

## POPULATION GENETIC THEORY OF THE ASSORTATIVE MATING, SEXUAL SELECTION AND NATURAL SELECTION OF THE TWO-SPOT LADYBIRD, *ADALIA BIPUNCTATA*

PETER O'DONALD, MATTHEW DERRICK, MICHAEL MAJERUS  
AND JACQUELINE WEIR

*Department of Genetics, University of Cambridge, Downing Street, Cambridge CB2 3EH, England*

Received 31.v.83

### 1. INTRODUCTION

The ladybird *Adalia bipunctata* is polymorphic for the colour and pattern of elytra and pronotum, with a wide range of distinct forms from all red to almost all black (Hodek, 1973). These forms are determined by a multiple allelic series comprising at least 11 alleles at a single locus. In general melanic forms (black with red spots) are dominant to non-melanic forms (red with black spots), but some distinct heterozygous phenotypes are formed in particular combinations with very rare alleles.

The different alleles and the phenotypes they produce are shown in fig. 1. Except for the  $s^a/s^d$  heterozygotes, all heterozygotes with distinct phenotypes carry one of the very rare alleles, not found in many natural populations. Eliminating the rare alleles, we see in fig. 1(ii) that a dominance hierarchy exists among the melanic and non-melanic forms, except that the most recessive allele  $s^d$  is co-dominant with the next most recessive allele  $s^a$ . Sexual selection maintains melanic phenotypes in some populations (Muggleton, 1979; O'Donald and Muggleton, 1979; Majerus, O'Donald and Weir, 1982a). Non-melanic phenotypes may show assortative mating (Majerus, O'Donald and Weir, 1982a).

Detailed studies of sexual selection and assortative mating are now being carried out on ladybirds from a large population at Keele in England. This population is polymorphic for the alleles  $s^{lu}$ ,  $s^m$ ,  $s^p$ ,  $s^t$  and  $s^a$ , which form a complete, ordered dominance hierarchy as shown in fig. 1(ii). These are the alleles normally found in English populations. The alleles  $s^m$  (*quadrifasciata*, Q),  $s^t$  (*typica*, T) and  $s^a$  (*annulata*, A) are common;  $s^{lu}$  (*sublunata*, L) is rare (about 0.7 per cent) and  $s^p$  (*sempustulata*, X) is fairly rare (about 5 per cent of the population). *Sublunata* (black with two red spots), *quadrifasciata* (black with four red spots) and *sempustulata* (black with six red spots) are melanic; *typica* (red with two black spots) and *annulata* (red with two extended black spots) are non-melanic. Majerus, O'Donald and Weir (1982a) observed mating pairs in both the wild population and in population cages at fixed phenotypic frequencies. They also carried out direct observations of matings in mating chambers. The frequencies of the matings of *quadrifasciata* and *typica* males are consistent with Karlin and O'Donald's model combining sexual selection with assortative mating of two phenotypes determined by two alleles one dominant to the other (Karlin and O'Donald, 1978). According to this model, 23 per cent of all females

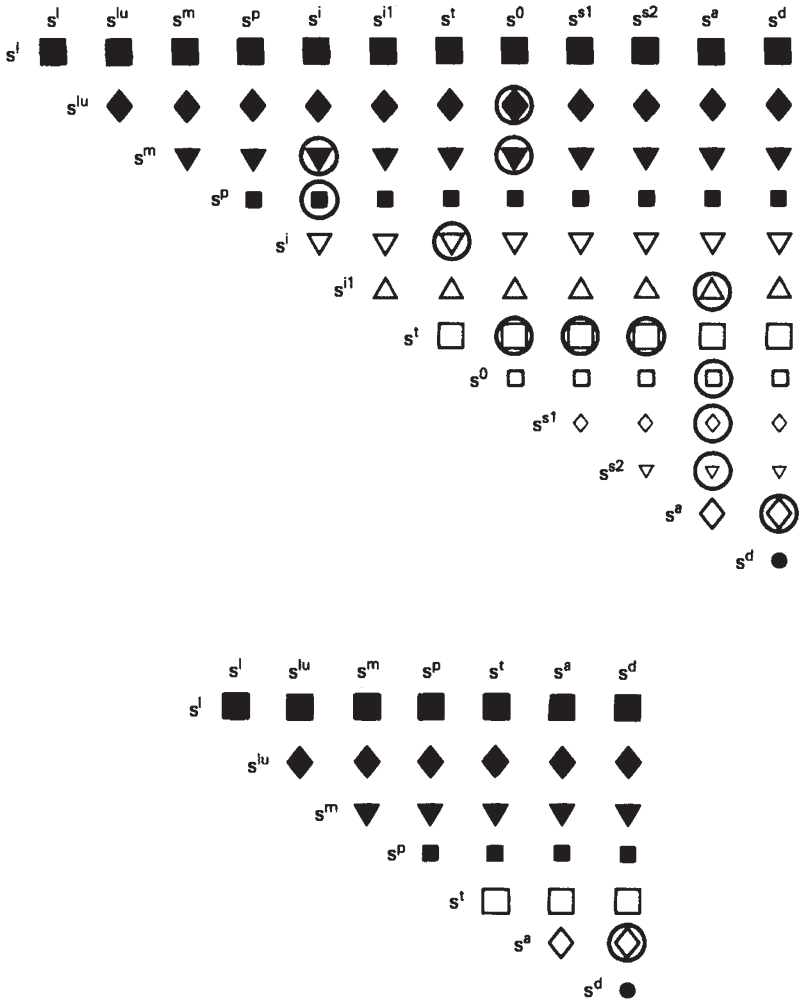


FIG. 1.

(i) The genetics of melanic and non-melanic phenotypes of *Adalia bipunctata*. (ii) The genetics of melanic and non-melanic phenotypes omitting alleles not found in European populations or found only very rarely.

In this figure, heterozygotes that are encircled have distinct phenotypes. The melanic forms, in order of decreasing melanistic pigment, are determined by the alleles  $s^l$ ,  $s^{lu}$ ,  $s^m$ ,  $s^p$  and  $s^d$ . The alleles  $s^i$ ,  $s^{i1}$ ,  $s^0$ ,  $s^{s1}$  and  $s^{s2}$  are very rare;  $s^0$ ,  $s^{s1}$  and  $s^{s2}$  are not found in European populations.

express a non-assorting preference for *quadrimaculata* males and 16 per cent of *quadrimaculata* females express an assorting preference for *quadrimaculata* males. These estimates of preference were based on the frequencies of matings in population cages and mating chambers in which only the two phenotypes *quadrimaculata* and *typica* were used. From the data of mating frequencies in the wild population at Keele, 21 per cent of females prefer melanic (M) males (M includes L, Q and X phenotypes). This estimate

from the wild population is completely consistent with population cage and mating chamber estimates of the non-assorting preference for *quadrifasciata*.

Later experiments proved that the female preference for *quadrifasciata* was genetically determined (Majerus, O'Donald and Weir, 1982b). In a selection line, the female preference increased rapidly reaching 55 per cent after four generations, but showed no increase in a control line. Female preference alone had been selected, for selected females tested with original unselected males showed the high level of preference, while unselected females tested with males from the selected line showed the original low level of preference.

