 0$, $I = 2$, and $\phi_2 = \phi_3 = 0.5$. The allele 0102 at DQA1 is found in complementary haplotypes with three different alleles at DQB1, and 0103 at DQA1 is found in complementary haplotypes with two different alleles at DQB1. Note 0101, 0201, and 0301 at DQA1 do not contribute to *I* because the 0101-0503, 0201-0201, and 0301-0301 haplotypes are not complementary. In Sauermann's (1998) study of rhesus monkeys, $I = J = 0$, and $K = 14$.

DETERMINISTIC THEORY OF SELECTION

Before considering the effects of mutation and genetic drift, I review and extend slightly results from the deterministic theory of overdominant selection and recombination. The selection model in this article is the symmetric overdominance model (Christiansen and Feldman 1975; Feldman *et al.* 1975), in which the relative fitness of doubly heterozygous individuals is 1, of doubly homozygous individuals is $1 - \alpha$, of individuals heterozygous only at the A locus is $1 - \beta$, and of individuals heterozygous only at the B locus is $1 - \gamma$. The recombination rate between the loci is *c*.

Two alleles per locus: This model with only two alleles at each locus has been analyzed by several authors beginning with Lewontin and Kojima (1960). We restrict our attention to the range of parameter values that seems most appropriate for the MHC region and assume α, β, and γ are all between 0 and 1 and further that $\alpha > \beta, \gamma$.

The quantity $\varepsilon = \beta + \gamma - \alpha$ plays an important role in the deterministic theory. If $\varepsilon > 0$, which is sometimes called *positive epistasis*, the fitnesses of the single homozy-

gotes are closer on average to the fitnesses of the double homozygotes than they are to the fitness of the double heterozygote. If $\varepsilon < 0$, *negative epistasis*, the opposite is the case. At present, there is no basis in saying whether loci in the MHC or other systems are better modeled by assuming positive or negative epistasis, but both existing deterministic theory and the new theory developed later for a finite population show that the sign of ε is important.

Lewontin and Kojima (1960) showed that, if $\varepsilon > 4c$, there is an equilibrium at which alleles at both loci are in frequency 1/2 and at which D is nonzero. Karlin and Feldman (1970) proved that this equilibrium is locally stable under the restricted ranges of values of α, β, and γ considered here. If $\varepsilon \leq 0$, no equilibrium with $D \neq 0$ exists for any $c > 0$.

More than two alleles per locus: Feldman *et al.* (1975) and Christiansen and Feldman (1975) studied this selection model when one locus has more than two alleles. These articles show that positive epistasis is necessary for the existence of stable equilibria with nonzero values of D. Christiansen and Feldman (1975, p. 195) found that, in a model with two alleles at one locus and m alleles at the other, stability of the equilibrium with only the haplotypes $A_1B_1, \ldots, A_1B_{m/2}, A_2B_{m/2+1}, \ldots, A_2B_m$ (m even) is possible only if $c < \varepsilon/(2m)$. The upper bound on c decreases inversely with the number of alleles at the B locus.

We can obtain a similar result for the generalization of the Lewontin and Kojima model to the case with K alleles at each locus. The symmetry of the problem suggests that the frequencies of all coupling haplotypes, $A_1B_1, \ldots, A_KB_K$, are the same, z, and the frequencies of all recombinant haplotypes (A_iB_m for $l \neq m$) are the same. There is always an equilibrium at which there is complete linkage equilibrium, $z = 1/K^2$, and it is easy to show that if $c < \varepsilon/(4(K-1))$, there is a second equilibrium at which there is permanent linkage disequilibrium:

$$z = \frac{1}{2K}\left(1 + \sqrt{1 - \frac{4c(K-1)}{\varepsilon}}\right). \qquad (1)$$

This result is similar to that of Christiansen and Feldman (1975) in that the condition becomes more restrictive as the number of alleles increases. In both models, a slight change in c or ε can eliminate the equilibrium at which permanent linkage disequilibrium is maintained, and the solution becomes more sensitive to changes in c or ε as the number of alleles increases.

Small recombination rates: Karlin and McGregor (1972) showed that the equilibrium solutions to a model of overdominant selection for small c are very close to those for the same model of selection with $c = 0$, and, if an equilibrium is locally stable for $c = 0$, then it will also be locally stable for sufficiently small values of c. The Karlin-McGregor theory can be applied to the equilibrium given by (1) to prove that it is locally stable for small values of c.

The Karlin-McGregor theory tells us that a thorough analysis of a deterministic model with $c = 0$ provides a complete understanding of the behavior of a deterministic model of selection with very small values of c. We will see that the Karlin-McGregor theory is a useful starting point for analyzing overdominant selection in finite populations, but that new results are found when genetic drift and mutation are included.

SSWM APPROXIMATION

We begin with the analysis of the case in which $c = 0$, for which an analytic approximation can be obtained, and then use simulations to determine the effects of small amounts of recombination. With $c = 0$, the model reduces to a model of selection at one locus, for which many general results are known (Nagylaki 1992).

Deterministic equilibria: Initially, we consider only those equilibria that are relevant when there is positive epistasis ($\varepsilon > 0$). With no recombination, a new allele at one locus will remain linked to the allele at the other locus that was on the chromosome on which it first appeared, and it cannot be joined by recombination to any other allele present when it appeared. Thus, if A_1 appears on a chromosome initially carrying B_1, and there are $m-1$ other alleles at the B locus, $B_2 \ldots B_m$, A_1 cannot later be on a chromosome carrying $B_2 \ldots B_m$. Mutation at B can then create new haplotypes carrying the mutant. For example, A_1B_{m+1} could be created by mutation from B_1 to B_{m+1}.

At a complementary equilibrium, all haplotypes are either multi-A haplotypes, multi-B haplotypes, or unique haplotypes. As before, I is the number of A alleles in multi-A haplotypes, J is the number of B alleles in multi-B haplotypes, K is the number of unique haplotypes, ϕ_i is the fraction of multi-A haplotypes that are class i, and ρ_j is the fraction of multi-B haplotypes that are class j.

