

## THE GENETIC BASIS OF HYBRID INVIABILITY IN THE GRASSHOPPER *PODISMA PEDESTRIS*

N. H. BARTON and G. M. HEWITT

*Dept. of Genetics, Downing St., Cambridge, CB2 3EH, U.K.; School of Biological Sciences, University of East Anglia, Norwich, NR4 7TJ, U.K.*

Received 3.iv.81

### SUMMARY

Samples were taken from five sites in a transect across the hybrid zone between two chromosomal races of the grasshopper *Podisma pedestris*. Crosses were set up between insects from the same population, and between populations spanning the zone, and the early viability of the offspring was measured. Hybrids between pure populations had reduced viability, and the viability of insects from the middle of the zone was still lower, showing that most (~87 per cent) of the inviability was due to the breakup of coadapted gene complexes. Although the total selection acting was strong (log. fitness reduced by  $S \sim 2.5$ ), it was spread over a region wider than the dispersal range (350 m vs. 20 m). Hence, the selection on each locus contributing towards the inviability is weak (~3 per cent). Many (~150) independent chromosome segments act cumulatively to produce inviability at this stage in the life history. The implications of these findings for models of divergence are discussed.

### 1. INTRODUCTION

IN order to speculate about the mechanisms by which the differences between species might have arisen, one first needs to know how different the species are. However, it has proved difficult to analyse the genetic basis of speciation; unless the barriers to gene exchange between isolated groups depend largely on preferential mating, the crosses required for a conventional genetic analysis will generally be unsuccessful. Where fertile crosses can be made, a wide array of recombinants are usually found, suggesting that many loci are involved (*e.g.*, Mather and Vines, 1951, with *Antirrhinum*; Grant, 1964, with *Gilia*). Studies with *Drosophila* have shown that hybrid sterility between sibling species is often caused by differences at several loci on each chromosome (Dobzhansky, 1936, with *D. pseudo-obscura*  $\times$  *persimilis*; Pontecorvo, 1943, with *D. simulans*  $\times$  *melanogaster*; see Dobzhansky, 1970, for a review).

Even where crosses are possible, studies such as these are limited by the difficulties of mapping polygenes, and can only give a lower limit to the number of genes involved. Electrophoretic methods promised a more general method for assessing the amounts of genetic differentiation during speciation; extensive studies have shown that genetic distances (Nei, 1972) between related species range from values similar to those between conspecific populations (~0.05) upwards (Avisé, 1976; Ayala, 1975). However, there is no clear relationship between electrophoretic divergence and taxonomic rank, morphological change, or speciation rate (*e.g.*, Turner, 1974; King and Wilson, 1975; Avisé and Ayala, 1976); thus, there is no reason to suppose that changes in abundant metabolic enzymes are typical of the rest of the genome. In any case, a measurement of the proportion of loci which have altered their allele frequencies is not necessarily relevant

to the central problem of how the isolation between divergent populations is maintained, and how it evolved. There is a crucial difference between the total divergence, and the amount of that divergence which actually produces the barriers between populations, and hence promotes speciation (see Futuyma, 1979, pp. 417–481).

This paper describes the results of a set of crosses designed to characterise the selectively important differences between two races of the alpine grasshopper *Podisma pedestris*, and so to address directly the problems outlined above. The races of *Podisma* were originally identified through the existence of a Robertsonian fusion between the sex chromosome and an autosome, which has become established in south-eastern France. (The southern race has therefore developed a neo-XY sex determining system, and will be referred to as the XY race; the northern, ancestral race still has the standard acridid XO system.) The chromosome types meet in a cline which runs east-west for about 130 km along a high mountain ridge (Hewitt, 1975). Since this cline is about 800 metres wide, much wider than the dispersal distance measured by mark-release-recapture experiments, selection on the chromosome difference must be weak ( $\sim 0.5$  per cent) (Barton and Hewitt, 1981a). However, there must be other differences between the two types, since F1 hybrids between populations on either side of the cline are less viable, yet chromosomal hybrids from mixed populations in the centre of the cline show no reduction in fitness; the bulk of the reduction in F1 viability is not caused by the chromosomal heterozygosity (Barton, 1980).

To investigate the hybrid inviability further, crosses were made between grasshoppers from five sites in a transect across the cline. The aim of the experiment was to determine, first, the total selection acting on the stages studied, second, the proportion of this selection which acts on F1s through heterozygote disadvantage, and the proportion which only acts on F2s, through the breakup of coadapted sets of alleles, and third, to find the width of the region of hybrid inviability. This last parameter is especially interesting, since by comparing the width to the dispersal distance, the average selection acting on each independent difference can be estimated. The ratio between the total selection and the selection on each locus then gives an estimate of the number of selected loci involved in isolating the two races.

## 2. METHODS

Samples of about 100 immature grasshoppers were collected from the five sites shown in fig. 1 in July, 1979. The positions of the sites relative to the centre of the chromosomal cline, and the corresponding frequency of the XY karyotype, are shown in table 1. The insects were brought back to the laboratory, and set up in single pair crosses as they became adult, using the rearing methods described in Barton (1980). Crosses were made between grasshoppers from within the same population ( $1 \times 1$ ,  $2 \times 2$ , etc.), and also between populations symmetrically spanning the cline ( $1 \times 5$ ,  $2 \times 4$ , etc.). The eggs from these crosses were kept at room temperature for 3 months, until they went into diapause, and then kept at  $0-4^\circ\text{C}$  for  $4\frac{1}{2}$  months to simulate winter conditions, and break the diapause. (The precise temperature during these months did not significantly affect the viability of the embryos.)

