Maternal-Fetal Interactions and the Maintenance of HLA Polymorphism

Philip W. Hedrick* and Glenys Thomson†

**Division of Biological Sciences, University of Kansas, Lawrence, Kansas 66045, and †Department of Genetics, University of California, Berkeley, California 94720*

Manuscript received August 24, 1987
Accepted January 25, 1988

ABSTRACT

There is some empirical evidence that a fetus with an HLA antigen not present in its mother has a higher survival than a fetus sharing antigens with its mother. We have developed both single locus and two-locus theoretical models to examine this mode of selection. First, this immunologically based model appears to have the potential to maintain many alleles at a single locus and to result in an excess of heterozygotes when selection is strong. Second, substantial gametic disequilibrium is maintained between alleles at two loci for this selection mode when recombination is that observed between HLA loci *A*, *B*, and *DR*. Overall, it appears that this mode of selection has the potential to strongly affect genetic variation in the HLA region.

THE major histocompatibility complex (MHC) loci in mammals are among the most polymorphic loci known. For example, the HLA (human leukocyte antigen) loci *A* and *B* have about 15 and 30 alleles, respectively, in many Caucasian populations (ALBERT, BAUER and MAYR 1984). In addition, the frequencies of these alleles are more even than expected from neutrality (HEDRICK and THOMSON 1983; HEDRICK *et al.* 1986; KLITZ *et al.* 1986), an observation consistent with the hypothesis that some form of balancing selection is important in maintaining variation at these loci. Furthermore, there is substantial statistical association of alleles at different HLA loci, gametic disequilibrium, more than expected from neutrality (HEDRICK and THOMSON 1986) and having a pattern of disequilibria consistent with selection in the HLA region (KLITZ and THOMSON 1987).

One possible mode of balancing selection at the HLA loci that was proposed two decades ago (CLARKE and KIRBY 1966; WARBURTON 1968) involves maternal-fetal interaction that results in a net heterozygote advantage. This hypothesis suggests that a fetus with an antigen not present in its mother may have a higher survival than a fetus sharing antigens with its mother. In other words, if the fetus is antigenically incompatible with the mother, *i.e.*, the mother would reject a skin graft from an individual of the fetal genotype (considering only the MHC loci), then the fetus would have higher viability. Although this is a controversial hypothesis (*e.g.*, MCLAREN 1975; GILL 1983), there is some recent evidence from humans that couples having a history of spontaneous abortions are more likely to share antigens at HLA loci than control couples, an observation consistent with this model (summarized by THOMAS *et al.* 1985).

Previously, CLARKE and KIRBY (1966) and WARBURTON (1968) showed that this mode of selection gives a stable polymorphism for two alleles and multiple alleles, respectively [see also HULL (1966)]. Here we first extend these findings for a single locus, giving equilibrium allelic and genotypic frequencies as well as the mating-type frequencies. Next we will develop two-locus, two-allele theory and give the equilibrium gametic frequencies. It is obvious from this theory that such maternal-fetal interactions can result in single-locus polymorphism and two-locus disequilibrium. Finally, we discuss the potential importance of this selection mode for maintenance of genetic variation at HLA loci.

SINGLE LOCUS

Let the frequency of allele A_i at the *A* locus be p_i and assume for the present that the genotypes occur in Hardy-Weinberg proportions. Examining the possible mating types and their progeny, there are three qualitatively different mating types or maternal-fetal combinations. Table 1 gives the different types of matings when there are two alleles at the *A* locus. The first type of mating, *e.g.*, $A_1A_1 \times A_1A_1$, occurs when the male has no alleles that are different from the female, *i.e.*, the parents share two antigens at locus *A*. As a result, all progeny have two alleles that are present in the female. The second type of mating, *e.g.*, $A_1A_1 \times A_1A_2$, occurs when the male shares one allele with the female but has one that is different. Therefore, half the progeny from this mating have an allele different from the mother and half do not. Note that the reciprocal of this mating type, $A_1A_2 \times A_1A_1$, has different consequences because the male has no alleles that are not present in the female. The

TABLE 1

Frequencies of different mating types and their progeny when there are two alleles

Female	Male	No. of antigens shared	Frequency of mating	Progeny A_1A_1	Progeny A_1A_2	Progeny A_2A_2
A_1A_1	× A_1A_1	2	p_1^4	$p_1^4(1-s)$		
	× A_1A_2	1	$2p_1^3p_2$	$p_1^3p_2(1-s)$	$p_1^3p_2(1)$	
	× A_2A_2	0	$p_1^2p_2^2$		$p_1^2p_2^2(1)$	
A_1A_2	× A_1A_1	1	$2p_1^3p_2$	$p_1^3p_2(1-s)$	$p_1^3p_2(1-s)$	
	× A_1A_2	2	$4p_1^2p_2^2$	$p_1^2p_2^2(1-s)$	$2p_1^2p_2^2(1-s)$	$p_1p_2^2(1-s)$
	× A_2A_2	1	$2p_1p_2^3$		$p_1p_2^3(1-s)$	$p_1p_2^3(1-s)$
A_2A_2	× A_1A_1	0	$p_1^2p_2^2$		$p_1^2p_2^2(1)$	
	× A_1A_2	1	$2p_1p_2^3$		$p_1p_2^3(1)$	$p_1p_2^3(1-s)$
	× A_2A_2	2	p_2^4			$p_2^4(1-s)$
				$p_1^2(1-s)$	$p_1p_2(2-s)$	$p_2^2(1-s)$

third type of mating, *e.g.*, $A_1A_1 \times A_2A_2$, occurs when both alleles in the male are different from those in the female. In this case, all progeny have an allele that is different from the mother.