The first step in the SSWM approximation is to find the equilibrium haplotype frequencies under selection alone. By symmetry, all the unique haplotypes will have the same equilibrium frequency, denoted by z, all class i multi-A haplotypes will have the same frequency, denoted by x_i, and all class j multi-B haplotypes will have the same frequency, denoted by y_j. These variables satisfy the condition that the haplotype frequencies sum to 1:

$$I\sum_{i\geq 2} i\phi_i x_i + J\sum_{j\geq 2} j\rho_j y_j + Kz = 1. \qquad (2)$$

The numbers of alleles are

$$k_A = K + I + J\sum_{j\geq 2} j\rho_j \qquad (3a)$$

and

$$k_B = K + J + I\sum_{i\geq 2} i\phi_i. \qquad (3b)$$

Given the selection parameters α, β, and γ, it is straightforward to show that

$$z = 1/\left[\alpha\left(I\sum_{i\geq 2}\frac{i\phi_i}{\alpha + (i-1)\gamma} + J\sum_{j\geq 2}\frac{j\rho_j}{\alpha + (j-1)\beta}\right) + K\right] \quad (4a)$$

$$x_i = \frac{\alpha z}{\alpha + (i-1)\gamma} \quad (4b)$$

$$y_j = \frac{\alpha z}{\alpha + (j-1)\beta}. \quad (4c)$$

Note that z is greater than both x_i and y_j and that x_i and y_j decrease with increasing values of their subscripts.

Invasion by new mutants: If $c = 0$, new haplotypes are created only by mutation, and under our assumption of the infinite alleles mutation model new haplotypes will carry a new allele at one locus. We consider in turn mutations at the A locus in each of the three kinds of haplotypes. First, consider a mutant A* to an A allele in a unique haplotype, say A_1B_1. When A^*B_1 is in very low frequency, ξ, it increases in frequency according to

$$\Delta\xi = (\alpha - \beta)z\xi, \quad (5)$$

which is positive under our restrictions on the values of α and β. If genetic drift is ignored, the A^*B_1 haplotype would increase in frequency until a new equilibrium is reached at which A_1B_1 and A^*B_1 are in equal frequency y_2, and the other frequencies are readjusted appropriately, thus reducing K by one and increasing J by one.

Next consider a mutation to an A that is in a class j multi-B haplotype. If one of the A's in the group $A_1B_1 \ldots A_jB_1$ mutates to A*, the frequency of A^*B_1, ξ will increase when rare according to

$$\Delta\xi = (\alpha - \beta)y_j\xi. \quad (6)$$

In this case, J will be unchanged but B_1 will move from class j to class $j + 1$, and the haplotype frequencies will change.

If the A that mutates is in a class i multi-A haplotype, say $A_1B_1 \ldots A_1B_i$, the frequency of the new haplotype, say A^*B_1, will increase when rare according to

$$\Delta\xi = (\alpha z - \beta x_i)\xi = (\alpha + (i-1)\gamma - \beta)x_i\xi. \quad (7)$$

Once A^*B_1 becomes common, then A_1B_1 will be at a disadvantage provided that

$$\varepsilon x_{i-1} > \beta(x_{i-1} - z), \quad (8)$$

where we interpret (8) as $\varepsilon > 0$ for $i = 2$ and note that the right-hand side is negative for $i > 2$. Condition (8) is obtained by finding the rate of increase of A_1B_i from a very low frequency once A^*B_1 and the $A_1B_1 \ldots A_1B_{i-1}$ have reached their new equilibrium frequencies. The general theory of one-locus models tells us that if a rare allele cannot increase in frequency when rare, it cannot be present at a stable polymorphic equilibrium (Nagylaki 1992). Therefore, if $\varepsilon > 0$, then after a new haplotype becomes common, selection will quickly reestablish a complementary equilibrium. At the new equilibrium, there is a new unique haplotype, A^*B_1, thereby increasing K by 1, and A_1 is in class $i - 1$. If $i = 2$, a second unique haplotype, A_1B_2, remains, with the result that K is increased by 2 and I is decreased by 1.

Similar results apply to mutations at the B locus in unique haplotypes, multi-A haplotypes, and multi-B haplotypes.

If $\varepsilon < 0$ and there is no genetic drift, invasion by a mutation will not preserve complementarity and the method described in the next section will not work. The approximate analysis can suggest what will happen when $\varepsilon < 0$ but at present an analytic treatment of that case does not seem possible because the solution for the deterministic equilibrium corresponding to (4) is difficult to obtain.

Finite population size: We now develop an approximation for a large finite population of constant size in which selection is strong enough that haplotype frequencies are nearly at their deterministic equilibrium values. The overall approach is to model the population as a large Markov chain with state variables I, J, K, ϕ_i, and ρ_j. We will see later that the dynamics of the multi-A and multi-B haplotypes can be analyzed separately, thus simplifying the analysis considerably. For reasons discussed in the previous sections, we restrict the parameters so that $1 > \alpha > \beta$, $\gamma > 0$ and $\varepsilon = \beta + \gamma - \alpha > 0$.

We assume that the population is of constant size N and that the mutation rates at the A and B loci are u and v, respectively. The probabilities that alleles in different types of haplotypes mutate are the frequencies of those haplotypes in the population, and the probability that a new haplotype increases in frequency and becomes common is approximately twice its deterministic rate of increase when rare.

Considering first the A locus, the probability that the copy of A that mutates is in a unique haplotype is Kz. The probability that the new haplotype will increase from one copy to become common is, from (5), $\sim 2(\alpha - \beta)z$. Therefore, the rate per generation at which mutations of this type create haplotypes that become common is approximately

$$\lambda_{A1} = (2Nu)(Kz)(2(\alpha - \beta)z) = 4NuK(\alpha - \beta)z^2, \quad (9a)$$

where λ denotes the probability of increase of I or J, the subscript A indicates that the increase is the result of a mutation at the A locus, and the subscript 1 indicates that the A that mutated was associated with only one B allele. The result of this transition is an increase in J by 1, a decrease in K by 1, an increase in ρ_2, and a corresponding decrease in ρ_j for $j > 2$. Similarly,

$$\lambda_{B1} = 4NvK(\alpha - \gamma)z^2, \quad (9b)$$

which is the probability of an increase in I by 1, a de-

crease in K by 1, an increase in ϕ_2, and a corresponding decrease in ϕ_i for $i > 2$.

