

THE MAINTENANCE OF HETEROZYGOSITY BY PARTIAL NEGATIVE ASSORTATIVE MATING

P. L. WORKMAN

Department of Genetics, University of Cambridge, Cambridge, England

Received July 8, 1964

NEGATIVE assortative mating, or disassortative mating, may be defined as a decrease in the frequency of matings between like zygotic or gametic types relative to the frequency expected under the assumption of random mating. Mathematical models based on a system of negative assortative mating have been described by WRIGHT (1921), HOGBEN (1946), DAHLBERG (1948), LI (1955), WATTERSON (1959), and NAYLOR (1962, 1963).

Mixed mating systems which included, in part, a system of disassortative mating would provide means for maintaining polymorphic variation in the absence of heterozygote advantage. Except for brief discussions by HOGBEN and by LI, the theoretical effect of such systems on the maintenance of heterozygosity has not previously been considered. There is, however, biological evidence which indicates that such a study would be of value. A tendency for disassortative mating has been observed in the white-throated sparrow, *Zonotrichia albicollis* (LOWTHER 1961), the marine copepod, *Tisbe reticulata* (BATTAGLIA 1963), the moths *Biston betularia* (KETTLEWELL 1956) and *Amathes glareosa* (KETTLEWELL and BERRY 1961), and the Scarlet Tiger moth, *Panaxia dominula* (KETTLEWELL 1942; SHEPPARD 1952). In particular, in a colony of *Panaxia* near Oxford there is a polymorphism for wing color pattern which has been studied intensively by FISHER and FORD (1947), SHEPPARD (1951, 1953, 1956), and SHEPPARD and COOK (1962). The polymorphism is controlled by a pair of alleles that give rise to three recognizable genotypes. An analysis of the selective forces acting on these genotypes has led to the suggestion that partial disassortative mating may be responsible for the maintenance of the polymorphism (SHEPPARD 1956; WILLIAMSON 1960; SHEPPARD and COOK 1962).

In the present paper, theoretical models will be developed for populations mating under (a) exclusive negative assortative mating, (b) mixed self-fertilization and negative assortative mating, and (c) mixed random mating and negative assortative mating. Particular emphasis will be placed on the equilibrium genotypic distributions and their stability, and on the levels of heterozygosity which these systems can maintain. Finally, a model of mixed random mating and negative assortative mating is used to consider the maintenance of the wing color pattern polymorphism in different artificial and natural colonies of *Panaxia dominula*.

General considerations: The models to be discussed will apply to zygotic dis-

assortative mating in either monoecious or dioecious diploid populations. We shall consider only a single locus at which there are two alleles, say A and a .

The systems of negative assortative mating included in the models are termed asymmetric systems (WATTERSON 1959) in that only one sex is assumed to contribute gametes in proportion to their frequencies in the adult population. That is, zygotes are produced by the union of two gametes; the first one is chosen, at random from one sex, and then the second gamete is chosen, also at random, from those individuals of the opposite sex with whom mating is permitted. Such a model would not apply to finite, monoecious populations composed of monogamous individuals, but it may be the most appropriate model for populations in which one of the sexes is either in excess, as in plant populations producing excess pollen, or is able to mate several times. One example of such a system is provided by those insect populations in which the male is mobile and polygamous, the female is essentially sessile and monogamous, and mating depends in part on female preference. Some of the difficulties involved in constructing models which assume equal contributions from both sexes have been discussed by WATTERSON (1959). It would also be possible to derive models in which the amount of disassortative mating in a mixed mating system is not constant, but depends, in part, upon the prevailing genotypic proportions. Such models will not be considered in the present paper.

The mathematical analysis of the models is based on the following procedure. We shall assume that there are constant probabilities, K , K_1 , and K_2 , for negative assortative mating, self-fertilization, and random mating, where $K + K_1 + K_2 = 1$. Let (u, v, w) and (u', v', w') represent the genotypic proportions of (AA, Aa, aa) in generations n and $n + 1$ respectively ($u + v + w = 1$). Then, depending on the system of negative assortative mating considered, general recurrence equations relating genotypic proportions in successive generations can be directly obtained. Since at equilibrium, the genotypic proportions in successive generations n and $n + 1$ remain constant, the equilibrium distribution, given, by (u_e, v_e, w_e) is determined by setting $(u', v', w') = (u, v, w) = (u_e, v_e, w_e)$ and solving the appropriate system of equations. The stability of the equilibrium distribution is then tested according to the method described by OWEN (1953). For example the recurrence equations which have been derived for three different systems of mixed negative assortative mating (to be discussed below), self-fertilization, and random mating are given in Table 1. To consider, say, mixed negative assortative mating and self-fertilization, set $K_2 = 0$, $K_1 = 1 - K$ and derive the pertinent equilibrium distributions as outlined above.

It will be assumed that the probability for negative assortative mating, K , is the same for all genotypes (or phenotypes) of the sex controlling the mating preference. This assumption greatly simplifies the mathematical analysis and it should not affect the general conclusions regarding the maintenance of heterozygosity.

