Female Choice and Variation in the Major Histocompatibility Complex

Philip W. Hedrick

Department of Zoology, Arizona State University, Tempe, Arizona 85287

Accepted for publication June 10, 1992

ABSTRACT

The cause of the high genetic variability in the major histocompatibility complex (MHC) is not entirely clear. Recently, two reports suggest that female mice prefer to mate with males different from them at the MHC. A model of female choice appropriate for those observations is developed here. Female choice can in fact reduce the observed proportions of homozygotes, maintain genetic polymorphism, influence mating-type frequencies and generate gametic disequilibrium.

THE major histocompatibility complex (MHC) is one of the most polymorphic regions in many vertebrates (e.g., Klein 1986). The most widely given reason for this high variability is that MHC alleles may confer resistance to various pathogens (e.g., Briles et al. 1983; Hill et al. 1991). In addition, it has been shown that different H2 types (the MHC region in mice) can be distinguished by mice and other trained animals (e.g., Beauchamp et al. 1990), suggesting that H2 types may be used in selecting mates.

Recently two reports indicate that female mice may prefer to mate with males different from them at H2 genes (EGID and Brown 1989; Potts, Manning and Wakeland 1991). In the female mating preference tests performed by EGID and Brown, in 29 out of 39 trials females mated first with males differing at the H2 locus, giving a mating preference of 1 to 0.344 for H2-different to H2-identical males. In the semi-natural populations examined by Potts, Manning and Wakeland a 27% deficiency of homozygotes was observed and attributed primarily to female mating preferences because other factors did not appear to completely explain this deficiency. Here I will develop and discuss a model of female choice that is appropriate for these MHC data.

MODEL

Let us assume that females preferentially mate with males that differ genetically from themselves. For example, if the female is a homozygote at a MHC locus and the males are identical at 2, 1 and 0 alleles, let the relative preferences of males by females be 1 - s, 1 - hs, and 1 respectively. From the experiment of EGID and Brown (1989), s would be 0.656. If the female is a heterozygote, then her preferences are 1, 1 - hs, and 1 - s assuming the male has 2, 1 or 0 alleles different from her alleles. Here if h = 0, males that differ at one allele are preferred by females as much as those that differ by two alleles. If h = 1, males must differ at both alleles to be preferred over males identical to the female. If h = 1/2, then males that differ at one allele are exactly intermediate in preference between those that differ by two or no alleles.

If we assume that there are a relatively large number of potential male mates, then the frequency of the various mating types when there are two alleles are given in Table 1 where

\[ w_{11} = 1 - s(P_{11} + P_{12}h) \]
\[ w_{12} = 1 - s \]
\[ w_{22} = 1 - s(P_{22} + P_{12}h) \]

and \( P_q \) is the frequency of genotype \( A_iA_j \). The \( w_q \) values assume the sum of the mating type frequencies for a given female are equal to her frequency, e.g., \( m_{11.11} + m_{11.12} + m_{11.22} = 1/P_{11} \) where, for example, \( m_{11.22} \) is the frequency of the mating type \( A_1A_1 \) (female) \( \times A_2A_2 \) (male). The genotypic frequencies in the progeny are found by summing the three columns, e.g.,

\[ P'_{11} = m_{11.11} + \frac{1}{2}m_{11.12} + \frac{1}{2}m_{11.22} + \frac{1}{4}m_{12.12}. \]

Such mating preferences may lead to deviations in genotypic frequencies from Hardy-Weinberg proportions. One way to measure this effect is by the deficiency in homozygotes

\[ d = (\sum p_i^2 - \sum P_q)/\sum p_i^2 \]

where \( p_i \) is the frequency of \( A_i \). From the data of Potts, Manning and Wakeland, \( d = 0.27 \). Another commonly used measure of this deviation is the fixation index

\[ F = 1 - \frac{1 - \sum P_q}{1 - \sum p_i^2} \]

(e.g., Hedrick 1990a). These two measures are related such that

\[ d = \frac{F(\sum p_i^2 - 1)}{\sum p_i^2} \]
The model can be directly extended to two loci by assuming that the mating preferences are as indicated above for each locus. The overall mating preference can then either be the product of the single-locus values (multiplicative), or the sum of the single-locus values (additive). The amount of association between alleles at the two loci can be measured by

\[ D_g = x_{ij} - p_i q_j \]  

(5)

where \( D_g \) is a measure of gametic disequilibrium for alleles \( A_i \) and \( B_j \), the frequencies of these alleles are \( p_i \) and \( q_j \), and \( x_{ij} \) is the observed frequency of gamete \( A_i B_j \) (e.g., Hedrick 1985).