For an A allele in a class j multi-B haplotype, the net frequency of such A alleles is $Jj\rho_j y_j$ and the probability of increase from a single copy is, from (6), $\sim 2(\alpha - \beta)y_j$. Therefore, the rate at which such haplotypes arise and become common is

$$\lambda_{Aj} = (2Nu)(Jj\rho_j y_j)(2(\alpha - \beta)y_j) = 4NuJ(\alpha - \beta)j\rho_j y_j^2. \quad (10)$$

In this case, there is no change in I, J, or K, but an increase in j by one. Similarly, the rate at which mutant B's in class i multi-A haplotypes increase and become common is approximately

$$\lambda_{Bi} = 4NvI(\alpha - \gamma)i\phi_i x_i^2. \quad (11)$$

We next consider transitions caused by mutations to new A alleles in class i multi-A haplotypes. The frequency of such haplotypes is $Ii\phi_i x_i$, and if a mutation occurs, the probability that the new haplotype will increase to become common is, from (7), $\sim 2(\alpha + (i - 1)\gamma - \beta)x_i$. If the new haplotype becomes common, the previously common haplotype from which it arose will become disadvantageous and quickly lost provided that $\varepsilon > 0$. Therefore, i will be reduced by 1 and K will be increased by 1. If $i = 2$, then K will be increased by 2 and I will be reduced by 1. The probability of this transition per generation is

$$\begin{aligned}\eta_{Ai} &= (2Nu)(Ii\phi_i x_i)2(\alpha + (i - 1)\gamma - \beta)x_i \\ &= 4NuI(\alpha + (i - 1)\gamma - \beta)\phi_i x_i^2. \end{aligned} \quad (12)$$

Similarly, the probability that a mutation of a B allele in a class j multi-B haplotype increases to become common is approximately

$$\eta_{Bj} = 4NvJ(\alpha + (j - 1)\beta - \gamma)j\rho_j y_j^2. \quad (13)$$

In this case, j would be reduced by 1 and K would be increased by 1 if $j > 2$, and K would be increased by 2 and J reduced by 1 if $j = 2$.

In a finite population, each haplotype is subject to stochastic loss because of genetic drift. The rate at which such loss occurs is, as in the one-locus model, the inverse of the average time to loss. For this model, the recursion equation for the deviation, f, of any haplotype from its equilibrium frequency is the same:

$$f' = (1 - \alpha)f. \quad (14)$$

Takahata's (1990) result for the average time to loss in a population of size N for the one-locus model tells us that the probability of loss per generation of a haplotype in frequency ξ because of genetic drift is approximately

$$\frac{\xi^2}{2N} e^{-S\xi^2} \sqrt{\frac{S^3}{2\pi}}, \quad (15)$$

where $S = 2N\alpha$. As discussed by Slatkin and Muirhead (1999), this formula differs by a factor of $\sqrt{2}$ from that used by Sasaki (1989, 1992). Although Sasaki's formula is correct for very large values of S, we found that Equation 15 is slightly more accurate when S is on the order of 100. The numerical results are nearly the same when Sasaki's formula is used instead of (15). Note that the exponential dependence on S/ξ^2 implies that slight differences in ξ result in very different rates of loss by genetic drift.

To compute the net rate of loss of each haplotype per generation, we multiply by the number of different haplotypes. There are K unique haplotypes and each is in frequency z, so the net rate of loss is

$$\zeta_z = \frac{Kz^2}{2N} e^{-Sz^2} \sqrt{\frac{S^3}{2\pi}}. \quad (16)$$

When such a loss occurs K is reduced by 1.

For the class i multi-A haplotypes, their total number is $Ii\phi_i$, and the frequency of each is x_i. Therefore, the probability of loss per generation is

$$\zeta_{xi} = Ii\phi_i x_i^2 e^{-Sx_i^2} \sqrt{\frac{S^3}{2\pi}}. \quad (17)$$

When such a loss occurs, i is reduced by 1; if $i = 2$, then I is reduced by 1 and K is increased by 1. Similarly, the rate of loss of class j multi-B haplotypes because of genetic drift is

$$\zeta_{yj} = Jj\rho_j y_j^2 e^{-Sy_j^2} \sqrt{\frac{S^3}{2\pi}}, \quad (18)$$

and when this loss occurs, j is reduced by 1, and if $j = 2$, J is reduced by 1 and K is increased by 1.

Calculating ϕ_i and ρ_j: We can find the stationary distributions of i, ϕ_i, by noting that, once a class $i = 2$ multi-A haplotype is created by mutation, subsequent changes in i can be modeled by a separate Markov chain. The value of i can increase because of mutation at the B locus, and the probability of increase per generation is

$$\Lambda_i = 4Nv(\alpha - \gamma)ix_i^2, \quad (19)$$

which is Equation 11 divided by $I\phi_i$. The value of i can decrease either because of mutation of one of the A alleles or because of genetic drift. The probability of loss by mutation is $\eta_{Ai}/(I\phi_i)$ and the probability of loss by drift is $\xi_{xi}/(I\phi_i)$. Thus, the net probability of a decrease in i is the sum,

$$M_i = 4Nu(\alpha + (i - 1)\gamma - \beta)ix_i^2 + \frac{ix_i^2}{2N} e^{-Sx_i^2} \sqrt{\frac{S^3}{2\pi}}. \quad (20)$$