Exclusive negative assortative mating. (i) *No dominance—genotypic disassortative mating:* On the assumption of no dominance, the three genotypes (AA, Aa, aa) will give rise to six different mating types, of which three are between

TABLE 1

General recurrence equations for genotypic frequencies at a single, diallelic locus in populations mating under mixed random mating, self-fertilization, and three systems of negative assortative mating

No dominance—System I

$$AA: u' = K \left[\frac{\frac{1}{2}uv(1+w)}{(1-u)(1-v)} \right] + K_1(u + \frac{1}{4}v) + K_2(u + \frac{1}{2}v)^2$$

$$Aa: v' = K \left[\frac{\frac{1}{2}v+uw(1+\frac{1}{2}v)}{(1-u)(1-w)} \right] + K_1(\frac{1}{2}v) + 2K_2(u + \frac{1}{2}v)(w + \frac{1}{2}v)$$

$$aa: w' = K \left[\frac{\frac{1}{2}vw(1+u)}{(1-w)(1-v)} \right] + K_1(w + \frac{1}{4}v) + K_2(w + \frac{1}{2}v)^2$$

No dominance—System II

$$AA: u' = K \left[\frac{uv(2-u)}{2(1-u)} + \frac{1}{4}v^2 \right] + K_1(u + \frac{1}{4}v) + K_2(u + \frac{1}{2}v)^2$$

$$Aa: v' = K \left[\frac{\frac{1}{2}v+uw(1+\frac{1}{2}v)}{(1-u)(1-w)} \right] + K_1(\frac{1}{2}v) + 2K_2(u + \frac{1}{2}v)(w + \frac{1}{2}v)$$

$$aa: w' = K \left[\frac{wv(2-w)}{2(1-w)} + \frac{1}{4}v^2 \right] + K_1(w + \frac{1}{4}v) + K_2(w + \frac{1}{2}v)^2$$

Complete dominance

$$AA: u' = 0 \qquad \qquad \qquad + K_1(u + \frac{1}{4}v) + K_2(u + \frac{1}{2}v)^2$$

$$Aa: v' = K \left[\frac{(u + \frac{1}{2}v)/(u+v)}{1} \right] + K_1(\frac{1}{2}v) + 2K_2(u + \frac{1}{2}v)(w + \frac{1}{2}v)$$

$$aa: w' = K \left[\frac{(\frac{1}{2}v)/(u+v)}{1} \right] + K_1(w + \frac{1}{4}v) + K_2(w + \frac{1}{2}v)^2$$

(u,v,w) and *(u',v',w')* represent genotypic proportions in generations *n* and *n+1*, respectively. $(u+v+w=1)$. K, K_1, K_2 are the proportions of negative assortative mating, self-fertilization, and random mating ($K+K_1+K_2=1$). The three systems of negative assortative mating are described in the text.

like genotypes ($AA \times AA, Aa \times Aa, aa \times aa$) and three between unlike genotypes ($AA \times Aa, AA \times aa, Aa \times aa$). There are then (relative to the asymmetric system already described) five different simple systems of negative assortative mating which can be considered depending on whether one, two, or all three of the matings between like genotypes are completely prohibited. For completeness, the different schemes, together with their possible equilibrium distributions are given in Table 2 in terms of the matings which are prohibited. Of these five systems, it can be shown that only two (I, II) lead to stable equilibrium states at which more than one genotype is present, and only one system (II) leads to an equilibrium at which all three genotypes are present. Systems IIIa,b, Va,b and IV (for $u_0 \neq w_0$) are similar to systems of random mating with selection favouring one or the other of the homozygotes and lead to fixation of one of the homozygous

TABLE 2

Systems of genotypic disassortative mating based on a single diallelic locus with alleles A and a

System	Matings not permitted	Equilibrium states			Character
		AA	Genotypic distribution Aa	aa	
I	AA×AA, Aa×Aa, aa×aa	1/2	1/2	0	stable($u_0 > w_0$)
		0	1/2	1/2	stable($u_0 < w_0$)
		1/4(5-√17)	1/2(-3+√17)	1/4(5-√17)	unstable($u_0 = w_0$)
II	AA×AA, aa×aa	1/4(5-√17)	1/2(-3+√17)	1/4(5-√17)	stable
IIIa	AA×AA, Aa×Aa	0	0	1	stable
IIIb	aa×aa, Aa×Aa	1	0	0	stable
IV	Aa×Aa	1	0	0	stable
		0	0	1	stable
		1/4	1/2	1/4	unstable
Va	AA×AA	0	0	1	stable
Vb	aa×aa	1	0	0	stable

genotypes. Here, and in the treatment of partial negative assortative mating, we shall consider only Systems I and II which result in the retention of heterozygous genotypes for all stable equilibrium distributions.

Under System I, sometimes termed complete negative assortative mating, the only matings permitted are between unlike genotypes ($AA \times Aa$, $AA \times aa$, $Aa \times aa$). The recurrence relations describing the change in genotypic frequencies under such a system can be obtained from Table 1, setting $K = 1$, and $K_1 = K_2 = 0$ in the top system of equations. For this system, there are three possible genotypic distributions at equilibrium (Table 2); $(1/2, 1/2, 0)$, $(0, 1/2, 1/2)$, and $(1/4 [5 - \sqrt{17}], 1/2 [-3 + \sqrt{17}], 1/4 [5 - \sqrt{17}])$, which correspond to $u_0 > w_0$, $u_0 < w_0$, and $u_0 = w_0$, respectively, where (u_0, v_0, w_0) represent any initial genotypic distribution. However, it can be shown that the last of the three distributions ($u_0 = w_0$), also derived by HOGBEN (1946), is unstable and in any finite population, owing to sampling variation making $u \neq w$, the genotypic proportions at equilibrium will be given by one of the other two distributions, both of which are stable.

