Errata

p. 11 last line. replace 'causes' with 'cause'
p. 22 line 18. replace 'McNair' with 'Macnair'
p. 31 2nd line of footnote. replace 'if' with 'of'
p. 47 footnote delete 'early'
p. 49 line 3. insert 'that is' after 'An allele'
p. 50 line 9 replace 'corky' with 'corkey'
p. 51 last paragraph, 2nd line. replace 'McNair' with 'Macnair'
p. 52 lines 4 and 11. replace 'McNair' with 'Macnair'
p. 71 line 1. insert after full stop. 'Relatedness is usually represented by r, here z is used to avoid confusion with the re-mating frequency r used in chapter 5.'
p. 73 last paragraph, 4th line. replace 'Graffen' with 'Grafen'
p. 73 line 1 replace 'Where' with 'where'
p. 96 line 5. replace 'equilibria' with 'equilibrium'
p. 103 3rd paragraph, line 2. replace 'analysis' with 'analyses'
p. 108 2nd to last line. insert 'of' between 'fixation' and 'alleles'
p. 108 line 2. replace 'phenomena' with 'phenomenon'
p. 135 3rd line. replace 'to' with 'from'
p. 140. legend to fig 5.7. replace 'top to bottom' with 'bottom to top'
p. 142 line 2. replace 'phenomena' with 'phenomenon'
p. 143 line 7. replace 'X Linked' with 'X-linked'
p. 145 line 2. insert '(W)' after cline width.
p. 160 line 5 from bottom. replace '3.2.4' with '5.2.4'
p. 196 append to footnote 'where p is the frequency or the fused chromosome'
p. 217 7th reference. replace 'Graffen' with 'Grafen'
Pre- and post-zygotic isolation in hybrid zones

Adam Herzl Leibowitz

The Galton Laboratory
Department of Genetics and Biometry
University College London

Abstract

This thesis considers three aspects of hybrid zones: modification through reinforcement; maintenance by assortative fertilization; and, the extent of morphometric variation across the *Podisma pedestris* zone.

The theory of speciation by reinforcement postulates that, in areas of hybridization, selection will favour the evolution assortative mating since assortment prevents wastage of gametes in the production of partially sterile or inviable hybrids. A model is considered that extends this hypothesis to the evolution of hybrid inviability. Reduced viability of partially sterile hybrids may be favoured if they compete for resources, or parental care, with their non-hybrid sibs. The model is analysed for monogamous and polygamous species and for different degrees of sib-competition. Such reinforcement can only occur if hybrid fertility is very low, and is less likely, but not impossible, in monogamous species.

A model is considered to determine the effect of assortative fertilization in the maintenance of clines between divergent populations. Assortment will produce frequency dependent selection, against the rare type, which will lead to stable clines if balanced by dispersal. When applied to the chromosomal cline in *Podisma pedestris* (in which assortment for karyotype has been observed), the model predicts a much narrower cline than is actually observed.

An analysis of morphometric variation across the *Podisma* hybrid zone is also presented. No difference in morphology is detected between karyotypes within the zone. However, a significant correlation between size and the frequency of the XO karyotype is shown. Whether this is due to genetic or environmental variation cannot be determined with the available data.
Acknowledgements

I would like to thank my supervisor, Dr. Nick Barton, for his advice and encouragement throughout the gestation of this thesis. I would also like to thank all those past and present members of the Galton Laboratory who have helped and encouraged me. In particular, I have benefited greatly from discussion and debate with Neil Sanderson, Jaques Gianino, Dot Currie, Shahin Rouhani, and, especially over the last few months, Jim Mallet and Andrew Pomiankowsi.

Dot Currie and Brian Connolly were extremely hospitable and put me up, and put up with me, in Edinburgh during much of the time I was working on this thesis, without complaining even when 'a few weeks' turned into a few months. Thanks to Amanda Bennet for help with the grasshopper collections, and to all those who helped in previous years; Steve Meluish, Dot Currie, Nuala Gregory, Brian Connolly, Ella Leibowitz, Amanda Bennet, and Katy Trent: oh well, at least you got a bit of a tan.

I suppose I should thank my flatmate, Richard Pollock, but I'm not sure what for. At least he has made me see that something interesting happens every day, even in Camberhell. However, no thanks are due to FisHead & Issac, who are both as mad as muons and deserve all they get.

Lastly, I would like to especially thank my parents, brother and sisters, who, over the last few years, have been a limitless source of support, encouragement, free meals and much much more.
3.4.1 Methods ................................................................................................................. 70
   Inclusive fitness and Hamilton's rule ................................................................. 70
   Allele frequency change at a single locus ......................................................... 71
   Inclusive fitness of alleles in the model of reinforcement ................................. 74
   Genotypes, gene frequencies and mean fitness over two generations .......... 75
3.4.2 Results: Single populations ............................................................................. 79
   3.4.2.1 Sib competition only ............................................................................... 79
       Polygamy ........................................................................................................ 80
       Monogamy ..................................................................................................... 80
   3.4.2.2 Competition between sibs and non-sibs ................................................. 85
       Polygamy ........................................................................................................ 85
       Monogamy ..................................................................................................... 89
   3.4.2.3 Modifiers that also reduce viability of non-hybrids ............................. 92
3.5 Modification of selection in structured populations ........................................ 94
   3.5.1 The importance of migration ................................................................. 94
   3.5.2 A simple model of migration ................................................................. 96
   3.5.3 Modification of selection in a two deme system ..................................... 99
       Polygamy ........................................................................................................ 99
       Monogamy ..................................................................................................... 101
   3.5.4 Modification in hybrid zones ................................................................. 101
3.6 Discussion ............................................................................................................. 103

Chapter 4 The hybrid zone in *Podisma pedestris* ........................................... 109
4.1 *Podisma pedestris* ......................................................................................... 109
4.2 A narrow hybrid zone ....................................................................................... 111
4.3 A narrow tension zone ...................................................................................... 113
4.4 Reproductive isolation between the races .................................................... 115
4.5 New work on the Podisma pedestris hybrid zone ....................................... 116

Chapter 5 The effect of homogamy on the chromosomal cline in
   *Podisma pedestris* ........................................................................................... 117
5.1 Introduction ........................................................................................................ 117
   5.1.1 Assortative fertilization in *Podisma pedestris* .................................... 117
   5.1.2 The consequences of homogamy ............................................................. 119
      5.1.2.1 Theory of assortative mating ........................................................... 119
      5.1.2.2 Assortative mating in hybrid zones ............................................... 121
      5.1.2.3 Clines maintained by homogamy .................................................... 122

-- 5 --
5.1.3 Assortative fertilization and the chromosomal cline in *Podisma pedestris*:
a paradox? ................................................................................................................... 123

5.1.4 Outline ................................................................................................................. 123

5.2 The model ............................................................................................................... 125

5.2.1 Describing populations of X linked genes .......................................................... 125

5.2.2 A model of assortative fertilization .................................................................... 128

5.2.2.1 Single populations ......................................................................................... 130

5.2.2.2 Approximations for weak assortment ......................................................... 138

5.2.2.3 Comparing the exact and approximate equations ........................................ 140

5.2.3 Assortative fertilization and clines .................................................................... 142

5.2.3.1 Computer simulations .................................................................................. 143

5.2.3.1.1 Simulation results - cline width ........................................................... 145

5.2.3.1.2 Simulation results - heterozygote deficit ............................................ 149

5.2.3.1.3 Simulation results - Gene frequency differences between the sexes 154

5.2.4 Extension to multiple mating ............................................................................ 154

5.3 Discussion .............................................................................................................. 158

5.3.1 Some assumptions of the model and their consequences ................................ 159

Equal degrees of homogamy in the two races ......................................................... 159

Mating frequency .................................................................................................. 160

Sperm precedence ................................................................................................. 162

5.3.2 Application to the *Podisma pedestris* hybrid zone ........................................ 162

5.3.2.1 Measured parameters .................................................................................... 163

5.3.2.2 Predictions from measured parameters. ...................................................... 164

Heterozygote deficit ............................................................................................... 166

Cline width ............................................................................................................. 167

5.3.2.3 Reconciling the theory and data .................................................................. 167

5.3.2.3.1 Cline width ............................................................................................ 168

5.3.2.3.2 Strength of selection. ............................................................................. 169

5.3.2.3.3 Gene flow .............................................................................................. 175

5.3.2.4 Understanding the *Podisma* chromosomal cline ...................................... 179

5.3.3 Assortative fertilization as a barrier to gene flow ............................................ 180

5.3.4 Conclusions ....................................................................................................... 180
Chapter 6 Morphometric variation across the hybrid zone in *Podisma Pedestris* ................................................................. 182

6.2 Methods & Analysis...................................................................................................................... 185

6.2.1 Sampling................................................................................................................................ 185
6.2.2 Dissection & karyotyping......................................................................................................... 185
6.2.3 Choice of variables.................................................................................................................. 186
6.2.4 Measuring grasshoppers........................................................................................................ 188
6.2.5 Analysis .................................................................................................................................. 189
   6.2.5.1 Missing values................................................................................................................ 191

6.3 Results ....................................................................................................................................... 192

6.3.1 Describing the data............................................................................................................... 192
6.3.2 Differences between XO and XY grasshoppers within populations?............................... 194
6.3.3 Examining between population differences ........................................................................ 195
   6.3.3.1 Racial differences?..................................................................................................... 195
   6.3.3.2 Clinal variation in morphology?................................................................. 195
   6.3.3.3 Clinal variation other than size?............................................................................ 199
6.3.4 Increased variation in the hybrid zone?.................................................................................. 202

6.4 Discussion ................................................................................................................................. 205

Chapter 7 Summary and conclusions............................................................................................ 209

References ........................................................................................................................................ 212
Tables & Figures

Chapter 1
Table 3.1 The effect of a modifier of fitness ................................................................. 55
Table 3.2 A summary of notation used in the two locus analysis .................................. 63
Table 3.3 Distributions of genotypes within sib groups. 2 loci, Polygamy ...................... 64
Table 3.4 Genotypes within sib groups. 1 locus. Polygamy ........................................ 77
Table 3.5 Genotypes within sib groups. 1 locus. Monogamy ....................................... 77
Table 3.6 Summary of notation used in the inclusive fitness analysis ......................... 78

Figure 3.1 Stages of the model .................................................................................... 58
Figure 3.2 $\Delta q_b$ vs s. Polygamy ........................................................................... 68
Figure 3.3 $\Delta q_b$ vs $p_{aq}$. Polygamy ...................................................................... 69
Figure 3.4 Conditions for the modifier to be favoured. Polygamy ......................... 69
Figure 3.5 $\Delta q_b$ vs s. Monogamy .......................................................................... 83
Figure 3.6 $\Delta q_b$ vs $p_{aq}$. Monogamy ................................................................. 83
Figure 3.7 Conditions for the modifier to be favoured. Monogamy and polygamy ... 84
Figure 3.8 $\Delta q_b$ vs s. Polygamy. More than 1 sib group per patch. $p_{aq} = 0.25$ .... 87
Figure 3.9 $\Delta q_b$ vs s. Polygamy. More than 1 sib group per patch. $p_{aq} = 0.15$ .... 87
Figure 3.10 Conditions for spread of the modifier. Polygamy. More than one sib
  group per patch ..................................................................................................... 88
Figure 3.11 $n^*$ vs s. Polygamy ............................................................................... 88
Figure 3.12 $\Delta q_b$ vs s. Monogamy. More than one sib group per patch. $p_{aq} = 0.15$. 90
Figure 3.13 $\Delta q_b$ vs s. Monogamy. More than one sib group per patch. $p_{aq} = 0.05$ ... 90
Figure 3.14 Conditions for the modifier to be favoured. Monogamy. More than
  one sib group per patch. ....................................................................................... 91
Figure 3.15 $n^*$ vs s. Monogamy. ........................................................................... 91
Figure 3.16 $a^*$ vs s. $p_a = \frac{1}{2}$. Polygamy ................................................................. 93
Figure 3.17 Two demes exchanging migrants. ............................................................. 98
Figure 3.18 Conditions for the spread of the modifier in a two deme system.
  Polygamy ............................................................................................................... 100

Chapter 4
Figure 4.1 *Podisma pedestris* karyotypes ............................................................... 110
Figure 4.2 Location of the *Podisma pedestris* hybrid zone ..................................... 112
Chapter 5
Table 5.1 Homogamy in *Podisma pedestris* ................................................................. 118
Table 5.2 Summary of notation used in the homogamy model ........................................ 127
Table 5.3 Genotypes of female progeny ........................................................................... 130
Table 5.4 Genotype frequencies of female progeny. Multiple mating ............................ 155
Table 5.5 Expectations for the *Podisma pedestris*cline................................................. 165
Figure 5.1 Description of female mating history ............................................................... 129
Figure 5.2 Allele frequency vs generation ...................................................................... 135
Figure 5.3 Heterozygote deficit vs generation .................................................................. 136
Figure 5.4 Allele frequency difference between sexes vs generation ............................ 136
Figure 5.5 Allele frequency difference between sexes vs allele frequency ..................... 137
Figure 5.6 Heterozygote deficit vs generation for $P = \frac{1}{2}$ ......................................... 137
Figure 5.7 Allele frequency vs generation. approximate and exact equations ............... 140
Figure 5.8 Heterozygote deficit vs generation. approximate and exact equations ......... 141
Figure 5.9 Allele frequency difference vs generation. approximate and exact equations 141
Figure 5.10 Allele frequency, heterozygote deficit and sex difference across a cline .... 146
Figure 5.11 Cline width vs $\sqrt{1/rx}$ .............................................................................. 147
Figure 5.12 Cline width vs $\sqrt{m/rx}$ ............................................................................. 148
Figure 5.13 Difference in width predicted by simulations and diffusion equation........... 148
Figure 5.14 $F_{mid}$ vs $rx$. measured after fertilization .................................................... 153
Figure 5.15 $F_{mid}$ vs $rx$. measured after migration ........................................................ 153
Figure 5.16 Change in gene frequency vs $P$ for different mating frequencies ............. 157
Figure 5.17 Proportion of egg pods laid by females that have mated at least twice ......... 174

Chapter 6
Table 6.1 Variables used for morphometric analysis ...................................................... 187
Table 6.2 Correlations between the 11 variables ............................................................. 192
Table 6.3 Principal components ..................................................................................... 193
Table 6.4 MANOVA test statistics .................................................................................. 194
Table 6.5 regression of the eleven variables vs logit($p$) ................................................. 197
Table 6.6 Principal components .................................................................................... 200
Table 6.7 Regression of principal components vs logit($p$) .......................................... 200
Table 6.8 Regression of sample variance vs $p(1-p)$ ...................................................... 203

Figure 6.1 Equipment for measuring grasshoppers ......................................................... 189
Figure 6.2 Sample means across the hybrid zone ........................................................... 198
Figure 6.3 Variation in size across the hybrid zone ......................................................... 202
Figure 6.4 Sample variances across the zone .................................................................. 204
Introduction: The background to hybrid zones, species and speciation.

1.1 Hybrid zones.

Hybrid zones are areas in which genetically distinct populations meet, mate and produce hybrids (Barton & Hewitt, 1985, 1989; Hewitt 1988). They consist of a cline, or set of clines in gene frequency between divergent populations. Such clines may vary widely in width: e.g. in hybrid zones between warning colour races of the butterfly *Heleconius erato* clines are ≈80km wide (Mallet, 1986b); ≈20km between *Mus musculus musculus* and *M. m. domesticus* (Hunt & Selander, 1973); ≈5 to 7 km between the toads *Bombina bombina* and *B. variegata* (Szymura & Barton, 1986, 1991); and only a few hundred metres in chromosomal hybrid zones of the grasshoppers *Podisma pedestris* and *Caledia captiva* (Barton & Hewitt, 1981a; Moran et al., 1980). They are typically narrow relative to the distribution of the species and may run for hundreds of kilometres across its range.

Hybrid zones have been found to subdivide many species of many genera; Barton & Hewitt (1985) review published work on 150 hybrid zones. The characters by which hybridizing taxa differ are many and varied. Of the 20 best studied hybrid zones listed by Barton & Hewitt (1985) 16 involve morphological characters; 13 allozymes; 12 behavioural and seven chromosomal differences. With the application of molecular techniques, less obvious differences between hybridizing taxa can be detected. For example, mtDNA differences are seen between hybridizing species of *Bombina* (Szymura et al., 1985) and the crickets *Gryllus firmus* and *G. pennsylvanicus* (Harrison et al., 1987); ribosomal DNA differences across the *Podisma* hybrid zone (Dallas et al., 1988) and Y chromosome sequence differences across the *Mus* hybrid zone (Vanlerberghe et al., 1986). While
morphological transitions across hybrid zones may be visible to the casual observer, molecular and chromosomal transitions are obviously more difficult to detect. One can therefore only guess at the number of apparently homogeneous species that are in fact subdivided by hybrid zones.

1.1.1 Origin and maintenance of clines and hybrid zones.

There are several possible explanations for the presence of a cline dividing a species range. It may have evolved, *in situ*, in response to variation in selection pressure, represent the continuing spread of a novel genotype through the species range or be the result of secondary contact between populations that have diverged in isolation.

Consideration of single locus models, in which selection varies at a sharp environmental discontinuity (Haldane, 1948; Slatkin, 1973) or along a gradient (Fisher, 1950; Clarke, 1966; Endler, 1977) have shown that stable clines may develop. Such clines are essentially the product of selection alone, but may be maintained in the presence of gene flow, provided that the spatial scale over which selection coefficients vary is greater than a certain minimum dependent on the strength of selection and the dispersal distance (Slatkin, 1973). Two of the classic examples of the action of natural selection, heavy metal tolerance and industrial melanism, show clinal variation of this form (McNeilly, 1968; Bishop, 1972).

It is likely that many of the hybrid zones studied are the result of secondary contact between previously isolated populations that have diverged at one or a number of loci (Barton & Hewitt, 1985; Hewitt, 1988). For instance, many hybrid zones run along regions previously covered with glaciers. When divergent populations come into secondary contact several outcomes are possible. If the populations differ only at neutral loci, steep clines will be initially set up, but will decay and get progressively wider as dispersal and hybridization causes the genotypes of each to diffuse into the range of the other, resulting...
eventually in a single homogeneous population. The width of a cline can be defined as the inverse of the maximum slope. After \( t \) generations of neutral diffusion through a uniform habitat, the expected width of a cline is given by \( W = \sqrt{2\pi \sigma^2 t} \) (Barton, 1979b), where \( \sigma^2 \) is the variance in parent-offspring distances along some axis, a standard measure of dispersal. Since many hybrid zones are very narrow, rough estimates of the time since secondary contact (e.g. the last ice age) and dispersal distances are sufficient to demonstrate that the observed clines are not due to neutral diffusion. For example, in *Podisma pedestris* \( \sigma \approx 20\text{m} \) and the races were last separated by glaciation some 8000 years ago. If the chromosomal cline were due simply to the diffusion of chromosomes it would be \( \approx 4.5\text{km} \) wide rather than the 800m actually seen. This clearly indicates that selection must be involved in its maintenance.