Equilibrium allelic frequencies: Let us now calculate the expected change in allelic frequency and the equilibrium allelic frequency from the frequencies given for the progeny in Table 1. Notice that all the progeny of a mating of a male with two antigens shared with a female have a fitness of $1 - s$ and all the progeny of a mating with no shared antigens have a fitness of 1. For a mating in which the male shares only one of his antigens with the female, half the progeny have a fitness of 1 and half $1 - s$.

Using these fitnesses and summing the three progeny columns in Table 1, then

$$\bar{w} = 1 - s(1 - p_1p_2) \tag{1a}$$

and

$$\Delta p_1 = \frac{p_1^2(1-s) + \frac{1}{2}p_1p_2(2-s) - p_1\bar{w}}{\bar{w}} = sp_1(1-p_1)(\tfrac{1}{2} - p_1)/\bar{w}. \tag{1b}$$

The only stable, polymorphic equilibrium occurs when $p_{1e} = \frac{1}{2}$.

If we carry out the same approach for three alleles, then

$$\underline{w} = 1 - s(1 - p_1p_2 - p_1p_3 - p_2p_3 - 3p_1p_2p_3) \tag{2a}$$

and

$$\Delta p_i = sp_i[(1-p_i)(\tfrac{1}{2} - p_i) - 3p_1p_2p_3]/\bar{w}. \tag{2b}$$

The only stable, polymorphic equilibrium occurs here when the term in brackets is zero for all three alleles. Therefore, at equilibrium

$$(1-p_1)(\tfrac{1}{2} - p_1) = (1-p_2)(\tfrac{1}{2} - p_2)$$

which is the quadratic

$$p_1^2 - \tfrac{3}{2}p_1 - p_2(p_2 - \tfrac{3}{2}) = 0$$

with solutions

$$p_{1e} = \tfrac{3}{4} \pm (p_2 - \tfrac{3}{4}).$$

Therefore, $p_{1e} = p_{2e}$ or $\frac{3}{4} - p_{2e}$. This holds for all alleles and given that $\Sigma p_i = 1$, then all $p_i = \frac{1}{3}$.

For k alleles

$$\bar{w} = 1 - s\left(1 - \sum_{i \neq j}\sum p_ip_j - 3\sum_{i \neq j \neq k}\sum\sum p_ip_jp_k\right) \tag{3a}$$

and

$$\Delta p_i = sp_i[(1-p_i)(\tfrac{1}{2} - p_i) - 3\sum_{i \neq j \neq k}\sum\sum p_ip_jp_k]/\bar{w}. \tag{3b}$$

Using the same logic as above, then $p_{ie} = 1/k$, *i.e.*, all alleles have the same equilibrium frequency that is equal to the reciprocal of the number of alleles.

We can also demonstrate that a new allele can always invade a given set of alleles. Let us define from (3b)

$$h(p_1) = \Delta p_1/sp_1. \tag{4a}$$

If $h(0) > 0$, then A_1 can always invade. For two alleles

$$h(0) = 1/[2(1-s)] > 0 \tag{4b}$$

for three alleles

$$h(0) = 1/\{2[1 - s(1 - p_2p_3)]\} > 0, \tag{4c}$$

and for four alleles

$$h(0) = \frac{\frac{1}{2} - 3p_2p_3p_4}{1 - s(1 - p_2p_3 - p_2p_4 - p_3p_4 - 3p_2p_3p_4)} > 0. \tag{4d}$$

Similarly, it can be shown that $h(0) > 0$ for any number of alleles.

As an illustration of the effect of this selection mode on allelic frequency, Figure 1 gives the change in the frequency of allele A_1 when it is below the equilibrium frequency for 2, 4, and 8 alleles. Here