Therefore, the Markov chain for changes in the value of i for $i \geq 2$ has a particularly simple form for which many analytic results are known. The values of M_i and Λ_i are such that i cannot increase indefinitely, so ultimate loss from $i = 2$ is certain. In that case, the theory summarized by Ewens (1979, Equation 2.137 with Λ_i and M_i replacing λ_i and μ_i) provides the sojourn times,

which are the average times spent in each value of i, $\bar{t}_{i2}$, given that $i = 2$ initially. The stationary distribution of i, ϕ_i, is then

$$\phi_i = \frac{\bar{t}_{i2}}{\sum_{i\geq 2}\bar{t}_{i2}}. \tag{21}$$

Although both Λ_i and M_i are proportional to x_i^2 and hence depend on I, J, and K, the factor x_i^2 cancels from the resulting expression for ϕ_i, which then depends only on the selection and mutation parameters and the population size. In principle, there is no upper bound on i, but in the numerical analysis of this model described below, relatively small upper bounds can be assumed because ϕ_i decreases very rapidly with i. A similar analysis yields ρ_j with appropriate redefinition of Λ_j and M_j.

Zero flux condition: Given that ϕ_i and ρ_j have reached their stationary distributions, equilibrium values of I, J, and K will be such that the fluxes into and out of the three categories of haplotypes are balanced. For multi-A haplotypes,

$$\lambda_{B1} = (\eta_{A2} + \zeta_{x2}) + \sum_{i\geq 2}\eta_{Ai}; \tag{22a}$$

for multi-B haplotypes,

$$\lambda_{A1} = (\eta_{B2} + \zeta_{y2}) + \sum_{j\geq 2}\eta_{Bj}; \tag{22b}$$

and for unique haplotypes,

$$\lambda_{A1} + \lambda_{B1} + \zeta_z = \eta_{A2} + \sum_{j\geq 2}\eta_{Ai} + \eta_{B2} + \sum_{j\geq 2}\eta_{Bj} + \zeta_{x2} + \zeta_{y2},$$

which, using (22a) and (22b), can be reduced to

$$\zeta_z = \sum_{i\geq 2}\eta_{Ai} + \sum_{j\geq 2}\eta_{Bj}. \tag{22c}$$

These three equations can be solved for I, J, and K for various values of the selection parameters, mutation rates, and population size. These equations can be shown to depend on the selection and mutation parameters scaled by N: $2N\alpha$, $2N\beta$, $2N\gamma$, $4Nu$, and $4Nv$. Although they appear difficult to solve, they have a simple structure that makes numerical analysis easy. Equations 22a–c form a linear homogeneous set of equations for I, J, and K. The general theory of such equations tells us that there is no nontrivial solution unless the determinant of the coefficient matrix is 0. The value of z determines the values of x_i, y_j, ϕ_i, and ρ_j, so the coefficient matrix depends only on z. The problem is then reduced to finding the value of z that makes that determinant 0. Once that value of z is found, any two of the equations plus the normalization condition, (3), form a linear inhomogeneous system for I, J, and K that is trivial to solve. The numerical analysis can be done in a few seconds using a Mathematica (Wolfram 1996) program that will be provided upon request to the author.

Numerical results for $c = 0$: Given the five combinations of parameters ($2N\alpha$, $2N\beta$, $2N\gamma$, $4Nu$, and $4Nv$), the approximate equations described in the previous section yield I, J, and K, as well as the x_i, y_j, ϕ_i, and ρ_j. Although there are many possible combinations of parameters, and many patterns in the results can be envisioned, the model's predictions are relatively simple: one of two patterns is found. If the single-locus homozygotes have roughly the same relative fitnesses, meaning that β and γ have approximately the same values, then almost all haplotypes will be unique and very few will be multi-A or multi-B haplotypes. Furthermore, the few multi-A and multi-B haplotypes that are present will almost all be in class 2. Some typical results are shown in Table 3, A–C. That pattern is found even if there are substantial differences in the mutation rates. We can see why by recalling the very strong dependence of rate of haplotype loss by genetic drift, ζ_{xi} and ζ_{yj} in Equations 17 and 18, on the haplotype frequencies. If $\varepsilon > 0$ and β and γ are comparable in magnitude, then (4b) and (4c) imply that x_i and y_j are less than z and decrease with i and j. That ensures that multi-A and multi-B haplotypes will be lost to drift at a much higher rate than will unique haplotypes, especially if i or j exceeds 2. If the mutation rates differ, that causes only a linear difference in the λ's, while there is highly nonlinear dependence of the ζ's on x_i and y_j. As a consequence, the numbers of alleles at each locus are roughly equal. This tendency does not depend on exact equality of β and γ, but only on the fact that they are of the same order of magnitude.

This explanation for the pattern found when β and γ are comparable in magnitude suggests when another pattern may be found. If γ is quite small, then (4) implies that x_i is nearly equal to z and decreases only slightly with increasing i. When γ is small, the condition $\varepsilon > 0$ requires that β then be nearly equal to α. In this case, the symmetry disappears and a substantial number of multi-A haplotypes can be maintained at equilibrium. At this equilibrium, there are almost no multi-B haplotypes and the number of alleles at B will exceed the number at A. A typical example is shown in Table 3D.

RECOMBINATION

We next consider small amounts of recombination. The Karlin and McGregor (1972) theory indicates that the stability of the equilibrium for $c = 0$ is important.

Stability of deterministic equilibria: The complementary equilibrium defined by (4) is feasible, meaning that all the haplotype frequencies are between 0 and 1 for all values of α, β, and γ between 0 and 1. But it is not necessarily stable to invasion by noncomplementary haplotypes created by recombination.

