POLYMORPHISM RESULTING FROM THE MATING ADVANTAGE OF RARE MALE GENOTYPES

BY WYATT W. ANDERSON

DEPARTMENT OF BIOLOGY, YALE UNIVERSITY, NEW HAVEN

Communicated by Theodosius Dobzhansky, April 4, 1969

Abstract.—An algebraic model is presented for selection by the mating advantage of rare genotypes. In accordance with the results of recent experiments, the selective advantage is formulated as inversely proportional to genotype frequency. The model applies to any number of alleles at an autosomal locus or at a sex-linked locus, and it provides for a different intensity of selection for each genotype in each sex. The frequency of an allele under this selection proceeds to a stable, nontrivial equilibrium when any one of the genotypes carrying the allele and any one of those carrying another allele possess some degree of the advantage which varies inversely with genotype frequency. The changes in gene frequency are not affected by the mating system, being, for instance, the same with random mating as with inbreeding. With low values for the selection parameters, genetic polymorphisms could be maintained with little selection apparent except at rather low or rather high gene frequencies. The rare genotype advantage may be a significant element among the mechanisms which maintain the extensive genetic polymorphisms of natural populations.

Introduction.—The success of male genotypes in mating has been found in several instances to depend on their relative frequencies. The rarer the male genotype becomes, the greater is its mating success. Petit1–2 first observed this phenomenon in mating tests with mutants of Drosophila melanogaster. Ehrman, Spassky, Pavlovsky, and Dobzhansky4 encountered this same rare genotype advantage when they set up experimental populations of D. pseudoobscura with different initial frequencies of chromosomal inversion types. Their initial sample showed a higher frequency of the inversion which was carried by the less-frequent genotype than was expected with random mating and no selection. The same effect was observed in mating chambers where females and males of two karyotypes were allowed to mate. Ehrman (see Petit and Ehrman4 for a review of the pertinent literature) has observed the rare genotype mating advantage in a wide variety of circumstances and for several species. Spiess6 studied the effect in both D. persimilis and D. pseudoobscura with a somewhat different technique. The rare genotype mating advantage does not seem to extend to females, or if present, it is very weak. Petit, Ehrman, and Spiess have all realized that the mating advantage of rare males could be a powerful force in maintaining genetic polymorphisms in nature. The theoretical behavior of genes under selection by the rare genotype advantage has not been investigated, however, partly owing to the complications of having a frequency-dependent selection which operates differently in the two sexes. It is the purpose of this article to propose an algebraic model for the rare male mating advantage and to examine the nature of the polymorphism which results.
Choice of Selective Values.—The foundation of a model for the mating advantage of rare males is the proper choice of selective values for the genotypes. Ehrman7 presented data which provide the material for determining how the selective values should be related to male genotype frequencies. Ehrman studied the frequency and preference of matings between strains of Drosophila pseudoobscura which contained the CH and AR inversions on the third chromosome. Each test involved males and females of AR/AR and CH/CH karyotypes. Experiments were run with frequencies of genotypes ranging from 0.1 to 0.9. We shall define the mating success of each male genotype to be the number of females inseminated by these males relative to the number of males of this genotype which were present. This measure of mating success is the selective value, since we are considering only the selection through mating advantage of less-frequent male genotypes. The values of the male-mating success in the two sets of data given by Ehrman7 are graphed in Figure 1 for various frequencies of the genotypes. They range from a high value of 3.6 to a common low value of approximately 0.6. Let us call the increase in mating success above the baseline value the selective advantage. A linear relation between selective advantage and male genotype frequency does not fit the experimental data well. Rather, the selective advantage of a male genotype increases dramatically as its frequency decreases. This suggests a selective advantage inversely proportional to genotype frequency. The selective value is then of the form $W_i = 1 + x/G_i$, where $x$ is a parameter of selection and $G_i$ is the frequency of the $i$th male genotype. Some time ago Wright8 briefly discussed this system of selective values. Each selective value in the model is a simple function of the genotype frequency, a hyperbola as illustrated in Figure 2 for several values of the selection parameter. The experimental data are clearly similar to those generated with the model of selective advantage inversely proportional to genotype frequency. The model thus seems appropriate.