Under System II, like homozygotes do not mate but the heterozygotes can mate at random giving four possible mating types ($AA \times aa$, $AA \times Aa$, $Aa \times aa$, $Aa \times Aa$). The appropriate recursions can be taken from the middle system of Table 1, for $K = 1$, $K_1 = K_2 = 0$. For this system there is only one equilibrium state possible, $(u_e, v_e, w_e) = (1/4 [5 - \sqrt{17}], 1/2 [-3 + \sqrt{17}], 1/4 [5 - \sqrt{17}])$ or approximately (.219, .562, .219), and it can be shown to be stable. It is interesting that these are the same values as given for the unstable equilibrium state of System I. However, the distributions coincide only at the equilibrium values as can be seen by substituting any initial distribution at which $u_0 = w_0$ into the recursion formulas.

(ii) *Complete dominance—phenotypic disassortative mating:* With complete

dominance at a diallelic locus there are only two phenotypes, say $A-$ and aa , and three mating types ($A \times A-$, $A \times aa$, $aa \times aa$). The only system of interest is one in which matings between like phenotypes are prohibited, permitting only the mating type $A \times aa$. Starting from any initial genotypic distribution a stable equilibrium is attained in only two generations at which $(u_e, v_e, w_e) = (0, \frac{1}{2}, \frac{1}{2})$ (HOGBEN 1946; LI 1955). The appropriate recursions can be obtained from the bottom system in Table 1 ($K = 1, K_1 = K_2 = 0$). Systems in which either ($A \times A-$) or ($aa \times aa$) matings, but not both, are prohibited correspond to those of genotypic disassortative mating, IIIa and Vb, in Table 2 and need not be considered. Thus, under all three systems discussed, the proportion of heterozygosity at equilibrium is either equal to or greater than the maximum level of heterozygosity attainable under random mating (i.e. 50 percent).

Mixed self-fertilization and negative assortative mating: We shall now consider models of mixed negative assortative mating, in proportion K , and self-fertilization, in proportion $1-K$, for each of the three systems discussed in the previous section. The appropriate recurrence equations for each of the models can be obtained by setting $K_2 = 0, K_1 = 1-K$ in the formulas given in Table 1. The formulas giving the equilibrium genotypic proportions, the character of the equilibrium, and any pertinent restrictions on the formulas are summarized in Table 3 for each of the three models.

The equilibrium distributions under System I reflect an interaction between two opposing forces, selfing, which eliminates the heterozygotes, and complete negative assortative mating which eliminates one of the homozygotes but retains 50 percent heterozygotes. Two different sets of formulas are required to specify the genotypic proportions at equilibrium, one set applicable when selfing pre-

TABLE 3

Equilibrium states under mixed self-fertilization (in proportion 1-K), and negative assortative mating (in proportion K)

Model	Genotype proportions at equilibrium			Restrictions on u_0, w_0, K	Character
	AA	Aa	aa		
No dominance— System I	$\frac{3+2K-\alpha}{4}$	$\frac{-1-2K+\alpha}{2}$	$\frac{3+2K-\alpha}{4}$	$0 \leq K \leq K^*$	stable
				$K^* < K \leq 1,$ $u_0 = w_0$	unstable
	$\frac{K-3K^2-\beta}{1-9K^2}$	$\frac{1+K}{1+3K}$	$\frac{K-3K^2+\beta}{1-9K^2}$	$K^* \leq K \leq 1$ $u_0 > w_0$	stable
	$\frac{K-3K^2+\beta}{1-9K^2}$	$\frac{1+K}{1+3K}$	$\frac{K-3K^2-\beta}{1-9K^2}$	$K^* \leq K \leq 1$ $u_0 < w_0$	stable
No dominance— System II	$\frac{3+2K-\alpha}{4}$	$\frac{-1-2K+\alpha}{2}$	$\frac{3+2K-\alpha}{4}$	none	stable
Complete dominance	$\frac{1-K}{2+6K}$	$\frac{2K}{1+3K}$	$\frac{1}{2}$	none	stable

$\alpha = \sqrt{4K^2+12K+1}$; $\beta = \sqrt{(3K-1)(6K^3-3K-1)}$; $6K^{*3}-3K^*-1=0$, or $K^* \cong .8365$

dominates, the other when the disassortative mating is more effective. The exact value of K at which the opposing forces are "balanced" (and the two systems produce identical equilibrium distributions), denoted by K^* , can be shown to be given by the solution of the equation $6K^* - 3K^* - 1 = 0$, or approximately, $K^* = .8365$. This equation for K^* can be obtained either from the stability conditions of the equilibrium distributions or from the formulas for the equilibrium distributions given in Table 3. For $K < .8365$, selfing determines the form of the equilibrium distribution; homozygous genotypes are maintained at equal proportions and as K tends to zero, the proportion of heterozygotes also tends to zero. When $K > .8365$, the negative assortative mating is more effective; the frequency of heterozygotes, at equilibrium, lies between .50 and, approximately .55, and as K tends to one, one or the other of the homozygotes tends to be eliminated. For $K = K^* = .8365$, as already noted, the two sets of formulas yield identical values for the equilibrium proportions.