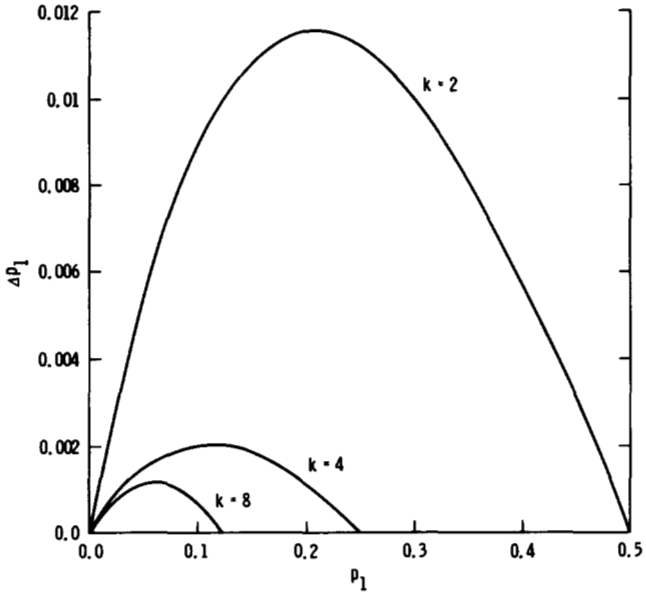

FIGURE 1.—Expected change in frequency for A_1, when the frequency of all other alleles is $(1 - p_1)/(k - 1)$ and $s = 0.2$.

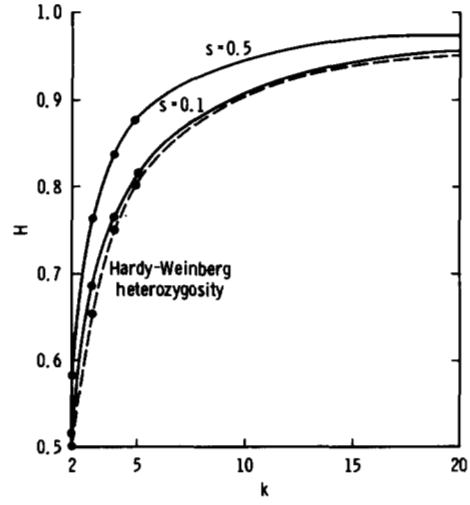

FIGURE 2.—Equilibrium heterozygosity for different numbers of alleles when $s = 0.1$ or 0.5 (*solid lines*) and when there are Hardy-Weinberg proportions (*broken line*).

the frequency of all the other alleles is assumed to be $(1 - p_1)/(k - 1)$ and $s = 0.2$. The change in allelic frequency is positive between 0 and p_{1e} and declines in magnitude as the number of alleles increases.

Genotypic frequencies: Let us relax the assumption that the genotypes occur in Hardy-Weinberg proportions. If we let P_{ij} be the frequency of the genotype with alleles i and j, then for two alleles, the frequencies of the genotypes after selection are

$$P'_{11} = (1 - s)p_1^2/\bar{w}$$

$$P'_{12} = (2p_1p_2 - \tfrac{1}{2}sP_{12})/\bar{w}$$

$$P'_{22} = (1 - s)p_2^2/\bar{w}$$

where

$$\bar{w} = 1 - sp_1^2 - sp_2^2 - \tfrac{1}{2}sP_{12}.$$

If we assume that the population is at equilibrium, then $P'_{12} = P_{12} = P_{12e}$ and

$$P_{12e} = (2p_{1e}p_{2e} - \tfrac{1}{2}sP_{12e})/(1 - sp_{1e}^2 - sp_{2e}^2 - \tfrac{1}{2}sP_{12e}).$$

Setting $p_{1e} = p_{2e} = \tfrac{1}{2}$, the equilibrium frequencies for these alleles, and rearranging this equation, we get the quadratic

$$sP_{12e}^2 - 2P_{12e} + 1 = 0.$$

and solving for P_{12e} we find

$$P_{12e} = \frac{1 - (1 - s)^{1/2}}{s}. \qquad (6a)$$

The equilibrium frequencies of the homozygotes are then

$$P_{11e} = P_{12e} = \tfrac{1}{2}(1 - P_{12e}).$$

For three alleles, using the same approach

$$P_{12e} = [2p_{1f}p_{2e} - \tfrac{1}{2}sP_{12e}(1 - p_{3e})]/\bar{w}$$

where

$$\bar{w} = 1 - s(p_{1e}^2 + p_{2e}^2 + p_{3e}^2) - \tfrac{1}{2}s[P_{12e}(1 - p_{3e}) + P_{13e}(1 - p_{2e}) + P_{23e}(1 - p_{1e})].$$

After setting $p_{1e} = p_{2e} = p_{3e} = \tfrac{1}{3}$, this becomes

$$sP_{12e}^2 - P_{12e} + \tfrac{2}{9} = 0$$

and

$$P_{12e} = [1 - (1 - \tfrac{8}{9}s)^{1/2}]/2s. \qquad (6b)$$

Repeating the same approach for four and five alleles, it is apparent that for k alleles

$$sP_{12e}^2 - \frac{2}{k - 1}P_{12e} + \frac{4}{k^2(k - 1)} = 0$$

and

$$P_{12e} = \frac{1}{s(k - 1)}\left\{1 - \left[1 - \frac{4s(k - 1)}{k^2}\right]^{1/2}\right\}. \qquad (6c)$$