![Figure 1](image-url)

**Fig. 1.**—Male mating success as a function of male genotype frequency in mating tests involving two homokaryotypes of Drosophila pseudoobscura. Data of Ehrman7: $A$, from her Table 1A; $B$, from her Table 1B.
Changes in Gene Frequency at an Autosomal Locus.—Let us consider first the consequences of selection involving two alleles or gene arrangements on an autosome. We shall assume that all the genotypes are distinct as far as the operation of selection is concerned. In order that the model be as general as possible we shall allow both sexes to possess an advantage when rare, but we shall provide for different degrees of selection in males and females. Thus, the situation where the mating advantage occurs only for rare male genotypes corresponds to the simpler case of zero advantage in the females. We are specifically interested in the selection conferred by male mating advantage, but the model is just as easily interpreted in terms of any other form of selective advantage—viability or female fertility, for example. The system of mating is not specified, since the model holds for any system. All that is required is that each genotype frequency at the beginning of a generation be some function of the gametic frequencies among the parents of this new generation, and that these functions be nonnegative and add to unity.

Let \( P_\varphi(T) \) and \( P_\sigma(T) \) be the frequencies of allele \( A \) in the eggs and sperm which combine to form the zygotes of generation \((T + 1)\). Let \( P(T + 1) \) be the frequency of allele \( A \) among the newly-formed zygotes of generation \((T + 1)\); \( P(T + 1) = \frac{1}{2} \left( P_\varphi(T) + P_\sigma(T) \right) \). \( Q_\varphi(T) \), \( Q_\sigma(T) \), and \( Q(T + 1) \) are the corresponding frequencies for allele \( B \); \( Q = 1 - P \) for each \( Q \). At generation \( T \), the model is as follows:

<table>
<thead>
<tr>
<th>Genotype</th>
<th>( AA )</th>
<th>( AB )</th>
<th>( BB )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f ), frequency among newly formed zygotes</td>
<td>( f_{AA}(T) )</td>
<td>( f_{AB}(T) )</td>
<td>( f_{BB}(T) )</td>
</tr>
<tr>
<td>Selective values in females</td>
<td>( 1 + \frac{u}{f_{AA}(T)} )</td>
<td>( 1 + \frac{2v}{f_{AB}(T)} )</td>
<td>( 1 + \frac{w}{f_{BB}(T)} )</td>
</tr>
<tr>
<td>Selective values in males</td>
<td>( 1 + \frac{r}{f_{AA}(T)} )</td>
<td>( 1 + \frac{2s}{f_{AB}(T)} )</td>
<td>( 1 + \frac{t}{f_{BB}(T)} )</td>
</tr>
</tbody>
</table>

An increased mating success for a genotype effectively increases the frequency of that genotype among the parents which contribute alleles to the next generation. The frequencies of allele \( A \) among the germ cells in generation \( T \) will be

\[
P_\varphi(T) = \frac{f_{AA}(T) + u + f_{AB}(T)/2 + v}{1 + u + 2v + w}
\]

and

\[
P_\sigma(T) = \frac{f_{AA}(T) + r + f_{AB}(T)/2 + s}{1 + r + 2s + t}.
\]

Now \( P(T) = f_{AA}(T) + f_{AB}(T)/2 \) and \( P(T + 1) = (P_\varphi(T) + P_\sigma(T))/2 \). Hence

\[
P(T + 1) = \frac{1}{2} \left( \frac{1}{K_1} + \frac{1}{K_2} \right) P(T) + \frac{1}{2} \left( \frac{u + v}{K_2} + \frac{r + s}{K_1} \right)
\]

where \( K_1 = 1 + r + 2s + t \) and \( K_2 = 1 + u + 2v + w \). The zygotic gene frequencies form a simple, first-order linear difference equation. For the rare genotype advantage, each selection parameter is greater than or equal to zero, and
Fig. 2.—Male mating success as a function inversely proportional to male genotype frequency.