The non-melanic phenotypes mate assortatively in the wild population at Keele: *annulata*  $\times$  *annulata* and *typica*  $\times$  *typica* matings occur more often than at random. Assortative mating is evidence of female choice, not male competition, since it is most unlikely that a male would compete only, or more strongly, for a female with the same phenotype as his own. Male ladybirds have not been observed to compete for females.

Only an elaborate genetical model can adequately describe the mating system of the two-spot ladybird. Karlin and O'Donald's (1978) original model of mixed sexual selection and assortative mating is too simple, for it assumes that only two alleles are segregating genetically. A later, completely general model of sexual selection incorporates any number of alleles showing both dominance and co-dominance in their phenotypic effects, but omits assortative mating (Karlin and O'Donald, 1981). Assortative mating greatly increases the complexity of the analysis' for the genotypes never occur at Hardy-Weinberg ratios; whereas, by sexual selection alone, Hardy-Weinberg ratios are finally reached at equilibrium, greatly simplifying the analysis.

The simplest model that may be sufficient to describe the sexual selection and assortative mating of the ladybirds must incorporate three alleles: an allele *S* determines a dominant melanic phenotype *M* (thus implying that the three melanic phenotypes *L*, *Q* and *X* are each equivalent to *M* and their corresponding alleles  $s^l$ ,  $s^m$  and  $s^p$  are each equivalent to the single allele *S*); the allele  $s^t$  determines the *typica* phenotype *T*, which is recessive to *M* but dominant to *A*; and the allele  $s^a$  determines the recessive *annulata* phenotype *A*. Non-assorting female preferences give rise to sexual selection in favour of *M* and perhaps also *T* and *A* males. Assorting preferences, expressed only by females who themselves possess the phenotype they prefer, give rise to assortative mating of *A*, *T* and perhaps *M* phenotypes. Thus in this paper, we extend Karlin and O'Donald's original model of mixed sexual selection and assortative mating to a genetical system of three alleles. Natural selection is also incorporated into the model, which is then fitted to new and much more extensive data on ladybird frequencies at Keele. We thus hope to provide an evolutionary explanation of the polymorphism of the two-spot ladybird.

## 2. A THEORETICAL MODEL

The three phenotypes *M*, *T* and *A* are determined by alleles *S*,  $s^t$  and  $s^a$ , with dominance order  $S > s^t > s^a$ . Females may express both non-assorting and assorting preferences for males of each phenotype thus giving

rise to sexual selection (by non-assorting preferences) and assortative mating (by assorting preferences).

(i) *Notation*

We use the following notation for frequencies and preferences.

<i>Genotypic frequencies</i>	<i>SS</i>	$u_{11}$
	$Ss'$	$u_{12}$
	$Ss^a$	$u_{13}$
	$s's'$	$u_{22}$
	$s's^a$	$u_{23}$
	$s^as^a$	$u_{33}$

*Phenotypic Frequencies*

Melanic	$\pi_1 = u_{11} + u_{12} + u_{13}$
Typica	$\pi_2 = u_{22} + u_{23}$
Annulata	$\pi_3 = u_{33}$

*Gene frequencies*

$$\begin{aligned}
 S: & \quad p_1 = u_{11} + (u_{12} + u_{13})/2 \\
 s': & \quad p_2 = u_{22} + (u_{12} + u_{23})/2 \\
 s^a: & \quad p_3 = u_{33} + (u_{13} + u_{23})/2 \\
 & \quad q_2 = u_{22} + \frac{1}{2}u_{23}
 \end{aligned}$$

*Mating preferences*

Phenotype	<i>M</i>	<i>T</i>	<i>A</i>
Non-assorting	$\alpha_1$	$\alpha_2$	$\alpha_3$
Assorting	$a_1$	$a_2$	$a_3$

The non-assorting preferences are expressed by all females regardless of phenotype or genotype. The assorting preferences are expressed only by females with the same phenotype as that they prefer.

(ii) *Recurrence relationships and equilibrium frequencies*

The matings are polygynous in ladybirds. The males who were chosen for preferential mating may also be chosen at random, in proportion to their population frequencies, by females that express no preference. Thus, consider the matings of  $SS\delta\delta \times SS\varphi\varphi$ . The genotype *SS* occurs at a population frequency of  $u_{11}$ . Of the females, proportions  $\alpha_1$  and  $a_1$  express non-assorting and assorting preferences for *M* males. *SS* males form a proportion  $u_{11}/\pi_1$  of the *M* males. Hence the proportion of preferential matings of  $SS\delta\delta \times SS\varphi\varphi$  is  $u_{11} (\alpha_1 + a_1) \cdot (u_{11}/\pi_1)$  and the proportion of

random matings is  $u_{11} \cdot u_{11} \cdot (1 - \theta - a_1)$  where  $\theta$  is the total proportion of females expressing non-assorting preferences. Thus we obtain the frequencies of matings shown in table 1, and hence the following recurrence equations of genotypic frequencies:

$$\begin{aligned}
 u'_{11} &= p_1^2((\alpha_1 + a_1)/\pi_1 + (1 - \theta - a_1)) \\
 u'_{12} &= p_1 u_{12}((\alpha_1 + a_1)/\pi_1 + (1 - \theta - a_1)) \\
 &\quad + p_1 q_2(\alpha_1/\pi_1 + \alpha_2/\pi_2 + 2 - 2\theta - a_1 - a_2) \\
 u'_{13} &= p_1[u_{13}((\alpha_1 + a_1)/\pi_1 + (1 - \theta - a_1)) \\
 &\quad + u_{33}(\alpha_1/\pi_1) + \alpha_3 + u_{33}(2 - 2\theta - a_1 - a_3) \\
 &\quad + u_{23}(\alpha_1/\pi_1 + \alpha_2/\pi_2 + 2 - 2\theta - a_1 - a_2)/2] \\
 u'_{22} &= u_{12}^2[(\alpha_1 + a_1)/\pi_1 + (1 - \theta - a_1)]/4 \\
 &\quad + q_2[u_{12}(\alpha_1/\pi_1 + \alpha_2/\pi_2 + 2 - 2\theta - a_1 - a_2)/2 \\
 &\quad + q_2((\alpha_2 + a_2)/\pi_2 + (1 - \theta - a_2))] \\
 u'_{23} &= u_{12}u_{13}((\alpha_1 + a_1)/\pi_1 + (1 - \theta - a_1)) \\
 &\quad + u_{23}(\alpha_1/\pi_1 + \alpha_2/\pi_2 + 2 - 2\theta - a_1 - a_2)/2 \\
 &\quad + u_{33}(\alpha_1/\pi_1 + 2 - 2\theta - a_1 - a_3)]/2 \\
 &\quad + q_2[(u_{13}(\alpha_1/\pi_1 + \alpha_2/\pi_2 + 2 - 2\theta - a_1 - a_2)/2 \\
 &\quad + u_{33}(\alpha_2/\pi_2 + 2 - 2\theta - a_2 - a_3) \\
 &\quad + u_{23}((\alpha_2 + a_2)/\pi_2 + (1 - \theta - a_2))] + \alpha_3 p_2 \\
 u'_{33} &= u_{13}^2[(\alpha_1 + a_1)/\pi_1 + (1 - \theta - a_1)]/4 \\
 &\quad + u_{13}u_{23}(\alpha_1/\pi_1 + \alpha_2/\pi_2 + 2 + 2\theta - a_1 - a_2)/4 \\
 &\quad + u_{13}u_{33}(\alpha_1/\pi_1 + 2 - 2\theta - a_1 - a_3)/2 \\
 &\quad + u_{23}^2[(\alpha_2 + a_2)/\pi_2 + (1 - \theta - a_2)]/4 \\
 &\quad + u_{23}u_{33}(\alpha_2/\pi_2 + 2 - 2\theta - a_2 - a_3)/2 + \alpha_3 p_3 \\
 &\quad + u_{33}^2(1 - \theta - a_3) + u_{33}a_3.
 \end{aligned}$$