All other heterozygotes have the same frequency at equilibrium. The total heterozygosity for k alleles is then

$$H = \frac{k(k - 1)}{2}P_{ije} \qquad (7a)$$

and the equilibrium frequency of the homozygotes is

$$P_{iie} = (1/k)(1 - H). \qquad (7b)$$

Figure 2 gives the equilibrium heterozygosity after selection for different numbers of alleles when $s = 0.1$ and $s = 0.5$. As a comparison, the Hardy-Weinberg proportion of heterozygotes, $(k - 1)/k$, is

also given in Figure 2. When $s = 0.5$, there is a substantial excess of heterozygotes for all numbers of alleles. For example, when there are five alleles, the observed heterozygosity is 0.877 while the Hardy-Weinberg heterozygosity is 0.8. When $s = 0.1$, there is much smaller excess, *e.g.*, with five alleles, the observed heterozygosity is 0.8125.

Mating type frequencies: First, assuming Hardy-Weinberg proportions, let us calculate the frequency of the different types of matings for different numbers of alleles. For two alleles, the frequency of matings where two antigens are shared is

$$M_2 = p_1^4 + p_2^4 + 2p_1^3p_2 + 2p_1p_2 + 2p_1p_2^3$$
$$= \sum_{i=1}^{2} p_i^4 + 4p_1^2p_2^2$$

For k alleles, the same approach gives

$$M_2 = \sum_{i=1}^{k} p_i^4 + 4\sum_{i<j}^{k}\sum^{k} p_i^2p_j^2. \quad (8a)$$

The frequency of a mating in which one antigen is shared, given that there are two alleles, is

$$M_1 = 2p_1^3p_2 + 2p_1p_2^3 + 2p_1^3p_2 + 2p_1p_2^3$$
$$= 4p_1p_2(p_1^2 + p_2^2).$$

For k alleles, this becomes

$$M_1 = 2\sum_{i=1}^{k} p_i^3(1 - p_i) + 2\sum_{i<j}\sum p_ip_j$$
$$\times [1 - 2p_ip_j - (1 - p_i - p_j)^2]. \quad (8b)$$

Finally, for the frequency of matings in which there are no shared antigens with two alleles is

$$M_0 = 2p_1^2p_2^2.$$

For k alleles, this becomes

$$M_0 = \sum_{i=1}^{k} p_i^2(1 - p_i)^2$$
$$+ 2\sum_{i<j}\sum p_ip_j(1 - p_i - p_j)^2. \quad (8c)$$

Figure 3 gives the proportions of the different mating types expected for different numbers of alleles using expressions (8a), (8b), and (8c) when all $p_i = 1/k$. When there are five alleles or less, the most common mating type is M_1, *i.e.*, when there is one allele shared. When there are more than five alleles, the mating types in which there are no shared alleles is most common, reaching a frequency of over 80 percent when there are 20 alleles. Remember that selection occurs against fetuses from mating types

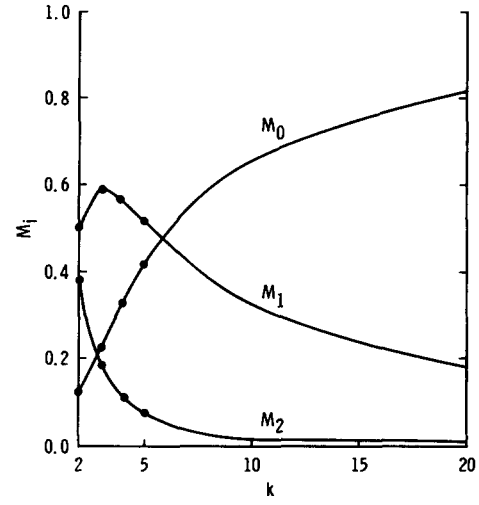

FIGURE 3.—Proportions of the mating types, M_i, where i is the number of alleles the male and the female share, and k is the number of alleles in the population.

M_1 and M_2 so that as M_0 becomes more common, the potential for selection is less.

TWO LOCI

Assume that a second locus B has alleles B_1 and B_2 with frequencies q_1 and q_2 and that gametes A_1A_2, A_1B_2, A_2B_1, and A_2B_2 have frequencies x_1, x_2, x_3, and x_4, respectively. Let c be the rate of recombination between the two loci, $D = x_1 - p_1q_1$, a measure of gametic disequilibrium between the loci, and G_{ij} be the frequency of the genotype composed of gametes i and j, *e.g.*, G_{13} is the frequency of genotype A_1B_1/A_2B_1.

Let us extend selection to include both loci. Because there are 100 mating types and 10 progeny types we will only describe the case in which the female is a double homozygote, A_1B_1/A_1B_1, a single heterozygote, A_1B_1/A_1B_2, or a double heterozygote, A_1B_1/A_2B_2 (all the other seven female genotypes fit into one of these categories). As a further shorthand, we will just give the male gamete rather than the complete male genotype. Table 2 gives the twelve different categories with the number of antigens shared between the female and the male gamete. For example, in the first row when a female A_1B_1/A_1B_1 receives a male gamete A_1B_1, *i.e.* both antigens in the male are in the female, all progeny are A_1B_1/A_1B_1 and share alleles with the mother at both loci so that we can designate the fitness in general as w_{AB} or specifically as $1 - t$. In the second row, one allele in the male gamete is shared with the female genotype, A_1, and one is not, B_2, making the fitness w_A or $1 - s$. As we will see below, it is useful to use the right-hand fitness parameterization given in Table 2.