The local stability of (4) can be determined by finding whether haplotypes of different kinds will increase in frequency when rare. If even one noncomplementary haplotype can increase, the equilibrium is unstable. There are nine kinds of noncomplementary haplotypes

TABLE 3

Comparison of simulation and analytic results for $\varepsilon > 0$

	I	J	K	k_A	k_B	ϕ_2	ρ_2	CF	FE
A.			$2N\alpha = 500$, $2N\beta = 2N\gamma = 300$, $4Nu = 0.05$, $4Nv = 0.2$						
Analytic	0.88	0.26	7.18	8.6	9.2	0.98	0.997	1	0
$2Nc = 0$	1.4	0.39	7.4	10.0	11.2	0.87	1.0	0.987	0.004
$2Nc = 0.1$	1.5	0.6	5.3	9.9	11.3	0.94	1.0	0.894	0.015
$2Nc = 1$	0.5	0.06	0.6	9.2	11.4	0.86	1.0	0.460	0.080
$2Nc = 10$	0	0	0.06	9.2	11.4	—	—	0.087	0.279
B.			$2N\alpha = 100$, $2N\beta = 2N\gamma = 60$, $4Nu = 0.05$, $4Nv = 0.2$						
Analytic	0.55	0.17	3.32	4.21	4.60	0.96	0.99	1	0
$2Nc = 0$	1.4	0.1	2.5	4.7	6.1	0.88	1.0	0.967	0.011
$2Nc = 0.1$	0.78	0.28	2.4	5.3	6.3	1.0	1.0	0.883	0.036
$2Nc = 1$	0	0	0	4.9	6.1	—	—	0.183	0.310
C.			$2N\alpha = 500$, $2N\beta = 2N\gamma = 450$, $4Nu = 0.05$, $4Nv = 0.2$						
Analytic	0.25	0.07	7.38	7.77	7.95	0.997	0.999	1	0
$2Nc = 0$	0.89	0.17	8.1	9.4	10.2	0.91	0.67	0.996	0.002
$2Nc = 0.1$	0.61	0.33	6.8	8.6	9.2	0.90	1.0	0.962	0.005
$2Nc = 1$	0.61	0.39	3.5	9.3	10.0	0.90	1.0	0.732	0.037
$2Nc = 5$	0.56	0	0.11	10.5	12.1	1	—	0.096	0.201
D.			$2N\alpha = 500$, $2N\beta = 475$, $2N\gamma = 50$, $4Nu = 4Nv = 0.2$						
Analytic	2.75	0.03	4.14	7.0	10.5	0.71	0.999	1	0
$2Nc = 0$	3.3	0	4.4	8.1	12.6	0.67	—	0.980	0.002
$2Nc = 0.1$	2.0	0.06	3.2	7.8	12.3	0.61	1.0	0.830	0.019
$2Nc = 1$	0.11	0	0.28	5.8	12.9	1.0	—	0.384	0.112
$2Nc = 5$	0	0	0	5.4	12.3	—	—	0.108	0.292

In all simulations $N = 5000$. Each set of parameter values was run for 200,000 generations and data were accumulated every 10,000 generations beginning in generation 30,000.

that can be created by recombination, depending on the common haplotypes the A and B alleles are in, but the analysis of each case is not difficult. To illustrate, let ξ be the frequency of a rare haplotype containing alleles at both loci that are otherwise found in unique haplotypes. For example, if the rare haplotype is A_1B_1, and, at the equilibrium being analyzed, A_1 is in a unique haplotype with some other B, say B_2, and B_1 is in a unique haplotype with some other A, say A_2, it is easy to show that A_1B_1will decrease in frequency if $\varepsilon > 0$ and will increase in frequency if the inequality is reversed.

In a similar way, the condition $\varepsilon > 0$ ensures that the recombinant haplotype will decrease in frequency when the A allele is otherwise in a unique haplotype and the B allele is otherwise in a class j multi-B haplotype, or when the B allele is in a unique haplotype and the A allele is in a class i multi-A haplotype. The same result also holds when the A allele of the rare haplotype is in a class i multi-A haplotype and the B allele is in a class j multi-B haplotype.

However, local stability is not guaranteed by $\varepsilon > 0$ for the other noncomplementary haplotypes. For example, if the B allele of the rare haplotype is in a unique haplotype and the A allele is in a class j multi-B haplotype, the rare haplotype will increase in frequency if

$$j > \frac{1}{\beta}\left(\frac{\alpha\gamma}{\alpha - \beta} + \beta - \alpha\right). \tag{23a}$$

Similarly, if the A allele of the rare haplotype is in a unique haplotype and the B allele is in a class i multi-A haplotype, then the rare haplotype will increase in frequency if

$$i > \frac{1}{\gamma}\left(\frac{\alpha\beta}{\alpha - \gamma} + \gamma - \alpha\right). \tag{23b}$$

These two conditions imply that $\varepsilon > 0$ is not sufficient for (4) to be locally stable. For (23a) and (23b) to be satisfied for $i \geq 2$ and $j \geq 2$, both β and γ must be close to α. For example, if $\alpha = 0.05$, $\beta = \gamma = 0.03$ (Table 3A), then the right-hand sides of (23a and b) are both 1.83. If, instead, $\beta = \gamma = 0.045$ (Table 3C), the critical values of i and j increase to 3.08. In general, larger values of ε permit local stability of (4) for larger values of i and j, but only with very large values of α are much larger values of i and j possible in a stable equilibrium. If β is close to α but γ is quite small, the range of values of i and j for which (4) is stable is also restricted. For example, with $\alpha = 0.05$, $\beta = 0.0475$, and $\gamma = 0.005$ (Table 3D), the right-hand sides of (23a) and (23b) are 1.56 and 2.05, respectively.

These results suggest that the complementary equilibria predicted when $c = 0$ are not of biological interest because they are not locally stable. Simulation results presented below show otherwise. With very low recombination rates, recombinant haplotypes that could destabilize the equilibrium appear only very infrequently, and their rate of increase when rare is very small, meaning that they would have a very small chance of becoming common. As a consequence, complementary equilibria will be found in finite populations even if they would be unstable in infinitely large populations.

SIMULATION RESULTS

The predictions of the approximate theory were compared with those from a computer simulation of the same model. The computer model used a rejection scheme to simulate overdominant selection. In all cases shown in the tables, a population size of $N = 5000$ was used. For each set of parameter values, samples from the population were taken every 10,000 generations, beginning in generation 30,000 and continuing to generation 200,000. Results presented in the tables are values averaged over those 18 samples. Other simulations showed that this sampling scheme was sufficient to obtain results close to those from simulations carried out for much longer times.