These equations can be used to give the gene frequencies for the next generation:

$$p'_1 = p_1[1 + \alpha_1/2\pi_1 - \theta/2 + (a_1 - a_2)\pi_2/2 + (a_1 - a_3)\pi_3/2] \quad (1)$$

and

$$\begin{aligned}
 p'_2 &= p_2[1 + \alpha_1/2\pi_1 - \theta/2 + (a_1 - a_2)\pi_2/2 + (a_1 - a_3)\pi_3/2] \\
 &\quad + q_2(\alpha_2/\pi_2 - \alpha_1/\pi_1 + a_2 - a_1). \quad (2)
 \end{aligned}$$

The equilibrium frequencies can be derived from equations (1) and (2) as follows. Since, at equilibrium,

$$\alpha_2/\pi_2 - \alpha_1/\pi_1 + a_2 - a_1 = 0 \quad (3)$$

and

$$\alpha_1/\pi_1 - \theta + (a_1 - a_2)\pi_2 + (a_1 - a_3)\pi_3 = 0, \quad (4)$$

TABLE 1  
Hypothetical frequencies of the matings

Mating	Frequency
$SS \times SS$	$u_{11}^2[(\alpha_1 + a_1)/\pi_1 + (1 - \theta - a_1)]$
$SS \times Ss'$	$2u_{11}u_{12}[(\alpha_1 + a_1)/\pi_1 + (1 - \theta - a_1)]$
$SS \times Ss^a$	$2u_{11}u_{13}[(\alpha_1 + a_1)/\pi_1 + (1 - \theta - a_1)]$
$SS \times s's'$	$u_{11}u_{22}[\alpha_1/\pi_1 + \alpha_2/\pi_2 + (2 - 2\theta - a_1 - a_2)]$
$SS \times s's^a$	$u_{11}u_{23}[\alpha_1/\pi_1 + \alpha_2/\pi_2 + (2 - 2\theta - a_1 - a_2)]$
$SS \times s^a s^a$	$u_{11}u_{33}[\alpha_1/\pi_1 + (2 - 2\theta - a_1 - a_3)] + \alpha_3 u_{11}$
$Ss' \times Ss'$	$u_{12}^2[(\alpha_1 + a_1)/\pi_1 + (1 - \theta - a_1)]$
$Ss' \times Ss^a$	$2u_{12}u_{13}(\alpha_1 + \alpha_1)/\pi_1 + (1 - \theta - a_1)]$
$Ss' \times s's'$	$u_{12}u_{22}[\alpha_1/\pi_1 + \alpha_2/\pi_2 + (2 - 2\theta - a_1 - a_2)]$
$Ss' \times s's^a$	$u_{12}u_{23}[\alpha_1/\pi_1 + \alpha_2/\pi_2 + (2 - 2\theta - a_1 - a_2)]$
$Ss' \times s^a s^a$	$u_{12}u_{33}[\alpha_1/\pi_1 + (2 - 2\theta - a_1 - a_3)] + \alpha_3 u_{12}$
$Ss^a \times Ss^a$	$u_{13}^2[(\alpha_1 + a_1)/\pi_1 + (1 - \theta - a_1)]$
$Ss^a \times s's'$	$u_{13}u_{22}[\alpha_1/\pi_1 + \alpha_2/\pi_2 + (2 - 2\theta - a_1 - a_2)]$
$Ss^a \times s's^a$	$u_{13}u_{23}[\alpha_1/\pi_1 + \alpha_2/\pi_2 + (2 - 2\theta - a_1 - a_2)]$
$Ss^a \times s^a s^a$	$u_{13}u_{33}[\alpha_1/\pi_1 + (2 - 2\theta - a_1 - a_3)] + \alpha_3 u_{13}$
$s's' \times s's'$	$u_{22}^2[(\alpha_2 + a_2)/\pi_2 + (1 - \theta - a_2)]$
$s's' \times s's^a$	$u_{22}u_{23}[(\alpha_2 + a_2)/\pi_2 + (1 - \theta - a_2)]$
$s's' \times s^a s^a$	$u_{22}u_{33}[\alpha_2/\pi_2 + (2 - 2\theta - a_2 - a_3)] + \alpha_3 u_{22}$
$s's^a \times s's^a$	$u_{23}^2[(\alpha_2 + a_2)/\pi_2 + (1 - \theta - a_2)]$
$s's^a \times s^a s^a$	$u_{23}u_{33}[\alpha_2/\pi_2 + (2 - 2\theta - a_2 - a_3)] + \alpha_3 u_{23}$
$s^a s^a \times s^a s^a$	$u_{33}(\alpha_3 + a_3) + u_{33}^2(1 - \theta - a_3)$

then, given that  $\pi_3 = 1 - \pi_2 - \pi_1$ , we obtain

$$\alpha_1/\pi_1 + \pi_2(a_3 - a_2) + \pi_1(a_3 - a_1) - \theta + a_1 - a_3 = 0. \quad (5)$$

Re-arranging (3), we have

$$\pi_2 = \alpha_2 \pi_1 / (\pi_1(a_1 - a_2) + \alpha_1).$$

When this equation is put into (5) it gives a cubic equation in the equilibrium frequency  $\pi_1^*$ :

$$\pi_1^{*3}(a_1 - a_2)(a_3 - a_1) + \pi_1^{*2}[\alpha_1(a_3 - a_1) + \alpha_2(a_3 - a_2) - \theta(a_1 - a_2) - (a_1 - a_2)(a_3 - a_1)] - \pi_1^* \alpha_1(\theta - 2a_1 + a_2 + a_3) + \alpha_1^2 = 0. \quad (6)$$

Hence for  $\pi_2^*$  and  $\pi_3^*$

$$\pi_2^* = \alpha_2 \pi_1^* / (\pi_1^*(a_1 - a_2) + \alpha_1) \quad (7)$$

$$\pi_3^* = 1 - \pi_2^* - \pi_1^*. \quad (8)$$

Two cases of these results may apply to the ladybirds.