Using these fitness values, the genotype frequencies after selection (after loss from maternal-fetal

TABLE 2

Examples of the three female genotypes; single homozygote, single heterozygote, and double heterozygote, the number of antigens male gametes share with them and the possible progeny genotypes and their fitness

Female	Male	No. of antigens shared	Progeny	
			Genotypes	Fitness
A_1B_1/A_1B_1	× A_1B_1	2	A_1B_1/A_1B_1	w_{AB} $1-t$
	× A_1B_2	1	A_1B_1/A_1B_2	w_A $1-s$
	× A_2B_1	1	A_1B_1/A_2B_1	w_B $1-s$
	× A_2A_2	0	A_1B_1/A_2B_2	1 1
A_1B_1/A_1B_2	× A_1B_1	2	A_1B_1/A_1B_1, A_1B_2/A_1B_1	w_{AB} $1-t$
	× A_1B_2	2	A_1B_1/A_1B_2, A_1B_2/A_1B_2	w_{AB} $1-t$
	× A_2B_1	1	A_1B_1/A_2B_1, A_1B_2/A_2B_1	w_B $1-s$
	× A_2B_2	1	A_1B_1/A_2B_2, A_1B_2/A_2B_2	w_A $1-s$
A_1B_1/A_2B_2	× A_1B_1	2	A_1B_1/A_1B_1, A_2B_2/A_1B_1, A_1B_2/A_1B_1, A_2B_1/A_1B_1	w_{AB} $1-t$
	× A_1B_2	2	A_1B_1/A_1B_2, A_2B_2/A_1B_2, A_1B_2/A_1B_2, A_2B_1/A_1B_2	w_{AB} $1-t$
	× A_2B_1	2	A_1B_1/A_2B_1, A_2B_2/A_2B_1, A_1B_2/A_2B_1, A_2B_1/A_2B_1	w_{AB} $1-t$
	× A_2B_2	2	A_1B_1/A_2B_2, A_2B_1/A_2B_2, A_1B_2/A_2B_2, A_2B_2/A_2B_2	w_{AB} $1-t$

interaction) are then

$$
\begin{aligned}
G'_{11} &= (x_1 - cD)^2(1-t)/\bar{w}\\
G'_{12} &= \{(x_1 - cD)\,[x_2(1 - sq_2 - tq_1) + cD(1-t)]\\
&\quad + (x_2 + cD)\,[x_1(1 - sq_1 - tq_2)\\
&\quad - cD(1-t)]\}/\bar{w}\\
G'_{22} &= (x_2 + cD)^2(1-t)/\bar{w}\\
G'_{13} &= \{(x_1 - cD)\,[x_3(1 - sp_2 - tp_1) + cD(1-t)]\\
&\quad + (x_3 + cD)\,[x_1(1 - sp_1 - tp_2)\\
&\quad - cD(1-t)]\}/\bar{w}\\
G'_{23} &= \{(x_2 + cD)\,[x_3(1 - sx_1 - sx_4 - tx_2)\\
&\quad + cD(1-t)] + (x_3 + cD)\\
&\quad \times [x_2(1 - sx_1 - sx_4 - tx_3) + cD(1-t)]\}/\bar{w}\\
G'_{14} &= \{(x_1 - cD)\,[x_4(1 - sx_2 - sx_3 - tx_1) \qquad (10)\\
&\quad - cD(1-t)] + (x_4 - cD)\\
&\quad \times [x_1(1 - sx_2 - sx_3 - tx_4) - cD(1-t)]\}/\bar{w}\\
G'_{24} &= \{(x_2 + cD)\,[x_4(1 - p_2s - p_1t) - cD(1-t)]\\
&\quad + (x_4 - cD)\,[x_2(1 - sp_1 - tp_2)\\
&\quad + cD(1-t)]\}/\bar{w}\\
G'_{33} &= (x_3 + cD)^2(1-t)/\bar{w}\\
G'_{34} &= \{(x_3 + cD)\,[x_4(1 - sq_2 - tq_1) - cD(1-t)]\\
&\quad + (x_4 - cD)\,[x_3(1 - sq_1 - tq_2)\\
&\quad + cD(1-t)]\}/\bar{w}\\
G'_{44} &= (x_4 - cD)^2(1-t)/\bar{w}
\end{aligned}
$$

where $\bar{w}$ is the mean fitness, *i.e.*, the sum of all the right-hand expressions above excluding $\bar{w}$.