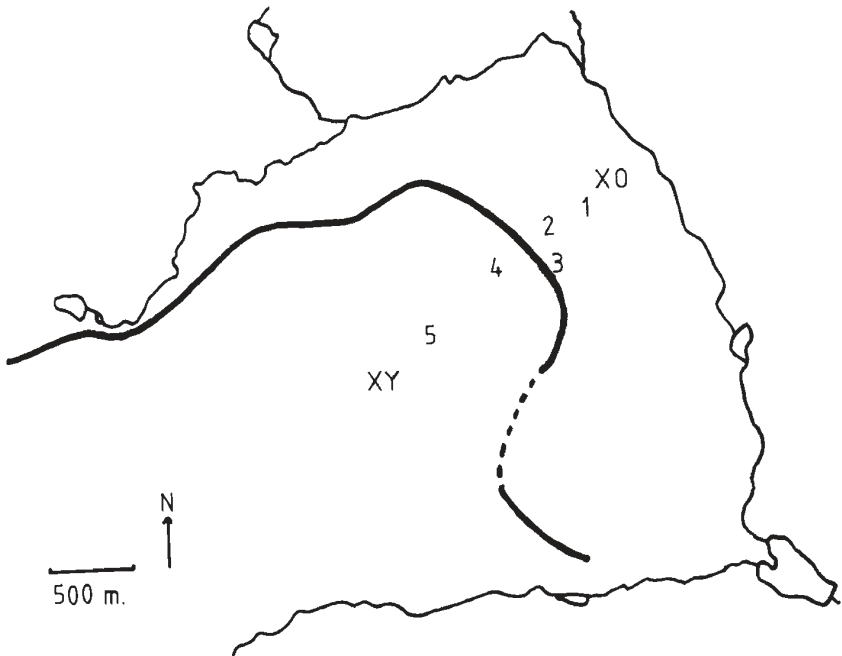


FIG. 1.—The position of the collecting sites relative to the chromosomal cline; the thick line shows the 50 per cent karyotype frequency contour. In dotted sections, no mixed populations have been found.

The following spring, the eggs from each cross were put into separate cages arranged in a Latin Square design, and hatched at 25°C; the egg pods were later dissected, so that the stage of development of unhatched embryos could be determined. Nymphs were counted after the first two weeks, and weekly thereafter.

### 3. RESULTS

There was no significant difference between the various classes of cross in the number of eggs laid, the number of eggs per pod, the date of laying or the time the eggs took to hatch after being brought out of the cold. The fraction of eggs which hatched, and survival though the first fortnight did not differ between reciprocal and symmetric classes (e.g.,  $1 \times 5$ ,  $5 \times 1$ ;  $2 \times 2$ ,

TABLE 1

*The distance of each collecting site from the centre of the chromosomal cline, and the expected karyotype frequency at that point (from data in Barton and Hewitt, 1981a)*

Site	Position/metres	% XY
1	-600	5%
2	-140	33%
3	50	56%
4	200	73%
5	820	98%

TABLE 2a

The fraction of eggs which hatched, and the viability in the first two weeks; for comparison, results from previous crosses (Barton, 1980) are included

Cross	No. of crosses	Previous hatch rates (E → H)	Hatch rates (E → H)	Early viability (H → 2nd week)	Overall viability (E → 2nd week)
1 × 1, 5 × 5	25	0.53	0.77	0.75	0.57
2 × 2, 4 × 4	31	~	0.62	0.65	0.40
3 × 3	27	0.23	0.58	0.42	0.24
2 × 4, 4 × 2	17	~	0.66	0.46	0.30
1 × 5, 5 × 1	17	0.34	0.67	0.62	0.46

TABLE 2b

The fraction of eggs developing up to the following stages:—UD—little or no embryonic material, BD—stopped developing at or before diapause, PD—developed past diapause but failed to hatch, R—died and rotted just after hatching, H—hatched successfully. Means in both a and b are weighted by number of crosses, rather than number of eggs

Cross	H	R	PD	BD	UD	No. of eggs
1 × 1, 5 × 5	0.77	0.05	0.01	0.06	0.11	635
2 × 2, 4 × 4	0.62	0.05	0.01	0.10	0.22	643
3 × 3	0.58	0.06	0.03	0.14	0.19	857
2 × 4, 4 × 2	0.66	0.07	0.02	0.07	0.18	404
1 × 5, 5 × 1	0.67	0.13	0.02	0.11	0.17	524

4 × 4), but table 2a shows that there were substantial differences between the five main types of cross. (In calculating the mean viability, all crosses were given equal weight, regardless of the number of eggs they produced, since the error in determining survival rates for each cross was smaller than the variation between crosses.) Crosses between sites on either side of the cline (1 × 5, 5 × 1) showed more inviability at all stages than pure crosses (1 × 1, 5 × 5), and crosses from the centre showed still greater inviability. Crosses involving the intermediate sites two and four were, as expected, intermediate in viability. The overall effects, from hatching to the end of the second week, were significant at the 5 per cent level using the Kolmogorov-Smirnov test, and the viability of grasshoppers derived from the centre of the cline was significantly reduced at both the first two stages. However, more detailed comparisons were not significant when restricted to a single stage; the overall effects were much more reliable. Only three points in the life history have been given, both because there was heavy mortality later on, so that the numbers became too small to analyse properly, and also because no differential mortality between hybrids and non-hybrids was apparent between the second week and the emergence of adults in the fifth week. The following analysis will therefore only deal with the viabilities of the embryo, and in the first two weeks of nymphal life.

These experiments confirm the results of Barton (1980), and show that the incompatibilities between the races coincide with the chromosomal difference to within about 100 metres. However, detailed comparisons are difficult, because the overall hatch rate in these crosses was higher than before (0.77 instead of 0.53 in pure crosses). This presumably reflects

some difference in laboratory conditions between the two sets of experiments; nevertheless, the relative inviabilities are not significantly different.

Dissection of unhatched eggs showed the stage of development at which the embryo died (table 2b). Selection against crosses spanning the cline acts at all stages, whilst the reduced viability in the centre of the cline operates at early stages. However, there were too few crosses to put any confidence in these observations; the nature of mortality in previous crosses was not significantly different, but the above patterns were not discernible.