*Case (i):* Sexual selection of melanic and *typica* males. If  $\theta = \alpha_1 + \alpha_2$  ( $\alpha_3 = 0$ ), then  $\pi_1$  is given by one of two equations, either

$$\pi_1^* = \alpha_1 / (a_3 - a_1), \quad \text{hence } \pi_2^* = \alpha_2 / (a_3 - a_2), \quad (9)$$

or

$$\pi_1^{*2}(a_1 - a_2) + \pi_1^*(\theta - a_1 + a_2) - \alpha_1 = 0. \quad (10)$$

The polymorphic equilibria given by the solutions of the recurrence equations are generally globally stable. Usually only one of the solutions of (6) gives rise to applicable values of  $\pi_1^*$ ,  $\pi_2^*$  and  $\pi_3^*$  (i.e., values for

which  $\pi_1^* + \pi_2^* + \pi_3^* = 1$ ). Sometimes two such applicable solutions have been found to exist. If so, they have been mutually exclusive: one solution has admitted a polymorphism of all three phenotypes ( $\pi_1^*, \pi_2^*, \pi_3^* > 0$ ); the other solution has entailed the absence of one phenotype, usually *annulata*, from the polymorphism ( $\pi_3 = 0$ ). Given the presence or absence of *annulata*, either one or other polymorphic equilibrium has been found to be globally stable.

*Case (ii)*: Sexual selection of only melanic males. If sexual selection favours melanic males only, and hence  $\theta = \alpha_1, \alpha_2 = \alpha_3 = 0$ , then there are always two alternative polymorphic equilibria: either

$$\pi_1^* = \alpha_1 / (a_3 - a_1), \quad \pi_2^* = 0, \quad \pi_3^* = 1 - \pi_1^* \quad (11)$$

or

$$\pi_1^* = \alpha_1 / (a_2 - a_1), \quad \pi_2^* = 1 - \pi_1^*, \quad \pi_3^* = 0 \quad (12)$$

The equilibria reached depend on the magnitudes of the assorting parameters  $a_2$  and  $a_3$ , always provided that  $a_2, a_3 > a_1$  which is the condition for the existence of the polymorphism. If  $a_3 > a_2$ , the equilibrium frequencies are given by (11): if  $a_2 > a_3$ , they are given by (12). This is the expected result: by assortative mating alone, the phenotype with the greatest assorting preference in its favour replaces the others. Assortative mating gives a roughly constant mating advantage to the preferred males: the female preference for a particular phenotype is expressed only by females who possess that phenotype; hence it is expressed in strict proportion to the frequencies of the males preferred, thus giving them a constant mating advantage. But the mating advantage of sexual selection is frequency-dependent (O'Donald, 1980a): a constant proportion of all females express the non-assorting preference; when preferred males are rare, a relatively greater proportion of females prefer them than when they are common. Sexual selection can thus balance assortative mating to produce a stable polymorphism.

Consider *Case (ii)* in which melanic males alone are sexually selected ( $\theta = \alpha_1, \alpha_2 = \alpha_3 = 0$ ). The melanics reach equilibria at frequencies of either

$$\pi_1^* = \alpha_1 / (a_3 - a_1)$$

or

$$\pi_1^* = \alpha_1 / (a_2 - a_1).$$

These values represent polymorphisms if  $\alpha_1 < (a_3 - a_1)$  and  $\alpha_1 < (a_2 - a_1)$ . The effects of assortative mating give a constant overall advantage of *typica* or *annulata* depending on whether  $a_2 > a_3$  or  $a_3 > a_2$ . If  $\pi_1$  falls below  $\pi_1^*$ , the sexual selective advantage conferred by  $\alpha_1$  increases, and  $\pi_1$  increases again. If  $\pi_1$  rises above  $\pi_1^*$ , the sexual advantage decreases, and with it,  $\pi_1$ . The polymorphic equilibrium is thus maintained by the balance between the constant mating advantage of assortative mating and the frequency-dependent mating advantage (or "rare male effect") of sexual selection.

So far we have neglected the effects of natural selection of the phenotypes. This is easily incorporated into the recurrence equations and will be analysed numerically in section 4.

## 3. FIELD DATA ON PHENOTYPIC FREQUENCIES OF LADYBIRDS

In 1982, two-spot ladybirds were sampled in large numbers on the campus of Keele University, Staffordshire, England. The phenotypes were recorded both of mating pairs and of individual ladybirds. Table 2 shows the total numbers of each phenotype sampled in 1982. Samples were taken in May, June, July and August. The frequencies show no significant differences between months. The monthly totals were therefore added together to give the grand totals of table 2. These frequencies are similar to those of a much smaller sample taken in the previous year (Majerus, O'Donald and Weir, 1982*a*).

The field data corroborate the results of the population cage and mating chamber observations in the laboratory (Majerus, O'Donald and Weir, 1982*a, b*). In these experiments, only *quadrimaculata* and *typica* phenotypes were used for mating. Their mating frequencies were fitted to Karlin and O'Donald's model of mixed sexual selection and assortative mating (Karlin and O'Donald, 1978), giving the following maximum likelihood (ML) estimates of mating preferences for *quadrimaculata*:

$$\hat{\alpha}_1 = 0.230 \quad (\text{non-assorting preference for Q})$$

$$\hat{\alpha}_1 = 0.162 \quad (\text{assorting preference for Q}).$$

TABLE 2

*Phenotypes of individual ladybirds and mating pairs at Keele University in 1982*

(i) Numbers of individual ladybirds						
Phenotypes	Females			Males		
L	47			57		
Q	2921			2907		
X	414			446		
T	4556			4422		
A	2027			2083		
Totals	9965			9915		

(ii) Numbers of mating pairs						
Females	Males					Totals
	L	Q	X	T	A	
L	0	3	1	3	0	7
Q	3	167	27	156	55	408
X	1	18	9	22	9	59
T	10	263	39	228	65	605
A	4	98	9	59	79	249
Totals	18	549	85	468	208	1328

In this table phenotypes are symbolised as follows:

L	<i>sublunata</i>
Q	<i>quadrimaculata</i>
X	<i>sexpustulata</i>
T	<i>typica</i>
A	<i>annulata</i>



In the model, a fixed proportion of the females are assumed to express a preference. Suppose Q males alone mate preferentially without assortment. The matings would take place at frequencies

$$\alpha_1 + v(1 - \alpha_1) \quad \text{for Q males}$$

$$(1 - v)(1 - \alpha_1) \quad \text{for T males}$$

where  $v$  is the frequency of Q in the cage or mating chamber. This is constant since males were sampled with replacement of the same phenotypes. At maximum likelihood.

$$\hat{\alpha}_1 = \frac{n_Q(1 - v) - n_T v}{n(1 - v)}$$

where  $n = n_Q + n_T$ , the sum of Q and T males taking part in the matings.

In more realistic models of mating behaviour, the females express their preferences with probabilities that depend on their chances of encountering the males they prefer (O'Donald, 1980*a*). The males' frequencies determine the probabilities that females encounter particular male phenotypes and hence express a mating preference for them. In general, a female expresses a preference with a probability that is a function of male frequency (O'Donald, 1980*b*). If the male frequencies are kept constant, the same proportions of females express their preferences. The simple ML estimate is thus an estimate of the same relative proportion of females with a preference, regardless of whether preference is expressed as some function of male frequency. This was a crucial consideration in the design of the experiments to select for female choice: by always using the same ratio of male phenotypes, the same relative proportion of females express their preference and hence the ML estimate measures the relative level to which the preference has been selected in each generation.