Let us first consider the situation in which there is no recombination, $c = 0$, because we can obtain some analytical results for this case. With $c = 0$, the genotypic frequencies become

$$
\begin{aligned}
G'_{11} &= x_1^2(1-t)/\bar{w}\\
G'_{12} &= x_1x_2(2 - s - t)/\bar{w}\\
G'_{22} &= x_2^2(1-t)/\bar{w}\\
G'_{13} &= x_1x_3(2 - s - t)/\bar{w}\\
G'_{23} &= x_2x_3[2 - 2x_1s - 2x_4s - x_2t - x_3t]/\bar{w} \qquad (11)\\
G'_{14} &= x_1x_4[2 - 2sx_2 - 2sx_3 - x_1t - x_4t]/\bar{w}\\
G'_{24} &= x_2x_4(2 - s - t)/\bar{w}\\
G'_{33} &= x_3^2(1-t)/\bar{w}\\
G'_{34} &= x_3x_4(2 - s - t)/\bar{w}\\
G'_{44} &= x_4^2(1-t)/\bar{w}.
\end{aligned}
$$

The frequency of gamete A_1B_1 after selection is

$$x'_1 = G'_{11} + \tfrac{1}{2}(G'_{12} + G'_{13} + G'_{14})$$

so that

$$
\begin{aligned}
\bar{w}x'_1 &= x_1^2(1-t) + \tfrac{1}{2}x_1x_2(2 - s - t)\\
&\quad + \tfrac{1}{2}x_1x_3(2 - s - t)\\
&\quad + \tfrac{1}{2}x_1x_4(2 - 2sx_2 - 2sx_3 - x_1t - x_4t).
\end{aligned}
$$

This expression simplifies to

$$
\begin{aligned}
\bar{w}x'_1 &= \tfrac{1}{2}x_1(2 - s - t) + \tfrac{1}{2}x_1^2(s - t)\\
&\quad + \tfrac{1}{2}x_1x_4[s(x_1 + x_4) - (s - t)(x_2 + x_3)].
\end{aligned}
$$

The expressions for the other gametes are

$$\bar{w}x_2' = \tfrac{1}{2}x_2(2 - s - t) + \tfrac{1}{2}x_2^2(s - t) + \tfrac{1}{2}x_2x_3[s(x_2 + x_3) - (s - t)(x_1 + x_4)]$$

$$\bar{w}x_3' = \tfrac{1}{2}x_3(2 - s - t) + \tfrac{1}{2}x_3^2(s - t) + \tfrac{1}{2}x_2x_3[s(x_2 + x_3) - (s - t)(x_1 + x_4)]$$

$$\bar{w}x_4' = \tfrac{1}{2}x_4(2 - s - t) + \tfrac{1}{2}x_4^2(s - t) + \tfrac{1}{2}x_1x_4[s(x_1 + x_4) - (s - t)(x_2 + x_3)]$$

and the mean fitness is

$$\bar{w} = 1 - \tfrac{1}{2}(s + t) + \tfrac{1}{2}[(x_1 - x_4)^2 + (x_2 - x_3)^2](s - t) + (2s - t)[x_1x_4(x_1 + x_4) + x_2x_3(x_2 + x_3)].$$

Because of the symmetry of the selection model, we can assume $x_1 = x_4$ and $x_2 = x_3$. Furthermore, at equilibrium $p_1 = p_2 = q_1 = q_2 = \tfrac{1}{2}$ so that $x_1 = \tfrac{1}{4} + D$ and $x_2 = \tfrac{1}{4} - D$ making $x_2 = \tfrac{1}{2} - x_1$. With these substitutions, then

$$\bar{w}x_1' = \tfrac{1}{2}x_1(2 - s - t) + x_1^2(2s - t)x_1. \quad (12)$$

The first solution of this expression is

$$x_1 = 0 \text{ with } D = -\tfrac{1}{4}. \quad (13a)$$

Furthermore, assuming that $x_1 \neq 0$, then

$$\bar{w} = 1 - \tfrac{1}{2}(s + t) + x_1^2(2s - t)$$

so that by substitution

$$x_1^2(2s - t) = (2s - t)\,2(x_1^3 + x_2^3).$$

Assuming that $2s \neq t$, then

$$x_1^2 = 2x_1^3 + 2(\tfrac{1}{2} - x_4)^3$$

and

$$(4x_1 - 1)(2x_1 - 1) = 0.$$

The solutions of this expression are then

$$x_1 = \tfrac{1}{2} \text{ with } D = \tfrac{1}{4} \quad (13b)$$

$$x_1 = \tfrac{1}{4} \text{ with } D = 0. \quad (13c)$$

Using standard stability techniques (e.g., FELDMAN, FRANKLIN and THOMSON 1974), it can be shown that when $t \geq 0$ the equilibrium with $D = 0$ (expression 13c), *i.e.*, $x_1 = x_2 = x_3 = x_4 = \tfrac{1}{4}$ is stable if and only if $s < t/2$. The equilibria with $D = \tfrac{1}{4}$, i.e., $x_1 = x_4 = \tfrac{1}{2}$ and $x_2 = x_3 = 0$ and $D = -\tfrac{1}{4}$, *i.e.*, $x_1 = x_4 = 0$ and $x_2 = x_3 = \tfrac{1}{2}$ when $t \geq 0$ are stable if and only if $s > t/2$.