#### 4. INTERPRETATIONS

These differences in viability can be explained by a simple model in which there are a number of genes which are fixed for different alleles in the two races. At some of these loci, there is heterozygote disadvantage, giving lower F1 viability, whilst others interact epistatically, so that recombination between coadapted pairs reduces viability still further in the centre of the hybrid zone, where the grasshoppers have been interbreeding for many generations. Both these schemes can maintain clines, whose width will depend on the selection on the locus concerned (Bazykin 1972, 1973). For the moment, we will assume that the clines coincide, that selection acts with equal strength on all loci, so that all the clines have the same width, that the alleles *only* affect early viability, having no other effects on fitness, and that linkage can be neglected. The essential parameters at the level of single loci are the selection on each locus ( $s$ ), and the width of the clines ( $w$ ); if  $w$  is defined as the inverse of the maximum slope in allele frequency, these are related by  $w = 2\sigma\sqrt{2}/s$ , where  $\sigma$  is the root-mean-square distance between parent and offspring (Bazykin, *op. cit.*). Since sites 1 and 5 are outside the chromosomal cline, and since we will show that the region of inviability is narrower than this cline, these populations will be taken to be homozygous for the relevant alleles. The only information about width comes from sites 2 and 4, and the state of these populations can be described by the proportion of heterozygotes ( $h$ ). (We neglect the slight asymmetry of the collecting sites.) With this simple model (defined in table 3), the allele frequency,  $p$ , varies with position,  $x$ , as:

$$p = \frac{1}{2}(1 + \tanh(2x/w)).$$

TABLE 3

<i>Definition of the basic model</i>	
<i>Heterozygote disadvantage (<math>\alpha n</math> loci)</i>	
Genotype	Fitness
+ / +	1
+ / -	$1 - s$
- / -	1
<i>Epistasis (<math>(1 - \alpha)n</math> loci)</i>	
+ + / + +	1
+ - / + + etc.	$1 - s/2$
+ - / - + etc.	$1 - s$
- - / - + etc.	$1 - s/2$
- - / - -	1

Hence, the heterozygosity,  $h = 2p(1-p)$ , is given by:

$$h = \frac{1}{2}(1 - \tanh^2(2x/w)).$$

This describes the effect of selection on single loci; to interpret the observations, we must extrapolate to find the viability of whole insects. Two new parameters are needed; the number of loci ( $n$ ), and the proportion of loci ( $\alpha$ ) giving heterozygous disadvantage, rather than epistasis. If all the loci, or pairs of interacting loci, act independently, the overall inviability will be obtained by multiplying the individual effects together—the log. fitnesses will add. The total selection,  $S$ , can be defined as the reduction in log. fitness of an individual heterozygous at all loci, and with all coadapted pairs the wrong way round. Since there are  $\alpha n$  loci giving heterozygous disadvantage, and  $(1-\alpha)n/2$  coadapted pairs,  $S$  is given by:

$$S = s \left( n\alpha + (1-\alpha) \frac{n}{2} \right) = \frac{ns}{2} (1 + \alpha).$$

The log. viabilities expected from this model are shown in table 4; they depend on  $\alpha$ ,  $h$  and  $S$ .

TABLE 4

*Mean log. viabilities expected with the model defined in table 3. See text*

Cross	$-\ln(W/W_0)$
1 × 1, 5 × 5	0
2 × 2, 4 × 4	$hS$
3 × 3	$\frac{1}{2}S$
2 × 4, 4 × 2	$(\alpha(1-h) + h(1-\alpha))S$
1 × 5, 5 × 1	$\alpha S$

This model has been fitted to the data by assuming that the arc-sin transformed viabilities are normally distributed, and then finding the maximum likelihood estimates of  $\alpha$ ,  $h$  and  $S$ . (Kolmogorov-Smirnov tests show no significant deviations from normality.) Since  $h$  is related to the width of the underlying clines, and thus to the selection on each locus, an estimate of the number of loci involved can be derived from these three parameters. Support limits have also been calculated for log. likelihoods 2 units below the maximum; in large samples, these correspond to 95 per cent confidence limits (Edwards, 1972). This analysis has only been applied to the viability from egg to the end of the second week since data on the separate components are more variable. The variance of the transformed viabilities is 0.059.

Table 5 shows that strong selection ( $S = 2.5$ ) is acting on these early stages of development. Since the viability of offspring from the centre is lower than the viability of F1s, even though the former are only half as heterozygous, most of the selection must act through epistasis ( $\alpha = 0.13$ ); in fact, the heterozygote disadvantage is only just significant. The region of inviability is narrower than the chromosomal cline (350 m. vs. 800 m.), but is still a good deal greater than the dispersal distance of  $\sim 21$  metres in a generation measured by mark-release-recapture experiments (Barton

TABLE 5

Maximum likelihood estimates of the parameters defined in tables 3 and 4, together with 2-unit support limits

		M.L.E.	Support limits
Total selection	$S$	2.5	1.2-3.5
Fraction due to heterozygote disadvantage		0.13	0.01-0.37
Heterozygosity at sites 2 and 4	$h$	0.22	0.09-0.36
Cline width/metres	$w$	350	230-570
Selection on each locus	$s$	0.029	0.011-0.068
Number of loci	$n$	150	65-560

and Hewitt, 1982). Selection on each locus must therefore be weak ( $s = 2.9$  per cent), and many loci must be involved ( $n = 150$ ).

In reaching these conclusions, we made many assumptions about the way selection acts to produce hybrid inviability. We will now look at a range of elaborations around the basic model, to see whether the estimates are robust.

## 5. ELABORATIONS

### (i) *Heterogeneous selection*

Selection pressures on different loci, or pairs of loci, are likely to vary widely; the observed drop in viability at the hybrid zone is the sum of a set of troughs, each with its own depth and width. The average selection on each locus was obtained by comparing the average width with the dispersal range. We must find how a set of troughs with different widths are combined to give this average.