If, after secondary contact, alleles in one population have higher fitness than the other, they will spread in a 'wave of advance' behind a travelling cline with a constant velocity dependent on the strength of selection (Fisher, 1937). The dynamics of such a process are essentially the same as the spread of a novel genotype from its area of origin. This is well illustrated by the spread of the gene for warfarin resistance in rats (Greaves *et al.*, 1977). Movement of hybrid zones is, however, unlikely to be observed. If the selective advantage of one race over the other is only slight, the rate of spread will be slow and many generations of study would be required for any movement to be detected. Alternatively, if selection is strong, spread will be rapid; advantageous alleles are likely to have already passed through the hybrid zone.

**Tension zones**

If hybrids between divergent populations have lower fitness than either of the parentals, stable clines may be maintained since, dispersal into the area of contact is opposed by selection against hybrids; e.g. hybrids have lower fertility, viability or reproductive success than either parental (Bazykin, 1969, 1973; Barton, 1979b). Cline
width therefore depends on the relative strengths of selection and dispersal. Such clines have been called 'tension zones' (Key, 1968; Barton & Hewitt, 1985). As they are not confined to environmental transitions, tension zones are mobile and will move so as to minimize their length. If there is variation in population density across the cline, there will be a net flux of individuals causing the cline to move towards areas of low population density (Barton, 1979b). Such clines will become trapped by troughs of low density, or by barriers to dispersal (Barton, 1979b). For example, the fine scale position of the hybrid zone in *Podisma pedestris* can be largely explained by variation in population density and local barriers to dispersal (Nichols & Hewitt, 1986, 1988; Jackson, 1992).

Selection against hybrids may be due to heterozygote disadvantage at single loci (Bazykin, 1969; Barton, 1979b), or to an interaction between different alleles fixed at different loci in the hybridizing populations (Bazykin, 1973). Tension zones may also be maintained by selection acting directly on parental genotypes. For example, the hybrid zones between differently warning coloured *Heliconius* butterflies are maintained by frequency dependent selection against rare morphs (Mallet, 1986b, Mallet & Barton 1989a, 1989b). Individuals moving from one side of the hybrid zone to the other carry warning colours that are not recognised as such by the local predators and so suffer higher mortality.

Clines maintained by selection against hybrids are the consequence of partial postzygotic isolation between the hybridizing populations. Partial pre-zygotic isolation may also maintain clines since assortative mating is likely to lead to frequency dependent selection. Of the 20 well studied hybrid zones reviewed by Barton & Hewitt (1985), 12 connect populations with behavioural differences that may contribute to the maintenance of the zone (e.g. differences in mating call or habitat choice may lead to assortative mating). This effect has only rarely been considered. For example, a difference in the direction of coiling in *Partula suturalis* hybrid zones leads to a degree of assortative mating that can largely explain the maintenance of the cline (Johnson, 1982; Johnson *et al.*, 1990). Studies of the toads *Bombina bombina* and *B. variegata* show that despite the large differences in mating...
call between the 'species', selection directly on the call components is at most weak, and
not much greater than on marker enzymes (Sanderson et al., 1992). The effect of
homogamy in the maintenance of clines is considered in more detail in Chapter 5.

At equilibrium, the width of a simple tension zone is proportional to $\alpha \sqrt{s}$, where $s$
is the selection against hybrids. The constant of proportionality differs with different modes
of selection; for a cline across a uniform habitat, with heterozygote disadvantage, $w =$
$\sqrt{8\sigma^2/s}$ for an autosomal locus (Bazykin, 1969) and $w = \sqrt{12\sigma^2/s}$ for an X linked locus.
If cline widths and dispersal rates are known, these equations can be used to estimate
selection. They are however sensitive to errors in the estimates of dispersal distance and to
the assumptions of uniform density and dispersal.

Hybrid superiority

It is also possible that hybrid zones may be the result of hybrids having higher
fitness than either of the two parentals, but only in areas of intermediate habitat between
those to which the parentals are adapted (Moore, 1977). This has been advocated as an
explanation of some of the hybrid zones in vertebrates. However, the data available allows
for alternative explanations - i.e. tension zones (Barton & Hewitt, 1985; Hewitt, 1988). In
general, one would expect that areas in which hybrids are favoured would vary in width
and be patchily distributed and that the position and shapes of hybrid zones would reflect
this distribution. As with other environmentally determined hybrid zones, under the hybrid
superiority model, one would not expect clines in different characters to coincide (see
below). It is therefore unlikely that hybrid superiority is a major factor in the maintenance
of hybrid zones.
1.1.2 Hybrid zones vs single locus clines

The possibilities discussed above are examples of the way in which clines in single characters may be maintained. Some hybrid zones appear to involve transitions in only one or a few characters. For example, the *Heliconius* hybrid zones appear to involve only the few loci controlling warning colouration; there are no electrophoretic differences between the races and no evidence of reproductive incompatibility (Mallet, 1989; Mallet *et al.*, 1990). In most others, many different characters change across the hybrid zone (Barton & Hewitt, 1985; Harrison, 1990). For example, the *Bombina bombina/variegata* hybrid zone involves transitions in allozyme frequency, belly markings and colouration, morphology, habitat preference, mtDNA and mating call (Szymura & Barton, 1986, 1991; Szymura *et al.*, 1985; Sanderson *et al.*, 1992).

The widths of clines in individual characters within a hybrid zone may differ considerably. Clines in allozyme frequency in the *Mus* hybrid zone are \(\approx\)20km wide, whereas those in Y-chromosome markers are an order of magnitude narrower (Hunt & Selander, 1973; Vanlerberghe *et al.*, 1986). Similarly, in the hybrid zone between the grasshoppers *Chorthippus parallelus parallelus* and *C. p. erythropus*, clines in some song characteristics are \(\approx\)20km wide whereas stridulatory peg number and nucleolar organizer regions change over 4km and 0.5km respectively (Butlin & Hewitt, 1985a, 1985b; Hewitt *et al.*, 1988). Differences in width between the various clines that make up a hybrid zone indicate that selection varies between characters. However, if there are several coincident clines, dispersal of parental genotypes into the centre generates linkage disequilibrium between the genes involved. Selection at one locus will therefore affect alleles at others. In this case, estimates of selection based on measurements of cline width and dispersal are, in fact, estimates of this net selection. However, provided that selection acts at only a few loci, and recombination between them is high, the clines at different loci are approximately independent.

If selection against hybrids is due to the combined effects of selection at many loci, clines in marker loci (e.g. allozymes) are not the simple sigmoid curve seen when selection
acts primarily at single loci. Rather, two regions of the cline can be identified: the central portion characterised by a sharp step in frequency, flanked by shallow tails of introgression (Barton, 1983). Because of the linkage disequilibria maintained by dispersal, neutral loci embedded in strongly selected genomes will appear to be under selection. Stepped clines in neutral or weakly selected marker loci can therefore be taken as evidence that many loci contribute to the selected differences between hybridizing taxa. Analysis of the clines in such marker loci can be used to estimate the number of selected loci by which the populations differ and the amount of selection acting upon them. This methodology has, thus far, been applied only to the Bombina hybrid zone in which it has been shown that at least 55 genes of weak effect contribute hybrid unfitness (Szymura & Barton, 1986, 1991).

1.1.3 Environmentally determined hybrid zones or tension zones?

Many hybrid zones connect ecologically differentiated taxa (Harrison, 1990). This may be taken as evidence that they are maintained by variation in selection pressure. However, Barton & Hewitt (1985, 1989) argue convincingly that most hybrid zones are in fact tension zones maintained by a balance between dispersal and selection against hybrids. Their arguments are three fold. First, there is often evidence that hybrids have reduced fitness. Evidence may be indirect such as the increased frequency of parasitism or developmental and meiotic abnormalities in the hybrid populations or more direct such as reduced viability and/or fertility of F₁ hybrids and backcrosses or the reduced frequency of hybrids in successive age classes.

Second, as has already been noted, many hybrid zones involve changes in many different characters. In most cases, the clines in the different characters coincide and are of different width. For coincident clines to be maintained by variation in selection along an environmental gradient, requires that the effect of the gradient is the same for each character. It is unlikely that this would be the case unless the clines represent transitions in
co-adapted characters. Coincident clines in neutral characters are expected between populations that have come into secondary contact. However, in this case one would expect all clines to be of equal width. Variation in the width of clines within a single hybrid zone indicates that selection must be acting on at least some of the characters. Further evidence of selection can in principle come from the shape of the cline itself. The shape of any single cline reveals little about the forces important for its maintenance (Endler, 1977). The shapes of clines produced by neutral diffusion, response to an environmental gradient and simple tension zones are subtly different (Barton & Clark, 1990; Barton & Jackson, 1992), but it is unlikely that sample sizes in any survey would be large enough to distinguish them. On the other hand, tension zones maintained by selection at many loci, produce distinctly different (stepped) clines at marker loci which can be distinguished from simple clines.

The final argument for the prevalence of tension zones, is that even if hybrid zones are initially established in response to environmental gradients, they are likely to evolve into tension zones. Clines initially at single loci may be steepened by the spread of modifiers (Clarke, 1966; Endler, 1977). The same modifiers are unlikely to be favoured on both sides of the cline since they may interact with both the environmental gradient and the primary loci. Such co-adaptation may lead to a degree of genetic incompatibility between the hybridizing populations.

The coincidence between the position of a hybrid zone and an environmental discontinuity can only be taken as weak evidence that the discontinuity is the cause of the zone. Since tension zones are mobile and likely to become trapped at environmental discontinuities (Barton, 1979b), the same association between habitat and position of the zone is expected, whether the zone is due to adaptation to different environments or to incompatibility between populations that have come into secondary contact.

1.1.4 Motivation for hybrid zone research

There are several reasons for the current interest in hybrid zones, not least of which, as pointed out by Barton & Hewitt (1985), is simply that such a widespread
phenomenon deserves study, and hopefully an explanation. More incisive reasons arise from the light hybrid zones can shed on evolutionary processes. Hybridization between divergent populations can reveal much about the nature of the differences between them; for example, the number of genes involved and the magnitude of their effects.

Hybrid zones are of particular interest to students of speciation. A central tenet of neo-Darwinism is that gene flow binds populations together and so prevents divergence and speciation (Mayr, 1963). However, hybrid zones are seen to connect widely divergent populations, which nevertheless remain distinct. This raises the possibility that speciation may occur even in the presence of gene flow. An alternative view is that hybridization between divergent populations is actually one of the driving forces behind speciation. Before these ideas can be considered a precise definition of what is meant by the terms 'species' and 'speciation' is required.

### 1.2 Species and speciation.

Species are best defined in terms of what it is that maintains the similarity between individuals of a species and keeps separate species distinct. These similarities and differences are maintained by patterns of gene flow. Members of a species are similar because they share a common gene pool distinct from that of other species. Species are therefore fundamental units of the biological world, since they evolve independently: an evolutionary innovation in an isolated gene pool cannot spread into another. This view is summarised in the so called 'Biological Species Concept': species are "groups of actually or potentially interbreeding natural populations, reproductively isolated from other such groups", (Mayr, 1963). The mechanisms by which species are kept isolated are varied and are commonly classified as being either pre- or post-zygotic. Pre-zygotic isolating mechanisms prevent formation of hybrid zygotes (e.g. habitat, ethological or physical differences between species mean that hybridization is not possible, or incompatibilities between gametes prevent fertilization). If species are isolated by post-zygotic isolating
mechanisms, hybrids may be formed but gene flow is prevented by the elimination of hybrids (e.g. hybrid sterility, inviability or F$_2$ break down). Any one or combination of these mechanisms may provide full reproductive isolation between species. Such a view of species does not fit comfortably with traditional notions of species as distinct phenotypes, since reproductive isolation may evolve without any corresponding morphological divergence by which species may be recognised.

Although widely accepted as a useful definition, the biological species concept is not without its critics. One minor criticism concerns terminology. Mayr (1963) gave the title 'biological' to his species concept not because it refers to biological species but because the criteria used in defining a species are biological (reproductive isolation), and meaningless in any other context. Paterson (1978, 1982) rightly argues that other species definitions may also be 'biological' in the sense used by Mayr. The distinguishing feature of Mayr's definition is reproductive isolation. For this reason Paterson argues that the terminology should be altered to the 'Isolation Species Concept' or ISC.

Sokal & Crovello (1970) argue that the ISC is neither necessary nor sufficient for a proper understanding of species and speciation: unnecessary since speciation may be understood by simply considering the evolution of populations between which there are varying degrees of gene flow, and insufficient since in practice reproductive isolation cannot easily be tested. If one's task is to describe the extent and nature of organic diversity, one must utilize easily measurable parameters and so rely on morphological characters which have no role in the ISC. One obvious failing of the ISC is that it can only apply to sexually reproducing species: a strict application of the ISC would designate all asexual individuals as separate species. Conversely, populations that are completely isolated by geographical barriers to gene flow are not granted species status, despite the fact that their evolutionary fates are independent.

Another argument levelled at the ISC is that the terminology of 'isolating mechanisms' is misleading (Paterson, 1978, 1982). Use of the term suggests that the
features that provide reproductive isolation have evolved in order to fulfil this function. This is particularly true when the evolution of pre-zygotic isolation is considered (see Chapter 2). Paterson argues that, as speciation most often occurs in geographically isolated populations, the evolution of intrinsic isolating mechanisms is largely irrelevant to the process of speciation and should not therefore be included in any species definition. Paterson advocates as an alternative, the 'recognition species concept': species are groups of individuals that share common species specific mate recognition and fertilization systems. It is argued that this is superior to the ISC since it forces attention to be focused on processes that actually guide the evolution of mating systems. This definition is, in practice, very similar to the ISC. A common fertilization system is simply an isolating mechanism viewed from a slightly different perspective. A major failing of the recognition concept is that it takes no account of post-zygotic barriers to gene flow.

The debate over species concepts is largely fruitless. It is unlikely that taxonomists will ever embrace wholeheartedly a genetical species concept that cannot be applied easily in the field. However, the ISC remains the most useful definition for the study of speciation since it sets out clearly the problem to be explained: how are intrinsic barriers to gene flow between species established? Worries over the implications of different terms may be well founded, but should be easily placated by careful use of language. Theories of speciation based on the ISC will not be wholly satisfying since they only address the question of reproductive isolation in sexually reproducing organisms. Whether or not reproductive isolation is accompanied by recognizable phenotypic change or ecological divergence, that allows coexistence in sympatry, are related but separate questions. Throughout this thesis 'speciation' will refer to the evolution of complete reproductive isolation.
Chapter 1

1.2.2 Evolution of reproductive isolation.

Under the isolation concept of species, the process of speciation is clearly defined as the evolution of reproductive isolation. Framed in these terms speciation appears to be unlikely since any change that produces reproductive isolation is likely to be at a selective disadvantage when rare. For example, in the context of pre-zygotic isolation, an individual with a novel mating preference is unlikely to find a mate. Likewise, for post-zygotic isolation, a novel genotype, which on recombination with its ancestor produces inviable or sterile progeny, is also at a selective disadvantage. A useful metaphor is the idea of the 'adaptive landscape' introduced by Wright (Wright, 1932; Provine, 1986). The possible states in which a population may exist can be visualized as a graph of population mean fitness against allele frequencies, or the means of quantitative traits. Allele frequencies may vary at many loci and many traits may contribute to mean fitness: the landscape is therefore a multidimensional surface with many peaks and troughs. Since natural selection increases population mean fitness, evolution moves populations 'uphill' to a local peak. Species may be considered as groups of populations sitting at different 'adaptive peaks' which are stable to the invasion of foreign alleles.

The problem of speciation is that, in order to move from one adaptive peak to another, a population must apparently traverse the valley of low mean fitness which separates them. Until the low point of this valley is reached, such movement is opposed by natural selection. Three genetic models of speciation have been proposed that overcome this problem: changes in the landscape occur allowing adaptive divergence; populations may diverge along different 'neutral' pathways; or divergence may take place by genuine 'peak shifts' brought about by drift and selection.

Adaptive divergence.

Populations almost certainly do not live on static adaptive landscapes. Any real population is likely to exist in a complex environment which is, on an evolutionary time scale, likely to change; either because of physical changes or because its predators, prey
and competitors are themselves evolving. Different populations of a single species are likely to experience different selection pressures and adapt to them. Dobzhansky (1937) and Muller (1942) considered how such adaptation may lead to the evolution of reproductive isolation. In the simplest case, they considered two populations, both initially fixed for the genotype AABB, but experiencing different selection pressures so that one becomes fixed for aaBB and the other for AAbb. Reproductive isolation results if there is epistasis such that individuals carrying both of the new alleles have reduced fitness (i.e. reduced fertility, viability or reproductive success). However, since both of the transitions (A→a, B→b) were in response to selection, no valley has been traversed and reproductive isolation is a fortuitous consequence of adaptation. Reproductive isolation may, in principle, result from adaptation even if populations experience the same selection pressures. The response to selection will depend on the available genetic variation and so different populations may respond in different ways.

Although it may seem obvious that reproductive isolation will evolve as a pleiotropic effect of adaptation, there are few examples where this is clearly seen to be the case (Coyne, 1992). One clear example is seen in the monkey flower *Mimulus guttatus* in which the gene conferring copper tolerance in some populations is lethal when present in inter-population hybrids (McNair & Christie, 1983). It is interesting that in this example a single gene produces strong reproductive isolation. In general, one would expect that selection would elicit a polygenic response and therefore many genes would be involved. The *Mimulus* example may be exceptional; genetic analysis of components of reproductive isolation have shown that, in general, several to many\(^1\) genes contribute to isolation between species (Coyne, 1992).

\(^1\) The vagueness of these terms is due to the fact that, in general, the few generations of crossing involved in a typical genetic analysis of reproductive isolation cannot distinguish between 9 and 900 genes (Coyne, 1992).
'Neutral' divergence.