RESULTS

The mating preference model as outlined above is a symmetrical balancing-selection model in which there is identical selection against all homozygotes and identical selection favoring all heterozygotes [see Karlin (1969) for a discussion of general negative-assortative mating models and their consequences]. In cases like this, there is only one stable polymorphic equilibrium and that occurs when all alleles are equally frequent. Because we are concerned with more complicated factors such as genotypic frequencies or two loci here, we will consider only the symmetrical preference model. Because of the complex nature of this model, the following will rely primarily on numerical results.

Comparison to heterozygous advantage: First, it is instructive to compare this female choice model, a type of negative-assortative mating limited to one sex, with the heterozygous advantage model, the simplest type of balancing selection. One way is to find the selective value from the symmetrical heterozygous advantage model (assuming the fitnesses of homozygotes are 1 and that of heterozygotes are 1) which minimizes the difference in the expected change in allelic frequency for given values of female choice (after Hedrick 1972). For example, Figure 1a gives the \( \Delta p_1 \) values when there are two alleles for \( s = 0.5 \) and \( h = 0, \frac{1}{2}, \) and 1 (solid lines) and the closest heterozygous advantage values (these are \( t = 0.256, 0.186 \) and 0.085 for \( h = 1, \frac{1}{2} \) and 0). The shape of the curve for \( h = \frac{1}{2} \) is very close to that of a heterozygous advantage model (so \( t = 0.186 \) is not shown) while for \( h = 1, \) the \( \Delta p_1 \) values are larger near the equilibrium and for \( h = 0 \) they are larger near zero and unity. The shape of the curve for \( h = 0 \) is such that the retention of genetic variation in a finite population is greater than the most similar heterozygous advantage value while that for \( h = 1 \) is less (see Hedrick 1972).

To obtain similar comparisons when there are more alleles, the change in allele frequency for \( A_i \) below the equilibrium is compared when the frequency of the other alleles are all \((1 - p_i)/(n - 1)\) and \( n \) is the number of alleles. This is comparable to the situation which may occur if the population is at equilibrium for \( n - 1 \) alleles and a new allele is introduced because of mutation or gene flow and is increasing in frequency toward the equilibrium.

When there are four alleles, the \( \Delta p_1 \) curves are very similar for different levels of \( h \) (Figure 1b). Furthermore, the heterozygous advantage model curves that have the closest fit to these curves are indistinguishable on these plots (\( t \) varies from 0.251 for \( h = 1 \) to 0.274 for \( h = 0 \)). However, when there are ten alleles (Figure 1c), the \( \Delta p_1 \) values vary considerably from \( h = 0 \) (\( t = 0.117 \)) to that for \( h = 1 \) (\( t = 0.424 \)). Again there is a close fit of the heterozygous advantage curve to these respective values.

The basis for the switch in magnitude of \( \Delta p_1 \) values as a function of \( h \) for 2 alleles \( h = 0 \) is the largest while for 10 alleles \( h = 1 \) is the largest is the frequency of various mating types for different numbers of alleles. If we classify the frequency of mating types that differ by 2, 1 and 0 alleles as \( m_2, m_1 \) and \( m_0 \), then these frequencies (assuming no mating preference) are
Female Choice and MHC

Figure 1.—The change in the frequency of $A_t$ for three levels of $h$ when $t = \frac{1}{2}$ and the other alleles have a frequency of $(1 - p_0)/n$ for (a) 2 alleles, (b) 4 alleles and (c) 10 alleles. The value of $t$ indicated is that for the heterozygote advantage model which gives the most similar curve to the given mating preference curve.