The log. viability is the sum of contributions from loci  $i$ :

$$L(x) = \sum L_i(x).$$

We measure the reduction in log. viability at the centre,  $\sum L_i(0)$ , and also the width of the function  $L$ . Roughly speaking, this width can be measured by the curvature at the centre, and the average selection by the ratio between dispersal rate and width<sup>2</sup>; thus:

$$\frac{1}{W^2} = \frac{-1}{8L_0} \frac{\partial^2 L}{\partial x^2}, \quad s = \frac{8\sigma^2}{W^2} = \frac{-\sigma^2}{L_0} \frac{\partial^2 L}{\partial x^2}.$$

Consider first loci which give heterozygote disadvantage,  $s_i$ ; let the allele frequencies be  $p_i, q_i$ :

$$\begin{aligned} L_i &= \ln(1 - 2s_i p_i q_i) \sim -2s_i p_i q_i \quad \text{for } s_i \ll 1 \\ \therefore \frac{\partial^2 L_i}{\partial x^2} &= -2s_i \frac{\partial^2}{\partial x^2} (p_i q_i) = 2s_i \left( \frac{\partial^2 p_i}{\partial x^2} (p_i - q_i) + 2 \left( \frac{\partial p_i}{\partial x} \right)^2 \right) \\ &= 4s_i \left( \frac{\partial p_i}{\partial x} \right)^2 \quad \text{at } x = 0, \quad \text{since } \frac{\partial^2 p_i}{\partial x^2} = 0. \end{aligned}$$

Now, the maximum slope,  $(\partial p_i / \partial x)_0$ , is, by definition, the inverse of the width, which is  $w_i = 2\sigma\sqrt{2/s_i}$ . Hence:

$$\frac{\partial^2 L_i}{\partial x^2} = \frac{4s_i}{w_i^2} = \frac{s_i^2}{2\sigma^2}$$

$$\therefore s = -\frac{\sigma^2}{L_0} \frac{\partial^2 L}{\partial x^2} = \frac{(\sum s_i)}{(\sum s_i)}$$

This formula can also be shown to describe genes which interact epistatically; the only difference is that the sum must be taken over pairs of loci, not the loci themselves. The estimate of the number of loci is therefore:

$$n = \frac{2S}{s(1+\alpha)} = \frac{2(\sum s_i)^2}{(1+\alpha)(\sum s_i^2)}$$

This estimate is weighted towards loci which produce strong selection. For example, the chromosomal cline by which the two races were first detected is 800 m. wide, implying selection of only 0.5 per cent, compared with an average of 2.9 per cent. This difference therefore only contributes  $\sim 0.3$  of a locus to the estimate of  $n \sim 150$ . Of course this bias towards selected differences was intentional; the experiment was designed to count the number of differences which actually produce isolation.

#### (ii) Selection on different stages of the life cycle

In section 4, only the viability from the egg to the second week was analysed. However, fitness has many other components, and the method could be extended to any of these, to give estimates of the amount of selection ( $S$ ), and the number of loci ( $n$ ). For example, data on survival up to hatching give  $S = 0.6$ , and  $n = 84$ , whilst data on survival from hatching through the first fortnight give  $S = 1.6$  and  $n = 50$ . (These estimates are crude, since the data on individual components are very variable.) If one assumes that each locus only affects one stage, the interpretation is clear; 84 loci affect hatching success, and 50 affect early viability. However, the effect of a locus on fitness might be spread over a number of stages, in which case the proper interpretation is that there are  $\sim 134$  loci (or  $\sim 150$ , if one uses the more reliable overall estimates), each with an effect 84/134 on hatching, and 50/134 on early viability. In principle, one can test whether all loci affect the different stages in the same proportion; in the example above, more loci affect hatching, and yet the selection on hatching success is weaker ( $S = 0.6$  vs. 1.6). This suggests that different loci act on different stages.

However the genes act, study of later stages (*e.g.*, fertility, mating success) can only increase these limited estimates of  $S$  and  $n$ , which are based on early stages; one can only guess at how many loci have been missed, since this depends on how rapidly genes affecting different stages diverge.

#### (iii) Staggered clines

If the underlying clines did not coincide precisely, the width of the region of inviability would be an overestimate of the true width, and so



the selection on each locus would be underestimated. In the extreme case, the inviability might be spread over just a few strongly selected and sharp clines, centred a few dispersal ranges apart from each other. However, this possibility is unlikely on theoretical grounds; clines at different loci will attract each other because they are influenced in the same way by population structure, so that secondary contact might have brought all clines to the same place (Barton, 1979a); the clines may interact epistatically; and linkage disequilibrium will pull them together (Slatkin, 1975). In any case, the observation that the viability cline coincides with the chromosomal cline to within a few hundred metres suggests that the genetic components of the hybrid zone coincide exactly.

(iv) *Interactions between more than two loci*

So far, we have assumed that the reduced viability at the centre is due to the breakup of coadapted *pairs* of alleles. In fact, larger sets of alleles may be coadapted with each other, and this may affect our interpretation. The argument in this paper is essentially a comparison between the genetic load due to hybridisation and the selection on each locus; it is known that if fitness is determined by a large array of genes acting in concert (as in the truncation models of Milkman, 1967; Sved *et al.*, 1967; and others), large amounts of selection can act without producing excessive load. Similarly, selection on coadapted pairs produces the same load as equal selection on a single locus through heterozygote disadvantage, even though twice as many loci are involved. So, might the two races be differentiated by many more than  $\sim 150$  loci?

In this context, truncation models, where only those with the lowest "hybridity" survive, are unlikely—there is no opportunity for competition between developing embryos. We will consider a more plausible model, similar to that of Wright (1935), in which a large number ( $k$ ) of loci act additively to determine some quantitative trait,  $\alpha$ . This trait is under disruptive selection, so that fitness is greatest at extreme values ( $\alpha = \pm 1$ , say). The reduction in the fitness of intermediate phenotypes can be described by some function  $f$ , which we will assume is symmetrical, has maximum value  $f(0) = 1$ , and is zero at  $\alpha = \pm 1$  (see fig. 2):

$$-\ln(W/W_0) = s * f(\alpha).$$

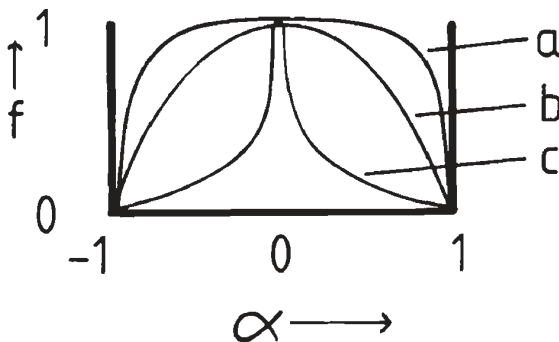


FIG. 2.—Possible relations between fitness and phenotype for the model described in 5(iv).