In the field data, frequencies are not strictly constant, for mating males are sampled to some extent without replacement. Fortunately this can have only a very small effect on the frequencies of the males from which the females make their choices. At any one moment only a small proportion of males are mating. Our sample, too, is very large. The effect of non-replacement will be very slight. It will reduce the estimate of preference; but even if no males had been replaced, this would produce only a five per cent increase in the estimated preference, *i.e.*, from 20 per cent to 21 per cent.

The sexual selection acts only on the males. A comparison of the phenotypic frequencies of mating and non-mating males gives the contingency  $\chi^2$  for 4 degrees of freedom  $\chi_4^2 = 114.7324$  while for females  $\chi_4^2 = 2.5774$ ,  $P = 0.631$ . The frequencies are similar in mating and non-mating

Component of variation	Value of $\chi^2$	Degrees of freedom	Value of $P$
Within melanic phenotypes (L, Q, X)	4.6413	2	0.0982
Within non-melanic phenotypes (T, A)	0.3619	1	0.5475
Between melanic and non-melanic phenotypes	109.7292	1	—
Total variation	114.7324	4	—

females, but very significantly different in males. The  $\chi^2$  for males can be analysed into the independent components shown in the table.

The melanic males mate more often than the other phenotypes. The three melanic phenotypes, L, Q and X, gain roughly the same mating advantage; the non-melanic phenotypes suffer a very similar mating disadvantage.

We can estimate the preference for melanics from the following table

	Number of males		
	Mating	Non-mating	Total
Melanics	652	3410	4062
Non-melanics	676	6505	7181

This gives the estimate

$$\hat{\alpha}_1 = \frac{652 \times 7181 - 676 \times 4062}{1328 \times 7181} = 0.203$$

which does not differ significantly from the estimate

$$\hat{\alpha}_1 = 0.230$$

obtained from the mating chamber experiments (Majerus, O'Donald and Weir, 1982*b*). However the value for the field data,  $\hat{\alpha}_1 = 0.203$ , will be a slight under-estimate, since some males must have been sampled without replacement.

The assorting preferences can be estimated from the data of numbers of matings. If the melanic phenotypes L, Q and X are included in the single class M, the data form the following  $3 \times 3$  table of frequencies:

Females	Males			Totals
	M	T	A	
M	229	181	64	474
T	312	228	65	605
A	111	59	79	249
Totals	652	468	208	1328

For this table

$$\chi^2_4 = 65.2450$$

showing that the assortative mating is highly significant. Table 3 shows an analysis of this  $\chi^2$  into four independent additive values by Kimball's formulae (Kimball, 1954). The assortative mating of *annulata* and *typica* is highly significant. There is also some evidence that melanic females prefer *typica* rather than *annulata* males, as shown in 3(ii): *typica* males are chosen relatively more often by melanic females than by non-melanic females.

TABLE 3

Analysis of  $\chi^2$  of data of matings of ladybirds in the sample from Keele 1982

(i) Difference in preference of T and A females for T and A males

		Females		
Males		T	A	
T		228	59	$\chi_1^2 = 57.6917$
A		65	79	

(ii) Difference in preference of M and Nm females for T and A males

		Females		
Males		M	Nm	
T		181	287	$\chi_1^2 = 3.9213$ $P = 0.0477$
A		64	144	

(iii) Difference in preference of T and A females for M and Nm males

		Females		
Males		T	A	
M		312	111	$\chi_1^2 = 3.4506$ $P = 0.0632$
Nm		293	138	

(iv) Difference in preference of M and Nm females for M and Nm males

		Females		
Males		M	Nm	
M		229	423	$\chi_1^2 = 0.1814$ $P = 0.670$
Nm		245	431	

In this table, T and A represent *typica* and *annulata*, as elsewhere, M represents melanic and Nm non-melanic. Each of the four  $2 \times 2$  tables are independent. Values of  $\chi^2$  have been calculated by Kimball's formulae which give strictly additive values that sum to the total value of  $\chi_4^2$  for the  $3 \times 3$  table.

The assortative mating of *typica* and *annulata* could be produced either by an assorting preference of *typica* for *typica* or by an assorting preference of *annulata* for *annulata* or by some combination of both preferences. If *annulata* females prefer *annulata* males, their preference is estimated to be

$$\hat{a}_3 = 0.441 \quad (a_2 = 0).$$

If *typica* females prefer *typica* males, their preference is

$$\hat{a}_2 = 0.6184 \quad (a_3 = 0).$$

These values of  $a_3$  and  $a_2$  are sufficient to fit the data of table 3(i). But *typica* does not mate assortatively in the absence of *annulata*. For example,

in the population cage experiments using only Q and T phenotypes, T does not assort with T in the matings. Therefore, the assortative mating of A with A and T with T is most probably determined by an assorting preference solely of A females for A males, giving estimates

$$\hat{a}_2 = 0$$

$$\hat{a}_3 = 0.44.$$

Our field and laboratory data have thus shown:

(i) that 20–23 per cent of females express a non-assorting preference for melanic males; and

(ii) that 44 per cent of *annulata* females express an assorting preference for *annulata* males. The population cage data also showed that an estimated 16 per cent of melanic females may express an assorting preference for melanic males, although this result was not quite statistically significant.

Given these estimates, we can now ask whether they are sufficient to explain the polymorphism of the ladybirds at Keele; if not, what other factors must be postulated and what their magnitudes must be in order to explain the apparently stable polymorphic frequencies in the population.

#### 4. RESULTS OF COMPUTATIONS

We have computed both equilibrium frequencies and rates of approach to equilibrium. For some parameter sets, two different equilibria can exist. If so, one of the equilibria always has the *annulata* frequency equal to zero. The equilibria we have found are all globally stable. But in spite of large parameter values, such as those we estimate for the ladybird's mating preferences, the general rate of approach to equilibrium is usually very slow, even when a long way off. The point of equilibrium would easily be disturbed by genetic drift in small populations or occasionally bottlenecks in population size. The frequencies would then remain displaced from equilibrium for a long time. Fortunately the Keele population is very large and dense—we sampled over 22,000 ladybirds in 1982—and genetic drift should not be a significant factor in determining the phenotypic frequencies.

##### (i) *Hypothetical example giving two equilibria*

Consider an example of case (i), section 2, in which  $\alpha_3 = 0$ . Let the parameter values be

$$\alpha_1 = \alpha_2 = 0.1$$

$$a_1 = 0.3$$

$$a_2 = 0.2$$

$$a_3 = 0.5$$

From equations (9) and (10), we find

		Equilibrium phenotypic frequencies		
		melanic	typica	annulata
		M	T	A
either		0.5	0.3333	0.1667
or		0.6180	0.3820	0.0

The first set of equilibrium frequencies (with all three phenotypes present) is reached from a very wide range of initial frequencies. The equilibrium thus appears to be globally stable. The second set of equilibrium frequencies (with *annulata* absent,  $\pi_3^* = 0$ ) is trivial in the sense that it is apparently reached only when *annulata* is absent initially. Fig. 2 shows the approach to equilibrium from different sets of initial frequencies.