In the calculation of the various statistics shown, it was necessary to choose a minimum number of copies of a haplotype counted as a common haplotype. The results shown in Tables 3 and 4 are for the entire population. Haplotypes were counted in the analysis only if >100 copies were present in the population. That number is less than the number of haplotypes expected at the deterministic equilibrium frequencies Equation 4, which were in the range 0.01–0.04 for the parameter values used in Tables 3 and 4. As discussed below, the choice of that minimum number affects the results somewhat but does not change the overall patterns found. In Table 5, samples of 253 chromosomes were chosen at random without replacement from each data set for comparison with the data shown in Table 1, and in those cases a minimum value of 2 was used, as in the analysis of the data in Table 1.

Positive epistasis: If $\varepsilon > 0$, the simulation results can be compared to predictions from the analytic theory. Some results are shown in Table 3. For $c = 0$, the simulation results agree roughly with the analytic predictions. When the threshold haplotype number is 100, there tend to be slightly more multi-A and multi-B haplotypes than predicted. If that threshold is raised to 200, the simulation results are slightly different. For example, in Table 3A, if a threshold value of 200 copies is used, the averages of I, J, and K in the simulations for $c = 0$ are 0.89, 0.39, and 8.2, respectively.

When c is of the same order of magnitude as u and v, the simulation results are still similar to analytic predictions. That is consistent with the Karlin and McGregor (1972) theory. When c is an order of magnitude larger than u and v, the simulation results no longer agree with the analytic theory. Fewer complementary haplotypes remain. Nevertheless, there is still substantial linkage disequilibrium present, many of the possible haplotypes are still not found in appreciable frequencies, and many complementary haplotypes are still present.

Negative epistasis: If $\varepsilon < 0$, the analytic theory no longer provides an approximate solution but it does suggest what will be seen. Even without recombination, some mutations in multi-A or multi-B haplotypes can generate noncomplementary haplotypes that will not be eliminated by selection. For example, if for a class 2 multi-A haplotype, A_1B_1 and A_1B_2, an A_1 in A_1B_1 mutates to A*, A^*B_1 has a chance of increasing in frequency to become common but A_1B_1 will not be eliminated. However, A_1B_1 is still liable to be lost because of drift, so complementarity can be restored. Although the solution for the deterministic equilibrium in that case cannot be obtained in a simple form, it is easy to see that the frequency of A_1B_1 at the deterministic equilibrium

TABLE 4

Simulation results for $\varepsilon < 0$

	I	J	K	k_A	k_B	ϕ_2	ρ_2	CF	FE
A.			$2N\alpha = 500$, $2N\beta = 2N\gamma = 50$, $4Nu = 0.05$, $4Nv = 0.2$						
$2Nc = 0$	2.8	0.56	2.4	8.8	12.9	0.57	0.94	0.911	0.002
$2Nc = 0.1$	0.61	0.28	0.89	7.7	12.1	0.69	1.0	0.595	0.019
$2Nc = 1$	0	0	0	6.3	9.9	—	—	0.164	0.112
$2Nc = 5$	0	0	0	6.2	8.7	—	—	0.057	0.432
B.			$2N\alpha = 500$, $2N\beta = 400$, $2N\gamma = 50$, $4Nu = 4Nv = 0.1$						
$2Nc = 0$	2.7	0.28	4.8	8.6	12.2	0.71	1.0	0.988	0.005
$2Nc = 0.1$	1.9	0.11	3.1	7.4	12.1	0.69	1.0	0.854	0.018
$2Nc = 1$	0.06	0	0.28	6.6	12.1	1.0	—	0.342	0.123
$2Nc = 5$	0	0	0.56	5.9	11.8	—	—	0.100	0.293

is lower than the frequencies of either A^*B_1 or A_1B_2 because it forms single homozygotes with both of them, and hence A_1B_1 will have a higher rate of loss by drift. The magnitude of this effect cannot be predicted but we can expect that substantial complementarity can be found if $\varepsilon < 0$ and that a higher degree of complementarity will be found if ε is negative but small in absolute value. These patterns are seen in Table 4 for $c = 0$.

For cases with $\varepsilon < 0$ and $c > 0$, the Karlin and McGregor (1972) theory is not useful for predicting patterns found in the simulations. If $\varepsilon < 0$, a completely complementary deterministic equilibrium with linkage disequilibrium is unstable to invasion by recombinant haplotypes, so the Karlin-McGregor theory predicts that such an equilibrium should be structurally unstable to increases in c from 0. That is not seen in the simulation results. Extensive linkage disequilibrium is found for small values of c even when ε is relatively large in absolute value ($\varepsilon = -0.04$ in Table 4A). The loss of complementarity with increasing c is more rapid than in the cases with $\varepsilon > 0$, because selection is no longer acting to maintain it. But the observation of substantial linkage disequilibrium and relatively high complementarity does not imply that ε is positive.

Small sample sizes: The simulation results shown in Tables 3 and 4 were chosen to illustrate the patterns found in the whole population. I do not consider the general sampling problem but present results for samples of the same size as in Table 1, to determine what kinds of selection are consistent with these observations. We can restrict our attention to cases that exhibit some of the general features of the data in Table 1, particularly the difference in the numbers of alleles at the two loci, the relatively small number of haplotypes found, the high degree of complementarity, and the presence of multi-A but no multi-B or unique haplotypes. None of the cases in which $\varepsilon > 0$ and β and γ are comparable in magnitude, including those shown in Table 3, A–C, predict those patterns. The cases shown in Table 3D and in Table 4 are similar enough to provide an adequate comparison with Table 1.