Nei (1976) has pointed out that the Dobzhansky/Muller model may operate without selection. If, in the absence of the allele b, the transition A→a is a neutral substitution (and likewise, B→b is neutral in the absence of a) reproductive isolation may evolve through mutation and drift alone. However, since this model relies only on mutation and drift, the time required for the accumulation of incompatibility can be very large (Nei et al. 1983, Bengtsson & Christiansen 1983). In the landscape metaphor, the evolutionary pathway connecting the two populations can be thought of as a high U-shaped ridge between two peaks; only transitions along the direct line from one peak to the other are opposed by selection. 'Neutral' divergence such as this is, however, unlikely to provide permanent isolation, unless hybrids between the two populations have zero fitness. Upon hybridization, the original AABB genotype will be produced and favoured by selection in the F2 or backcross generations and so isolation will be rapidly eroded (Barton, 1988).

Sexual selection may also result in divergence along neutral pathways. Models by Lande (1981) and Kirkpatrick (1982) have shown that the coevolution of female preferences and costly male traits moves populations to some point on a line of neutral equilibrium for the two characters, along which they may then drift up and down. A small population may therefore acquire extreme values of preference and trait, effectively isolating it from other such populations.

There is still much debate about driving force behind sexual selection, especially the origin of female preferences (Kirkpatrick, 1987; Kirkpatrick & Ryan, 1991; Maynard Smith, 1991). The debate is primarily about whether female preferences are the result of non-adaptive processes (as in the Lande & Kirkpatrick models) or because choice increases offspring fitness. However, both non-adaptive and 'good genes' models of sexual selection may result in divergence and reproductive isolation. Fisher (1930) originally suggested that, in order for sexual selection to take place, the trait favoured by females must initially be correlated with fitness in natural selection but that this correlation is lost during the course of evolution. The Lande and Kirkpatrick models have shown that this
requirement is unnecessary: all that is required is that there is initial linkage disequilibrium between the characters, which may be brought about by drift. Since no correlation with fitness is required, any trait may become the focus of sexual selection so long as there exists genetic variation for the trait and the female preference. 'Good genes' models of sexual selection specifically require that traits under sexual selection are not entirely arbitrary: they must be indicators of genetic value (Maynard Smith, 1991). Any such trait may be favoured. There are reasons why the same characters might be selected in different populations: recently isolated populations are likely to share the same genetic variation for preferences and traits and it is likely that some traits are better than others as indicators of fitness. However, under neither model of sexual selection is parallel evolution of isolated populations specifically predicted.

**Peak shifts**

A population may occasionally shift rapidly from one adaptive peak to another through a process combining drift and selection: a 'peak shift'. Drift may move a population away from its local adaptive peak and into the attractive domain of another; i.e. to the bottom of the valley between them. From this point, selection will rapidly move the population 'uphill' to a new peak. However, models of this process have shown that single peak shifts are unlikely to be major causes of reproductive isolation (Barton & Charlesworth, 1984; Lande, 1985). Essentially, the problem is that the average time between transitions is expected to be very large and increases with the depth of the valley between the peaks. Since only deep valleys (low hybrid fitness) lead to strong reproductive isolation, speciation by peak shifts is likely to be very rare. Also, since peak shifts are initiated by drift, they are only likely to occur in small demes or in continuous populations with small neighbourhood size (Lande, 1979, 1985; Rouhani & Barton, 1987).
1.3 Hybrid zones as barriers to gene flow: partial reproductive isolation

A hybrid zone represents an absolute barrier to the spread of alleles directly involved in its maintenance. If one were to consider only the spread of these genes, the hybridizing populations would be considered as separate species under the ISC. However, a hybrid zone is only a partial barrier to other genes.

Consider a neutral allele that is present on only one side of a single locus tension zone formed by secondary contact. When the cline was initially established, the allele in question will have been exclusively associated with the selected allele in one race. Because of this linkage disequilibrium, selection at the primary locus also has an effect on neutral alleles and impedes their spread. A neutral allele cannot pass through the hybrid zone until recombination has separated it from its original genetic background. Once this has occurred, the likely fate of any particular 'escapee' allele is that it will be lost from the population through drift. However, unless linkage between the selected and marker loci is particularly close, recombination within the hybrid zone generates many 'escapees', at least some of which will not be lost. Single locus clines, therefore, represent only minor barriers to gene flow, and the presence of a cline contributes little to reproductive isolation between hybridizing populations (Barton, 1979c).

A hybrid zone maintained by selection at many loci is a far stronger barrier to gene flow (Barton, 1979c,1983; Barton & Bengtsson, 1986). To diffuse through a multilocus hybrid zone, a foreign allele must recombine out of one and into another complex genome, before being eliminated from hybrid populations by the hitch-hiking effect of selection at other loci. The barrier strength, B, of a hybrid zone can be defined as the length of uniform habitat which it would take a neutral allele the same time to diffuse across as does the hybrid zone. The barrier strength increases with the strength of selection against hybrids, and the number of genes contributing to this and with decreasing recombination between selected and introgressing genes (Barton, 1983; Barton & Bengtsson, 1986).
Barrier strength can be estimated from the shapes of the clines in marker loci across hybrid zones, and from this the fitness of hybrids may be estimated (e.g. Szymura & Barton, 1986, 1991). The barrier strength can in turn be related to the time it would take a neutral allele to cross the zone. For example, in the *B. bombina/variegata* hybrid zone, $B \approx 51\text{km}$, in comparison to the physical width of only $\approx 6\text{km}$ (Szymura & Barton, 1991). The dispersal rate in *Bombina* is $\sigma \approx 0.99\text{km}$, so the hybrid zone would impede the spread of neutral alleles for $\approx 2700$ generations. Given that this is approximately the same as the age of the hybrid zone itself, it obviously acts as a major factor in maintaining the distinction between the 'species', at least with respect to neutral alleles (Barton & Hewitt, 1989).

If selection against hybrids reduces the population density, more individuals will migrate into a tension zone than out of it. It has been suggested that this 'hybrid sink' effect will increase the barrier to gene flow represented by a hybrid zone (Hall, 1973). The effect is only possible if population density is sensitive to the action of selection, i.e. selection must be 'hard' (Wallace, 1975). However, the magnitude of this effect is small unless selection is very strong (Barton, 1980b). The effect of physical barriers is likely to be more important. The mobility of tension zones means that they are likely to stabilize in areas of low population density and/or physical barriers to dispersal (e.g. across rivers and streams or areas of inhospitable habitat). In this way, a tension zone may represent a barrier to gene flow greater than that due to selection alone.

Hybrid zones are only minor barriers to advantageous alleles, since the selection due to disequilibrium in the zone is opposed by selection directly on the introgressing allele. Also, once recombined away from its original background, the introgressing allele is less likely to be lost through drift. For example, the 2700 generations it would take for a neutral allele to penetrate the *Bombina* hybrid zone is reduced to only $\approx 200$ generations for an allele with a 1% selective advantage (Barton & Hewitt, 1989). Even if selection against hybrids is very strong, a hybrid zone may represent a barrier no greater than, for example, a large river to a favoured allele.
Chapter 1

Many apparently stable hybrid zones connect what have been previously described as separate species, whereas others are between 'sub-species' or 'races'. Given the potential for gene flow across a hybrid zone, the classification of these taxa is clearly somewhat arbitrary. Unless hybrids have zero fitness, any pair of hybridizing populations must be considered as belonging to the same species.

1.4 Hybrid zones as stages in the speciation process.

Hybrid zones have been explicitly implicated in several models of speciation. The model of speciation by reinforcement has perhaps been the most popular (Dobzhansky, 1940). Hybridization often produces progeny of low fitness; therefore, in areas where hybridization takes place, one might expect natural selection to favour processes that reduce the frequency of such matings. In this manner, post-zygotic isolation may be reinforced by pre-zygotic isolation. Careful consideration of this model has, however, shown that reinforcement is at best unlikely (see Chapter 2). In other models of speciation, hybrid zones are implicated either because of their effects as barriers to gene flow or because variation across a hybrid zone represents variation in the genetic background against which future evolution takes place.

Barriers

Since the presence of a hybrid zone impedes gene flow, populations spanning a zone may diverge and so accumulate incompatibilities that may contribute to reproductive isolation. This role of hybrid zones has been especially stressed by White in his model of 'stasipatric' speciation (White, 1968). Prompted by studies of chromosomal hybrid zones in Australian grasshoppers, White proposed that speciation may occur in the following manner: 1) a new chromosomal rearrangement becomes locally established either within or on the edge of a species range. 2) the rearrangement spreads throughout part of the species range forming hybrid zones where it meets the ancestral arrangement. 3) because of meiotic
incompatibilities, heterokaryotypes have reduced fitness and so the hybrid zone acts as a barrier to gene flow, allowing the populations it separates to diverge, eventually acquiring full reproductive isolation.

Although the model was proposed specifically in relation to chromosomal rearrangements, it may equally apply to the effect of local fixation of any gene or combination of genes that produce partial reproductive isolation. There are several problems with this model. A new chromosomal arrangement that causes reduced fertility in heterozygotes is not easily fixed in a population since it is selected against when rare. This initial hurdle may be overcome by drift in small demes (Wright, 1941). However, the probability of local fixation is dependent on the strength of selection against heterokaryotypes: rearrangements that will provide a significant barrier to gene flow are extremely unlikely to become established even in very small demes (Bengtsson & Bodmer, 1976; Lande, 1979). Migration from other demes also greatly reduces the chance of establishment (Lande, 1979; Barton & Rouhani, 1991). The same arguments also apply to the local establishment by drift of new multilocus genotypes or 'adaptive peaks' that are reproductively isolated from the original (Barton & Charlesworth, 1984; Rouhani & Barton, 1987).

Once locally established, the stasipatric model requires that the new arrangement will spread through a portion of the species range. The evidence from present hybrid zones indicates that most are, in fact, stable. Unless the new arrangement has some selective advantage (or is subject to meiotic drive, as suggested by White) there is little reason for it to spread. The most important weakness of the original stasipatric model is however its assumption that karyotype changes drive speciation. It was discussed above how a single locus cline presents only a minor barrier to gene flow: the same is true for single chromosomal rearrangements. In well studied orthopteran hybrid zones there is little evidence that the chromosomal component contributes anything more than a minor amount to the barrier to gene flow (Shaw, 1981). Although there is evidence of an association between the rate of chromosomal evolution and speciation (Bush et al., 1977; Bengtsson,
1980) there is little to suggest that one drives the other. It is as likely that factors promoting karyotype evolution (e.g. small deme size, Wilson et al., 1975) also promote speciation.

Although single chromosomal clines cause only minor barriers to gene flow, other hybrid zones, maintained by strong selection at many loci represent much larger barriers (see above). However, since the barrier strength depends on the nature of introgressing alleles, only certain types of divergence are likely to be increased by the presence of a hybrid zone. Different 'neutral' alleles may easily accumulate on either side of the zone but these cannot contribute to reproductive isolation unless interactions between such alleles at different loci reduce hybrid fitness. Since generally advantageous alleles are not significantly impeded by the presence of a hybrid zone, adaptations arising on one side of a zone are likely to spread through it. Alternatively, if adaptations are sensitive to the genetic background in which they occur, adaptive divergence across hybrid zones is to be expected perhaps leading to 'parapatric speciation'.

**Genetic background: parapatric speciation.**

It has already been noted that clines may develop *in situ* across a species range in response to variation in selection pressure, provided that the 'grain' of the variation is not too fine (Clarke, 1966; Slatkin, 1973; Endler, 1977). This theoretical conclusion, and the observation of such clines (e.g. McNeilly 1968; Bishop, 1972) clearly demonstrates that complete geographic isolation is not necessary for adaptive divergence. Clarke (1966) and Endler (1977) proposed models in which the initial divergence promotes further divergence which increases isolation. Divergence is promoted since modifiers that increase the fitness of individuals on either side of the cline are likely to spread. Since the genetic background in which such modifiers evolve is different on either side of the cline, different coadapted modifiers may become established on either side. Incompatibility between these modifiers increases isolation across the cline. The same arguments apply equally well to tension zones produced after secondary contact of divergent populations: initial divergence in allopatry alters the subsequent course of evolution even when the populations re-establish
contact and share the same environment. It is important to note that the role of hybrid zones in this parapatric model of speciation is quite different to the barrier role they are supposed to play in stasipatric speciation: the very presence of a hybrid zone alters the 'environment' in which evolution takes place.

The consequence of runaway sexual selection in a cline is similar to that of coadaptation. The outcome of (cost free) female choice on a polygenic male character, for which there is an optimum size under natural selection, is that the population moves to some arbitrary point on a line of 'neutral equilibria' for female preference and male trait (Lande, 1981). Provided that the male optimum is the same in all areas, gene flow ensures that the same equilibrium is reached in all populations. In the absence of sexual selection, a step in the male optimum causes a cline to be established in response to migration. In this situation, sexual selection exaggerates the difference across the cline and may lead to rapid divergence (Lande, 1982).

1.5 Thesis overview.

Three separate pieces of work are presented in this thesis. Two studies are relevant to the hybrid zone in *Podisma pedestris*, whilst the third is a consideration of one of the possible courses of evolution in hybrid zones.

Chapter 3 considers the possibility of, and conditions for, an adaptive increase in post-zygotic isolation in hybrid zones. This is an extension of the more familiar question of adaptive increase in pre-zygotic isolation - speciation by reinforcement. In order to set the present work in its proper context, a review of reinforcement is given in Chapter 2. Chapters 5 and 6 concern the hybrid zone in the alpine grasshopper *Podisma pedestris*. The current understanding of this hybrid zone is summarised in Chapter 4. Chapter 5 considers the effect of observed patterns of non-random fertilization in the maintenance of the chromosomal cline between the two races. Chapter 6 is a consideration of morphometric variation across the zone.
Chapter 2

Reinforcement of post-zygotic isolation

2.1 Speciation by reinforcement

Consider two, previously isolated, races that have diverged such that on secondary contact, hybrids between them show reduced fitness. Consider also that in areas of sympatry there is complete random mating between them (i.e. divergence has affected post- but not pre-zygotic isolation). Those individuals that happen to mate with members of their own race will have higher fitness than those that mate with members of the other race, since the fitness of their progeny will be greater. It seems obvious that selection should favour genes that promote assortative mating. In this way natural selection could increase the barriers to gene flow between divergent populations, perhaps eventually leading to complete isolation and so to new species.

The idea that pre-mating isolation will evolve in response to hybridization is commonly attributed to A.R. Wallace (1889) and as such has been termed the 'Wallace effect', (Grant, 1966). Although considered by Fisher (1930), the idea became widely known after promotion by Dobzhansky (1940, 1951). Blair gave it the name "reinforcement" (Blair, 1955) - post-zygotic isolation is reinforced by the evolution of pre-zygotic isolation.

1 It has been pointed out that this may be an unjustified attribution (Littlejohn, 1981; Cronin, 1991). Wallace was primarily concerned with the evolution of hybrid sterility and believed that it was an adaptive response to hybridization. See the introduction to Chapter 3 for further discussion of this point.
Although reinforcement has been widely accepted as a mode of speciation, the reason for this is perhaps more to do with Dobzhansky's advocacy of the idea and a general desire to give natural selection a direct role in speciation, than the weight of evidence or the plausibility of the model. Another possible reason that the hypothesis has been accepted uncritically lies in what Paterson (1982) identified as a problem with the isolation species concept. By defining a species in terms of the means by which it is isolated from other species, one is easily misled into thinking that those mechanisms must have evolved for that purpose. For example, Dobzhansky (1940) states that "the very fact that isolating mechanisms are as diversified as they are is strong evidence for the prevention of interbreeding being an essential characteristic of the process of speciation." There are many reasons why characters leading to reproductive isolation may become diversified, direct selection for isolation is just one of them. Dobzhansky (1940) describes hybridization as being a "challenge" to the species that may be overcome by adaptation. This is certainly the case. However, much of the recent theoretical work (discussed below) has shown that natural selection is not always able to meet this challenge and is constantly hampered by other forces.

Before discussing in detail the reinforcement hypothesis one must consider more precisely what it represents. A distinction has been made between the evolution of premating isolation between taxa that are completely isolated, and those that are only partially isolated by post-mating barriers. The former situation is an interaction between what are already species (under the ISC) and so cannot be a process of speciation. Butlin (1987a, 1987b, 1989) has suggested that the term reinforcement should be reserved for the latter situation since it may be a process of speciation. The former situation should, he suggests, be referred to as reproductive character displacement, since it is exactly analogous to the displacement of ecological characters in response to competition between species (Brown & Wilson, 1956). This distinction is not universally accepted (for example Spencer et al., 1986, 1987). The objections of Spencer et al. are two-fold. First, there is a disagreement on the definition of species. If one adheres to the recognition concept of species, then what
Butlin defines as reproductive character displacement is a process of speciation. The second is that it is irrelevant to distinguish the two since both are a response to the same basic selective pressure; the only difference is that selection is stronger when post-zygotic isolation is absolute. Although they are both responses to the same selection pressure, only reinforcement (Butlin's definition) leads to an increase in reproductive isolation. Since I take the isolation concept to be my working definition of species, I will stick to Butlin's classification.

Reproductive character displacement is an inherently more likely process since one of the forces acting against the modification of isolation cannot apply. Namely, introgression of assortative mating alleles from one taxa to the other cannot take place. This point is discussed more fully below.

There are several situations in which reinforcement may be possible: in a cline maintained by environmental variation; in a tension zone; or in areas where there are broad areas of 'overlap with hybridization'. The first possibility is essentially the same as the second. The shape of an environmentally determined cline is similar to that of a tension zone (Barton & Hewitt, 1989). Further, it is likely that environmentally determined clines evolve into tension zones since adaptation to different environmental conditions is likely to also lead to genetic incompatibilities between populations (Barton & Hewitt, 1985). Reinforcement is also implicated in models of sympatric speciation, in which disruptive selection (e.g. for host specialization) is reinforced by the evolution of reproductive isolation (Maynard Smith, 1966a). However, reinforcement is not a necessary step in models of sympatric speciation (Tauber & Tauber, 1989; Diehl & Bush, 1989). I will concentrate on reinforcement in tension zones.
2.2 Difficulties on theory

Several theoretical studies have shown that an adaptive increase of post-mating isolation is at least possible (Crosby, 1970; Endler, 1977; Caisse & Antonovics, 1978; Felsenstein, 1981; Spencer et al., 1986; Sanderson, 1989). However, all have shown that the conditions required for the reinforcement are restrictive. Four specific objections to the reinforcement hypothesis have been raised. These are discussed below.