Under mixed selfing and negative assortative mating by System II, the two systems are not in opposition. Both the disassortative mating and the selfing of the heterozygotes produced by the disassortative mating lead to equal proportions of homozygotes. Only one set of equations (Table 3) is needed to specify the equilibrium state which includes homozygotes in equal proportions and heterozygotes in a proportion determined by the amount of negative assortative mating ($0 \leq v_e \leq .562$).

If one allele is dominant, for all values of K there will be 50 percent of the recessive homozygotes at equilibrium. The proportion of heterozygotes at equilibrium reflects only the relative proportion of negative assortative mating and varies accordingly between 0 and $\frac{1}{2}$.

There is considerable evidence that predominantly selfed populations possess much genetic heterogeneity (e.g. JAIN and ALLARD 1960; ALLARD and JAIN 1962). Although there is no evidence which suggests that the proportion of outcrossing in such populations derives from negative assortative mating rather than from random mating, a comparison of the relative effects of these two mixed mating systems on the maintenance of heterozygosity may be of interest. Table 4 shows the equilibrium proportions of heterozygotes expected in populations mating by predominant self-fertilization and (a) random mating, (b) negative assortative mating with complete dominance, and (c) negative assortative mating with no dominance, for different amounts of selfing ($1-K$). The equilibrium proportion of heterozygotes under mixed selfing and random mating depends both on the amount of selfing and on the initial allelic frequencies (ALI and HADLEY 1955; WORKMAN and ALLARD 1962) and is at a maximum when the initial gene frequency, say p_0 , equals .50 (see Table 4). Two interesting observations can be taken from Table 4. First, if the amount of selfing exceeds 50 percent, under partial negative assortative mating, the heterozygosity at equilibrium is approximately the same whether dominance is complete or absent. Second, when the amount of selfing is at least 80 percent, partial negative assortative mating can maintain $1\frac{1}{2}$ to 2 times as much heterozygosity as can be maintained by equivalent amounts of random mating under maximal conditions ($p_0 = .5$), and 4 to 5

TABLE 4

Equilibrium proportions of heterozygotes in populations in which mating is by a mixture of self-fertilization, in proportion 1-K, and (a) random mating, (b) negative assortative mating with complete dominance, and (c) negative assortative mating with no dominance

1-K	(a)*			(b)	(c)†
	p ₀ =.1	p ₀ =.3	p ₀ =.5		
.50	.120	.280	.333	.400	.414
.60	.103	.261	.286	.364	.369
.70	.083	.194	.231	.316	.314
.80	.060	.140	.167	.250	.243
.90	.033	.076	.091	.154	.148
.95	.017	.040	.048	.087	.084
.96	.013	.032	.038	.071	.070
.97	.010	.024	.029	.055	.054
.98	.007	.017	.020	.038	.037
.99	.004	.008	.010	.019	.019

* p₀ is the frequency of A in the initial population. The frequency of heterozygotes at equilibrium, for (a), is given by v_e = 4p₀q₀(t)/(1+t) where t is the proportion of random mating (WORKMAN and ALLARD 1962).

† The equilibrium proportions of heterozygotes for Systems I and II coincide for the amounts of selfing (1-K) considered.

times as much when one allele has an initial frequency of .10. For example, 2 percent of negative assortative mating produces as much heterozygosity (v_e = .038) as 10 percent random mating when p₀ = .10. If the maintenance of genetic heterogeneity at certain loci were advantageous for a predominantly selfed population, the evolution of a system of partial negative assortative mating would be the most efficient method for obtaining the desired variability.

Mixed random mating and negative assortative mating: Under the assumptions outlined in the previous sections, we shall now consider three models for a population in which there is a constant proportion, K, for negative assortative mating, and a constant proportion, 1-K, for random mating. The appropriate recurrence equations are given in Table 1, for K₁ = 0, K₂ = 1-K.

Since the equilibrium conditions for mixed random mating and complete negative assortative mating (System I) are rather complex, they will be considered in some detail. By the symmetry of the recurrence equations for u' and w' (see Table 1) it is obvious that there will exist one equilibrium state at which the proportions of the homozygous genotypes will be equal (u_e = w_e). The genotypic proportions at this equilibrium are given by

$$u_e = w_e = \frac{5 - \sqrt{9 + 8K}}{4(2 - K)}, \text{ and } v_e = \frac{-(2K + 1) + \sqrt{9 + 8K}}{2(2 - K)} \tag{1}$$

It can be shown that this equilibrium distribution will be stable only when both -6u_e³+9u_e² - 6u_e+1 ≤ 0, and -6w_e³+9w_e² - 6w_e+1 ≤ 0, or approximately, when both u_e and w_e are greater than .23845, and 0 < K ≤ .327.