Table 5 shows that many of the features found in the data in Table 1 can be reproduced by this model. Values of c of the same order of magnitude as the mutation rates give the closest agreement. If ε is positive (Table 5A) or only slightly negative (Table 5C), strong asymmetry in selection on the single homozygotes ($\beta \gg \gamma$) is required, but if ε is much <0, then a modest asymmetry in mutation rates (Table 5B, with $v = 4u$) produces similar results even with $\beta = \gamma$. In all these cases, we see differences in the numbers of alleles maintained, 0 or nearly 0 values of J, relatively small values of ϕ_2 (indicating relatively large values of ϕ_i for $i > 2$), and values of both CF and FE that are roughly comparable to those found in the data. The simulation results differ from Table 1 by predicting nearly the same values of I and K, while in Table 1, $I = 2$ and $K = 0$.

For the data of Sauermann (1998) for the rhesus monkey, the pattern is very different. All 14 haplotypes are unique haplotypes, so $K = 14$, $I = J = 0$, and there is perfect complementarity (CF = 1) and no extra haplotypes (FE = 0). That pattern is found if ε is positive and relatively large, which means that β and γ are comparable in magnitude and each nearly as large as α, and if c is very small.

DISCUSSION AND CONCLUSIONS

This article has three goals. The first is to present new ways of quantifying the pattern of linkage disequilibrium between two multiallelic loci when there is obvious sig-

TABLE 5

Simulation results for samples of 253 chromosomes

	I	J	K	k_A	k_B	ϕ_2	CF	FE
A. Table 3D			$2N\alpha = 500$, $2N\beta = 475$, $2N\gamma = 50$, $4Nu = 4Nv = 0.2$					
$2Nc = 0$	3.2	0	4.4	8.2	12.6	0.68	0.98	0.003
$2Nc = 0.1$	1.8	0.06	2.8	7.8	12.6	0.65	0.82	0.021
$2Nc = 1$	0	0	0.11	6.0	13.1	—	0.28	0.145
B. Table 4A			$2N\alpha = 500$, $2N\beta = 2N\gamma = 50$, $4Nu = 0.05$, $Nv = 0.2$					
$2Nc = 0$	2.8	0.5	2.4	8.8	13.3	0.54	0.90	0.018
$2Nc = 0.1$	0.61	0.28	0.89	7.9	12.3	0.61	0.57	0.069
$2Nc = 1$	0	0	0	6.3	10.0	—	0.147	0.268
C. Table 4B			$2N\alpha = 500$, $2N\beta = 400$, $2N\gamma = 50$, $4Nu = 4Nv = 0.1$					
$2Nc = 0$	2.7	0.4	4.5	8.8	12.2	0.73	0.97	0.007
$2Nc = 0.1$	1.7	0.06	2.4	8.2	13.5	0.72	0.80	0.024
$2Nc = 1$	0.06	0	0.22	6.9	12.5	1	0.29	0.146

These results were obtained by randomly sampling 253 chromosomes from each data matrix in the corresponding cases in Tables 3 and 4.

nificant nonrandom association between them. The second is to develop an approximate analytic model that accounts for selection and mutation at two multiallelic loci in a finite population. The third is to determine the extent to which a simple highly symmetric model of selection can account for patterns found in data from the DQA1 and DQB1 loci in the MHC regions of humans and nonhuman primates.

Quantifying the extent of disequilibrium when there is strong balancing selection requires new approaches because it is not enough to show that there is significant disequilibrium. In many cases, linkage disequilibrium is so apparent that statistical tests are unnecessary. The problem instead is to distinguish between the many kinds of patterns that could be found. The approach taken here is based on the theoretical result that with very low recombination rates, selection, mutation, and drift will together create and maintain complementarity, with relatively few haplotypes found in appreciable numbers and many or most of them having Lewontin's $D' = 1$. That theoretical result suggests that FE and CF indicate how close to complete complementarity a particular data set is. If high complementarity is found, then the numbers of alleles at one locus in haplotypes with each allele at the other locus, measured by I, J, K, ϕ_i, and ρ_j, are of interest because their values are predicted by the analytic model.

The analytic model predicts the simulation results with reasonable accuracy when there is positive epistasis ($\varepsilon > 0$) and low rates of recombination. When there is positive epistasis, one of two patterns is found. If the fitnesses of the single-locus homozygotes are comparable (β and γ have similar magnitudes), then at an equilibrium, most haplotypes are unique and roughly equal numbers of alleles are found at both loci even if mutation rates differ. Only if one of the single-locus homozygotes is much less fit than the other ($\beta \gg \gamma$, but still with $\varepsilon > 0$) are the results asymmetric, with an appreciable number of multi-A or multi-B haplotypes and a substantial difference in the numbers of alleles at the two loci.

If $\varepsilon < 0$, the analytic model can no longer be used but the development of that model suggests that nearly complete complementarity will be found for small enough recombination rates. Simulation results confirmed that prediction, but with two differences from the cases with $\varepsilon > 0$. Recombination is more effective in destroying complementarity and differences in the mutation rates can more easily lead to asymmetry of the results.

Both the analytic and simulation models are not intended to provide a complete description of DQA1 and DQB1 or any other pair of loci. The model was designed to have as few parameters as possible. Selection is highly symmetric, with the many possible genotypes grouped into only four fitness classes. There are no specific interactions among alleles, as has been demonstrated for DQA1 and DQB1 by Kwok *et al.* (1993), no allowance for the possibility that some alleles may carry intrinsic disadvantages, as modeled by Slatkin and Muirhead (1999), and no possibility that mutation rates to different alleles or classes of alleles at each locus might differ. The goal in comparing data to theory is to see how good the predictions of this simple model are and where apparent differences lie.

For Sauermann's (1998) data for the rhesus monkey, the theory fits the data well provided that selection is strong, the fitnesses of the single-locus homozygotes are comparable and nearly as small as the fitnesses of the double homozygotes, and the recombination rate is low, not much larger than the mutation rates to functionally different alleles at each locus. The agreement between theory and data does not mean that the model's assumptions are valid for DQA1 and DQB1 in rhesus monkeys, only that no additional assumptions are necessary to account for the observations.