1) It is not necessarily true that there is a large fitness premium to be gained by assortative mating (Butlin, 1989). Although there are several cases in which it has been shown the the low fitness of interspecific hybrids is largely due to major effects of single genes (e.g. Orr, 1992; Pantazidis & Zouros, 1988; Coyne & Charlesworth, 1986; or as might be expected with chromosomal rearrangements, White, 1973), it is most frequently found that a medium to large number of genes are involved (Coyne, 1992). Studies of hybrid zones have shown that a large number of genes with small effect are responsible for the reduced fitness of hybrids. For example, in the hybrid zones between the two chromosomal races of the grasshopper Podisima pedestris and between the 'species' of toads Bombina bombina and B. variegata, selection against hybrids is strong (~ 50%). The number of loci contributing to this selection is large: ~ 150 in Podisma (Barton & Hewitt, 1981a, 1981b) and ~ 300 in Bombina (Szymura & Barton, 1986, 1991) but selection on each is weak.

In the absence of linkage disequilibria between these selected genes there can be little advantage to be gained through mate choice in the centre of the hybrid zone. In this situation the selected genes will be segregating independently of each other and so there simply are not any pure race individuals to choose. Although there is a large amount of variation in fitness across a tension zone variation within any one population is limited.

However, linkage disequilibrium between selected genes in a hybrid zone is maintained by the diffusion of pure race genotypes into the zone (Barton, 1982, 1983). If overall selection against hybrids is strong, significant amounts of disequilibrium will be
maintained in the centre of the zone. In this situation, the net selection on each allele is increased since they do not segregate independently. In the centre of a tension zone produced by selection at many loci the advantage to be gained by mate choice is, therefore, larger than suggested by the selection on individual genes but less than that predicted by the reduced fitness of F1 hybrids.

If the hybridizing races have different habitat requirements that are patchily distributed a mosaic hybrid zone will result (Harrison & Rand, 1989). The properties of such a zone depend on both the selection against hybrids and distribution of patch types. This patchy structure means that a population of essentially 'pure race' individuals may be found in the centre of a hybrid zone. The opportunity for interaction between pure race individuals is therefore greater than in a conventional hybrid zone. The chance of reinforcement in a mosaic hybrid zone is also increased since the edges of each patch are in effect mini hybrid zones. The entire zone therefore constitutes many 'trials' in which reproductive isolation may evolve (Harrison, 1990; Harrison & Rand, 1989).

2) Paterson (1978, 1982) has argued that reinforcement is an unlikely outcome of hybridization since strong selection against hybrids in the absence of pre-mating isolation leads to an unstable population, analogous to the dynamics of underdominance at a single locus (Hartl & Clark, 1989). Sampling drift would mean that the unstable equilibrium is never actually maintained. Paterson has argued that a more likely outcome is that one race will eliminate the other. What is not taken into account is that such a situation is readily stabilised by a balance between migration and selection (as evident in the many examples of hybrid zones, Barton & Hewitt, 1985; Hewitt, 1988). A potentially more serious problem is that if one race has higher fitness than the other, one would expect the hybrid zone between them to be constantly retreating into the range of the less fit (Bazykin, 1969). Again, one would expect that one or other of the races would eventually be eliminated. This scenario is based on a naive view of population structure. If population density is completely uniform throughout the species range, the race with the higher fitness will always eliminate the other. Real populations are not so simple. Variation in population
density causes tensions zones to move. In the absence of fitness differences between races they will always come to rest in areas of low population density. A deep trough in population density (for example, an area of very inhospitable habitat) will hold back a moving hybrid zone for a long period of time during which reinforcement may occur (Barton, 1979b). In the case of a hybrid zone tied to environmental gradient, neither of these two arguments apply.

3) Areas of hybridization usually only represent a small proportion of a species range. Most evolution of the species therefore takes place in the absence of hybridization. Mutations that alter the mating scheme are more likely to arise outside areas of hybridization and diffuse into it than vice versa. However, the selection pressure for modification is only present in areas of hybridization. The selection pressure for modification increases with the strength of selection against hybrids. Large areas of hybridization (in which reinforcement may be possible) are only found if selection against hybrids is weak, in which case there is little selection for modification of mating systems.

Several authors have pointed out that an alteration of the mating scheme is likely to be selected against outside areas of hybridization (Moore, 1957; Paterson, 1982; Barton & Hewitt, 1985). This is because mating systems are necessarily under stabilizing/frequency dependent selection: they consist of co-adapted sets of signals. A rare signal will be recognised by only a small proportion of the population and so will suffer a frequency dependent disadvantage. Also, it may be that there is a general selection pressure against the elaboration of the mating system. An elaborate system of signals and responses is likely to take longer to complete than a simple one. 'Choosy' individuals may therefore take longer to find a mate or may be more prone to predation in the process (e.g. Tuttle & Ryan, 1981). In areas of hybridization the extra costs of a change in the mating system may be outweighed by the advantage of producing only pure race progeny.

Obviously, stabilizing selection on the mating system is not an insurmountable hurdle to evolution - mating systems do evolve. However, in order to evolve they must
overcome this stabilizing selection. This may be brought about by chance (drift), runaway sexual selection, or because the new mating system is actually adaptive. In the context of reinforcement, a change in the mating system is only adaptive in areas of hybridization. Areas outside the hybrid zone will act as a source of wild type alleles. Migration into the zone may, therefore, swamp modifier alleles that are favoured there. Likewise, the areas outside the zone represent a 'sink' for alleles altering the mating system. An allele favoured in the zone will be selected against once it is carried out. Sanderson (1989) has shown that only small levels of selection against the modification of mating behaviour outside a hybrid zone are sufficient to prevent it increasing within the zone. In particular, selection against reinforcement must be small relative to the selection against hybrids.

Dobzhansky (1940, 1951) supposed that reinforcement would work by the sequential fixation of genes that increase assortative mating. However, as each allele is fixed, the chance of further reinforcement is reduced since the effective selection against hybrids is reduced. Also, as alleles are fixed, the amount of selection against modification outside the hybrid zone that will prevent reinforcement becomes smaller. Thus, if dependent on the combined effect of many genes of small effect, full reproductive isolation can only be approached asymptotically.

4) Recombination can seriously impede reinforcement. Essentially, the reinforcement hypothesis is that there is strong selection for assortative mating if there is strong selection against hybrids. Assortative mating for a particular racial character alone is not enough to promote speciation. If selection against hybrids is absolute (i.e. the hybridizing taxa are already "good species") then any gene which promoted assortative mating may be favoured within a hybrid zone. However, if selection against hybrids is not absolute (which is by definition required for reinforcement to occur), the crucial point is the maintenance of linkage disequilibrium between alleles promoting assortative mating and those contributing to hybrid inferiority, rather than simply the promotion of assortative mating for some racial character. Recombination will continually erode the association between assortative mating alleles and those causing hybrid inviability (Felsenstein, 1981). If linkage between loci
promoting assortment and selected ones is weak, alleles will rapidly introgress from one race to another. Once this happens, the cue for assortative mating looses its value as an indicator of the suitability of a potential mate. Eventually, the two loci should reach linkage equilibrium and there will be no benefit in assortative mating.

Felsenstein (1981) has made a distinction between one and two allele models of reinforcement. Reinforcement based on two alleles promoting assortative mating will be easily halted by recombination. On the other hand if the same allele promotes assortative mating in both races, recombination is not a problem. A plausible one allele system that causes assortative mating in two races is given by Rice (1984). He suggested that in phytophagous insects, a single gene that caused imprinting on the host plant on which an insect developed, would promote assortative mating between two host races. In general, sympatric divergence will be more likely if mating and selection are closely connected (Diehl & Bush, 1989).

In summary, the arguments presented above indicate that reinforcement is only possible if there is strong selection against hybrids, only weak stabilizing selection on the mating system, and close linkage between the alleles that promote assortment and those that contribute to reduced hybrid fitness.

Arguments 2 and 3 above apply equally to models of reinforcement and to those of reproductive character displacement. Arguments 1 and 4 apply only to reinforcement. If there is no gene flow between taxa, there can be no recombination. The problem of recombination is one of the biggest obstacles to reinforcement. Since this is not applicable to reproductive character displacement, hybridization presents a very different 'challenge' when it is between species, to when it is between only partially reproductively isolated populations. The retention of the difference in terminology when referring to these two processes reminds one of this difference.

The discussion of reinforcement addresses only the question of whether divergence in mating patterns are directly favoured by natural selection in order to prevent
Chapter 2

hybridization. It does not address the question of whether speciation occurs through divergence in reproductive characters by means other than as a pleitropic effect of adaptation. The mating systems of isolated populations may diverge through sexual selection in such a way that on secondary contact they are reproductively isolated. Such divergence may occur because of positive feedback between female mating preferences and male traits, i.e. Fisher's "runaway" process, (Fisher, 1930). Lande (1981) and Kirkpatrick (1982) have shown that a system of female preferences and male traits is inherently unstable and that the outcome of sexual selection is to a degree arbitrary. Lande (1982) extended his analysis to consider sexual selection in an environmentally determined cline. He showed that the runaway process is still possible in the presence of gene flow. If the optimum male phenotype is different on either side of a cline, it is likely to trigger a bout of runaway sexual selection in different directions on each side. In this model the presence of the cline promotes divergence but it is not through a process of reinforcement.

2.3 Empirical evidence for reinforcement

The theoretical conclusion that reinforcement is unlikely to be a major cause of speciation is backed up by a paucity of clear examples from the field (for example, Barton and Hewitt (1985) could find only three clear examples in a survey of over 150 tension zones). Grant (1966) suggests four sources of data that may provide evidence for reinforcement: 1) artificial selection experiments; 2) direct observation of temporal changes in isolating mechanisms; 3) geographic variation in isolating mechanisms within species, and 4) a comparison of pre-zygotic isolation between sympatric and allopatric species pairs.

Although attempts to increase pre-zygotic isolation by artificial selection have been successful (e.g. Koopman, 1950; Paterniani, 1969; Koepfer, 1987), they can only be taken as very weak evidence that natural selection has done the same in the wild. All that these experiments show is that the requisite genetic variation exists; whether natural selection acts upon it is an entirely different matter. Slightly more convincing is the
experiment of Thoday and Gibson (1962), in which assortative mating evolved after only 12 generations of disruptive selection on sternopleural bristle number in *Drosophila melanogaster*.

Perhaps the most convincing evidence for reinforcement would be direct observation of a decrease in hybridization over time in sympatric populations. Such data are obviously hard to come by. Jones (1973) reports that, at a site in an area of sympatry between the toads *Bufo americanus* and *B. woodhousii fowleri* no hybrids were found, whereas ≈8% of the population at the same site were hybrids thirty years previously (Blair, 1941). Similarly, in the centre of a hybrid zone between the orioles *Icterus galbula* and *I. bullockii*, which differ in plumage characters, a bimodal distribution of these characters was observed, whereas twenty years earlier a continuous distribution had been found (Corbin & Sibley, 1977). This change they attribute to the evolution of assortative mating. Both of these examples are flawed in that they rely on a comparison of only two data points. Furthermore, both can be explained by alternative processes. The durations of the breeding seasons of the toad species are temperature dependent: Loftus-Hills (1975) points out, in addition to doubts over the correct identification of hybrids, that only a few years of unusual weather conditions prior to the survey could mean that there was no opportunity for hybridization. In the case of the orioles, migration into the zone by pure-race individuals would produce a bimodal distribution (Butlin, 1987a). In addition, a second survey failed to show a bimodal distribution or any evidence of assortative mating (Rising, 1983).

The most frequently cited evidence for reinforcement is an increased difference in mating characteristics between divergent populations in areas of sympatry. This pattern has been shown for advertisement song characteristics in the frogs *Gasterophryne olivacea* and *G. carolinensis* (Blair, 1955; Loftus-Hills & Littlejohn, 1992); *Litoria ewingi* and *L. parawingi* (Littlejohn & Watson, 1983) and in flowering time in a hybrid zone between green and blue forms of the bishop pine *Pinus muricata* (Millar, 1983). More direct tests involve comparing mating preferences (in laboratory conditions) of individuals from
sympatric populations to those of individuals from allopatric populations brought together. Such studies have shown that pre-mating isolation between *Drosophila mojavensis* and *D. arizonensis* is greater when individuals are taken from sympatric rather than allopatric populations (Wasserman & Koepfer, 1977). Assortative mating in the field for elytral spot size in a wide cline of this character in the soldier beetle *Chauliognathus pennsylvanicus* is greatest in populations from the centre of the cline (McLain, 1985). However, in a subsequent study, no consistent pattern of assortment could be found in the hybrid zone (McLain, 1988).

A fundamental problem with the analysis of geographic variation in reproductive characters is that the expected pattern of increased divergence in sympathy is not necessarily predicted by the reinforcement hypothesis. Unless the change to the mating system is selected against outside the zone of hybridization, the pattern of increased divergence in the zone will be transient: the initial divergence in the centre is expected to spread into the pure populations (Caisse & Antonovics, 1978). If reinforcement is rapid, the pattern of geographic variation in reproductive characters left behind will be indistinguishable from that left behind by divergence in allopatry or parapatry.

A second problem is that looking for divergence in mating characters presupposes that one knows in advance which characters have responded to selection for reinforcement. For example, song characteristics are often used in such analyses whereas other cues, such as pheromonal or visual signals may be more important. Ritchie *et al.* (1992) have suggested an alternative approach that comes closer to testing the precise predictions of the reinforcement hypothesis. This predicts that there is a fitness premium to be gained by non-random mating. One may not know the cues used but by pairing individuals at random, choice can be prevented. A comparison of the fitness of progeny from random and non-randomly mated parents should reveal any benefit to non-random mating. Since it is likely that there are benefits to non-random mating even in non-hybridizing populations (e.g. Partridge, 1980), a comparison between hybrid and non-hybrid populations must be made. When this methodology was applied in the hybrid zone
between the grasshoppers *Chorthippus parallelus parallelus* and *C. p. erythropus*, no increased benefit to non-random mating in hybrid populations was discovered (Ritchie *et al.*, 1992). Even if such an effect had been found, there would be a problem in interpretation. Hybridization and recombination between pure race genotypes increases the variation in fitness. There is, therefore, a greater opportunity for female choice of 'good genes' which may have nothing to with the avoidance of hybridization.

A fourth method described by Grant (1966) is a comparison of the strength of isolation between sympatric and allopatric species pairs. If reinforcement has been responsible for increased pre-zygotic isolation, one would expect greater isolation between sympatric than between allopatric species pairs within a genera. A problem with this method is that the sample of sympatric species pairs is not random. The sample can only include those species pairs that have coexisted from first coming into contact until the present. Species pairs that are strongly isolated are more likely to coexist and so bias the sample in favour of reinforcement (Templeton, 1981; Butlin, 1987a). This bias may be overcome by comparing the rate of increase of isolation in sympatric and allopatric species pairs. In an analysis of 119 pairs of *Drosophila* species for which estimates of genetic distance and data on pre- and post-zygotic isolation in laboratory conditions was available, it has been shown that the rate of increase of pre and post-zygotic isolation was comparable for allopatric pairs (Coyne & Orr, 1989). However, for sympatric species pairs, the rate of increase was greater for pre- than for post-zygotic isolation.

### 2.4 Variations on a theme: Responses to hybridization other than mate choice.

Judging by the discussion presented above one might conclude that an increase in pre-zygotic isolation is the only possible adaptive response to hybridization. Perhaps the easiest response to imagine is that genes that ameliorate selection against hybrids would be favoured. If, for example, hybrid sterility were the result of incompatibility at a single
locus, a back mutation in one of the hybridizing populations would be favoured by selection. Alternatively, modifier alleles that ameliorate selection might be favoured. Several studies have revealed natural variation for genes that rescue the inviable hybrids between Drosophila species (Watanabe, 1979; Hutter & Ashburner, 1987; Pantazidis & Zouros, 1988). However, these examples of 'hybrid rescue' do not actually increase the fitness of hybrids since they remain sterile, and there is no evidence that they are the result of selection for increased hybrid fitness.

The races of the common shrew Sorex araneus differ in a number of robertsonian fusions. In the hybrid zone between these karyotypic races there is an increased frequency of acrocentric chromosomes relative to either of the pure races (Searle, 1986). It is suggested that this increased frequency is as a result of selection to increase the fertility of hybrids by reducing the frequency complex heterozygotes involving different metacentric chromosomes which suffer reduced fertility through non-disjunction (Searle, 1986, Hatfield et al., 1992).

F1 hybrids between the grasshoppers Chorthippus parallelus and C.p.erythropus have very low fertility, due in part at least to dysfunctional testes (Hewitt et al., 1987). In the centre of the hybrid zone between these taxa Ritche et al. (1992) found no evidence of testes dysfunction. They suggest that a likely explanation for this result is that alleles that reduce testes dysfunction have been favoured by selection. It is ironic that this discovery was made during the course of an investigation looking for reinforcement. An alternative explanation may be that testis dysfunction is due only to heterozygous disadvantage at several loci. If this is the case F1 hybrids will have lower fitness than subsequent generations since in these some of the loci will be homozygous.

Modifiers that increase the fitness of hybrids but that have no effect on the fitness of non-hybrids could be easily favoured by selection in a hybrid zone. It is possible that such alleles would have a negative effect on fitness in non-hybrids. Sanderson (1989) has
shown that the conditions for the spread of alleles that enhance hybrid fitness are the same as those that reduce the amount of hybridization that takes place.

In the context of hybrid zones, another possible response to hybridization is modification of migration rates. Reduced migration would mean that individuals are less likely to pair with an 'inappropriate' mate (in the context of tension zones) or land up in an environment to which they are not adapted (in the context of environmentally determined clines). In simple models of structured populations with spatially varying selection pressures, genes that reduce migration will always be favoured (Balkau & Feldman, 1973), though these models included no cost to reduced migration which may arise from increased inbreeding and competition (Bengtsson, 1978).

Both amelioration of selection and migration modification can fall into the one allele class of modifiers (Felsenstein, 1981). If this is the case, then introgression of alleles from one race to another does not impede modification. Theory predicts that these two possibilities are at least as likely a response to hybridization as is reinforcement. The actual outcome of evolution will depend on the availability of genetic variation.

Reinforcement through any of the means discussed above leads to a net reduction of selection against hybrids (either directly or because fewer hybrids are produced). There are certain circumstances in which it may be possible for natural selection to actually decrease the fitness of hybrids between species. This may come about essentially through a process of kin selection (Maynard Smith, 1966b; Coyne, 1974). Inviability may be favoured if the early death of hybrids increases the fitness of non-hybrid siblings. This possibility is considered in detail in the following chapter.
Chapter 3

'Reinforcement' through increased post-zygotic reproductive isolation

3.1 Introduction

Relatively little attention has been given to possible modes of evolution of post-zygotic reproductive isolation, other than as a by-product of the divergence of populations. In particular, the possibility that post-zygotic isolation may be adaptive has been largely ignored. Indeed, it has been claimed that it is impossible for natural selection to improve post-zygotic reproductive isolation (Muller, 1942; Mecham, 1961; Grant, 1963; Littlejohn, 1981). An increase in the post-zygotic isolation between species can only come about by a decrease in fitness of hybrid individuals. Since adaptation implies an increase in fitness it seems that increased hybrid inviability or sterility cannot be adaptive. In this chapter a model of speciation by reinforcement through increased post-zygotic reproductive isolation is considered.