An individual with  $m$  "+" alleles and  $(2k - m)$  "-" alleles will have  $\alpha = (m - k)/k$ . If there are  $N$  of these sets of  $k$  loci, the reduction of log. viability of an individual at the centre, with  $\alpha = 0$ , will be  $Ns^*$ . Since we defined  $S$  as twice the loss of fitness of individuals in the centre, our measure of  $S$  will be  $2Ns^*$ .

The curvature of the viability trough can be found if we assume that linkage disequilibrium is negligible; the selection on each locus then depends on the change in fitness when one allele is changed for another:

$$s(p) = \frac{\Delta W}{\bar{W}} \approx \Delta(\ln(W/W_0)) \approx -\frac{s^*}{k} \frac{\partial f}{\partial \alpha}.$$

To find the slope of the cline at the centre, we take the equation for selection and gene flow in a continuous habitat (Nagylaki, 1974):

$$0 = \frac{\sigma^2}{2} \frac{\partial^2 p}{\partial x^2} - \frac{s^*}{k} pq \frac{\partial f}{\partial \alpha}.$$

At any given point in the cline, the average value of  $\alpha$  is  $(p - q)$ . Transforming to this new variable, multiplying by  $(\partial \alpha / \partial x)$ , and integrating by parts:

$$\begin{aligned} 0 &= \left( \frac{\sigma^2}{4} \frac{\partial^2 \alpha}{\partial x^2} - s^* \frac{(1 - \alpha^2)}{4k} \frac{\partial f}{\partial \alpha} \right) \frac{\partial \alpha}{\partial x} \\ \therefore 0 &= \frac{\sigma^2}{8} \left( \frac{\partial \alpha}{\partial x} \right)^2 + \frac{s^*}{4k} \int (1 - \alpha^2) \frac{\partial f}{\partial \alpha} d\alpha \\ \therefore 0 &= \frac{\sigma^2}{8} \left( \frac{\partial \alpha}{\partial x} \right)_0^2 - \frac{s^*}{4k} \left\{ 1 - 2 \int_0^1 \alpha f d\alpha \right\}. \end{aligned}$$

Since  $(\partial \alpha / \partial x) = 2(\partial p / \partial x) = 2/w$ , the above formula gives the width of the clines produced by this selection scheme. Our estimate of selection is based on this width, and so:

$$s = \frac{8\sigma^2}{w^2} = \frac{4s^*}{k} \left\{ 1 - 2 \int_0^1 \alpha f d\alpha \right\}.$$

We must now compare the estimate of the number of loci (made assuming coadapted pairs),  $n = 2S/s$ , with the actual number of loci,  $n^* = Nk$ :

$$\frac{n^*}{n} = \frac{Nks}{2S} = \frac{ks}{4s^*} = \left\{ 1 - 2 \int_0^1 \alpha f d\alpha \right\}.$$

This ratio depends on the shape of fitness function,  $f$ , shown in fig. 2. If fitness is very sensitive to deviations from the optimum (curve  $c$ ), then the estimate will be almost correct; if the fitness function is quadratic ( $f = 1 - \alpha^2$ ; curve  $b$ ), then we will overestimate the number of loci by a factor of 2; whilst if fitness is low unless  $\alpha \sim \pm 1$  (curve  $a$ ), the error will be large. However, such extreme schemes do not seem plausible, and so the estimate of  $n \sim 150$  derived in the last section is probably of the right order of magnitude regardless of the underlying pattern of epistasis.

(v) *Linkage*

If selection acts to maintain clines at two loci, linkage disequilibrium will build up between them simply because alleles from the same side of the cline will tend to be found together (Li and Nei, 1974). This disequilibrium will sharpen the clines, and so reduce the estimate of the number of loci involved (Slatkin, 1975). The figure obtained from the crosses thus refers to the number of independent chromosome segments, not the number of genetic loci. However, the total selection detected here is rather less than the total recombination rate. John and Hewitt (1970) observed an average of 1.7 chiasmata on each of the 11 chromosome arms; the average map length of an arm is therefore  $\sim 0.85$  Morgans, compared with average selection of  $2.5/11 = 0.23$  per arm. Unless strong selection is acting at other stages, disequilibrium between selectively independent loci should be low.

Disequilibrium between coadapted pairs acts rather differently. If linkage is tighter than epistasis ( $r < s$ ), selection will only act against rare recombinants which will be present at a frequency  $\sim r/s$ . The effective selection on a closely linked pair will therefore be  $(r/s) s = r$ , and so be independent of the strength of epistasis; the figure of 3 per cent for the average selection on each locus would refer to the average recombination rate between linked pairs. It is unlikely that two randomly chosen genes should be linked this closely, though there may be some tendency for coadapted genes to be linked, both because of "sieving" effects, whereby only closely linked pairs can become established, and because of the evolution of modifiers in the (presumably brief) period when the alleles were polymorphic (see Turner, 1977, for a discussion based on *Heliconiid* Mullerian mimics). The possibility that coadapted pairs are tightly linked could be tested by comparing the fitness of F2 hybrids from crosses between populations spanning the cline with that of individuals from the centre which had been recombining for much longer.