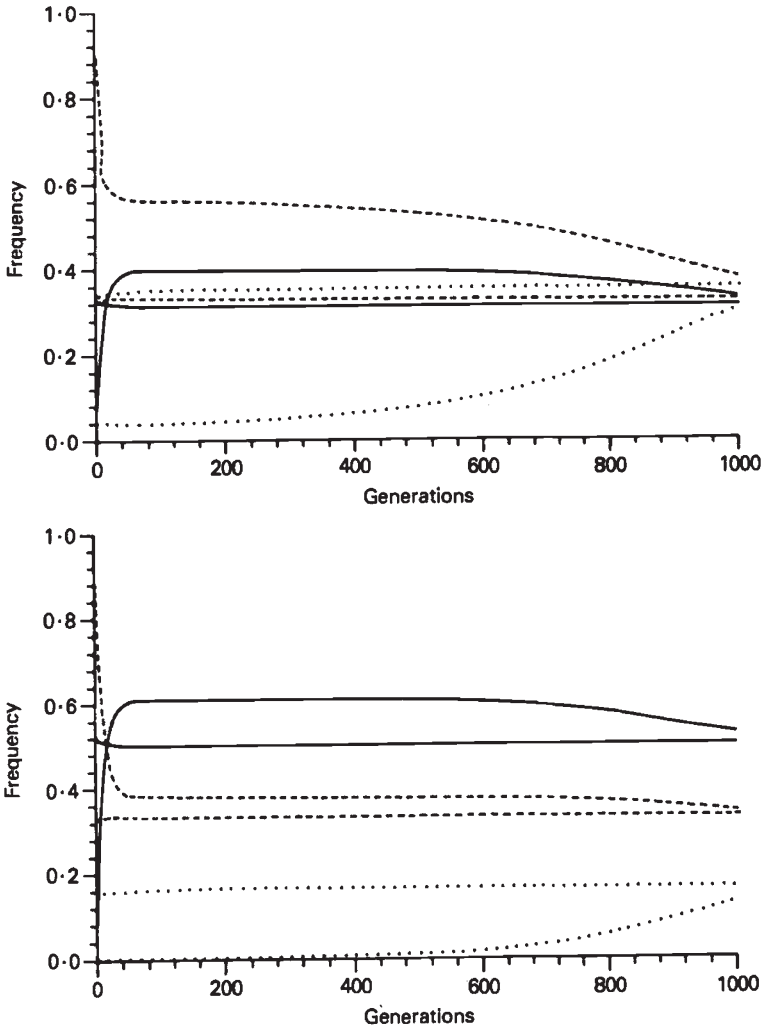


FIG. 2. Approach to equilibrium at different initial frequencies when  $\alpha_1 = \alpha_2 = 0.1$ ,  $a_1 = 0.3$ ,  $a_2 = 0.2$ ,  $a_3 = 0.5$ .

Graph 2(i). Gene frequency against generation.

Graph 2(ii). Phenotype frequency against generation.

<i>Key</i>	
Solid line	Melanic frequency
Dot	<i>Typica</i> frequency
Dot (widely spaced)	<i>Annulata</i> frequency

(ii) *Computations using estimated parameters*

In section 3, we have estimated at a high level of significance that

$$\hat{\alpha}_1 = 0.23 \text{ (from population cage or mating chamber experiments)}$$

or

$$\hat{\alpha}_1 = 0.203 \text{ (from field data of samples from Keele)}$$

and

$$\hat{\alpha}_3 = 0.44 \text{ (from field data of samples from Keele).}$$

In addition, the estimate

$$\hat{a}_1 = 0.162$$

was nearly significant at the 5 per cent level. As parameter values, these estimates are not sufficient to maintain a stable polymorphism. They represent a particular set of values of parameters for *case* (ii), section 2, in which  $\alpha_2 = \alpha_3 = 0$ . In this case, only the melanics are subject to frequency-dependent sexual selection. Only two phenotypes can then be maintained by the balance between assortative mating and sexual selection. To maintain all three phenotypes, at least two must be subject to frequency-dependent sexual or natural selection. If the other sexually selected phenotype is *typica*, we should have the parameter set of *case* (i), in which  $\alpha_1, \alpha_2 \neq 0, \alpha_3 = 0$ . The field data show that this is the only possible assumption: relative to *annulata*, *typica* has a slight mating advantage; although not significant, this suggests that *typica* may also be the object of female mating preference as well as the melanic phenotypes. Unfortunately, no such preference was observed in the population cage experiments.

For *case* (i), the melanic frequency is always given by

$$\pi_1^* = \alpha_1 / (a_3 - a_1)$$

where  $\alpha_1 = 0.23$  or  $0.203$ ,  $a_1 = 0$  or  $0.162$ ,  $a_3 = 0.44$ . Values of  $\alpha_2$  determine the relative frequencies of *typica* and *annulata*. The frequency of *annulata* is reduced in a simple inverse linear relationship to the value of  $\alpha_2$ : as we should expect, the greater the non-assorting preference for *typica*, the lower the frequency of *annulata*. Fig. 3 shows this relationship for four sets of values of  $\alpha_1$  and  $a_1$ . In all sets,  $\alpha_3 = 0.44$ . Taking

$$\alpha_1 = 0.23 \text{ or } 0.203$$

$$a_1 = 0.162$$

$$a_2 = 0.0$$

$$a_3 = 0.44$$

we see that to maintain *typica* at a frequency of 0.452 in the population (*i.e.*, the proportion of individual males and females observed in the sample from Keele), we should require a non-assorting preference for *typica* of

$$\begin{aligned} \alpha_2 &= \pi_2^* / (a_3 - a_2) \\ &= 0.199. \end{aligned}$$

So high a preference ought to have been detected in the population cage experiments, but none was found. Although it is unlikely, therefore, that

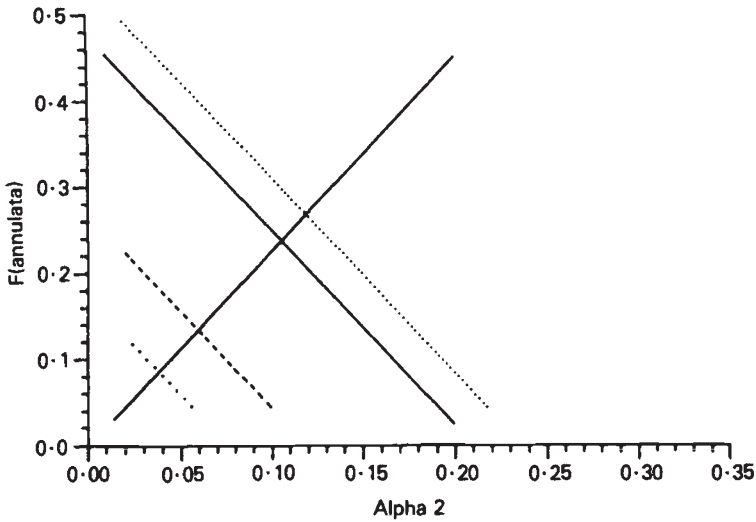


FIG. 3. Relationship between frequency of *annulata* and *typica* and value of  $\alpha_2$  (non-assorting preference for *typica*).