The idea that natural selection might play a direct role in the evolution of post-zygotic isolation was first introduced by Wallace (1889). He suggested that, given that there had been some initial divergence in allopatry such that the hybrids between to incipient species had reduced fertility, natural selection would favour those hybridizing populations which had the lowest fertility. The logic behind this argument is a classic example of group selection: sterility would be favoured since those hybridizing populations which showed hybrid sterility would be able to speciate faster and so out compete those that did not. The advantage gained through hybrid inferiority is to the population as a whole. It has been clearly shown that group selection of this sort is at most a weak
evolutionary force (Maynard Smith, 1976) since it will always be opposed by individual selection. It is therefore unlikely that Wallace's model is workable.

Wallace tried, unsuccessfully, to convince Darwin that hybrid sterility was adaptive. Though ultimately unconvinced, Darwin was obviously troubled by the 'sterility problem'. Darwin considered that the origin of hybrid sterility was simply an incidental consequence of differential adaptation of species. Darwin did not consider the possibility that hybrid sterility may be adaptive until prompted to do so by Wallace (Cronin, 1991). He did not equate the origin of sterility with speciation. His interest in sterility was as material in the argument against special creation. Darwin noted that the association between sterility and species status was not a rigid one: there is much variation in the degree of sterility between species pairs. If interspecific sterility had been granted by a creator in order to maintain the integrity of His creations how was it that not all species pairs were completely sterile? (while some 'varieties' were). It was not until the 4th edition of the Origin that the question 'is sterility adaptive?' arose. The conclusion he reaches is a firm no, based on three lines of reasoning. First, geographically isolated species pairs may be sterile when crossed. If these species pairs have never been in contact there will have been no opportunity for selection to increase sterility. Second, Darwin argued that the pattern of sterility between species was too unsystematic, and in cases too incomplete to have been the result of natural selection. This in itself is a weak argument and a surprising one for him to have made. It was, after all, Darwin who argued that the inconsistencies of nature and the imperfect adaptation of organisms revealed the action of natural selection. The third argument was that individual selection would always oppose a decrease in fitness, even if such a decrease would benefit the species as a whole. It is this argument that has led modern workers to claim that post-mating isolation mechanisms cannot be improved by

---

1 Darwin to Wallace: 'I do not think that I will grapple with the sterility argument till my return home; I have tried once or twice and it has made my stomach feel as if it had been placed in a vice.' For a commentary on the Darwin/Wallace debate on hybrid sterility see Cronin 1991.
selection. The Darwin/Wallace dialogue concerned only hybrid sterility. They did not focus their attention on hybrid inviability. Perhaps if they had done so they may have anticipated one of the major developments in biology of this century.

In the early 1960's W.D. Hamilton initiated something of a revolution in evolutionary biology (Hamilton, 1963, 1964a, 1964b). Fisher (1930) and Haldane (1955) had previously noted how apparently mal-adaptive behaviour could be favoured by selection if the altruism was directed towards close kin. Hamilton formalised and greatly extended this idea into his theory of inclusive fitness (the exact definition of this quantity and its application is discussed in section 3.4). The evolution of altruism is possible since closely related kin are likely to also carry the genes for the altruistic act. Evolution proceeds through a process of 'kin selection'.

Hamilton's theory of kin selection has been very successful in helping to understand a diverse set of phenomena such as: sterile casts in social hymenoptera (Hamilton, 1964b); cooperative breeding (Emlen, 1991); variation in response to parasitism (McAllister & Roitberg, 1987); and many others. Most of the applications of the theory of kin selection have concerned the evolution of behavioural interactions within groups of related individuals. However, it can equally well be applied to the evolution of non-behavioural traits such as distastefulness as an anti-predator strategy and aposomatic colouration (e.g. Harvey & Greenwood, 1978).

It was first explicitly suggested by Maynard Smith (1966b) that an increase in postzygotic reproductive isolation could be favoured by natural selection through a process of kin selection\(^1\). He considered a situation in which two races interbreed producing viable but sterile hybrids. If there is multiple mating then individual families will consist of a mixture of hybrid and pure race individuals. If the members of a single family are in

---

\(^1\) Ideas essentially the same as those of Coyne and Maynard Smith were previously expressed in an unpublished paper by W.D. Hamilton in the early 1950's (Cronin 1991).
competition with each other for limited resources (for example parental investment) then a hybrid (whose fitness is otherwise zero) can increase its fitness by reducing its viability. By reducing its viability it will reduce its consumption of resources which will be made available for its pure race sibs. In the extreme a form of 'adaptive suicide' may result.

The same suggestion was also made independently by Coyne (1974). Coyne emphasised that this situation is essentially the same as one in which an individual parent may increase its fitness by diverting resources away from sterile hybrid progeny and to fertile pure-race progeny. There is a conflict of interest between parent and offspring. Genes which cause a parent to sacrifice hybrid progeny may be favoured if the chance of producing a second, non-hybrid offspring is big enough. If the fitness of hybrids is zero, then such genes should always be favoured since there is then no cost involved in infanticide. Likewise, genes that cause sterile hybrids to be inviable may be favoured if the death of hybrids occurs before parental investment in progeny is complete, so allowing investment to be directed towards non-hybrid progeny.

Through the mechanism of kin selection Coyne and Maynard Smith have indicated that there may indeed be a direct role for natural selection in the origin of post-mating reproductive isolation. Given the enthusiasm with which similar ideas concerning the evolution of pre-zygotic reproductive isolation were received it is perhaps surprising that this hypothesis has received little attention. The reason for this may lie in the rather restrictive circumstances in which this process may operate, identified by both Maynard Smith and Coyne. Both these authors point out that the crucial requirements of the theory are that there is multiple mating and sib-competition. They pointed out that these conditions are most often met in flowering plants: the process of pollination often produces multiple paternity; and the limited dispersal of seeds implies that progeny are often in competition with their sibs and their parent (either through limited dispersal or directly in competition for provisions from the parent plant).
A general requirement of this model is that selection within sib-groups is 'soft' (Wallace, 1975). A gene that reduced the viability of hybrids could only be favoured if the viability reduction is only manifest when in competition with other individuals. An allele, unconditionally lethal when in a hybrid individual, could not be favoured since although it may increase the average fitness of a family (by eliminating partially sterile hybrids) it would decrease the total fitness by reducing the family size. An example of soft selection would be a situation in which a parent could successfully produce only a fixed number of fledglings, irrespective of the number initially laid and their fitness.

A general prediction of this theory is that, when the requisite conditions are met one might expect to observe that hybrid sterility promotes the evolution of hybrid inviability. That is, natural selection may shift the barrier to gene flow between incipient species from late to early in the life cycle. This is exactly analogous to the predictions of the classic reinforcement hypothesis: a shift in the barrier from post- to pre-zygotic reproductive isolation. The hypothesis may even be extended to include hybrid sterility. It is at least possible that the same processes could lead to reduction of hybrid fertility. If the barrier to gene flow between species is not manifest until the F$_2$ generation and the hybrid and non-hybrid 'grandchildren' of a single parent are in competition, then selection may favour hybrid sterility.

### 3.1.1 Empirical evidence for an adaptive decrease in hybrid fitness

Is there any evidence for adaptive increases in post-zygotic isolation? To determine if reinforcement has increased post-mating isolation the same methods can be applied as for the classical reinforcement hypothesis. There have been many studies investigating possible cases of reinforcement and reproductive character displacement (see Chapter 2) but fewer focusing on variation in post-zygotic barriers. Perhaps part of the reason for this (apart from the excessive popularity of theories of reinforcement) is that it is much easier to
identify divergence of pre-zygotic isolation than of post-zygotic isolation. The former can be achieved by careful measurement of mating signals whereas the latter requires measurements of fitness. Only three studies have shown any effect of hybridization on increased post-zygotic isolation.

The best documented case is that of the 'corkey' alleles in cotton species *Gossypium hirsutum* and *G. barbadense* (Stephens, 1946, 1950). Hybridization between these species normally produces luxuriant but partially sterile progeny. There is also hybrid breakdown in the F₂'s and backcrosses. In some areas of sympathy, hybridization produces stunted infertile progeny - the 'corky' syndrome. This is due to an interaction between alleles at a single locus, ckX (in *G. hirsutum*) and ckY (in *G. barbadense*). These alleles are found at high frequencies only in areas of sympathy and so seem likely candidates for adaptive responses to hybridization.

The ranges of the tree frogs *Hyla ewingi* and *H. vereauxi* overlap in south eastern Australia. In a comparison of sympatric and allopatric interspecific crosses, Watson & Martin (1968) found that the frequency of developmental abnormalities was greatest in the progeny of sympatric crosses. This was interpreted as a pleiotropic effect of ecological character displacement in these populations. The observation is also however consistent with a selected increase in inviability. Coyne (1974) is sceptical that this can be the case since there is no evidence of parental care in these species. However, as pointed out above, parental care is not required by the hypothesis; only that there is soft selection between sibs which is perfectly plausible for tadpoles.

Interspecific crosses between plant species of the *Gilla* family produce sterile hybrids. Grant (1966) investigated the ease of hybridization (after artificial pollination) between these species and found that greater numbers of mature seeds are produced in crosses between sympatric species than are produced between allopatric ones. The barriers to hybridization found in these crosses may be acting either pre-zygotically (e.g. pollen tube growth) or after fertilization but there is evidence in at least some crosses that the
barrier is partially post-zygotic. It is suggested that since the reproductive success of the female parent plant is limited more by investment in provisioning seeds than in the production of ovules, increased post zygotic barriers to hybridization may have been favoured by selection.

In the three examples given above there is no evidence that divergence of isolating mechanisms has taken place in the presence of gene flow between the species. They should therefore be considered as examples of character displacement rather than reinforcement.

Coyne & Orr (1989) found that among *Drosophila* species pairs hybrid inviability and hybrid sterility evolved at the same rate. If hybrid sterility were reinforced by inviability then one would expect inviability to evolve faster in sympatric populations. There is no evidence that this is the case in this set of species. However, it would be interesting to compare the rate of accumulation of hybrid inviability between allopatric and sympatric species pairs with a particular degree of hybrid sterility - for example complete sterility. This comparison is not given in this paper, perhaps because controlling for the degree of sterility would reduce the sample to just a few species pairs.

3.1.2 Theory

Maynard Smith (1966b) and Coyne (1974) presented only verbal models for their theory of reinforcement. McNair (1987) considered a simulation model in which hybrid sterility is caused by unspecified genetic differences between the hybridizing species and juvenile inviability by an interaction between different genes in the two species. He showed that provided hybrids were completely sterile, the inviability alleles will always increase in frequency. If sterility is not absolute, introgression of the inviability alleles from one species to the other may halt their spread. Wallace (1988) considered, in outline, a model similar to that of McNair but also considered one in which a single, dominant allele causes the early death of hybrids.
The same distinction between character displacement (a result of competition between species and, therefore, in the absence of gene flow) and reinforcement (divergence in the presence of gene flow) may be applied to the modification of post-zygotic reproductive isolation. The models of McNair and Wallace are primarily of reproductive character displacement.

3.1.3 Aims

The aim of this chapter is to construct and analyse an analytic model of speciation by reinforcement through increased post-mating isolation. Several factors will be considered: the initial selection against hybrids; the degree of sib-competition; and variation in the mating system. The key question in all the analysis is 'When can genes which decrease the fitness of hybrids be favoured by natural selection?'. The models differ from those of McNair and Wallace in that analytic solutions are reached which relate the conditions for an adaptive increase in post-zygotic isolation to the degree of sib-competition and the strength of selection against hybrids (i.e. they are of reinforcement rather than reproductive character displacement).

A two locus model of hybrid inviability and modification is considered. The model is analysed in two different ways. First an attempt is made to solve the problem by the application of standard population genetic techniques, i.e. to track changes in the frequency of all possible genotypes through the life cycle and to determine the change in frequency of the modifier allele. This turns out to be a rather inefficient approach. Meaningful expressions can only be gained under very restricted circumstances.

The second approach is to treat the problem as a specific case of the evolution of altruism. In doing so, the standard methods of behavioural ecology are applied: a consideration of the difference in inclusive fitness between altruistic and non altruistic individuals. This methodology is discussed more fully in section 3.4 below. This approach
turns out to be far more flexible. In particular, the simple model is easily extended to investigate different mating schemes and different levels of sib competition.

The models are only analysed in detail for single populations. A brief consideration is given to a two deme model. The effects of gene flow on modification in a hybrid zone are discussed.
3.2 A model of reinforcement

In order to analyse the reinforcement hypotheses described above four factors must be considered: The genetic basis of selection against hybrids; the genetic basis of modification of this selection; the population structure; and the mating patterns. The model analysed in the following sections is described below.

3.2.1 Hybrid Sterility

Hybrid sterility is considered to be caused by heterozygote disadvantage at a single locus, (referred to below as the primary locus). The three genotypes at this locus, A_1A_1, A_1A_2, A_2A_2, have relative fitness: 1, 1-s, and 1 respectively. The frequency of the A_1 allele is p_a, and that of A_2 is q_a (p_a+q_a = 1). The genotype frequencies at this locus are u, v and w. The dynamics of system of under-dominance at a single locus such as this are well understood (Hartl & Clark, 1989). There are no stable, polymorphic equilibria. The only equilibria are at p_a=0,1 or \( \frac{1}{2} \). Selection causes the most common allele to become fixed. Stable, polymorphisms may be maintained by a balance between selection and migration, as is frequently observed in hybrid zones.

In nature hybrid sterility is most often found to be a result of interactions at more than one locus (Coyne, 1992). The simple one locus model is used here largely for the sake of simplicity and will be applicable if genes of major effect or chromosomal rearrangements are involved. Such a system of under dominance of fitness is only exactly equivalent to reduced fitness on hybridization through epistasis when selection is absolute (s=1). The differences that this may make to the conclusions drawn are considered in the discussion.
3.2.2 Early inviability

Two alleles, at a second locus are considered. The wild type allele ($B_1$, frequency $p_b$) has no effect. The modifier allele, ($B_2$, frequency $q_b$) interacts with the primary locus. When occurring in an individual that is homozygous at the primary locus it has no effect. In a heterozygous individual it reduces juvenile viability by an amount $a$ (Table 3.1). The modifier and primary loci are un-linked.

When both loci are considered, there are ten different possible genotypes, (coupling and repulsion double heterozygotes are considered as separate classes). For the two locus analysis (section 3.3) the frequencies of these these ten genotypes are tracked throughout the life cycle. For the inclusive fitness analysis, only the genotype frequencies at the primary locus are explicitly considered.

<table>
<thead>
<tr>
<th>Primary locus</th>
<th>Modifier locus</th>
<th>$B_1B_1$</th>
<th>$B_1B_2$</th>
<th>$B_2B_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_1A_1$</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$A_1A_2$</td>
<td></td>
<td>1</td>
<td>1-a</td>
<td>1-a</td>
</tr>
<tr>
<td>$A_2A_2$</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 3.1 The effect of a modifier of fitness.

Figures in the table are the relative early viability of the genotypes.
3.2.3 Life cycle, sib competition & population structure

I will consider a somewhat idealised life cycle (laid out in Fig. 3.1). Generations are assumed to be discrete and non-overlapping. Adults mate at random and produce offspring which stay together in sib groups throughout their early life. The environment may be considered as a number of discrete patches into which adults lay their eggs. Each adult deposits all its young into a single patch but each patch may be colonised by more than one sib group. When the two loci are explicitly considered (section 3.3) only the situation of one family per patch is analysed. Each sib group is of equal size. There is no competition for patches, and no variation in patch quality. On reaching adulthood individuals leave the patches and form a single homogeneous population. Selection against hybrids (heterozygotes at the primary locus) is only manifest in this panmictic population. The action of the modifier is restricted to the early part of the life cycle and so acts only within patches. Each patch contributes equally to the adult population.

3.2.4 Mating

Two, extreme mating schemes are considered. In the first instance it is assumed that each individual mates a very large (effectively infinite) number of times and produces one progeny per mating. In this case each female is fertilized by all the male gametes in the population in proportion to their frequency. When considering the explicit two locus model, there are, therefore, ten possible types of sib group corresponding to the ten possible maternal genotypes.

At the other extreme I consider that there is strict monogamy. If one were to consider all ten genotypes under monogamy there would be 50 different types of sib groups (corresponding to the 50 different possible matings). Monogamy has only been analysed using the inclusive fitness methodology. With this analysis only the genotype
frequencies at the primary locus are considered and so only six different mating types need be enumerated.

There are two main differences between these mating schemes that may be important. With polygamy all members of a sib group are half sibs, whereas with monogamy they are full sibs. Relatedness between sibs is higher under monogamy and so one might expect evolution through kin selection to be more likely. However, with polygamy the variation in fitness (at the primary locus) within families is greater than it is under monogamy. The chance that one's sib has higher fitness than oneself is greater under polygamy.

Another difference is that with monogamy there are only a certain number of different possible sib groups (as defined by the distribution of genotypes that they contain). These are determined only by the genotypes of the parents. A change in the allele frequency at the primary locus causes a change in the frequency of these sib groups but no change in their content. With polygamy, a change in allele frequency alters the distribution of adult genotypes (the founders of sib-groups) and also a change in the composition of those sib groups.
Figure 3.1 Stages of the model.
3.3 Analyzing the model I: A two locus analysis.

3.3.1 Methods

The model described above can be completely defined by tracking the frequencies of all possible genotypes throughout the life cycle shown in Fig. 3.1. The effect of two distinct rounds of selection are considered. In round one, selection acts within sib-groups at the modifier locus, in round two, there is selection against heterozygotes at the primary locus. In round one, selection will act to reduce the frequency of the modifier allele. Here selection eliminates modifiers that are carried by heterozygotes at the primary locus. In the second round of selection, homozygotes at the primary locus are favoured. Amongst these there will be a higher than average frequency of the modifier allele since those modifiers associated with heterozygotes are eliminated in the first round of selection. Thus the second round increases the frequency of the modifier. In order to determine if the modifier is favoured over an entire generation the change in its frequency must be determined between the population immediately before round one and after round two. To do this the genotype frequencies must be traced through the stages outlined below. A summary of the notation used in this analysis is given in Table 3.2.