Figure 2.—The frequency of the three types of matings, those differing by zero alleles ($m_0$), one allele ($m_1$) and two alleles ($m_2$) when there are $n$ alleles and there is no selection.

given in Figure 2 for different numbers of alleles, assuming all alleles have frequencies of $1/n$. When there are two alleles, the frequency $m_0$ is the highest and $m_2$ the lowest. As a result, selection is very effective if $h = 0$ because this differentiates between the mating types that have frequencies $m_1$ and $m_0$. On the other hand, if $h = 1$, selection is much less effective because only 0.125 of the matings are the $m_2$ type. When there are high numbers of alleles, the situation is completely reversed with $m_0$ being very low. As a result, selection is most effective for $h = 1$ which differentiates between mating types that have frequencies $m_2$ and $m_1$.

If mating preference occurs, then the frequencies of these mating types are altered. As an example, the three parts of Figure 3 gives the frequencies of the three mating types at equilibrium when $h$ is 0, $\frac{1}{2}$ or 1 for four alleles and different levels of preference. When $h = 0$ or $\frac{1}{2}$, the mating type with frequency $m_1$ is the most common and $m_2$ the second most common frequency for all levels of preference. But when $h = 1$ and the preference level is high, the mating type with frequency $m_2$ becomes much more frequent than the one with frequency $m_1$.

Deficiency of homozygotes: When there is strong balancing selection, there can be a sizable deficiency of homozygotes and an excess of heterozygotes (Hedrick 1990b). For example, for the mating-preference model discussed here, Figure 4a gives the deficiency in homozygosity ($d$) at equilibrium for four alleles when $h = 0, \frac{1}{2}$ or 1. As expected, this deficiency increases as preference increases and does so almost linearly for the lower ranges of $s$. When $s = 0.5$, the deficiency is about 22.0%, 16.7% and 12.5% for $h$
Figure 3.—The frequency of the three types of matings, those differing by zero alleles ($m_0$), one allele ($m_1$) and two alleles ($m_2$) for (a) $h = 0$, (b) $h = \frac{1}{2}$ and (c) $h = 1$ when there are four alleles and mating preference $s$.

Figure 4.—The reduction in homozygosity ($d$) for (a) different $s$ and $h$ levels when there are four alleles and (b) when $h = \frac{1}{2}$ and there are 2, 3, 4, 5 or 10 different alleles.

values of 1, $\frac{1}{2}$ and 0, respectively. The basis for the higher reduction when $h = 1$ is that in this case $m_2$ type matings are favored over all others and these matings produce only heterozygous progeny. Figure 4b gives the reduction in homozygosity at equilibrium when $h = \frac{1}{2}$ for 2, 3, 4, 5 and 10 alleles. The reduction increases as the number of alleles increases and increases the most from two to three alleles. As the number of alleles increases, the amount of reduction decreases somewhat as each allele is added.

Also indicated on Figure 4a by the broken horizontal line is the value where $d = 0.27$, the level observed by Potts, Manning and Wakeland. In this case, only when $s = 0.57, 0.72,$ and 1.0 for $h = 1, \frac{1}{2}$ and 0 is the deficiency as large as 27%. Notice that the value of $s$ necessary to cause this 27% reduction for $h = \frac{1}{2}$
is similar to the value of 0.656 from the experiments of EGID and BROWN (1989).

 POTTS, MANNING and WAKELAND (1991) gave the homozygote deficiency observed in nine different populations which varied from 0.12 to 0.44 with the overall value of 0.27. In Table 2, these values are given along with the extent of female preference necessary to cause this deficiency for three levels of $h$. If $h = \frac{1}{2}$, the average preference necessary is 0.69 and ranges from 0.40 to 0.98. This is quite similar to the preference value of 0.656 estimated from the experiments of EGID and BROWN (1989). If $h = 0$, then four of the observed deficiencies cannot be explained completely by just female preferences. For $h = 1$, the extent of preference necessary is somewhat less, with a mean value of 0.58.