(vi) *Maternal effects and sterility*

The reduction in viability of offspring from the middle of the zone might be because their mother was heterozygous rather than because their own genomes had been broken up by recombination; this parental effect might be passed on through sterility (which could only account for the embryos which failed to develop) or by the influence of the maternal cytoplasm. Maternal effects would make the estimate of the proportion of loci acting through heterozygous disadvantage ( $\alpha = 0.13$ ) too low. Since the estimate of the number of loci depends on  $\alpha$  ( $n = 2S/s(1 + \alpha)$ ), maternal effects or sterility would make  $n$  rather too high. However, the error would only be serious if most of the early inviability were due to maternal heterozygosity.

(vii) *Assortative mating*

Potentially, the largest source of error would be strong assortative mating between the races. The low viability in these crosses might just result from the forced mating when single pairs were set up. In nature, there might be fewer hybrids, hence weaker selection on each allelic

difference, and a broader cline in inviability than is suggested by the strong selection seen in the laboratory. In the most extreme case, there might be a single pair of loci, at which recombinant gametes produce very unfit organisms. The broad cline, which has been taken to imply selection of only ~3 per cent on each difference, might be due instead to a ~3 per cent rate of hybridization. However, this extreme possibility cannot hold, since the chromosomal genotypes are found in Hardy-Weinberg proportions (5:19:16 in a sample of 40 females, implying less than 30 per cent assortment at 95 per cent significance levels; Barton, 1980). It is possible that a few sets of loci are involved, and although there is strong assortment for these, the chromosomal difference is only weakly correlated with them; however, it is even harder to see how assortative mating for several genetic differences could evolve than it is for simple single locus assortment (see Discussion).

## 6. DISCUSSION

To summarise, the observation that severe hybrid inviability is spread over a region much wider than the individual dispersal range implies that the inviability is caused by many weakly selected differences. The ratio between the total selection and the individual selection, deduced from the cline width, gives an estimate of ~150 for the number of loci involved; this estimate is probably too low, since other loci may well affect other components of fitness, or be relatively weakly selected.

This method for assessing the differentiation between divergent populations seems robust to the various assumptions made about the exact way selection works. In principle, it could be applied to analyse hybrid zones in a variety of organisms. However, the only other zone we know of where hybrid inviability has been detected and mapped is that between the frogs *Pseudophryne bibroni* and *P. semimarmorata* (Woodruff, 1979). These were originally recognised by differences in adult colour, which intergrade in a cline about 3 km wide. Egg masses collected from within a 5 km zone which spans the colour cline show about 20 per cent more inviability than egg masses outside the zone. This is wider than the colour cline, and so selection on each locus affecting embryonic viability must be weaker than that on the colour loci. However, without data on the dispersal ability of these frogs no quantitative interpretation is possible (see Barton and Hewitt, 1981*b*).

We will now consider what these observations imply about the present status of the two groups of grasshoppers, and about the way the differences between them might have arisen. First, we may contrast the divergence of many genes to produce hybrid inviability with the lack of electrophoretic divergence; 20 enzyme loci have been studied, six of them polymorphic, but no difference between the two races has been detected (Halliday, Barton, and Hewitt, unpublished). It has been argued that the lack of a strict correlation between electrophoretic divergence and speciation shows that isolation must be due to "regulatory" loci, or to DNA sequences that do not code for proteins (Wilson *et al.*, 1974; Hedrick and McDonald, 1980). However, this argument is not supported by the *Podisma* data, for several reasons. First, although the conservation of these 20 enzymes may suggest that they diverge more slowly than other loci, this does not mean

that the relevant divergence is at utterly different sites. The inviability may be caused by quite ordinary changes at loci coding for "structural" proteins which happen not to be involved in the biochemical steps mediated by the enzymes usually studied. We would not argue that "regulatory" evolution is unimportant, but simply that electrophoretic data tell us nothing about such evolution, since they themselves are only a limited sample of structural proteins.

Second, even if these 20 enzymes are typical, it is not surprising that no differences were detected. On the basis of estimates of the number of lethal loci in *Drosophila* of  $\sim 10,000$ , a figure which may be too low if many genes can only give weakly selected mutations, one would only expect  $(150/10,000) \times 20 = 0.3$  of an enzyme locus to have differentiated so as to contribute towards the inviability.

Of course, many loci might have fixed alleles with no selective effect by drift, or advantageous alleles might have arisen in one race but not the other. If this sort of divergence, irrelevant to the development of isolation, had been extensive, some of the enzymes should have changed. However, even if they had changed, the differences would have faded away by now—advantageous alleles would have swept forward, whilst neutral alleles would form shallow, decaying clines which would not have been detected in our survey. This decay would only be prevented if the hybrid zone were a strong barrier to gene flow. Calculation of the strength of multilocus barriers is complicated, since, in general, many interlocus interactions must be taken into account. However, in this case the total selection acting ( $S \sim 2.5$ ) is less than the total amount of recombination, as judged from the average of 18.5 chiasmata per cell (John and Hewitt, 1970 and 5v). Therefore, linkage disequilibrium may be ignored, and the single locus results of Barton (1979*b*) may be used. If each locus is selected by  $s = 3$  per cent, then after  $t = 10,000$  generations (the time since the two races met after the last glaciation), the flow of neutral alleles will be blocked over a region of chromosome  $2\sqrt{s/t} = 0.5$  centimorgans long. Since there are roughly 150 selected loci, only 75 centimorgans will be blocked, and even within these regions, mildly advantageous alleles will still be able to flow from one race to the other.

So, although there is striking postmating isolation, the two races can exchange alleles freely at most loci. It has been suggested that under these circumstances there will be selection for premating isolation, to allow the completion of speciation, and eventual sympatry (Wallace, 1889; Dobzhansky, 1937). However, the Hardy-Weinberg proportions of the karyotypes imply that such isolation has not developed in *Podisma* (see 5vii). Indeed, few hybrid zones show assortative mating, and in those that do, there is little evidence that it arose by reinforcement in the zones themselves (see Barton and Hewitt, 1981*b*). This may be because although absolute assortment would be at a strong advantage, and so could be selected quite rapidly (see, for example, Thoday and Gibson, 1962), it would be hard for assortment to evolve gradually. The *Podisma* zone involves many loci and initially there could only be assortment for one of these. The greatest advantage of a modifier producing partial premating isolation would therefore be  $\sim 3$  per cent, and this advantage could easily be swamped by pleiotropic effects of the change in mating preference. (This argument also holds if the modifier induces assortment for a polygenic character.)