Stage 1: Initialization. The initial genotype frequencies \((x_1, 1 \ldots x_1, 10)\) are defined as those expected after one round of selection at the primary locus, given that they are originally in Hardy-Weinberg and linkage equilibrium.
Here, $1 - 2p_a q_a s$, is the population mean fitness. The population must be initiated in this way since the outcome of the model depends on the distribution of sib-group types. Selection alters the allele frequency at the A locus. Hence the initial allele frequency at locus A is given by

$$p_{a,i} = x_{1.1} + x_{1.2} + x_{1.3} + \frac{1}{2} (x_{1.4} + x_{1.5} + x_{1.6} + x_{1.7})$$  \hspace{1cm} (3.2)

Only if $p_a = 0.5$ does this initialization procedure not alter allele frequency. The initial modifier frequency is $q_b$. As this locus is in complete linkage equilibrium with the primary one, selection at the first does not alter its frequency at this stage.

Stage 2: patch formation. Each female founds a patch with her offspring. There will be ten different types of patch, corresponding to the ten diploid genotypes. The distribution of genotypes within each patch differs between the female genotypes. There is random, multiple mating. The distribution of genotypes in each patch is a function of the females
genotype and the frequencies of male gamete genotypes. It is assumed that each female mates a very large number of times. The frequencies of each gamete genotype in each patch is equal to the gamete frequencies in the population as a whole. These gamete frequencies are:

\[
\begin{align*}
A_1 B_1 & \quad g_1 = x_{1,1} + \frac{1}{2} \left( x_{1,2} + x_{1,4} + \frac{1}{2} \left( x_{1,5} + x_{1,6} \right) \right) \\
A_1 B_2 & \quad g_2 = x_{1,3} + \frac{1}{2} \left( x_{1,2} + x_{1,7} + \frac{1}{2} \left( x_{1,5} + x_{1,6} \right) \right) \\
A_2 B_1 & \quad g_3 = x_{1,8} + \frac{1}{2} \left( x_{1,4} + x_{1,9} + \frac{1}{2} \left( x_{1,5} + x_{1,6} \right) \right) \\
A_2 B_2 & \quad g_4 = x_{1,10} + \frac{1}{2} \left( x_{1,7} + x_{1,9} + \frac{1}{2} \left( x_{1,5} + x_{1,6} \right) \right)
\end{align*}
\] 

(3.3)

The distribution of families and the frequencies of genotypes within each are given in Table 3.3. The frequency of genotype \(j\) in a patch founded by a female of genotype \(i\) is denoted below as \(y_{i,j}\).

**Stage 3: Selection within patches.** Each patch is treated as a single population and the genotype frequencies are altered by selection. Selection acts against individuals, heterozygous at the A locus which also carry the modifier allele \(B_2\). The frequency of genotype \(j\) in patches founded by a female of genotype \(i\) is given by

\[
y'_{ij} = y_{ij} \cdot \frac{\omega_{1,j}}{\bar{w}_i}
\] 

(3.4)

Where \(\omega_{1,j}\) is the relative fitness of genotype \(j\) in this first round of selection (given in Table 3.2). \(\bar{w}_i\) is the mean fitness of individuals in a patch founded by a female of genotype \(i\).

\[
\bar{w}_i = \sum_{j=1}^{10} y_{i,j} \omega_{1,j}
\] 

(3.5)
Stage 4: Individuals leave their patches. Individuals leave their patches and form a single homogeneous population. Each patch contributes equally irrespective of the patch mean fitness. Genotype frequencies in this homogeneous population, \( (x_{2,1} \ldots x_{2,10}) \) are given by the sum of the products of the genotypes frequency within each patch type, multiplied by the frequency of each patch type. i.e.

\[
x_{2,j} = \sum_{i=1}^{10} x_{1,i} y_{i,j}^\prime
\]  

(3.6)

Stage 5: Selection against heterozygotes. In the final stage selection acts, in the homogeneous population, to reduce the frequency of heterozygotes at the primary locus (A). The frequency of genotypes after selection are given by:

\[
x_{3,i} = x_{2,j} \cdot \omega_{2,j} / \bar{w}_2
\]  

(3.7)

Where, \( \omega_{2,i} \) is the relative fitness of genotype i in this second round of selection (given in Table 3.2), and \( \bar{w}_2 \) is the population mean fitness at this stage.

\[
\bar{w}_2 = \sum_{i=1}^{10} x_{2,i} \omega_{2,i}
\]  

(3.8)

The frequency of the modifier allele after this round of selection is given by

\[
q_{b,3} = x_{3,3} + x_{3,7} + x_{3,10} + \frac{1}{2}(x_{3,2} + x_{3,5} + x_{3,6} + x_{3,9})
\]  

(3.9)

Over one complete generation the change in frequency of the modifier is given by

\[
\Delta q_b = q_{b,3} - q_{b,1}
\]  

(3.10)
Table 3.2 A summary of notation used in the two locus analysis

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Frequency</th>
<th>Fitness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>initial</td>
<td>after selection at modifier locus</td>
</tr>
<tr>
<td>$A_1B_1/A_1B_1$</td>
<td>$x_{1,1}$</td>
<td>$x_{2,1}$</td>
</tr>
<tr>
<td>$A_1B_1/A_1B_2$</td>
<td>$x_{1,2}$</td>
<td>$x_{2,2}$</td>
</tr>
<tr>
<td>$A_1B_2/A_1B_2$</td>
<td>$x_{1,3}$</td>
<td>$x_{2,3}$</td>
</tr>
<tr>
<td>$A_1B_1/A_2B_1$</td>
<td>$x_{1,4}$</td>
<td>$x_{2,4}$</td>
</tr>
<tr>
<td>$A_1B_1/A_2B_2$</td>
<td>$x_{1,5}$</td>
<td>$x_{2,5}$</td>
</tr>
<tr>
<td>$A_1B_2/A_2B_1$</td>
<td>$x_{1,6}$</td>
<td>$x_{2,6}$</td>
</tr>
<tr>
<td>$A_1B_2/A_2B_2$</td>
<td>$x_{1,7}$</td>
<td>$x_{2,7}$</td>
</tr>
<tr>
<td>$A_2B_1/A_2B_1$</td>
<td>$x_{1,8}$</td>
<td>$x_{2,8}$</td>
</tr>
<tr>
<td>$A_2B_1/A_2B_2$</td>
<td>$x_{1,9}$</td>
<td>$x_{2,9}$</td>
</tr>
<tr>
<td>$A_2B_2/A_2B_2$</td>
<td>$x_{1,10}$</td>
<td>$x_{2,10}$</td>
</tr>
<tr>
<td>Sib group</td>
<td>progeny</td>
<td>Frequency</td>
</tr>
<tr>
<td>-----------</td>
<td>--------</td>
<td>-----------</td>
</tr>
<tr>
<td>$A_1B_1/A_1B_1$</td>
<td>$x_{1,1}$</td>
<td>$x_{1,1}$</td>
</tr>
<tr>
<td>$A_1B_1/A_1B_2$</td>
<td>$x_{1,2}$</td>
<td>$x_{1,2}$</td>
</tr>
<tr>
<td>$A_1B_2/A_1B_2$</td>
<td>$x_{1,3}$</td>
<td>$x_{1,3}$</td>
</tr>
<tr>
<td>$A_1B_1/A_2B_1$</td>
<td>$x_{1,4}$</td>
<td>$x_{1,4}$</td>
</tr>
<tr>
<td>$A_1B_2/A_2B_2$</td>
<td>$x_{1,5}$</td>
<td>$x_{1,5}$</td>
</tr>
<tr>
<td>$A_1B_2/A_2B_1$</td>
<td>$x_{1,6}$</td>
<td>$x_{1,6}$</td>
</tr>
<tr>
<td>$A_1B_2/A_2B_2$</td>
<td>$x_{1,7}$</td>
<td>$x_{1,7}$</td>
</tr>
<tr>
<td>$A_2B_1/A_2B_2$</td>
<td>$x_{1,8}$</td>
<td>$x_{1,8}$</td>
</tr>
<tr>
<td>$A_2B_1/A_2B_2$</td>
<td>$x_{1,9}$</td>
<td>$x_{1,9}$</td>
</tr>
<tr>
<td>$A_2B_2/A_2B_2$</td>
<td>$x_{1,10}$</td>
<td>$x_{1,10}$</td>
</tr>
</tbody>
</table>

| Fitness (round 1), $w_{ij}$ | $1$ | $1$ | $1$ | $1$ | $1$ | $1-a$ | $1-a$ | $1-a$ | $1$ | $1$ | $1$ |

Table 3.3 Distributions of genotypes within sib groups. 2 loci, Polygamy.
3.3.2 Results

The calculations listed above were manipulated algebraically using the 'Mathematica' system (Wolfram, 1988). Without making some simplifying assumptions the expression for $\Delta q_b$ is too complex to be interpretable. An interpretable expression can be arrived at by considering only the initial change in frequency for a rare ($q_b << 1$) modifier of small effect ($a << 1$). By making these assumptions it may be assumed that terms in $q_b^2$, $a^2$ and higher are of negligible size and so may be removed from the expression. (In 'Mathematica', this is easily achieved by approximating the exact expression for $\Delta q_b$ as a Taylor series around $q_b = 0, a = 0$.) After much rearrangement, the change in frequency of the modifier allele can be expressed as:

$$\Delta q_b = \frac{-a q_b p_a q_a \left(12 - 2s(27+20p_a q_a) + 15s^2(1+4p_a q_a) - 6p_a q_a s^3(7+4p_a q_a) + 28p_a^2 q_a^2 s^4 \right)}{8 \left(1 - 2p_a q_a s\right) \left(1 - 6p_a q_a s + 2p_a q_a s^2(1+2p_a q_a) - 2p_a^2 q_a^2 s^3 \right)}$$

(3.11)

Part of the reason why such a complicated expression is derived is because the model depends on the distribution of genotypes at the beginning of each generation and not just the gene frequencies. This starting distribution is defined as that expected after selection at the primary locus. Thus, included in Eq. 3.11 are the effects of two rounds of selection at the primary locus. $p_a$ in Eq. 3.11 actually refers to gene frequency in the previous generation.

Although complex, some sense may be made of Eq. 3.11. Clearly, when the modifier is absent from the population ($q_b = 0$), or has no effect ($a = 0$), there is no change in its frequency. The modifier only has effect when carried by a heterozygote at the modifier locus, thus when either allele is fixed at this locus ($p_a q_a = 0$), there is no change at the modifier locus.
The rate of change in frequency of the modifier, independent of the strength of its effect \( \Delta q_b/(a q_b) \) is illustrated in Figs. 3.2 and 3.3. When selection at the primary locus is weak (\( s \) small) the modifier is always selected against, \( (\Delta q_b < 0) \). The rate of change in the frequency of the modifier is dependent on the gene frequency at the primary locus \((p_a q_a)\). When \( p_a q_a \) is small, there can only be a small number of heterozygotes in the population and so the opportunity for selection on the modifier is small. The maximum opportunity for selection is when \( p_a q_a = \frac{1}{4} \) (i.e. \( p_a = \frac{1}{2}; q_a = \frac{1}{2} \)). Hence, the largest rate of change - either positive or negative, depending on the magnitude of \( s \) - is found when \( p_a = \frac{1}{2} \).

From Fig. 3.2 it can be seen that the magnitude of \( s \) for which there is no net selection on the modifier locus is a function of gene frequency at the primary locus, (i.e. the point at which \( \Delta q_b = 0 \)). This point I define as \( s^* \); the strength of selection above which the modifier will be favoured. This could be found by solving Eq. 3.11 equal to zero in \( s \). However, since this is a quartic function, no simple solution is obtainable. For particular values of \( p_a \) solutions may be found. For example, when \( p_a = \frac{1}{2} \) (the only polymorphic equilibrium in \( p_a \)), Eq. 3.11 reduces to:-

\[
\Delta q_b = \frac{a q_b (7s - 6)}{8 (2-s)}
\]

(3.12)

The solution of which gives \( s^* = 6/7 \). For other values of \( p_a \) the expression for \( \Delta q_b \) remains a quartic with four solutions. Only the first of these solutions lies within the possible range of \( 0 > s <= 1 \). The values of \( s^* \) is shown in Fig. 3.4 as a function of gene frequency. (This graph was produced by solving Eq. 3.11 numerically in Mathematica for a range of particular values of \( p_a \)).

The key point to note from Fig. 3.4 is that for all gene frequencies selection must be strong in order for the modifier to be favoured. A decrease in the value of \( p_a q_a \) decreases -- 66 --
Chapter 3

$s^*$. However, Fig. 3.3 illustrates that the actual rate of change in frequency of the modifier increases with $p_aq_a$. This second point is explained by the fact that the modifier is only influenced by selection when in a heterozygous individual; when $p_aq_a$ is low there can be only a very few heterozygotes and so the opportunity for selection is low. The first point becomes clearer when the problem is considered in terms of the spread of an altruistic trait (section 3.4). The decrease in $s^*$ with $p_aq_a$ follows the increase in the chance that the sibling of a heterozygous individual is homozygous, and so of higher fitness.

The dynamics of the primary locus are not considered here. In the absence of the modifier it would show the standard dynamics of under dominance for two alleles at a single locus (Hartl & Clark, 1989). There is an unstable equilibrium at $p_a = \frac{1}{2}$. For allele frequencies other than this the most frequent allele increases to fixation. If the modifier is fixed in the population the dynamics are essentially unchanged; the overall selection against heterozygotes is simply increased. For intermediate frequencies of the modifier, if selection is strong enough that it is increasing in frequency, as the modifier increases so will the overall selection on the primary locus. However, as selection on the modifier is much weaker than selection directly on the primary locus, one of the alleles at the primary locus will become fixed before the modifier does. In this situation, selection on the modifier ceases. The population is initiated in linkage equilibrium. Since the modifier has equal effect in both homozygotes linkage disequilibrium cannot be built up during the course of evolution.

The conditions given above are only the conditions that the modifier be favoured; not that it rises to fixation. It may be that the allele reaches some equilibrium value as is the case in other models of the evolution of altruism (Charlesworth, 1978). In order to determine if this will happen one needs to consider the complete dynamics of the system irrespective of allele frequency. However, the system is too complex for this to be done analytically. Although it is not possible to identify any internal equilibria, a minimum requirement for the fixation of the allele is that it continues to increase even when it is common. This is the same as asking: Can a wild-type allele invade a population fixed for
the modifier? This question can be easily answered with reference to the equations given above. It is simplest to consider Eq. 3.12 (the rate of change for \( p_a = \frac{1}{2} \)). This is for a rare modifier allele. Replacing \( q_b \) with \( p_b \) and changing the sign of \( a \) gives an expression for the rate of change of wild-type allele invading a population fixed for the modifier (the sign must be changed since the wild-type allele has opposite effect). Under the conditions that the modifier is favoured, this new expression is always negative. Wild type alleles cannot invade, and so the modifier must continue to increase in frequency when sufficiently common.

![Figure 3.2 Δq_b vs s. Polygamy.](image)

Rate of change in frequency of the modifier \([Δq_b/(a q_b)]\) vs selection strength against heterozygotes at the primary locus, \( s \). From the two locus analysis.

Curves shown are for \( p_aq_a = 0.1, 0.15, 0.2, \) and 0.25.
Figure 3.3 $\Delta q_b$ vs $p_a q_a$: Polygamy
Rate of change in frequency of the modifier $[\Delta q_b / (a q_b)]$ vs allele frequency (measured as $p_a q_a$) the primary locus (measured as $p_a q_a$). From the two locus analysis. Curves shown are for $s = 0.8, 0.9$ and $1.0$

Figure 3.4 Conditions for the modifier to be favoured. Polygamy.
$s^*$ vs $p_a q_a$, $s^*$ is the level of selection against heterozygotes at the primary locus above which the modifier will be favoured. For $s < -0.8$ the modifier is always selected against; for $s > 0.6$ the modifier is always favoured. With intermediate values of $s$ the fate of the modifier is dependent on the allele frequency at the primary locus.
3.4 Analyzing the model II: Inclusive fitness analysis

3.4.1 Methods

Inclusive fitness and Hamilton's rule

In this section I re-analyse the model already considered in the previous section. Here I treat the evolution of decreased hybrid viability as a specific example of the evolution of altruism. It will be seen that, for the simple model this analysis gives exactly the same results as does the two locus analysis. The advantage of this approach is that the simple model can be easily extended to consider other factors.

The standard methodology for such an analysis is to consider the relative 'inclusive fitness' of individuals that express the altruistic trait. Hamilton (1963) defined inclusive fitness as

The animal's production of adult offspring......stripped of all components......due to the individual's social environment, leaving the fitness he would express if not exposed to any of the harms or benefits of that environment...and augmented by certain fractions of the harm and benefit the individual himself causes to the fitnesses of his neighbours. The fractions in question are simply the coefficients of relationship'.

The concept of inclusive fitness is used to determine if a novel, altruistic trait may spread in a population. An altruistic trait is one which incurs a 'cost', (C) to its donor and some 'benefit', (B) to its recipient. Costs and benefits are both measured as an increase or decrease in the reproductive success of the individuals involved.

In the absence of selection and population growth, the average inclusive fitness of all individuals can be defined as 1 (since on average, each individual produces one offspring only). The donor of an altruistic act has an inclusive fitness $1 + zB - C$; where $C$ is the cost of the act; $B$, the benefit gained by the recipient and $z$ the relatedness between
donor and recipient. One will expect that the behaviour which confers the highest inclusive fitness will be favoured. Thus, altruism will be favoured if $zB - C > 0$ (Hamilton's rule).

The basic idea behind the notion of inclusive fitness is that benefiting a relative is like benefiting oneself since relatives share genes. The degree of relatedness between any two individuals may be defined as the total proportion of genes they share that are identical. Alternatively, one may consider a single locus only. Relatedness is the proportion of alleles at this locus that two individuals have in common ($0, \frac{1}{2}$ or $1$). It is only this second meaning of relatedness that is of importance to inclusive fitness theory for single gene modifiers.

The concept of inclusive fitness is generally held to be useful because it allows one to consider the evolution without explicit reference to the genes involved. However, it can be easily applied with reference to particular genes. If this is done, the logic behind the application of Hamilton's rule becomes clearer. Also, something can be said of the actual dynamics of altruistic genes. The methodology followed is discussed in the following section.