not all are zero. Beginning from gene frequency $P(0)$, the frequency of allele $A$ at generation $T$ is

$$P(T) = \left[\frac{1}{2} \left(\frac{1}{K_1 + 1} + \frac{1}{K_2} \right) \right] T \left( P(0) - P_B \right) + P_B,$$

where

$$P_B = \frac{K_1(u + v) + K_2(r + s)}{K_1(u + 2v + w) + K_2(r + 2s + t)}.$$

The fact that the difference equation (1) is of the form $P(T + 1) = C \ P(T) + B$, with $C > 0$, requires that the $P$'s converge monotonically to a stable equilibrium at $P_B$, beginning from any initial gene frequency. The conditions under which this equilibrium is nontrivial, that is, neither 0 nor 1, are found by considering the restrictions on $P_B$ for which $0 < P_B < 1$. For $P_B$ to be a nontrivial equilibrium, it is necessary and sufficient that the selection parameter in either sex for any one of the genotypes carrying allele $A$ be greater than zero, and that likewise the selection parameter in either sex for any one of the genotypes carrying allele $B$ be greater than zero.

The mating system does not affect the changes in gene frequency. The functions which express the genotype frequencies of a new generation in terms of the gametic frequencies in the previous generation do not enter the basic recurrence relation (1). The mating system may even differ from generation to generation without altering the changes in gene frequency. As a simple example, consider the case where selection is the same in the two sexes. With random mating, $f_{AA}(T) = P(T)^2$, $f_{AB}(T) = 2P(T)Q(T)$, and $f_{BB}(T) = Q(T)^2$. Thus,

$$P(T + 1) = \frac{P(T)^2 + P(T)Q(T) + u + v}{1 + u + 2v + w} = \frac{P(T) + u + v}{1 + u + 2v + w}.$$

With inbreeding, $f_{AA}(T) = P(T)^2(1 - F) + P(T)F$, $f_{AB}(T) = 2P(T)Q(T)$.
(1 - F), and \( f_{AB}(T) = Q(T)^2(1 - F) + Q(T)F \). Thus,

\[
P(T + 1) = \frac{P(T)^2(1 - F) + P(T)F + P(T)Q(T)(1 - F) + u + v}{1 + u + 2v + w} = \frac{P(T) + u + v}{1 + u + 2v + w}
\]

The changes in gene frequency are the same with random mating and with inbreeding. The same is true for any mating system. This remarkable independence of gene frequency changes from the mating system derives from the particular functions of genotype frequencies which were chosen as the selective values.

The polymorphism due to rare genotype advantage, according to the model presented above, does not require heterozygote superiority at all. Consider the case where \( s = 0 \) and \( v = 0 \). The heterozygotes are always less fit than either of the two kinds of homozygotes; yet a stable equilibrium is possible at any gene frequency, depending only on the values of \( u, w, r, \) and \( t \).

**Extension to Many Alleles at an Autosomal Locus.**—The model for the rare genotype advantage is readily extended to any number of alleles. Suppose there are \( K \) alleles at an autosomal locus. Let the frequency of the genotype carrying alleles \( A_i \) and \( A_j \) among the newly formed zygotes of generation \( T \) be \( f_{ij}(T) \), and let the female selective value of this genotype be \( 1 + x_{ij}/f_{ij}(T) \), the male selective value \( 1 + y_{ij}/f_{ij}(T) \). Proceeding as before, we find

\[
P_{ij}(T + 1) = \frac{1}{2} \left( \frac{1}{1 + x} + \frac{1}{1 + y} \right) P_{ij}(T) + \frac{1}{2} \left( \frac{x_i}{1 + x} + \frac{y_i}{1 + y} \right)
\]

where \( x_i = \Sigma x_{ij}, y_i = \Sigma y_{ij}, x = \Sigma x_{ij}, \text{ and } y = \Sigma y_{ij} \). For each allele we have an equation which is identical to the earlier equation (1), except that the constants are given in a more general form. All the earlier results stand and need not be repeated. The equilibrium frequencies are \( P_{ij}(E) = [x_i(1 + y) + y_i(1 + x)]/[y(1 + x) + x(1 + y)] \). The equilibrium frequency for any allele \( A_i \) will be greater than zero if and only if the selection parameter in either of the sexes for any one of the genotypes carrying allele \( A_i \) is greater than zero. At equilibrium, all alleles satisfying this condition will be present.