Two loci: First, when the mating preferences are additive over loci and $h = \frac{1}{2}$ or 1, even when $s$ is at a maximum, no gametic disequilibrium is generated at the equilibrium. When $h = 0$, only under the most extreme conditions is there some disequilibrium generated. For example, if $s = 0.99$ at both loci, then only if the rate of recombination ($c$) is equal to or less than 0.004 is there gametic equilibrium generated. The lack of gametic disequilibrium with additive preferences is reminiscent of two-locus fitness models with additivity over loci and does not generate gametic disequilibrium (BODMER and FELSENSTEIN 1967).

When the preferences are multiplicative over loci, then gametic disequilibrium can be generated for all levels of $h$. Figure 5 gives the level of recombination necessary to generate gametic disequilibrium for the range of female preferences. Notice that for the level of preference measured by EGID and BROWN, $s = 0.656$, that disequilibrium is generated for all $h$ values when $c < 0.008$. For example, H2-D and H2-K are estimated to be approximately 0.4 map units apart (KLEIN 1986), given $s = 0.656$ and $h = \frac{1}{2}$, then $D = 0.180$. Because all allele frequencies are 0.5 in this symmetrical model, this value is 72.0% of the maximum disequilibrium possible. When $s$ is large, then $h = 1$ has the largest ability to generate disequilibrium. This occurs because, for example, with $D > 0$, mating types $A_1A_1B_1B_1$ (female) $\times$ $A_2A_2B_2B_2$ (male) and vice versa, are in higher frequency and they produce only coupling gametes, $A_1B_1$ or $A_2B_2$.

**DISCUSSION**

There always has been strong debate over the selective mechanisms maintaining genetic variation at MHC loci. Because a basic function of MHC is to present foreign peptides to T cells, selective explanations that include pathogen resistance are generally favored, particularly with the recent report of malarial resistance associated with antigen HLA-Bw53 (and a class II haplotype) in humans (HILL et al. 1991). However, even on pathogen-resistance, some (e.g., BODMER 1972) feel that selection is frequency-dependent while others (e.g., HUGHES and NET 1988) feel that it is due to intrinsic heterozygous advantage.

In fact, there is evidence that selection may act in several other ways on MHC variants [for reviews, see HEDRICK, THOMSON and KLITZ (1987) and HEDRICK et al. (1991)]. First, MHC variants in humans (HLA) are known to be associated with many diseases (Tiwari and Terasaki 1985), mostly autoimmune in nature that potentially can reduce viability or fecundity. However, this effect may not have been very large in humans when life expectancies were shorter and probably would not have an effect like balancing selection. Second, there is some evidence that recurr-
rent spontaneous abortion occurs more frequently among couples that share HLA antigens (e.g., Thomas et al. 1985) although there are contradictory reports (e.g., Oksenberg et al. 1984). If this type of selection is important, it can both cause maintenance of genetic variation and generate gametic disequilibrium (Hedrick and Thomson 1988). Third, the $I$ locus in mice that causes segregation distortion is tightly linked to the H2 region (Silver 1985). In this case, however, selection should actually act to reduce genetic variation and gametic disequilibrium (Hedrick 1988). Finally, there has been previous evidence that mate selection by males in mice is influenced by MHC types (e.g., Yamazaki et al. 1983). In particular, male mice generally appear to prefer females that are different from themselves at the H2 region. This mate selection by males appears to be caused by imprinting of progeny on parents or sibs (Yamazaki et al. 1988), a phenomenon that can also maintain genetic variation (P. W. Hedrick, in preparation).

The recent reports of female choice based on MHC type discussed here provide another potential selective force to maintain genetic variation. In general, this selection is somewhat like heterozygous advantage but its effect is approximately half that of heterozygous advantage, as might be expected, because female choice operates only in one sex. For example, when $h = \frac{1}{2}$, each different allele in males is given a selective preference effect, a preference value of 0.5 gives an expected change in allelic frequency that is closely approximated by a heterozygous advantage selection coefficient of 0.186, 0.264 and 0.256 for 2, 4 and 10 alleles, respectively.