**Allele frequency change at a single locus**

The change in frequency of an allele over one generation may be described in terms of the difference between its average fitness and the average fitness of the population, i.e. the average excess of the allele (Fisher, 1930)

$$\Delta q_b = \frac{q_b(w_b - \bar{w})}{\bar{w}}$$

(3.13)

where $\bar{w}$ is the population mean fitness and $w_b$ the average fitness of the $b$ allele. $w_b - \bar{w}$ is the average excess of the $b$ allele. For example, consider there are only two alleles ($b$ and $B$) at a locus, $bb$ homozygotes have fitness $w_{bb}$ and heterozygotes have fitness $w_{bB}$. Each homozygote contributes $w_{bb}$, and each heterozygote $\frac{1}{2}w_{bB}$ to the average fitness of the $b$
allele. The average fitness of the b allele is the average of \( w_{bb} \) and \( \frac{1}{2} w_{Bb} \) weighted by the frequency of each genotype. The population mean fitness may be defined as the weighted sum of the average fitness of the allele (\( \bar{w} = p_b w_B + q_b w_b \)). Eq. 3.13 rearranges to

\[
\Delta q_b = \frac{p_b q_b (w_b - w_B)}{\bar{w}}
\]  

(3.14)

Clearly, allele b will increase in frequency if its average fitness is greater than that of B.

In the above derivation, the fitness of an individual allele varies only according to whether it is found in a homozygous or heterozygous state. Since I wish to consider a modifier allele, fitness will vary also according to the genotype at the primary locus. By considering only the initial change in frequency of a rare modifier one source of variation is removed. When rare the modifier will only be present in a heterozygous state. Hence, the average fitness of the modifier allele is the average of its fitness when expressed against a background of all the possible genotypes at the primary locus. By assuming that the modifier allele is initially in linkage equilibrium with the primary locus, the frequency with which the modifier is found with particular genotype at the primary locus is simply the frequency of that genotype.

The fitness of the modifier allele, when expressed in a particular individual, is dependent on the genotypes of its sibs. The appropriate measure of fitness is therefore the alleles inclusive fitness. The change in frequency of the modifier allele, when rare, is therefore given by Eq. 3.14 but with \( w_b \) and \( w_B \) replaced with the average inclusive fitnesses.

\[
\Delta q_b = \frac{p_b q_b (\bar{w}_{1m} - \bar{w}_{1+})}{\bar{w}}
\]  

(3.15)
Where \( \hat{w}_{im} \) and \( \hat{w}_{I^+} \) are the average, inclusive fitness of the modifier and the wild-type allele. Inclusive fitness is defined for an allele in the same way as it is for an individual. \( \hat{w}_{I^+} = 1 \) and \( \hat{w}_{im} \) equals the average of \( 1 - C + zB \). C and B are as before, the cost to the donor and the benefit to the recipient of the altruistic act caused by the modifier. When considering a particular locus, relatedness has an obvious meaning. The relatedness is specifically the probability that the recipient carries the modifier allele, given that the donor does. By considering only the initial change in frequency of a rare modifier allele it can be assumed that for full-sibs \( z = \frac{1}{2} \). For half sibs \( z = \frac{1}{2} \) if inherited through the maternal parent and zero through the paternal, giving an average of \( \frac{1}{4} \). This is a quite different interpretation of relatedness than considering it as a measure of genetic similarity. Here only similarity at one particular locus is important.

Eq. 3.15 can be simplified further. By definition \( \hat{w} = q_b \hat{w}_{im} + p_b \hat{w}_{I^+} \). When the modifier is rare \( q_b \) is very small and so \( \hat{w} \approx p_b \hat{w}_{I^+} \). Since \( \hat{w}_{I^+} = 1 \), \( \hat{w} \approx p_b \) which cancels giving

\[
\Delta q_b = q_b (\hat{w}_{im} - \hat{w}_{I^+})
\]

(3.16)

The sign of the difference in inclusive fitness between the alleles \( \hat{w}_{im} - \hat{w}_{I^+} \), indicates the direction of the change in gene frequency, giving rise to Hamilton's rule: The modifier will be favoured if \( \hat{w}_{im} - \hat{w}_{I^+} > 0 \), or \( zB - C > 0 \). The magnitude of the difference determines the rate of change. Graffen (1991) states that Eq.3.16 does not, in general, describe the change in allele frequency. However, it does give an approximation for a dominant allele under the assumption that the allele is rare. It is an approximation since it ignores the small proportion of homozygotes and the small chance that non-relatives share the modifier allele. If one were to consider the fate of a new mutation in an otherwise homozygous, wild-type population Eq. 3.16 describes exactly its fate. How the average inclusive fitness of the modifier allele is determined is described below.
Inclusive fitness of alleles in the model of reinforcement

Recall that I am considering the fate of a modifier allele that reduces the early viability of heterozygotes only. For simplicity I will consider that the effect of the allele is to reduce viability to zero (i.e. causes the early death of its carrier). I begin by considering polygamy only and patches founded by one female only. Patches are equivalent to single families. The distribution of family types and of genotypes within families are given in tables 4 and 5. In the tables \( p^i, q^i, u,v \) and \( w \) are allele and genotype frequencies at the primary locus after one round of selection.

The cost, \( C \), of the action of the modifier to its carrier is its lost chance of future reproduction, \( (1-s)/w' \), where \( w' \) is the population mean fitness in this population. This cost is the same across all families. The cost of a modifier allele carried by a homozygote is zero since it has no effect. The place of the 'altruist' in the population will be taken by another individual from the patch. This will be either a homozygote (with probability \( F_{\text{hom}} \)) or a heterozygote (probability \( F_{\text{het}} \)). The benefit gained by the recipient is \( 1/w' \) for a homozygote and \( (1-s)/w' \) for a heterozygote. The chance that the beneficiary also carries the modifier allele is \( z \) (the relatedness). Therefore the indirect benefit gained by the action of the modifier allele, \( zB \) is:

\[
zB = \frac{z}{w'} \left( F_{\text{hom}} + F_{\text{het}}(1-s) \right) = \frac{z}{w'} \left( 1 - F_{\text{het}}s \right)
\]

Note that it is assumed that the modifier is expressed in at most one individual per patch (penetrance is low). Consider modifier alleles at a low frequency randomly distributed across genotypes in the adults. It may end up in either a homozygote or a heterozygote progeny (i.e. any of the cells in tables 4 and 5). \( F_{\text{het}} \) varies across these cells and in consequence so does \( B \). The average inclusive fitness of the modifier allele is therefore the average of \( 1 + zB - C \) over all progeny genotypes and patch types.
\[
\hat{w}_{im} = 1 + \sum_{\text{families, progeny}} \text{frequency (} zB - C \text{)}
\]  
(3.18)

The genotype frequencies in Eq.18 are those of the progeny. For example, with multiple mating, the modifier will be found in a heterozygous offspring of an A1A1 mother with a frequency \( u q_a' \). The cost incurred will be \( (1-s)/\hat{w}' \), and the indirect benefit \( \frac{1}{4} (1-q_a')/\hat{w}' \).

Genotypes, gene frequencies and mean fitness over two generations

In order to calculate the average inclusive fitness of the modifier allele as described above, population parameters (genotype and gene frequencies and mean fitness) must be known over two generations. To define the distribution of patch types the distribution of parental genotypes (\( u, v \) and \( w \)) are required. These are the frequencies after the effect of selection and are given by the following, standard expressions (Hartl & Clark, 1989)

\[
\begin{align*}
\text{u} & = \frac{p_a^2}{1 - 2 p_a q_a s} \\
\text{v} & = \frac{2 p_a q_a (1-s)}{1 - 2 p_a q_a s} \\
\text{w} & = \frac{q_a^2}{1 - 2 p_a q_a s}
\end{align*}
\]

(3.19)

To determine the distribution of genotypes in the progeny produced by these parents with polygamy the allele frequencies after selection are required:

\[
\begin{align*}
p_a' & = \frac{p_a(1 - q_a s)}{1 - 2 p_a q_a s} \\
q_a' & = \frac{q_a(1 - p_a s)}{1 - 2 p_a q_a s}
\end{align*}
\]

(3.20)
The costs and benefits are fractions of the population mean fitness in the progeny generation. Given that there is random mating, population mean fitness in the following generation is given by

\[ W' = P_a^2 + Q_a^2 + 2 P_a Q_a(1-s) \]

\[ \frac{1 - 6 P_a Q_a s + 2 P_a Q_a s^2 (1+2P_a Q_a) - 2 P_a^2 Q_a^2 s^3}{(1 - 2P_a Q_a s)^2} \]  (3.21)

In this way mean fitness in one generation can be expressed as a function of selection and the allele frequency in the previous generation. Note that in the equations given above, the effect of the modifier itself is not included. Selection at the modifier locus will alter the the distribution of genotype frequencies. For the sake of simplicity this effect is ignored. All of the analysis presented below concerns only the initial changes in frequency of a rare modifier. Since only rare modifiers are considered their effect on genotype frequencies will be negligible.
### Table 3.4 Genotypes within sib groups. 1 locus. Polygamy

Distribution of family types and genotypes within families with multiple mating. $p_a'$ and $q_a'$ are allele frequencies at the primary locus. $u$, $v$ and $w$ are genotype frequencies after selection (given by Eq.s 3.19 & 3.20)

<table>
<thead>
<tr>
<th>Genotype of mother</th>
<th>Frequency</th>
<th>$A_1A_1$</th>
<th>$A_1A_2$</th>
<th>$A_2A_2$</th>
<th>$F_{hom}$</th>
<th>$F_{het}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_1A_1$</td>
<td>$u$</td>
<td>$p_a'$</td>
<td>$q_a'$</td>
<td>0</td>
<td>$p_a'$</td>
<td>$q_a'$</td>
</tr>
<tr>
<td>$A_1A_2$</td>
<td>$v$</td>
<td>$\frac{1}{4}$</td>
<td>$\frac{1}{2}$</td>
<td>$\frac{1}{4}$</td>
<td>$\frac{1}{2}$</td>
<td>$\frac{1}{2}$</td>
</tr>
<tr>
<td>$A_2A_2$</td>
<td>$w$</td>
<td>0</td>
<td>$p_a'$</td>
<td>$q_a'$</td>
<td>$q_a'$</td>
<td>$p_a'$</td>
</tr>
</tbody>
</table>

### Table 3.5 Genotypes within sib groups. 1 locus. Monogamy

As for Table 3.4 but with monogamy.
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_1$,$A_2$</td>
<td>Alleles at the primary locus</td>
</tr>
<tr>
<td>$p_a,q_a$</td>
<td>Frequency of the $A_1$ and $A_2$ alleles</td>
</tr>
<tr>
<td>$u$</td>
<td>Frequency of $A_1A_1$ genotypes</td>
</tr>
<tr>
<td>$v$</td>
<td>Frequency of $A_1A_2$ genotypes</td>
</tr>
<tr>
<td>$w$</td>
<td>Frequency of $A_2A_2$ genotypes</td>
</tr>
<tr>
<td>$s$</td>
<td>Selection against $A_1A_2$ genotypes</td>
</tr>
<tr>
<td>$s^*$</td>
<td>Selection required in order for the modifier to be favoured</td>
</tr>
<tr>
<td>$W$</td>
<td>Population mean fitness.</td>
</tr>
<tr>
<td>$\bar{w}_{t+}$</td>
<td>Average inclusive fitness of the wild type allele</td>
</tr>
<tr>
<td>$\bar{w}_{lm}$</td>
<td>Average inclusive fitness of the modifier allele</td>
</tr>
<tr>
<td>$z$</td>
<td>Relatedness between donor and recipient of an altruistic act</td>
</tr>
<tr>
<td>$B$</td>
<td>Benefit gained by the recipient of an altruistic act</td>
</tr>
<tr>
<td>$C$</td>
<td>Cost incurred by the donor of an altruistic act</td>
</tr>
<tr>
<td>$F_{het}$</td>
<td>Proportion of heterozygotes (at the primary locus) within a patch</td>
</tr>
<tr>
<td>$F_{hom}$</td>
<td>Proportion of homozygotes (at the primary locus) within a patch</td>
</tr>
<tr>
<td>$p_a,q_a^*$</td>
<td>allele frequency above which the modifier cannot be favoured</td>
</tr>
<tr>
<td>$n$</td>
<td>Number of (equal sized) sib groups per patch</td>
</tr>
<tr>
<td>$n^*$</td>
<td>Maximum number of sib groups per patch that will allow the spread of the modifier</td>
</tr>
</tbody>
</table>

**Table 3.6** Summary of notation used in the inclusive fitness analysis.
3.4.2 Results. Single populations

A summary of the symbols used in this analysis is given in Table 3.6.

3.4.2.1 Sib competition only

Polygamy

The average, inclusive fitness of the modifier allele, \( \bar{w}_{1m} \) is calculated as the average value of \( 1 - C + zB \) across all the cells of Table 3.4. The difference between this and the average inclusive fitness of the wild type allele \( \bar{w}_{1+} \) is given by:

\[
\bar{w}_{1m} - \bar{w}_{1+} = \frac{-2p_a'q_a'(1-s) + z(2p_a'q_a' - (q_a'^2 u + 4v + p_a'^2 w))}{w'}
\]  

Note that this difference is dependent on \( u, v \) and \( w \); the genotype frequencies after one round of selection. \( p_a' \) and \( q_a' \) are the allele frequencies after selection. The above expression can be simplified by replacing gene and genotype frequencies with expressions that relate them to selection and the gene frequency in the previous generation (Eqs. 3.19 and 3.20). This produces an apparently more complex equation but is in fact a simplification in that it reduces the number of variables. By making these substitutions and with reference to Eq. 3.16 an expression for the change in frequency of the modifier allele may be derived:

\[
\Delta q_a = - q_a p_a q_a x \\
\frac{(4(1-s)(1-2p_a q_a s)(1-s(1-p_a q_a s))-z(4-s(5+12p_a q_a s)+s^2(1+20p_a q_a s)-2p_a q_a s^3(3+4p_a q_a s)+4p_a^2 q_a^2 s^4))}{2(1-2p_a q_a s)(1-6p_a q_a s+2p_a q_a s^2(1+2p_a q_a s)-2p_a^2 q_a^2 s^3)}
\]

(3.23)
Chapter 3

With patches colonised by only one female, \( z = \frac{1}{4} \). Putting this into Eq. 3.23 and rearranging yields an expression identical to Eq. 3.11 (that derived from the two locus analysis) but without the factor \( a \) (recall that in the present analysis it is assumed that the effect of the modifier is the death of its carrier, \( a=1 \)). The rate of change of frequency of the modifier and the conditions for its spread are, therefore, the same as were given in section 3.3.2.

The maximum selection required for the spread of the modifier is at \( p_a q_a = \frac{1}{4} \) (or \( p_a = 1/2 \), the only polymorphic equilibrium in single populations). For such populations, \( s^* = \frac{6}{7} \), the same value as was arrived at through the complete two-locus model in section 3.3. It is interesting to note that in this particular case, \( s^* \) can be found in a far less tortuous manner by noting that \( F_{het} = \frac{1}{2} \) for all patches (see Table 3.4) and so the distribution of patch types is not required. One need only solve Hamilton's condition \( zB > C \) or \( \frac{1}{4} \left( 1 - \frac{1}{2} s \right) > (1-s) \) giving \( s > \frac{6}{7} \).

**Monogamy.**

When matings are strictly monogamous the conditions for the spread of the modifier are very different from those under polygamy. The difference in average inclusive fitness between the modifier and the wild type allele is given by the average of \( zB - C \) over the cells of Table 3.5:-

\[
\frac{w_{a}' - w_{1+}}{w_{1m} - w_{1+}} = -\frac{2 p_a' q_a' (1-s) + z \left( 2p_a' q_a' - s \left( p_a' q_a' + u w \right) \right)}{w'}
\]  

(3.24)
Substituting in Eqs. 3.19 and 3.20 for $p^a_1, q^a_1, u, v$ and $w$ into Eq. 3.24, and putting this difference in average inclusive fitness into Eq. 3.16 gives an expression for the change in the frequency of the modifier over one generation.

$$\Delta q_b = \frac{-q_b p_a q_a (s-1) \left(2 \left(1-s(1-p_a q_a s)\right) + z \left(s(1+p_a q_a (1-s)-2)\right)\right)}{(1-6p_a q_a s + 2 p_a q_a s^2 (1+2p_a q_a) - 2 p_a^2 q_a^2 s^3)} \quad (3.25)$$

For patches colonised by single females ($z=1/2$) and so

$$\Delta q_b = \frac{-q_b p_a q_a \left(1-s\right) \left(2 - 3s + p_a q_a s + 3p_a q_a s^2\right)}{2\left(1-6p_a q_a s + 2 p_a q_a s^2 (1+2p_a q_a) - 2 p_a^2 q_a^2 s^3\right)} \quad (3.26)$$

This rate of change is illustrated with respect to allele frequency at the primary locus and selection strength in Figs 3.5 and 3.6. One difference between polygamy and monogamy is immediately apparent. Under monogamy it is only ever possible for the modifier to be favoured in families founded by at least one heterozygous parent (rows 2, 4, and 5 of Table 3.5). As selection increases, although the modifier would be favoured in these patches, the proportion of the population accounted for by these patches decreases (since heterozygotes are eliminated by selection at the primary locus). When selection is absolute ($s=1$) only patches consisting entirely of homozygous progeny (in which the modifier does not act) or entirely of heterozygous progeny (in which there is no benefit to be gained) exist, and so the modifier cannot be favoured. With polygamy, there is a single threshold selection strength above which the modifier will be favoured.

When the population is at intermediate allele frequency ($p_a = \frac{1}{2}$ or $p_a q_a = \frac{1}{4}$) the modifier is always selected against (Fig. 3.6). At first this seems surprising. In the context of a hybrid zone one might expect the strongest selection for modification to be found in the centre. In populations in which $p_a q_a = \frac{1}{4}$, if selection is strong enough for the modifier to be
favoured in patches in which it is possible for it to be favoured, then the frequency of those patches is low (since they are all founded by heterozygotes).

The conditions for the spread of the modifier are defined by the minimum selection required and the maximum 'permissible' allele frequency at the primary locus (i.e. the allele frequency at the primary locus above which the modifier is selected against). These are found by solving Eq. 3.26 = 0. Unlike the situation with polygamy, here exact expressions are obtainable.

\[
S^* = \frac{1}{6p_{aqa}} \left(3 - p_aq_a - \sqrt{p_{aqa}(p_{aqa}^2 - 30) + 9}\right)
\]

\[
P_{a\theta a}^* = \frac{3s - 2}{3s^2 + s}
\]

These conditions are illustrated in Fig. 3.7.