When $t = 2s$, then the expression for $\bar{w}x_1'$ becomes

$$(x_1/4)(4 - 3t) = (x_1/4)(4 - 3t).$$

As a result, there is a neutral curve and whatever the initial gametic frequencies are, they remain there.

TABLE 3

Recombination level (c) necessary to generate an equilibrium with $D \neq 0$ for given selective values, below these values, $D = 0$ equilibrium is present

s	$t > 2s$	$t = 2s$	$t = 1 - (1 - s)^2$	$t = 3/2s$	$t = s$
0.1	—[a]	—	<0.002	<0.008	<0.014
0.2	—	—	<0.008	<0.018	<0.031
0.4	—	—	<0.047	<0.055	<0.080

[a] — = Only $D = 0$ equilibrium present.

TABLE 4

Level of disequilibrium D expected for given s and t values for the recombination amount between HLA loci *A*, *B*, and *DR*

	$t = 3/2s$			$t = s$		
c	$s = 0.1$	$s = 0.2$	$s = 0.4$	$s = 0.1$	$s = 0.2$	$s = 0.4$
0.008 (*A-B*)	0.0	±0.183	±0.231	±0.162	±0.216	±0.240
0.010 (*B-DR*)	0.0	±0.161	±0.226	±0.131	±0.206	±0.234
0.018 (*A-DR*)	0.0	0.0	±0.205	0.0	±0.162	±0.221

Now let us assume that $c > 0$. In this case, we must iterate the expressions for the genotypic frequencies given above. Using the results from $c = 0$ as a background, we can organize the results, given that there is recombination, in a similar manner. When $t > 2s$, then the only equilibrium present is $D = 0$. In addition, when $t = 2s$, the only equilibrium is for $D = 0$, unlike the $c = 0$ case. If $t < 2s$, then there are $D \neq 0$ equilibria if the recombination is low enough.

Table 3 gives several such cases, including $t = \tfrac{3}{2}s$ and $t = s$. The middle column gives the multiplicative case, *i.e.*, if $w_{AB} = (1 - s)^2 = 1 - t$ so that $t = 1 - (1 - s)^2$. For example, in the case analogous to multiplicative fitness values and assuming $s = 0.1$ (making $t = 0.19$), then if $c < 0.002$ there are $D \neq 0$ equilibria. The least restrictive situation here is when $s = t$. For example, when $t = s = 0.1$, then if $c < 0.014$, there are $D \neq 0$ equilibria.

How large is the disequilibrium generated by these selection and recombination values? Table 4 gives the D values for the map distances between the three HLA loci *A*, *B*, and *DR*. Note that because $p_1 = q_1 = \tfrac{1}{2}$, and $D = \tfrac{1}{4}D'$, where D' is the normalized disequilibrium measure of LEWONTIN (1964), and is the proportion of the maximum disequilibrium possible. For example, when $t = \tfrac{3}{2}s$, $s = 0.2$, and $c = 0.008$, then $D = \pm 0.183$ ($D' = \pm 0.732$). For these parameters which are not much larger than necessary for $D \neq 0$ equilibria given $c = 0.008$, there is 73.2 percent of possible disequilibrium generated. From Table 4 and other simulations, it appears that when the $D \neq 0$ equilibrium are present, then generally the extent of disequilibrium is large.

TABLE 5

Frequency of mating types in which the male parent shares 0, 1, or 2 alleles with the female parent for HLA-A, or -B in a Danish sample (LARSEN and HANSEN 1987)

	M_0	M_1	M_2
A	0.514	0.452	0.034
B	0.699	0.288	0.013
A or *B*	0.384	0.422	0.172

DISCUSSION

A balancing selection model based on the immunological hypothesis developed to explain recurrent spontaneous abortion appears to have the potential to maintain a large amount of genetic polymorphism. As with incompatibility systems in plants, the expected change in allelic frequencies and the equilibrium allelic frequency decline as the number of alleles increases. As a result, genetic drift should become a stronger influence on allelic frequencies when there are more alleles (*e.g.*, WRIGHT 1965). An excess of heterozygotes is predicted by this selection mode and occurs when there is strong selection. However, with weaker selection the genotypes are virtually in Hardy-Weinberg proportions.

There are three classes of mating types under this scheme, *i.e.*, the male parent shares 0, 1, or 2 alleles with the female. As shown in Figure 3, the mating type with no shared alleles is most frequent when there are many alleles (with a relatively even distribution) in the population and the mating type with two shared alleles most common when there are few alleles. Table 5 gives the expected frequencies of the mating types based on the observed haplotypic frequencies for HLA-*A* and *B* with 12 and 20 alleles, respectively, in a Danish sample (LARSEN and HANSEN 1987). Notice that the proportion of matings that share no antigens is largest for both loci individually and that only a small proportion share two antigens. When both loci are considered simultaneously, 61.6% of the matings share one or more antigens (2.2% shared three or four antigens).