**Changes in Gene Frequency at a Sex-Linked Locus.**—Suppose that the alleles \( A \) and \( B \) are sex linked. The model for females will be exactly that outlined for autosomal loci. For males, the model is altered as follows:

**Genotype**

<table>
<thead>
<tr>
<th>( f ), frequency among newly formed zygotes</th>
<th>( A )</th>
<th>( B )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P_{ij}(T - 1) )</td>
<td>( 1 - P_{ij}(T - 1) )</td>
<td></td>
</tr>
</tbody>
</table>

**Selective values in males**

| \( 1 + \frac{r}{P_{ij}(T - 1)} \) | \( 1 + \frac{t}{1 - P_{ij}(T - 1)} \) |
The effective frequencies of allele $A$ among the germ cells at generation $T$, after all selection, are

$$P_\varphi(T) = \frac{P_\varphi(T - 1) + r}{1 + r + t} \quad \text{and} \quad P_\varphi(T) = \frac{1/2(P_\varphi(T - 1) + P_\varphi(T - 1)) + u + v}{1 + u + 2v + w}.$$  

We need consider only $P_\varphi(T)$, since $P_\varphi(T)$ is a simple linear function of $P_\varphi(T - 1)$. Substituting for $P_\varphi(T - 1)$ in our last equation, and letting $K_3 = 1 + r + t$ and $K_4 = 1 + u + 2v + w$, we obtain the second-order linear difference equation

$$P_\varphi(T) = \frac{P_\varphi(T - 1)}{2K_4} + \frac{P_\varphi(T - 2)}{2K_4K_3} + \frac{2K_3(u + v) + r}{2K_4K_4}.$$

Let $X_1 = 1/4K_4(1 + \sqrt{1 + 8K_4/K_3})$ and $X_2 = 1/4K_4(1 - \sqrt{1 + 8K_4/K_3})$. It is easily shown that $0 < X_1 < 1$ and $-1 < X_2 < 0$. The solution to the difference equation (2) is

$$P_\varphi(T) = \left[\frac{(P_\varphi(0) - P_\varphi(E))X_2 - (P_\varphi(1) - P_\varphi(E))}{X_2 - X_1}\right]X_1^T + \left[\frac{(P_\varphi(1) - P_\varphi(E)) - (P_\varphi(0) - P_\varphi(E))X_1}{X_2 - X_1}\right]X_2^T + P_\varphi(E),$$

where $P_\varphi(E) = \frac{2(u + v)K_3 + r}{2K_4K_4 - K_3 - 1}$.

Since $|x_1| < 1$ and $|x_2| < 1$, it is easily seen that the gene frequencies $P_\varphi(T)$ converge to $P_\varphi(E)$, from any initial frequency. $P_\varphi(E)$ is a point of stable equilibrium. The necessary and sufficient conditions for the equilibrium to be nontrivial are exactly the same as for an autosomal locus. Now, $|X_1| > |X_2|$. The convergence of the sequence $\{P_\varphi(T)\}$ is, therefore, dominated by the behavior of the sequence $\{C_1X_1^T\}$, which converges monotonically to zero. For $T$ sufficiently large, convergence will be monotonic. The convergence of $\{P_\varphi(T)\}$ will be monotonic from $T = 0$ for most choices of the selection parameters.