If we assume that u_e ≠ w_e, then the recurrence equations (Table 1) yield the relations

$$\begin{aligned} u_e^2(3v_e-2)+u_e(3v_e^2-5v_e+2)+(v_e-2v_e^2) &= 0, \text{ and} \\ w_e^2(3v_e-2)+w_e(3v_e^2-5v_e+2)+(v_e-2v_e^2) &= 0, \end{aligned} \quad (2)$$

or, supposing that $u_0 > w_0$,

$$u_e = \frac{1-v_e - \sqrt{(3v_e-2)(3v_e^3+3v_e-2)}}{2} \quad (3)$$

$$w_e = \frac{1-v_e + \sqrt{(3v_e-2)(3v_e^3+3v_e-2)}}{2(3v_e-2)}. \quad (3')$$

In addition, we find that at equilibrium,

$$3(1-K)v_e^3 + (3+9K)v_e^2 - (10+4K)v_e + 4 = 0. \quad (4)$$

Now, we must have both $0 \leq u_e \leq 1$ and $0 \leq w_e \leq 1$, and, as can be seen by an examination of (3) and (3'), this will happen only when $(3v_e^3+3v_e-2) \leq 0$, or, approximately, when $v_e \leq .5231$. Substitution of $v_e = .5231$ into (4) shows that an equivalent restriction is given, approximately, by $K \geq .327$. Then, for $K > .327$, there is an equilibrium at which $u_e \neq w_e$, $v_e < .5231$, and it can be shown to be stable. The genotypic proportions at this equilibrium can be obtained from (3), (3'), and (4). Although equation (4) has two real roots between 0 and 1, for $.327 \leq K \leq 1$, only one of them will satisfy the above restriction on v_e .

As was observed for the model of mixed selfing and negative assortative mating by System I, the form of the equilibrium distribution for this system is also determined by the value of K . For $0 < K < .327$, random mating is the more effective force, and at equilibrium, homozygotes are in equal proportions and the proportion of heterozygotes varies between .50 and .5231. When $K = .327$, the two systems are balanced, and the equilibrium proportions given by (1) or by (3), (3') and (4) are equal. For $.327 < K \leq 1$, negative assortative mating is more effective; the proportion of heterozygotes varies between .5231 and .50, but as K tends to one, one or the other of the homozygous genotypes tends to be eliminated.

For mixed random mating and negative assortative mating by System II it can be shown that there is only one equilibrium distribution possible and it is stable for all values of K . The genotypic proportions at equilibrium are given by

$$u_e = w_e = \frac{5 - \sqrt{9+8K}}{4(2-K)}, \text{ and } v_e = \frac{-(2K+1) + \sqrt{9+8K}}{2(2-K)} \quad (5)$$

Although these are the same formulas as given by (1) for System I, the two systems coincide only at the equilibrium state.

If A is dominant over a , there is only one equilibrium state and it is stable. The equilibrium proportions, also derived by LI (1955), are given by (for $K > 0$)

$$u_e = \frac{3+K-2\sqrt{2+2K}}{2(1-K)}, \quad w_e = 1/2, \quad v_e = \frac{-(K+1) + \sqrt{2+2K}}{(1-K)} \quad (6)$$

The most striking feature of these systems is the maintenance, at equilibrium, of significant amounts of heterozygosity, independent of the value of K . For System I, v_e lies between .50 and .5231 and maximum heterozygosity is attained when $K = .327$. Under System II, $.50 \leq v_e \leq .562$, and the maximum occurs when

$K = 1$. When there is dominance, v_e varies between .414 and .500 with the maximum at $K = 1$.

If in the initial population one of the alleles, say A , is at a high frequency, and K is small, the rate of approach to equilibrium will be very slow. For example, suppose that the initial distribution is given by $(AA, Aa, aa) = (.90, .05, .05)$. With $K = .05$, after ten generations the distribution under System I will be (.64, .34, .03) and after 50 generations, (.41, .46, .11). The expected equilibrium distribution is given by (.246, .508, .246). Under the same initial conditions and complete dominance, the distribution after 50 generations will be (.136, .478, .386) and at equilibrium (.08, .42, .50). Therefore, the introduction of an allele, say a , causing disassortative mating, into a random mating population homozygous for A will have little immediate effect unless either K is large, or the new allele is introduced at a high frequency.

Nonrandom mating in Panaxia dominula: In a colony of *Panaxia dominula*, near Oxford, the wing-color pattern polymorphism is determined by two alleles which produce three recognizable genotypes: the common homozygote, *dominula*, the heterozygote, *medionigra*, and the rarer homozygote, *bimacula*. The rarer gene, present in *medionigra* and *bimacula*, is called the *medionigra* gene and occurs at approximately stable frequencies which vary between .01 and .08 in different artificial and natural colonies (SHEPPARD and COOK 1962; WILLIAMSON 1960; FORD 1964).

The analysis of experimental and artificial populations of *Panaxia* has demonstrated three major forces which can have a significant effect on the frequencies at this locus: disassortative mating, differential male fertility, and differential viability between the egg and adult stages (SHEPPARD 1951, 1952, 1956; SHEPPARD and COOK 1962). The mating behaviour was analyzed by an experimental system under which, if matings occurred at random, the numbers of matings between like genotypes would be equal to the number of matings between unlike genotypes. In a total of 199 observed matings, 73 were between like types and 126 between unlike types. This difference is highly significant ($\chi^2=14.12$, $P<0.001$). The tendency for disassortative mating appears to be stronger for the homozygotes, but the numbers involved are too small to be conclusive. The disassortative mating can be considered asymmetric since preliminary observations by SHEPPARD (1951) suggested that the "disassortative pairing is controlled by the female". Together, these observations indicate that a model of mixed random mating and partial negative assortative mating, by either System I or II, might be appropriate for a description of the mating system.

In other experiments, the fertility of both *medionigra* and *bimacula* males was found to be about 75 percent that of the *dominula* males. The data were highly significant for the *medionigra* males ($\chi^2=12.9$, $P<0.001$). No significant differences in female fertility were found although the *dominula* female appeared to have the highest fertility. Observations on artificial colonies of *Panaxia* indicate that both *medionigra* and *bimacula* have approximately 50 percent lower viabilities between the egg and adult stage than does *dominula* (SHEPPARD and COOK 1962).