Paradoxically, although the load from hybridisation is greatest when many genes are involved, since the load is spread over a wider area, this is also the situation where it is hardest for assortment to evolve and reduce the load.

Thus we know that although many genes have diverged, this divergence is not yet great enough to affect evolution at many loci, and will probably not be reinforced in the hybrid zone itself. Still, these polygenic changes might have been more extensive, and could have led to the isolation of the two types into separate species. If we want to know how species form (or at least, how they could form), we must ask how the differentiation in *Podisma* arose in the first place. Many models of speciation have been proposed, but here we will only contrast mechanisms driven by continuous selection in an essentially infinite population, with those relying on drift to knock a small population from one stable state to another, *via* a brief period of instability. Thinking heuristically in terms of Wright's adaptive topography, we can imagine on the first scheme that a peak, which carries both populations, splits into two incompatible peaks perhaps because of environmental change, and so separates the populations from each other. The alternative is that the topography remains fixed, but a population jumps by chance from one peak to another. Some mechanisms fall between these two extremes—for example, the chance fixation of incompatible mutations—but the classification seems the most useful (see Templeton, 1980).

So, the question is whether we can use data such as those for *Podisma* to distinguish these models of speciation. Unfortunately, we think not. All we can see are a certain set of differences, and it is hard to see whether drift or selection is the more likely cause. In a small population, a large number of weakly selected changes would accumulate faster than a few strongly selected shifts, since the rate of fixation by drift falls exponentially with the selection on each locus, whilst the time taken rises only linearly with the number of diverging loci. However, demes fixed for weak peak shifts are much more likely to be swamped by immigration from outside—Karlin and McGregor (1972) showed that for the differentiation to survive the proportion of the deme replaced by immigrants must be less than  $s/8$ , (see Lande, 1979). It is not clear whether drift can fix a lot of weak incompatibilities, as in *Podisma*, more easily than a single strong change.

Turning to selective mechanisms, we can apply Haldane's (1957) load argument to show that substitutions at many loci must take longer to accumulate than substitutions at only a few, regardless of the selective force driving each change. Haldane suggested a rough figure of 30 selective deaths per capita per fixation, which implies that ~4500 deaths were needed to substitute the ~150 new alleles in *Podisma*. In fact, the number of deaths may have been greater, since our estimate of 150 is biased towards the strongly selected loci, and so neglects the load due to any weaker changes. Selectively driven divergences of the two races of *Podisma* would have taken many millenia.

Comparing the two possibilities, then, the observation that many loci have diverged may be more easily explained by drift than by selection. However, the present-day architecture of isolating barriers cannot clearly distinguish the original mechanism.

Despite this, we can still make two points. Firstly, all of the differences in *Podisma* coincide. This could either be because they all arose in the

same place, perhaps because the alleles in each race are strongly coadapted with each other, or because they have been brought together by secondary contact. Since we know that the races must have been separated 10,000 generations ago by glaciers, the latter explanation is much more plausible. In general, it is striking that hybrid zones, initially identified through only a few loci, should often prove to involve many other coincident changes (e.g., *Caledia captiva* (Shaw and Wilkinson, 1980), *Mus musculus/domesticus* (Hunt and Selander, 1973), *Bombina bombina/variegata* (Szymura and Farana, 1978); for a review, see Barton and Hewitt, 1981*b*). This implies that either the rate of production of new clines is slower than the rate at which clines are brought together by population restructuring, or that the differences often arise together, and stay together.

This brings us to the second point; though we can say little about how differences arose, theoretical arguments entail that they must have arisen in a series of small steps, in each of which only a few loci change (see Templeton, 1980). If drift is responsible, the simultaneous fixation of many alleles is far less likely than the fixation of any one of them even in minute populations. On the other hand, if the isolation were produced by selection, extreme coadaptation must be invoked to explain simultaneous changes. Generally, the cohesion of even sluggish species over wide areas argues that the development of isolation is usually slow and cumulative (Mayr, 1970; Ehrlich and Raven, 1969).

Although it is rash to speculate about past events, the nature of the isolation in *Podisma* supports the classic allopatric model of speciation by the gradual accumulation of incompatibilities by geographically separate (though not necessarily isolated) populations (see Lande, 1980, for a detailed exposition). It is striking that so much cryptic differentiation should underly an apparently minor chromosomal difference; perhaps many more undetected hybrid zones dissect apparently uniform groups into nascent species.

*Acknowledgements.*—N. Barton was supported by a NERC studentship, and later by an SRC postdoctoral fellowship. G. Hewitt was also supported by an SRC grant. We are most grateful to T. M. East and J. Spencer for their help in rearing these insects, to M. et Mme Aviotti for their hospitality in Casterino, to J. Barrett, J. McKenzie, M. Shaw, and G. Dover, for their comments on the manuscript, and to J. Hunt for her patient typing.

## 7. REFERENCES

- AVISE, J. C. 1976. Genetic differentiation during speciation. In *Molecular evolution*, ed. F. J. Ayala, pp. 106-122. Sinauer, Sunderland, Mass.
- AVISE, J. C., AND AYALA, F. J. 1976. Genetic differentiation in speciose versus depauperate phylads: evidence from California minnows. *Evolution*, 30, 46-58.
- AYALA, F. J. 1975. Genetic differentiation during the speciation process. *Evol. Biol.*, 8, 1-78.
- BARTON, N. H. 1979*a*. The dynamics of hybrid zones. *Heredity*, 43, 341-359.
- BARTON, N. H. 1979*b*. Gene flow past a cline. *Heredity*, 43, 333-339.
- BARTON, N. H. 1980. The fitness of hybrids between two chromosomal races of the grasshopper *Podisma pedestris*. *Heredity*, 45, 47-59.
- BARTON, N. H., AND HEWITT, G. M. 1981*a*. A chromosomal cline in the grasshopper *Podisma pedestris*. *Evolution*, 35, in press.
- BARTON, N. H., AND HEWITT, G. M. 1981*b*. Hybrid zones and speciation. In *Essays on evolution and speciation in honor of M. J. D. White*, eds. W. R. Atchley, and D. S. Woodruff. Cambridge University Press.