Because the *DR* locus has fewer alleles than *A* or *B*, more matings should share alleles at *DR* than at *A* and *B*. Higher antigen sharing could, therefore, possibly result in more selection from maternal-fetal interactions at *DR* than at *A* or *B*. In fact, it appears that the rate of recurrent spontaneous abortion is higher for *DR* than for *A* or *B* (THOMAS et al. 1985), consistent with this prediction.

In addition, this selection mode has the potential to generate gametic disequilibrium between alleles at HLA loci. Given that the effect of a second locus is less than that of the first, *i.e.*, $t < 2s$, then the known linkage between HLA loci can generate disequilibrium. For example, when $c = 0.008$, the recombination level between *A* and *B*, and $s = t = 0.1$, 65% of the disequilibrium possible is generated.

The other model suggested as an explanation for the high rate of recurrent spontaneous abortion in couples that share HLA antigens is the genetic hypothesis (*e.g.*, SCHACTER, WEITCAMP and JOHNSON 1984; HEDRICK 1988). This model assumes that there are recessive lethals in the HLA region that may become homozygous in the progeny of parents that share antigens. However, HEDRICK (1988) has suggested that this would seem to depend upon high disequilibrium between lethals and a number of HLA antigens. In addition, such a selection mode would not result in stable polymorphism or stable disequilibrium at the HLA loci.

There are a number of other modes of selection, including resistance to pathogens, segregation distortion, and non-random mating that have been suggested to be important for major histocompatibility complex loci [see HEDRICK, THOMSON and KLITZ (1987) for a review]. However, along with the pathogen resistance mode, selection involving maternal-fetal interaction shows the most promise towards explaining genetic variation in the HLA region.

This research was supported by National Institutes of Health grants GM 35326 and HD 12731. We thank B. KLITZ, E. LOUIS, H. PAYAMI, and W. ROBINSON for comments on the manuscript.

LITERATURE CITED

ALBERT E., M. BAUR and W. MAYR, 1984 *Histocompatibility Testing 1984*. Springer-Verlag, Berlin.

CLARKE B., and D. R. S. KIRBY, 1966 Maintenance of histocompatibility polymorphisms. Nature **211:** 999–1000.

FELDMAN M. W., I. FRANKLIN and G. THOMSON, 1974 Selection in complex genetic systems. I. The symmetric equilibria of the three-locus symmetric viability model. Genetics **76:** 135–162.

GILL T. J., 1983 Immunogenetics of spontaneous abortions in humans. Transplantation **35:** 1–6.

HEDRICK P. W., 1988. HLA-sharing, recurrent spontaneous abortion, and the genetic hypothesis. Genetics **119:** 199–204.

HEDRICK P. W., and G. THOMSON, 1983 Evidence for balancing selection at HLA. Genetics **104:** 449–456.

HEDRICK P. W., and G. THOMSON, 1986 A two locus neutrality test: Applications to humans, *E. coli*, and lodgepole pine. Genetics **112:** 135–156.

HEDRICK P. W., G. THOMSON and W. KLITZ, 1987 Evolutionary genetics and HLA: Another classic example. Biol. J. Linn. Soc. **31:** 311–331.

HULL P., 1966 Possible stability of polymorphism in a multiallele incompatibility system in the mouse. Genetics **54:** 1049–1053.

KLITZ W., and G. THOMSON, 1987 Disequilibrium pattern analysis. II. Application to Danish HLA-A and -B data. Genetics **116:** 633–643.

LARSEN S. D., and H. E. HANSEN, 1987 Antigen sharing in couples in a two-locus ABO-like system. Ann. Hum. Genet. **51:** 177–180.

LEWONTIN R. C., 1964 The interaction of selection and linkage. I. General considerations; heterotic models. Genetics **49:** 48–67.

MCLAREN A., 1975 Antigenic disparity: does it affect placental size, implantation or population genetics. In: *Immunobiology of Trophoblast. Clinical and Experimental Immunoproduction I*, Edited

by R. G. EDWARDS, C. W. S. HOWE and M. H. JOHNSON. Cambridge University Press, Cambridge.

SCHACTER B., L. R. WEITCAMP and W. E. JOHNSON, 1984 Parental HLA compatibility, fetal wastage, and neural tube defects: evidence for a *T/t*-like locus in humans. Am. J. Hum. Genet. **36:** 1082–1091.

THOMAS M. L., J. H. HARGER, D. K. WAGENER, B. S. RABIN and T. J. GILL, 1985 HLA sharing and spontaneous abortion in humans. Am. J. Obstet. Gynecol. **151:** 1053–1058.

WARBURTON F. L., 1968 Maintenance of histocompatibility polymorphism. Heredity **23:** 151–152.

WRIGHT S., 1965 The distribution of self-incompatibility alleles in populations. Evolution **18:** 609–679.

Communicating editor: W. J. EWENS