Under the assumption of random mating, the observed differences in fertility and viability would rapidly eliminate the *medionigra* gene and the populations would contain only *dominula* homozygotes, barring unknown selective forces which might be acting to maintain the polymorphism. Consequently, it has been suggested that the polymorphism might be maintained by a balance between the negative assortative mating and the different selective forces favouring *dominula* (e.g. WILLIAMSON 1960; SHEPPARD and COOK 1962; FORD 1964). In the following discussion, we shall derive a model based on the three selective forces studied by SHEPPARD. By choosing different values for the differential viabilities and by varying the amount of negative assortative mating, we can determine which combinations of these forces, if any, would be sufficient to maintain a stable polymorphism.

In order to construct a mathematically manageable model, we shall restrict the model to a consideration of the distribution of genotypic proportions around the equilibrium point and neglect the contribution from *bimacula*. Since the proportion of *bimacula* at equilibrium will be very small, this should not lead to any serious errors. The model will be based on the following considerations. We let (D, M) and (D', M') represent the proportions of *dominula* and *medionigra* in generations n and $n+1$, such that $D+M=1$. For *dominula* and *medionigra* respectively, the relative fertilities of males are assumed to be in the ratio 1:3/4; for the relative viabilities from egg to adult, we shall consider the ratios 1:4, 1:5, and 1:6, which span the range of relative viabilities observed by SHEPPARD. The relative probabilities for disassortative and random mating are in the ratio $K:1-K$ for *dominula* and $K':1-K'$ for *medionigra*. This will allow the different genotypes to have different tendencies for disassortative mating. In particular, if $K'=K$, the partial disassortative mating is by System I and if $K'=0$, it is by System II. The disassortative mating system will be asymmetric and depend upon female preference. Then, if the relative viabilities of *dominula* and *medionigra* are in the ratio 1: x , it can be shown that the equilibrium proportion of *medionigra* (M_e) will be given by

$$M_e = \frac{x[M_e^2(-8+6K+2K') + M_e(14-2K)+6K]}{x[M_e^2(-8+6K+2K') + M_e(14-2K)+6K] + M_e^2(5-10K+5K') + M_e(20K-18) + 16-10K}$$

and for particular values of x we find that

$$x = .4: -M_e^3(9-38K+29K') + M_e^2(46-64K+4K') + M_e(14K-52) + 12K = 0 \quad (8)$$

$$x = .5: -M_e^3(1-7K+6K') + M_e^2(7-11K+K') + M_e(K-9) + 3K = 0 \quad (9)$$

$$x = .6: -M_e^3(1-32K+31K') + M_e^2(24-46K+6K') - M_e(38+4K) + 18K = 0 \quad (10)$$

Now, the frequency of the *medionigra* gene, say q , is given by $q = \frac{1}{2}M$ ($D+M=1$). Since the value of q at equilibrium varies in different populations between .01 and .08, we need only consider values of M_e such that $.02 \leq M_e \leq .16$. If we set $K=K'=0$ in equations 8 to 10, we find that $M_e=0$ is the only possible solution for M_e in the range $0 \leq M_e \leq 1$. If, however, we let $K=K'=1$, then $M_e = 2/7, 1/3, \text{ and } 3/8$ for (8), (9), and (10), respectively, and these are the levels of heterozygosity which would be maintained by a system of complete negative assortative mating and the selective forces in the values as given. This guarantees that there will be values of K , between 0 and 1 which are sufficient to maintain M_e in the range $.02 \leq M_e \leq .16$.

First, we can compare the estimates of K required to maintain M_e for partial disassortative mating by System I ($K'=K$) and System II ($K'=0$). By substitution of these values for K' into (8 to 10) we find that the estimates of K coincide when $M_e = 0$ and when $M_e = .14, .167$, and $.193$ for (8), (9), and (10), respectively. Between these values the estimates of K differ by at most $.002$, and, for any given value of K , system I is found to produce slightly higher values of M_e than does System II. Above these values, for any K , System II yields somewhat higher values of M_e than does System I. Thus, for the range of values with which we are concerned, that is $.02 \leq M_e \leq .16$, the two systems produce almost identical results. This means that, under the model as derived, the *medionigra* female may mate either completely at random (System II) or exhibit a strong tendency for nonrandom mating (System I) without affecting the level of heterozygosity in the population. If the mating system in nature is determined by female preference, then we need only an estimate of the tendency for nonrandom mating by the *dominula* female in order to determine the effect of disassortative mating upon the maintenance of the polymorphism.

This result could have been predicted on theoretical grounds. As shown in section 4, the two systems, in the absence of selection, coincide at equilibrium for $0 \leq K \leq .327$ and diverge only when, under System I, the nonrandom mating dominates the form of the equilibrium distribution. For $K < .327$ System II always produced slightly higher amounts of heterozygosity. The effect of selection is to raise the value of K below which the systems coincide and, as can be shown, as the selection pressure against *medionigra* is increased, the systems agree over a wider range of values.