- BARTON, N. H., AND HEWITT, G. M. 1982. A measurement of dispersal in the grasshopper *Podisma pedestris*. In prep.
- BAZYKIN, A. D. 1972. The disadvantage of heterozygotes in a population within a continuous area. *Genetika*, 8, 162-167.
- BAZYKIN, A. D. 1973. Population genetic analysis of disrupting and stabilising selection. Part II. Systems of adjacent populations and populations within a continuous area. *Genetika*, 9, 156-166.
- DOBZHANSKY, T. 1936. Studies on hybrid sterility. II. Localization of sterility factors in *Drosophila pseudoobscura* hybrids. *Genetics*, 21, 113-135.
- DOBZHANSKY, T. 1937. Genetic nature of species differences. *Amer. Nat.*, 71, 404-420.
- DOBZHANSKY, T. 1970. *Genetics of the evolutionary process*. Ch. 10. Columbia University Press, New York.
- EDWARDS, A. W. F. 1972. *Likelihood*, Cambridge University Press.
- EHRlich, P. R., AND RAVEN, P. H. 1969. Differentiation of populations. *Science*, 165, 1228-1232.
- FUTUYMA, D. J. 1979. *Evolutionary Biology*. Sinauer, Sunderland, Mass.
- GRANT, V. 1964. The biological composition of a taxonomic species in *Gilia*. *Adv. Genet.*, 12, 281-328.
- HALDANE, J. B. S. 1957. The cost of natural selection. *J. Genetics*, 55, 511-524.
- HEDRICK, P. W., AND MCDONALD, J. F. 1980. Regulatory gene adaptation—an evolutionary model. *Heredity*, 45, 85-99.
- HEWITT, G. M. 1975. A sex-chromosome hybrid zone in the grasshopper *Podisma pedestris* (Orthoptera: Acrididae). *Heredity*, 35, 375-387.
- HUNT, W. G., AND SELANDER, R. K. 1973. Biochemical genetics of hybridization in European house mice. *Heredity*, 31, 11-33.
- JOHN, B., AND HEWITT, G. M. 1970. Inter-population sex chromosome polymorphism in the grasshopper *Podisma pedestris*. I. Fundamental facts. *Chromosoma*, 31, 291-308.
- KARLIN, S., AND MCGREGOR, J. 1972. Application of the method of small parameters to multi-niche population genetic models. *Theor. Pop. Biol.*, 3, 186-209.
- KING, M. C., AND WILSON, A. C. 1975. Evolution at two levels: molecular similarities and biological differences between humans and chimpanzees. *Science*, 188, 107-116.
- LANDE, R. 1979. Effective deme sizes during long-term evolution estimated from rates of chromosomal rearrangement. *Evolution*, 33, 234-251.
- LANDE, R. 1980. Genetic variation and phenotypic evolution during allopatric speciation. *Am. Nat.*, 116, 463-479.
- LI, W. H., AND NEI, M. 1974. Stable linkage disequilibrium without epistasis in subdivided populations. *Theor. Pop. Biol.* 6, 173-183.
- MATHER, K., AND VINES, A. 1951. Species crosses in *Antirrhinum*. *Heredity*, 5, 195-214.
- MAYR, E. 1970. *Populations, species and evolution*. Harvard Univ. Press.
- MILKMAN, R. D. 1967. Heterosis as a major cause of heterozygosity in nature. *Genetics*, 55, 493-495.
- NAGYLAKI, T. 1974. Conditions for the existence of clines. *Genetics*, 80, 595.
- NEI, M. 1972. Genetic distance between populations. *Am. Nat.* 106, 283.
- PONTECORVO, G. 1943. Viability interaction between chromosomes of *Drosophila melanogaster* and *Drosophila simulans*. *J. Genetics*, 45, 51-66.
- SHAW, D. D., AND WILKINSON, P. 1980. Chromosome differentiation, hybrid breakdown, and the maintenance of a narrow hybrid zone in *Caledia*. *Chromosoma*, 80, 1-31.
- SLATKIN, M. 1975. Gene flow and selection in a two locus system. *Genetics*, 75, 787-802.
- SVED, J. A., REED, T. E., AND BODMER, W. F. 1967. The number of balanced polymorphisms that can be maintained in a natural population. *Genetics*, 55, 469-481.
- SZYMURA, J. M., AND FARANA, I. 1978. Inheritance and linkage analysis of five enzyme loci in interspecific hybrids of toadlets, *Bombina*. *Bioch. Genet.*, 16, 307-319.
- TEMPLETON, A. R. 1980. Modes of speciation and inferences based on genetic distances. *Evolution*, 34, 719-729.
- THODAY, J. M., AND GIBSON, J. B. 1962. Isolation by disruptive selection. *Nature*, 193, 1164-1166.
- TURNER, B. J. 1974. Genetic divergence of Death Valley pupfish species: biochemical versus morphological evidence. *Evolution*, 28, 281-294.
- TURNER, J. R. G. 1977. Butterfly mimicry—the genetical evolution of an adaptation. *Evol. Biol.* 10, 163-206.
- WALLACE, A. R. 1889. *Darwinism*, Ch. 7. Macmillan, London.



- WILSON, A. C., SARICH, V. M., AND MAXSON, L. R. 1974. The importance of gene rearrangements in evolution: evidence from studies on rates of chromosomal, protein and anatomical evolution. *Proc. Nat. Acad. Sci. (U.S.A.)*, 72, 5061-5065.
- WOODRUFF, D. S. 1979. Postmating reproductive isolation in *Pseudophryne* and the evolutionary significance of hybrid zones. *Science*, 203, 561-563.
- WRIGHT, S. 1935. Deviation from an optimum. *J. Genet.*, 30, 243-256.