The model may be extended to include many alleles at a sex-linked locus. Let us use the terms defined in considering an autosomal locus, except that $y_t$ will be the selection parameter for the males carrying allele $A_t$, and $y = \Sigma y_t$. Proceeding as for the autosomal case, we find

$$P_{t\varphi}(T) = \frac{P_{t\varphi}(T - 1)}{2(1 + x)} + \frac{P_{t\varphi}(T - 2)}{2(1 + x)(1 + y)} + 2x_t(1 + y) + y_t.$$

For each allele we have a difference equation like equation (2), except that the constants are given in a more general form. The behavior of these equations has already been described. The gene frequencies $P_{t\varphi}(T)$ will converge to stable equilibria given by $P_{t\varphi}(E) = [2x_t(1 + y) + y_t] / [(1 + y)(1 + 2x) - 1]$. The necessary and sufficient conditions for the equilibria to be nontrivial are exactly the same as for an autosomal locus.
Refinements to the Model.—The model presented above will be satisfactory over a broad range of genotype frequencies, particularly for small values of the selection parameters. But when the frequency of a genotype becomes extremely low, its selective value becomes unreasonably large. If the selective values are defined as \( W_{ij} = 1 + x_{ij}/(a_{ij} + f_{ij}) \) where each \( a_{ij} \) is a constant, then this problem is overcome.

As the frequency of a genotype decreases, its selective value approaches \( 1 + x_{ij}/a_{ij} \). The \( a' \)s will be chosen so that the \( W' \)s never become unreasonably large; they will usually be quite small. The course of selection will then be very little altered from that given by the model in the preceding pages, except when genotype frequencies become extremely small. With choices of the selection parameters which now seem likely, the selective advantage of rare genotypes will still lead to stable, nontrivial equilibria.

The selective values may be further modified to include a component independent of genotype frequency by defining them as \( W_{ij} = K_{ij} + x_{ij}/(a_{ij} + f_{ij}) \), where each \( K_{ij} \) is a constant. Equilibria are again likely under a wide variety of circumstances, but the course of selection is somewhat more complicated than before. With dominance, so that the genotypes are no longer distinct under the selection, the model must be further modified. I plan to consider these modifications to the basic model in greater detail elsewhere.

Numerical Example and Discussion.—The gene frequency changes under selection by rare male mating advantage are graphed in Figure 3 for two alleles.

![Graph of gene frequency changes](image-url)

**Fig. 3.**—Changes in gene frequency due to selection by the mating advantage of rare male genotypes. \( r = 0.1, s = 0.05, \) and \( t = 0.3. \)
at an autosomal locus, with \( r = 0.1, s = 0.05, \) and \( t = 0.3. \) In line with the experimental evidence, there was assumed to be no selection among the females, so that \( u = v = w = 0. \) The courses of selection from a high and from a low initial frequency are shown.

The selection by rare genotype advantage can change gene frequencies quite rapidly if the selection parameters are even moderately large. On the other hand, if they are small, the selection will be detectably different from zero only when one genotype becomes rare. Equilibrium at any gene frequency is possible, even with small values of the selection parameters. Only the speed of gene frequency change depends on the absolute size of the selection parameters. With small values of the selection parameters, the selection would be scarcely apparent but yet might play an extremely important role in maintaining genes or chromosomal arrangements at low frequencies.

That natural populations are storehouses for genetic variants has been known since the pioneering analysis of Tshetverikov. More recent studies have further documented the extent of this genetic variability. It is possible that the mating advantage of rare male genotypes plays a significant role in maintaining genetic variability. This selection is, of course, only one of many mechanisms which maintain balanced polymorphisms. The importance of each type of selection depends on the particular circumstances in a given population; taken together, the various types assure a high level of genetic variability.

I am particularly thankful to Dr. Howard Levene for his thoughtful criticism of the manuscript. I am grateful to Drs. F. Ayala, Th. Dobzhansky, M. Dowler, L. Ehrman, K. Kojima, and W. Watt for their helpful comments on the paper.

5 Petit, C., and L. Ehrman, in *Evolutionary Biology*, vol. 3 (in press).