Hedrick (1990b) discusses possible explanations for the substantial deficiency in homozygotes observed in two studies of HLA (the human MHC), including heterozygous advantage, specific lethals in gametic disequilibrium with specific HLA variants, maternal-fetal interaction, and variable resistance to pathogens. Although there was not much evidence to discriminate between these alternatives, it appears particularly unlikely that the linked lethal explanation is of major importance.

As I have shown above, the 27% deficiency in homozygotes in mice observed by Potts, Manning and Wakeland (1991) can be explained by another mechanism, namely female choice of the magnitude observed by Egid and Brown (1989). For example, an average $s$ value of 0.69 when $h = \frac{1}{2}$ results in the deficiencies observed by Potts, Manning and Wakeland (see Table 2), quite close to the value of 0.656 from the data of Egid and Brown. Further, I should note it is not necessary to explain all of the Potts, Manning and Wakeland deficiency because 30% of the deficiency was explained by settlement patterns and a further 22% was explained by maternal-fetal interaction. In other words, the $s$ value in the seminatural experiments of Potts, Manning and Wakeland need only be approximately half as large as estimated in the laboratory experiments of Egid and Brown.

The female-choice model can be compared to the maternal-fetal interaction model discussed in detail by Hedrick and Thomson (1988). Both models have the potential to maintain polymorphism at MHC loci, cause deficiencies in homozygotes, and to generate gametic disequilibrium. However, for a given $s$ value in both models ($s$ ranges from 0 to 1 in both), then the maternal-fetal model appears always to have more effect. For example, if $s = \frac{1}{2}$ and given four alleles at equilibrium, then the maternal-fetal model can result in a deficiency of 52% while the female-choice model results in a deficiency of 17%. Furthermore, when $s = \frac{1}{2}$ for the maternal-fetal interaction model, gametic disequilibrium may be maintained between loci over 5 map units apart while for the female-choice model two loci must be generally less than 0.5 map unit apart to generate disequilibrium. Obviously, part of the lesser strength of the female-choice model is that it only functions in one sex but it also appears to have some intrinsic properties that make it less effective in influencing genetic variation.

Potts, Manning and Wakeland (1991) suggests that MHC may serve as a genetic incompatibility system to minimize genome-wide inbreeding. For example, if MHC can provide information about which males are closely related and which are not, then a mate choice based on MHC can reduce the level of inbreeding. To illustrate, let us assume that females only mate with males that are not identical to them at MHC. When there are two alleles of equal frequency in a random-mating population then $P_{11} + P_{12} + P_{22} = 0.375$ of the matings would be between identical individuals. On the other hand, among fullsibs $1 - P_{12} (1 - \frac{1}{2}P_{12}) = 0.594$ of the matings would be between identical individuals. If there are four equally frequent alleles, then only 0.031 of the random matings would be between identical individuals while 0.402 of the matings between fullsibs would be identical. In other words, if females mated preferentially with males not identical to them at MHC, then matings between fullsibs (or other close relatives) would be greatly reduced.

Such an outbreeding system would function optimally for alleles at loci tightly linked to and in gametic disequilibrium with MHC alleles. However, a survey of gametic disequilibrium in the HLA region showed that such associations had already fallen off between loci at opposite ends of the HLA region, only two map units away and showed no evidence of association outside the region (Hedrick, Thomson and Klitz 1986) and Hedrick et al. (1991). Further, it is note-
worthy that the highest rates of inbreeding documented in natural populations of mammals (RALLS, HARVEY and LYLES 1988 found more than 5% parent-offspring or full-sib mating in only three bird or mammal species), will result in only a relatively small level of inbreeding even at equilibrium (e.g., HEDRICK and COCKERHAM 1986). If MHC mate selection plays an important role in keeping inbreeding levels low, then it may have an important effect on genetic variation in many species.

This research was supported by National Institutes of Health grant GM35326. I thank JERRAM BROWN, DICK HUDSON and an anonymous reviewer for comments on the manuscript.

LITERATURE CITED