<i>Key</i>		
Solid line	$\alpha_1 = 0.23,$	$a_1 = 0.0$
Dot	$\alpha_1 = 0.20302,$	$a_1 = 0.0$
Dot (widely spaced)	$\alpha_1 = 0.23$	$a_1 = 0.162$
Dash	$\alpha_1 = 0.20302,$	$a_1 = 0.162$

(The extra solid line shows the variation of  $f(\text{typica})$  against  $\alpha_2$ ). In all cases  $a_3 = 0.44$ .

this preference actually maintains the *typica* phenotype at its present frequency, we shall assume for the moment that it does so: some factor, acting to produce the same frequency-dependent advantage as  $\alpha_2$ , must be responsible for maintaining the polymorphism with *typica*. But, given the value  $\alpha_2 = 0.199$ , we immediately see that the predicted frequency of melanics is much too high:

$$\begin{aligned} \pi_1^* &= \alpha_1 / (a_3 - a_1) \\ &= 0.827 (\alpha_1 = 0.23) \text{ or } 0.730 (\alpha_1 = 0.203). \end{aligned}$$

No polymorphic equilibrium would thus exist. Since  $a_1$  is not in fact statistically significant ( $P = 0.061$ ), we might assume that  $a_1 = 0$ , giving

$$\pi_1^* = 0.523 \text{ or } 0.461.$$

These values are still much too high since at Keele  $\hat{\pi}_1 = 0.3416$ .

In spite of the female preference in their favour, the melanics are seldom observed at frequencies much greater than 50 per cent anywhere. Their frequencies have declined in some areas when smoke emissions into the environment have been reduced (Creed, 1971). Natural selection appears to act against melanics in smokeless environments. Natural selection, which is thus dependent upon a physical factor in the environment, would not of course be frequency-dependent. It would reduce the population frequency

of melanics maintained at stable equilibrium: the frequency-dependent sexual selection maintains the polymorphism; the constant natural selection shifts the equilibrium point. For this reason, constant natural selection would not maintain *typica* in the population: *typica* must gain a frequency-dependent advantage, either by sexual selection (for example, when  $\alpha_2 \neq 0$  as we have assumed), or by some special form of frequency-dependent natural selection. In general, when selection acts on three phenotypes with ordered dominance, at least two of the phenotypes must gain some frequency-dependent advantage for a polymorphism of all three phenotypes to be stable.

(iii) *Natural selection of melanic phenotypes*

Since the estimated parameters of sexual selection and assortative mating predict a melanic frequency that is much too high, we must postulate additional selective factors to explain the population frequency. As we have seen in the previous section, good evidence supports the hypothesis that melanics suffer a selective disadvantage that is greater in environments less polluted by smoke (Creed, 1971). So we may ask: what constant selective disadvantage of melanics would be sufficient to maintain the observed population frequencies given the parameters of sexual selection and assortative mating that we have already estimated with considerable accuracy and consistency?

We use the following parameters to estimate the selection coefficient of melanics:

$$\begin{array}{ll} \alpha_1 = 0.23 \text{ or } 0.203 & a_1 = 0.162 \\ \alpha_2 = 0.199 & a_2 = 0.0 \\ \alpha_3 = 0.0 & a_3 = 0.44. \end{array}$$

Fig. 4 shows the relation between selection coefficient and melanic frequency given these estimates of sexual selection and assortative mating. Interpolating in fig. 4, we estimate the natural selection required to produce the observed melanic frequency:

- (i) when  $\alpha_1 = 0.23$ ,  $s = 0.179$ ;  
(ii) when  $\alpha_1 = 0.203$ ,  $s = 0.143$ .

The rates of approach to the population equilibrium appear to depend on the initial frequency of *annulata*. When *annulata* starts at a low frequency, it only increases in frequency very slowly. Only when it has eventually become more abundant, is a more rapid approach made to the final equilibrium. This result is shown in fig. 5(i) and (ii). Two different sets of initial gene frequencies were used in the computations:

- (i)  $p_1 = 0.33$ ,  $p_2 = 0.33$ ,  $p_3 = 0.34$   
(ii)  $p_1 = 0.05$ ,  $p_2 = 0.9$ ,  $p_3 = 0.05$ .

After starting at frequencies of set (i), virtual equilibrium has been reached at 300 generations with  $\pi_1 = 0.3415$ ,  $\pi_2 = 0.4516$ ,  $\pi_3 = 0.2069$  as observed in the Keele population. But starting with set (ii), we find after 300 generations that

$$\pi_1 = 0.3767, \quad \pi_2 = 0.5113, \quad \pi_3 = 0.1211.$$



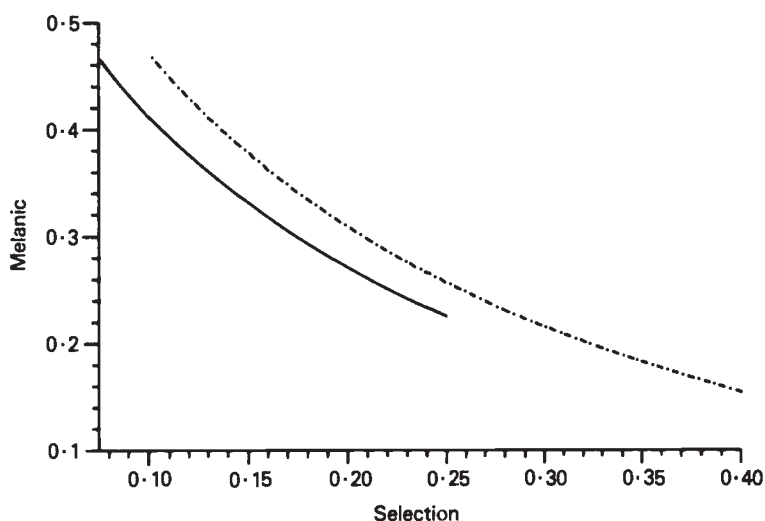


FIG. 4. Natural selection and melanic frequency.

Line of dot-dash  $\alpha_1 = 0.23$   
 Solid line  $\alpha_1 = 0.203$

The frequencies still have some way to go to the Keele equilibrium point. *Annulata*, having at last become more common, is now increasing more rapidly in frequency and starting to move towards equilibrium. Ultimately, the population stabilises at the frequencies observed at Keele.

## 5. CONCLUSIONS

In section 2, we formulated and analysed a model to describe the evolution of the melanic and non-melanic phenotypes of the two-spot ladybird. The model is quite general, however, and allows for sexual selection and assortative mating of three phenotypes determined by three alleles with ordered dominance. Females may express either non-assorting or assorting preferences for each of the three male phenotypes, thus giving rise to the sexual selection and assortative mating.