From equations 8 to 10 we can estimate, for given M_e and x , the amount of partial disassortative mating, K , required to maintain the polymorphism. Table 5 contains the estimates of K for System I ($K'=K$) for the three different relative viabilities, $1:x$ ($x = .4, .5, .6$). Since the values for System II are almost identical, they need not be presented. From Table 5 it is clear that the amount of nonrandom mating required to maintain the polymorphism at a given frequency depends to a great extent on the relative viabilities of *dominula* and *medionigra*. At a ratio of $1:4$ almost twice as much nonrandom mating is required as would be at a ratio of $1:6$.

With the use of the Table 5 it is possible to estimate amounts of negative assortative mating which would be required to maintain the polymorphism for different estimates of the viabilities. Two estimates of the viability ratio ($1:x$) have

TABLE 5

The proportions of disassortative mating (K), by System I, required to maintain different equilibrium proportions of var. medionigra (M_e)

x	M_e							
	.02	.04	.06	.08	.10	.12	.14	.16
.4	.083	.161	.234	.304	.370	.436	.500	.563
.5	.055	.115	.170	.222	.276	.333	.382	.435
.6	.042	.083	.124	.166	.209	.249	.292	.335

been obtained by SHEPPARD from studies of artificial colonies of *Panaxia*: 1:43, in 1952, from a colony at Hinksey; and, 1:58, in 1961, from a colony at Ness. Owing to the small sample sizes upon which these estimates were based, the standard errors are quite large. The approximate 95 percent confidence limits are: $.19 \leq x \leq .65$ for the estimate from Hinksey, and $.33 \leq x \leq .84$ for the estimate from Ness. The frequency of the *medionigra* gene at Hinksey is now .065 and appears to be relatively stable: .065 in 1959, .062 in 1960, and .073 in 1961 (SHEPPARD and COOK 1962). From Table 5 we see that for the ratios 1:43 and 1:58 about 45 and 26 percent nonrandom mating, respectively, would be required to maintain the polymorphism at Hinksey. The Hinksey colony was originally started in 1951 from 4000 fertile eggs of *dominula* \times *medionigra* matings giving an initial gene frequency of .25 for the *medionigra* gene. A model for which $(D_0, M_0, B_0) = (.50, .50, .00)$, $K = .25$, and $x = .58$, was found to attain a stable equilibrium gene frequency of about .06 in seven generations and consequently provides a good fit to the observed data.

It is difficult to relate these estimates of K to the nonrandom mating observed in the laboratory, although a crude estimate of K can be derived from the data. If U and L denote the proportion of unlike matings and like matings, respectively, then K can be estimated by $K = U - L = .266$, and $\sigma^2 K = 4\sigma^2 U = .00467$. This estimate is not corrected for the differences between the frequency of mating types in the laboratory and the probability of their occurrence in nature. Since each male can mate several times, even such a correction might not be meaningful. In the experimental studies by SHEPPARD, K varied between 0 and $4M$, depending on which mating types were permitted, so values of $K = 2M$ or $K = 3M$ might occur in the natural populations. The apparent agreement between the estimate of K derived from the experimental data, $K = .266$, with that used in the model which fit the Hinksey data, $K = .25$, could be merely coincidental.

The applicability of these estimates of K to the natural populations depends upon the accuracy of the estimates of the viabilities, which, as already noted, have large standard errors. Since, even in large populations, there can be considerable fluctuations in the estimate of the selective forces over a period of several years (ALLARD and WORKMAN 1963; WORKMAN and ALLARD 1964), an accurate estimate of the amount of disassortative mating in nature may require estimates of the viabilities in several different years. It is also possible that estimates of the relative viabilities in one environment may be inappropriate for a description of a different population. The present gene frequency at Cothill is about .02 (FORD 1964) which, for the estimated viabilities ($x = .43, .58$), would require only 15 or 9 percent disassortative mating. The tendency for disassortative mating is, presumably, related more to the genotype than to the selective forces in the environment and therefore it is more likely to be similar in different environments than are the intensities of the selective forces or even the nature of the selective forces. The different estimates of K from Cothill and Hinksey could therefore be indicative of differences in the relative viabilities of *medionigra* in the two colonies. Alternatively, one could argue that the tendency for disassortative mating is, in part, a function of the prevailing genotypic frequencies and

that it is affected by such factors as population size, population density and the vigour of the *medionigra* male which may vary according to its degree of adaptation to the different colonies. Although it is impossible to distinguish between these alternatives on the basis of the present data, it may be possible to design experiments which would consider the effects of differences in the relative frequencies, population densities, and the environmental conditions on the amount of disassortative mating. WILLIAMSON (1960) considered several systems which might lead to the establishment of the polymorphism at Cothill and concluded that the three forces studied by SHEPPARD were of the right type but that the estimated values of the forces were not of the right strength. In his model, K was taken to be, approximately, the value of M_e , and x was .5. As shown in Table 5, for $x = .5$, K must be approximately $2.75M_e$ in order to maintain the polymorphism. The present analysis suggests that the failure of these estimates to fit the data at Cothill may be due either to a difference in one or more of the selective forces, or to a difference in the amount of disassortative mating operating at Cothill and in the experimental colonies.

It does not seem possible at present to determine whether a model based on partial disassortative mating, differential male fertility, and differential viability will be sufficient to explain the wing color pattern polymorphism in *Panaxia*. It is hoped, however, that the present analysis has served both to illustrate the relationships between the postulated selective forces and to suggest those areas of research which would allow a critical evaluation of the model itself. The model has demonstrated that a system of mixed negative assortative mating and random mating can maintain heterozygosity even when there are strong selective forces favouring one of the homozygotes.