The fit of the model's predictions to the data in Table 1 is less good. Many of the features can be reproduced by the model but it does not seem possible to predict more multi-A haplotypes than unique haplotypes. That difference would probably require a more complex model of selection that embodied the functional restrictions of the type found by Kwok *et al.* (1993). Still, the parameter values needed to achieve the degree of asymmetry found in the data all have in common that one or both of the single-locus homozygotes must have a fitness comparable to the double homozygote. Not enough is known about the effect of these two loci on fitness to know whether that is true for DQA1 and DQB1 in humans, but at least the model does suggest that there are differences in the fitness interactions between these loci in humans and in the rhesus monkey.

For chimpanzees and gorillas, Gyllensten and Erlich (1993) found no significant linkage disequilibrium. They suggested that either the recombination rate is much higher or selection is weaker than in humans. The theory described here supports that view but shows that differences in either of these parameters would have to be by an order of magnitude or more. Small differences from the values in humans would not be sufficient to eliminate linkage disequilibrium. A difference in effective population size could not account for the patterns observed, because both the recombination rates and selection coefficients are scaled by the same factor, so the relative intensities of selection and recombination would remain the same.

These conclusions about the action of selection at DQA1 and DQB1 have to be regarded as tentative because there are too few population-level studies of haplotype frequencies in populations of humans and nonhuman primates to know whether those patterns are found in other populations as well. The data are discussed to illustrate the application of the theory and to show that population-level data can be used to test hypotheses about selection.

I thank S. McWeeney for providing information and references about linkage disequilibrium in DQA and DQB and M. W. Feldman, S. McWeeney, and G. Thomson for helpful discussions and advice. This research was supported in part by National Institutes of Health grant GM-40282.

LITERATURE CITED

Begovich, A. B., G. R. McClure, V. C. Suraj, R. C. Helmuth, N. Fildes *et al.*, 1992 Polymorphism, recombination, and linkage disequilibrium within the HLA class II region. J. Immunol. **148:** 249–258.

Bengtsson, B. O., and G. Thomson, 1981 Measuring the strength of associations between HLA antigens and diseases. Tissue Antigens **18:** 356–363.

Christiansen, F. B., and M. W. Feldman, 1975 Selection in complex genetic systems. IV. Multiple alleles and interactions between two loci. J. Math. Biol. **2:** 179–294.

Cullen, M., J. Noble, H. Erlich, K. Thorpe, S. Beck *et al.*, 1997 Characterization of recombination in the HLA class II region. Am. J. Hum. Genet. **60:** 397–407.

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

Feldman, M. W., R. C. Lewontin, I. R. Franklin and F. B. Christiansen, 1975 Selection in complex genetic systems. III. An effect of allele multiplicity with two loci. Genetics **79:** 333–347.

Fernandez-Viña, M. A., X. J. Gao, M. E. Moraes, J. R. Moraes, I. Salatiel *et al.*, 1991 Alleles at four HLA class II loci determined by oligonucleotide hybridization and their associations in five ethnic groups. Immunogenetics **34:** 299–312.

Gillespie, J. H., 1984 Some properties of finite populations experiencing strong selection and weak mutation. Am. Nat. **121:** 691–708.

Grahovac, B., R. I. Sukernik, C. O'Huigin, Z. Zaleska-Rutczynska, N. Blagitko *et al.*, 1998 Polymorphism of the HLA class II loci in Siberian populations. Hum. Genet. **102:** 27–43.

Griffiths, R. C., 1981 Neutral two-locus multiple alleles models with recombination. Theor. Popul. Biol. **19:** 169–186.

Gyllensten, U. B., and H. A. Erlich, 1993 MHC class II haplotypes and linkage disequilibrium in primates. Hum. Immunol. **36:** 1–10.

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

Karlin, S., and M. W. Feldman, 1970 Linkage and selection: two locus symmetric viability model. Theor. Popul. Biol. **1:** 39–71.

Karlin, S., and J. L. McGregor, 1972 Polymorphisms for genetic and ecological systems with weak coupling. Theor. Popul. Biol. **3:** 210–238.

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

Kwok, W. W., S. Kovats, P. Thurtle and G. T. Nepom, 1993 HLA-DQ allelic polymorphisms constrain patterns of class II heterodimer formation. J. Immunol. **150:** 2263–2272.

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

Lewontin, R. C., and K. Kojima, 1960 The evolutionary dynamics of complex polymorphisms. Evolution **14:** 458–472.

Marsh, S. G., 1998 HLA class II region sequences, 1998. Tissue Antigens **51:** 467–507.

Nagylaki, T., 1992 *Introduction to Theoretical Population Genetics*. Springer-Verlag, New York.

Peterson, A. C., A. Di Rienzo, A. E. Lehesjoki, A. de la Chapelle, M. Slatkin *et al.*, 1995 The distribution of linkage disequilibrium over anonymous genome regions. Hum. Mol. Genet. **4:** 887–894.

Sasaki, A., 1989 Evolution of pathogen strategies. Ph.D. Thesis, Biology, Kyushu University, Fukuoka, Japan.

Sasaki, A., 1992 The evolution of host and pathogen genes under epidemiological interaction, pp. 247–263 in *Population Paleo-Genetics*, edited by N. Takahata. Japan Scientific Society Press, Tokyo.

Sauermann, U., 1998 DQ-haplotype analysis in rhesus macaques: implications for the evolution of these genes. Tissue Antigens **52:** 550–557.

Slatkin, M., and C. A. Muirhead, 1999 Overdominant alleles in a population of variable size. Genetics **152:** 775–781.

Takahata, N., 1990 A simple genealogical structure of strongly balanced allelic lines and trans-species evolution of polymorphism. Proc. Natl. Acad. Sci. USA **87:** 2419–2423.

Vekemans, X., and M. Slatkin, 1994 Gene and allelic genealogies at a gametophytic self-incompatibility locus. Genetics **137:** 1157–1165.

Wolfram, S., 1996 *The Mathematica Book*. Wolfram Media Inc., Champaign, IL.

Communicating editor: N. Takahata