Given the experimentally determined parameters

$$\alpha_1 = 0.23 \quad (\text{non-assorting preference for melanics})$$

$$a_1 = 0.162 \quad (\text{assorting preference for melanics})$$

$$a_2 = 0.0 \quad (\text{assorting preference for } typica)$$

$$a_3 = 0.44 \quad (\text{assorting preference for } annulata)$$

we showed in sections 4(ii) and 4(iii) that to maintain the polymorphism at the observed frequencies at Keele, we must have parameter values

$$\alpha_2 = 0.199 \quad (\text{non-assorting preference for } typica)$$

$$s = 0.175 \quad (\text{melanic disadvantage in natural selection}).$$

This is consistent with the observation that melanics appear to suffer a disadvantage that increases as smoke pollution is reduced.

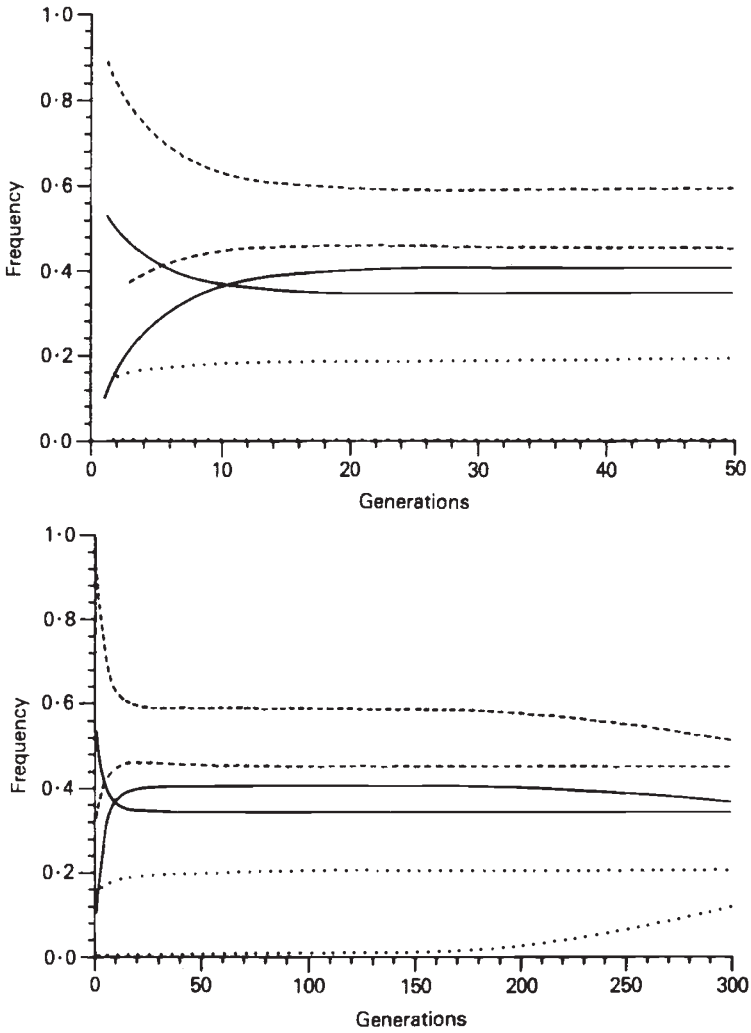


FIG. 5. Approach to the equilibrium of the Keele population.

- (i) Frequencies of the three phenotypes as they approach equilibrium over 50 generations.
- (ii) Frequencies of the three phenotypes over 300 generations.

The graphs have two sets of data plotted on them; the first set starts with the following gene frequencies:

$$p_1 = 0.33 \quad p_2 = 0.33 \quad p_3 = 0.34$$

and the second starts with:

$$p_1 = 0.05 \quad p_2 = 0.9 \quad p_3 = 0.05$$

*Key*  
 (for Graph 5(i) and Graph 5(ii))  
 Solid line                      Melanic frequency  
 Dot                                      Typica frequency  
 Dot (widely spaced)              Annulata frequency

The non-assorting preference for melanics has also been estimated from the relative frequencies of mating and non-mating males at Keele, giving

$$\hat{\alpha}_1 = 0.203$$

which is consistent with the experimentally determined value. Then we find

$$s = 0.143$$

to maintain the frequencies observed at Keele.

Unfortunately, as we have already stressed in section 4.(ii), no preference for *typica* was detected in our population cage experiments. Indeed, estimating the parameters by maximum likelihood, Majerus, O'Donald and Weir (1982a) found that

$$\hat{\alpha}_2 = 0.0.$$

Since we have had to postulate that

$$\alpha_2 = 0.199$$

to maintain *typica* at its observed frequency, the independently estimated value

$$\hat{\alpha}_2 = 0.0$$

refutes the simple model of sexual selection and assortative mating. Sexual selection certainly provides the necessary frequency-dependent advantage of the melanics, but not apparently that of *typica*. We must assume that some form of natural selection, acting with the same frequency-dependence as  $\alpha_2$ , determines the frequency of *typica*.

Experiments now in progress have been designed to give further independent estimates of  $\alpha_2$  and the assorting parameters  $a_2$  and  $a_3$  and thus independently test the assumptions of the model.

*Acknowledgments.* This work was done with the support of SERC research grant GR/B3080 to P. O'Donald, NATO research grant 102.81 to P. O'Donald and S. Karlin, a grant from the Wain Fund (ARC) to P. O'Donald and a Wain Fellowship to M. E. N. Majerus.

## REFERENCES

- CREED, E. R. 1971. Industrial melanism in the two-spot ladybird and smoke abatement. *Evolution*, 25, 290-293.
- HODEK, I. 1973. *Biology of Coccinellidae*. Academia, Prague.
- KARLIN, S. AND O'DONALD, P. 1978. Some population genetic models combining sexual selection with assortative mating. *Heredity*, 41, 165-174.
- KARLIN, S. AND O'DONALD, P. 1981. Sexual selection at a multiallelic locus with complete or partial dominance. *Heredity*, 47, 209-220.
- KIMBALL, A. W. 1954. Short-cut formulas for the exact partition of  $\chi^2$  in contingency tables. *Biometrics*, 10, 452-458.
- MAJERUS, M. E. N., O'DONALD, P. AND WEIR, J. 1982a. Evidence for preferential mating in *Adalia bipunctata*. *Heredity*, 49, 37-49.
- MAJERUS, M. E. N., O'DONALD, P. AND WEIR, J. 1982b. Female mating preference is genetic. *Nature*, 300, 521-523.
- MUGGLETON, J. 1979. Non-random mating in wild populations of polymorphic *Adalia bipunctata*. *Heredity*, 42, 57-65.
- O'DONALD, P. 1980a. *Genetic Models of Sexual Selection*. Cambridge University Press.
- O'DONALD, P. 1980b. A general analysis of genetic models with frequency-dependent mating. *Heredity*, 44, 309-320.
- O'DONALD, P. AND MUGGLETON, J. 1979. Melanic polymorphism in ladybirds maintained by sexual selection. *Heredity*, 43, 143-148.