I wish to thank PROFESSOR P. M. SHEPPARD, PROFESSOR J. M. THODAY, DR. A. R. G. OWEN, DR. J. B. GIBSON, and DR. M. H. WILLIAMSON for their helpful comments on this manuscript. A portion of the work was done during the tenure of a Staff Fellowship at the National Institutes of Health, Bethesda, Maryland.

SUMMARY

Theoretical models have been developed for populations mating under (a) exclusive negative assortative mating (b) mixed self-fertilization and negative assortative mating, and (c) mixed random mating and negative assortative mating. Particular emphasis was placed on the derivation of the genotypic distributions at equilibrium and the amount of heterozygosity which these mating systems can maintain. A model of mixed random mating and negative assortative mating was used to consider the maintenance of a polymorphism for wing color pattern which is found in certain colonies of the Scarlet Tiger moth, *Panaxia dominula*. The analysis showed that partial negative assortative mating can maintain a polymorphism even when one of the homozygotes possesses a large selective advantage.

LITERATURE CITED

- ALLARD, R. W., and S. K. JAIN, 1962 Population studies in predominantly self-pollinated species. II. Analysis of quantitative genetic changes in a bulk-hybrid population of barley. *Evolution* 16: 90-101.

- ALLARD, R. W., and P. L. WORKMAN, 1963 Population studies in predominantly self-pollinated species. IV. Seasonal fluctuations in estimated values of genetic parameters in lima bean populations. *Evolution* **18**: 470-480.
- ALI, M., and H. H. HADLEY, 1955 Theoretical proportion of heterozygosity in populations with various proportions of self- and cross-fertilization. *Agron. J.* **47**: 589-590.
- BATTAGLIA, B., 1963 Deviations from panmixia as a consequence of sex-determination in the marine copepod, *Tisbe reticulata*. (Abstr.) Proc. 11th Intern. Cong. Genet. **1**: 152.
- DAHLBERG, G., 1948 *Mathematical Methods for Population Genetics*. Interscience, New York.
- FISHER, R. A., and E. B. FORD, 1947 The spread of a gene in natural conditions in a colony of the moth *Panaxia dominula* L. *Heredity* **1**: 143-174.
- FORD, E. B., 1964 *Ecological Genetics*. Methuen, London.
- HOGBEN, L., 1946 *An Introduction to Mathematical Genetics*. Norton, New York.
- JAIN, S. K., and R. W. ALLARD, 1960 Population studies in predominantly self-pollinated species. I. Evidence for heterozygote advantage in a closed population of barley. Proc. Natl. Acad. Sci. U.S. **46**: 1371-1377.
- KETTLEWELL, H. B. D., 1942 A survey of the insect *Panaxia (Callimorpha) dominula* L. Proc. S. London Entomol. Nat. Hist. Soc., 1-49. — 1956 A resume of investigations on the evolution of melanism in the Lepidoptera. Proc. Roy. Soc. Lond. B. **145**: 297-303.
- KETTLEWELL, H. B. D., and R. J. BERRY, 1961 The study of a cline. *Heredity* **16**: 403-414.
- LI, C. C., 1955 *Population Genetics*. University of Chicago Press, Chicago, Illinois.
- LOWTHER, J. K., 1961 Polymorphism in the white-throated sparrow, *Zonotrichia albicollis* (Gmelin). *Canad. J. Zool.* **39**: 281-292.
- NAYLOR, A. F., 1962 Mating systems which could increase heterozygosity for a pair of alleles. *Am. Naturalist* **96**: 51-60. — 1963 A theorem on possible kinds of mating systems which tend to increase heterozygosity. *Evolution* **17**: 369-370.
- OWEN, A. R. G., 1953 A genetical system admitting of two distinct stable equilibria under natural selection. *Heredity* **7**: 97-102.
- SHEPPARD, P. M., 1951 A quantitative study of two populations of the moth *Panaxia dominula* (L.). *Heredity* **5**: 349-378. — 1952 A note on non-random mating in the moth *Panaxia dominula* (L.). *Heredity* **6**: 239-241. — 1953 Polymorphism and population studies. Symp. Soc. Exptl. Biol. **7**: 274-289. — 1956 Ecology and its bearing on population genetics. Proc. Roy. Soc. Lond. B. **145**: 308-315.
- SHEPPARD, P. M., and L. M. COOK, 1962 The manifold effects of the medionigra gene of the moth *Panaxia dominula* and the maintenance of a polymorphism. *Heredity* **17**: 415-426.
- WATTERSON, G. A., 1959 Non-random mating, and its effect on the rate of approach to homozygosity. *Ann. Human Genet.* **23**: 204-220.
- WILLIAMSON, M. H., 1960 On the polymorphism of the moth *Panaxia dominula* (L.). *Heredity* **15**: 139-151.
- WORKMAN, P. L., and R. W. ALLARD, 1962 Population studies in predominantly self-pollinated species. III. A matrix model for mixed selfing and random outcrossing. Proc. Natl. Acad. Sci. U.S. **48**: 1318-1325. — 1964 Population studies in predominantly self-pollinated species. V. Analysis of differential and random viabilities in mixtures of competing pure lines. *Heredity* **19**: 181-189.
- WRIGHT, S., 1921 Systems of mating. III. Assortative mating based on somatic resemblance. *Genetics* **6**: 144-161.