For comparison, the conditions for the spread of the modifier under polygamy are also shown in Fig. 3.7. At low or high allele frequencies at the primary locus \(p_{aqa}\) close to 0) less selection is required for the modifier to be favoured under monogamy than polygamy. At intermediate frequencies this difference is reversed. For both mating schemes and for all allele frequencies the modifier is only favoured if selection is strong. Comparison of Figs. 3.3 and 3.6 shows that the rate of change of frequency is at least an order of magnitude greater with polygamy than with monogamy. These figures also illustrate that in a population homozygous at the primary locus the modifier allele is not subject to selection since it only acts in heterozygotes.
Figure 3.5 $\Delta q_b$ vs $s$. Monogamy.
Change in frequency of the modifier ($\Delta q_b/q_b$) vs selection against heterozygotes at the primary locus ($s$). From the inclusive fitness analysis with monogamy and one family per patch. Curves shown are for $p_aq_a = 0.1, 0.15, 0.2, 0.25$

Figure 3.6 $\Delta q_b$ vs $p_aq_a$. Monogamy.
Change in frequency of the modifier ($\Delta q_b/q_b$) vs allele frequency at the primary locus (measured as $p_aq_a$). Curves shown are for $s = 0.9, 0.8, 0.7$
Chapter 3

Figure 3.7 Conditions for the modifier to be favoured. Monogamy and polygamy.

$s^*$ vs $p_{aq}$ from the inclusive fitness analysis with monogamy and one female per patch. Note with monogamy that for $p_{aq} = 0.25$, $s^* = 1$. Since $s$ must always be less than or equal to 1, this means that the modifier cannot be favoured.
3.4.2.2 Competition between sibs and non-sibs

The model is easily extended to allow for competition between sibs and non-sibs in the early stage of the life cycle. Each patch is considered to be founded by the progeny of \( n \) different females, thus there are \( n \) different, equal sized, sib groups per patch. The cost of an altruistic act (reduced viability) is the same as if there were only one group per patch. The indirect benefit, \( z_B \) is however reduced since the altruists place in the population may be taken by an unrelated individual. In this model there is no opportunity for the altruistic act to be directed exclusively at sibs.

A proportion \( \frac{1}{n} \) of the time, the beneficiary will be a sib of the altruist, and the indirect benefit gained is \( z_B \). The remaining \( 1 - \frac{1}{n} \) of the time a non-sib will gain so there is no benefit to the altruist (since it is assumed the relatedness of non-sibs is zero) Therefore the average, indirect benefit is \( \frac{1}{n} z_B \). The difference in average inclusive fitness between the wild-type and modifier alleles are calculated as in section 3.4.2.1 but using an average relatedness \( z = \frac{1}{2n} \) for monogamy and \( z = \frac{1}{4n} \) for polygamy.

**Polygamy**

The rate of change of frequency of the modifier allele is given by Eq. 3.23 with \( z=1/4n \). This is illustrated, in relation to \( n \) in Figs. 3.8 and 3.9. If selection is strong enough that the modifier may be favoured then increasing \( n \) decreases the rate at which the modifier increases. If selection is not strong enough for the modifier to be favoured, increasing \( n \) increases the rate of decline of the modifier.

The conditions for the modifier to increase in frequency are found as in section 3.4.2.1 but setting \( z \) to \( 1/4n \) and varying \( n \). These conditions are illustrated in Fig. 3.10. From this it can be seen that increasing the number of sib groups per patch increases the strength of selection required in order that the modifier be favoured. Similarly, for a
particular selection strength, increasing $n$ decreases the range of gene frequencies at the primary locus for which the modifier will be favoured. For higher values of $n$ the modifier will always be selected against. A critical number of sib groups per patch ($n^*$) may be defined above which the modifier will be selected against. $n^*$ is the solution of Eq. 3.23 in $n$.

$$n^* = \frac{(4-5s-12p_s q_a s + s^2 + 20p_s q_a s^2 - 6p_a q_a s^3 - 8p_a q_a s^3 + 4p_a q_a s^2 s^4)}{16(1-s)(1-2p_a q_a s)(1-s+p_a q_a s^2)}$$

(3.29)

$n^*$ is illustrated in relation to $s$ and $p_a q_a$ in Fig. 3.11. When selection against heterozygotes is very strong $n^*$ becomes very large. For example if $s=1$, $n^*$ is infinite. When the future fitness of a heterozygote is zero there is no cost involved in behaving altruistically. The chance that this act benefits a sib rather than a non-sib is of little importance so long as the chance of benefiting a sib is greater than zero the altruistic act may be favoured. The amount of variation at the primary locus ($p_a q_a$), has very little effect on $n^*$. 

-- 86 --
Figure 3.8 $\Delta q_b$ vs $s$. Polygamy. More than 1 sib group per patch. $p_a q_a = 0.25$. Change in frequency of the modifier ($\Delta q_b/q_b$) vs selection against heterozygotes at the primary locus ($s$). $n = 1,2,4,8$, from the top to the bottom curve.

Figure 3.9 $\Delta q_b$ vs $s$. Polygamy. More than 1 sib group per patch. $p_a q_a = 0.15$. Change in frequency of the modifier ($\Delta q_b/q_b$) vs selection against heterozygotes at the primary locus ($s$). From the inclusive fitness analysis with polygamy with more than one family per patch ($n = 1,2,4,8$, from the top to the bottom curve).
Figure 3.10 Conditions for spread of the modifier. Polygamy. More than one sib group per patch. $s^* \text{ vs } \paqa$. For polygamy and more than one family per patch.

Figure 3.11 $n^*$ vs $s$. Polygamy.

polygamy $n^*$ vs $s$ for $\paqa = 0.01$ and $\paqa = 0.25$. The modifier can only be favoured if $n < n^*$
Chapter 3

Monogamy

With monogamy, reducing the degree of sib-competition has a more critical effect than with polygamy. This is because the overall effective selection pressure on the modifier is much lower under monogamy. The change in frequency of the modifier allele is given by Eq. 3.25 with \( z = l/2n \). This is illustrated, in relation to \( n \) in Figs. 3.12 and 3.13. As \( n \) increases, the rate of change in the frequency of the modifier is reduced and \( s^* \) is increased.

The conditions for the increase of the modifier are given by the following critical values (found by solving Eq. 3.25 = 0 with \( z = l/2n \))

\[
s^* = \frac{1 - 4n + p_a q_a + \sqrt{(-4n + 1 + p_a q_a)^2 - 8p_q (1-4n)(1-2n)}}{2p_a q_a (1-4n)}
\] (3.30)

\[
p_{a q_a}^* = \frac{2 - 4n - s + 4ns}{s(1 - s + 4ns)}
\] (3.31)

\[
n^* = \frac{2 - s - p_a q_a s + p_a q_a s^2}{4 - 4s + 4p_a q_a s^2}
\] (3.32)

These are illustrated in Figs. 3.14 and 3.15. The effect of increasing \( n \) is more drastic than with polygamy. For large values of \( n \), the modifier can only be favoured if frequency at the primary locus is low since for high values, \( s^* > 1 \) and so impossible (Fig. 3.14). Unlike with polygamy variation at the primary locus has a large effect on \( n^* \). When \( n \) is greater than 3 or 4, the modifier can only be favoured (if selection is strong enough) in populations in which the allele frequency at the primary locus is very low (Fig. 3.15).
Figure 3.12 $\Delta q_b$ vs s. Monogamy. More than one sib group per patch. $p_{a|a} = 0.15$.
Change in frequency of the modifier ($\Delta q_b/q_b$) vs selection against heterozygotes at the primary locus (s). With monogamy and $n = 1$ (top curve) and $n = 2$ (bottom curve) families per patch.

Figure 3.13 $\Delta q_b$ vs s. Monogamy. More than one sib group per patch. $p_{a|a} = 0.05$
Change in frequency of the modifier ($\Delta q_b/q_b$) vs selection against heterozygotes at the primary locus (s). With monogamy and $n = 1$ (top curve) and $n = 2$ (bottom curve) families per patch. For $p_{a|a} = 0.05$. 
Figure 3.14 Conditions for the modifier to be favoured. Monogamy. More than one sib group per patch. $s^*$ vs $p_0q_a$. For monogamy $n = 1, 2, 4, 8$ (from bottom to top curve) families per patch.

Figure 3.15 $n^*$ vs $s$. Monogamy.
3.4.2.3 Modifiers that also reduce viability of non-hybrids

It may be more realistic to consider an allele that reduces the viability of both heterozygotes and homozygotes. One could imagine a modifier allele that reduces the fitness of hybrids to zero but also causes a small reduction in the fitness of pure race individuals. In a pure race population such an allele would obviously be selected against. In a hybrid population it may be favoured if the benefit gained by the sibs of heterozygotes outweighs the cost to both heterozygotes and homozygotes. The conditions for the spread of such an allele will obviously be more restrictive than those considered above.

When expressed in a hybrid individual, the cost to the carrier and benefit to the recipient of this allele are as in section 3.4.2.1 above. In non-hybrid individuals viability is reduced by a fraction α. Thus the cost is α and the indirect benefit αzB. The conditions most likely to promote reinforcement are when there is polygamy and high heterozygosity. For simplicity I consider only this situation and restrict the analysis to populations in which $p_a = \frac{1}{2}$. Under these circumstances, the rate of change in frequency of the modifier reduces a simple expression (c.f. Eq. 3.12)

$$\Delta q_b = q_b \frac{7s - 6 - \alpha (6 + s)}{8(2-s)} \quad (3.33)$$

In the absence of an effect in homozygotes the modifier is favoured if $s > 6/7$. For selection stronger than this, a small effect in homozygotes will prevent the modifier being favoured.

$$\alpha^* = \frac{7s - 6}{6 + s} \quad (3.34)$$

This critical amount is illustrated in Fig. 3.16. As s increases, so does $\alpha^*$. Note that only a small effect of the modifier in homozygotes relative to that in heterozygotes is required to prevent its spread. An unconditionally lethal allele ($\alpha = 1$) could never be favoured. With monogamy, the rate of increase in frequency of the modifier is an order of magnitude less
than with monogamy. One would expect a correspondingly smaller effect of the modifier in homozygotes to halt its spread.

Perhaps more important than its effect in single populations is the effect of a modifier that reduces fitness of homozygotes in a hybrid zone. In this situation the selection against the modifier outside the zone will have a large negative effect on its spread within the zone (Sanderson, 1989).

\[ \text{Figure 3.16} \quad \alpha^* \text{ vs } s. \quad p_a = \frac{1}{2}. \text{ Polygamy.} \]

$\alpha^* \text{ vs } s$, for a population with $p_a = \frac{1}{2}$. $\alpha$ is the reduction in fitness caused by the modifier in homozygotes. The modifier may be favoured if $\alpha > \alpha^*$.
3.5 Modification of selection in structured populations

3.5.1 The importance of migration

In the models described above I have only considered the conditions for modification of selection against heterozygotes in single populations. In all the models modification can only be favoured if there is polymorphism at the primary locus. The only polymorphic equilibrium at the primary locus is at $p_a = \frac{1}{2}$, which is unstable. In any real population sampling drift would mean that this equilibrium is never attained and so any polymorphism would be transient. For the modifier to become fixed it must arise in a polymorphic population (which is unlikely) and spread fast enough to become fixed before the one or other allele at the primary locus does. The time to fixation at the primary locus is a function of the selection coefficient, $s$. The rate of spread of the modifier is dependent on the magnitude of its effect, $\alpha$, and the amount of polymorphism and selection there is at the primary locus. Since $s$ must be large in order for the modifier to be favoured, the time to fixation will be much less for alleles at the primary locus than for the modifier. As the primary locus approaches fixation the rate of spread of the modifier slows down. It is, therefore, extremely unlikely that the modifier could become fixed in a single population.

A more realistic situation in which to consider modification of selection is in polymorphic populations stabilised by a balance between migration and selection. The original reinforcement hypothesis concerned modification of hybrid zones (Dobzhansky, 1940). Migration will have three effects relevant to the models presented here. These effects make the extension of the above analytical models to geographically structured populations, such as hybrid zones, less than straightforward.

1) Since the modifier is effectively neutral outside areas of hybridization its frequency will be determined by the mutation rate and drift only. This frequency is likely to be less than that within a hybrid zone where it may be favoured. Individuals moving into the area of hybridization will carry a lower frequency of the modifier allele and so exert a downward
force on the modifier's frequency, i.e. it is possible that migration will swamp the modifiers spread. If the modifier is selected against outside areas of hybridization (as would be the case if they reduced the viability of homozygotes - see section 3.4.2.3) this effect will be particularly strong (Sanderson, 1989). However, if the modifier is neutral outside areas of hybridization it will, in the long run, eventually spread throughout the species range.

2) Migration will alter the distribution of adult genotypes (at the primary locus), and hence patches, within any one location. Through the Wahlund effect, migration produces a deficit of heterozygotes. The distribution of genotypes after selection is therefore not related to gene frequency in the simple way shown in Eq. 3.19. Expressions for genotype frequencies will contain terms relating them to the heterozygote deficit. At each location this deficit will itself be a function of allele frequency at that location and also of allele frequencies in the locations from which immigrants originate. In the case of hybrid zones this deficit is comparable to that produced by migration and assortative fertilisation discussed in Chapter 5. With monogamy, the distribution of heterozygotes is crucial to the conditions for the spread of the modifier (see section 3.4.2.1.2). It is therefore likely that migration will have a severe limiting effect on modification through kin selection.

3) The costs and benefits to an individual of reduced viability are fractions of the population mean fitness. In structured populations mean fitness varies between locations. For example, a tension zone represents a trough of low population mean fitness. The costs and benefits of the modifier at the time of its action differ depending on how far the individual in which it is carried will move in later life.

The problem of the first effect can be easily avoided by, again, considering only the initial increase of a rare modifier and assuming that it is at the same low frequency throughout the population. The second two complicating factors are overcome by considering a very simple population structure: two demes, initially fixed for alternate alleles at the primary locus, exchanging migrants. At equilibrium such a system is symmetric: i.e. gene and
genotype frequencies in one deme are exactly reflected in the other. If it is assumed that the initial frequency of the modifier is the same in both demes one need only consider its spread in one. Also, since the system is symmetric, mean fitness is the same in both demes. Problem three above does not apply. Below I consider the spread of a modifier in a two deme model at a stable equilibria for the primary locus.

3.5.2 A simple model of migration
Two demes exchanging migrants

Consider two demes exchanging migrants (Fig. 3.17). If the system is set up such that the two demes are initially fixed for alternate alleles, gene and genotype frequencies will always be symmetric as shown in Fig. 3.17. Given this setup, the system can be entirely described by considering a single deme.

Population processes follow the order: fertilization; migration; selection. Fertilization is at random. Migration is at a rate \( m \) which is constant across demes and genotypes. After migration genotype frequencies in deme one will be as follows.

\[
\begin{align*}
{u_1}' &= u_1(1-m) + m u_2 = u + m(w-u) \\
{v_1}' &= v_1(1-m) + m v_2 = v + m(v-v) = v \\
{w_1}' &= w_1(1-m) + m w_2 = w + m(u-w)
\end{align*}
\]

(3.35)

After fertilization, the population is in Hardy-Weinberg. Therefore, after selection against heterozygotes, genotype frequencies reduce to:

\[
\begin{align*}
{u_1}'' &= \frac{(p_a^2 - m(p_a - q_a))/\tilde{W}} \\
{v_1}'' &= \frac{2p_a q_a(1-s)/\tilde{W}} \\
{w_1}'' &= \frac{(q_a^2 + m(p_a - q_a))/\tilde{W}}
\end{align*}
\]

(3.36)
Where $\tilde{W}$ is the population mean fitness and given by

$$\tilde{W} = u_1 + w_1 + v_1(1-s)$$  \hspace{1cm} (3.37)

In such a system there are nine possible equilibria of which four may be stable (Barton & Rouhani, 1991; Karlin & McGregor, 1971). Two of these are pairs of demes both fixed for the same allele and are of no interest here. The other two are at $\hat{p}_a = \frac{1}{2} \left( 1 \pm \sqrt{1 - \frac{4m}{s}} \right)$ in one deme and one minus this value in the second, (or more conveniently $\hat{p}_a A_a = \frac{m}{s}$). These equilibria exist if $s > 4m$, but are only stable if $s > 6m$. Putting $\hat{p}_a$ and $\hat{q}_a$ into Eqs. 3.36 and 3.37 gives the expressions for the equilibrium distribution of genotypes after selection and population mean fitness required for a consideration of selection modification.

$$\hat{u} = \frac{1 + \sqrt{1 - \frac{4m}{s}}}{4(1-2m)} \left( \sqrt{1 - \frac{4m}{s}} - 2 + 4m \right)$$

$$\hat{v} = \frac{2m(1-s)}{s(1-2m)}$$

$$\hat{w} = \frac{1 + \sqrt{1 - \frac{4m}{s}}}{4(1-2m)} \left( \sqrt{1 - \frac{4m}{s}} + 2 - 4m \right)$$

$$\hat{W} = 1-2m$$  \hspace{1cm} (3.38)

Note that there is no selection term in the equilibrium mean fitness. In the absence of migration, each deme would become fixed for one or other of the alleles and mean fitness would be 1 in each. Deviations from this are caused only by migration.
Figure 3.17 Two demes exchanging migrants.

Gene and genotype frequencies in each deme are shown.
3.5.3 Modification of selection in a two deme system.

By setting gene and genotype frequencies to their equilibrium values in Eqs. 3.22 (for polygamy) and 3.24 (for monogamy), expressions for change in frequency of the modifier can be derived. Since only equilibrium conditions at the primary locus are considered, the resulting expressions are simpler than those for single populations despite the added migration factor. Note that the expressions below only apply to the initial increase of the modifier. As the modifier increases in frequency it will alter the net selection on the primary locus and so alter the equilibria.

Polygamy

Substituting genotype frequencies and mean fitness into Eq. 3.22 and rearranging yields:

\[
\Delta q_b = \hat{p}_a \hat{q}_a \left( \frac{-m(4(1-s)(1-2m) - z(4-12m-s(1-8m+s)))}{2s(1-2m)^2} \right) \tag{3.40}
\]

The minimum condition for the spread of the modifier is that there is polymorphism at the primary locus (\(\hat{p}_a \hat{q}_a > 0\)), which is only achieved if \(m > 0\). Also, selection must be greater than a certain minimum \(s^*\), found by solving Eq. 3.40 = 0.

\[
s^* = 8(m + 2n(1-2m) - 1 \cdot \sqrt{17-32n(3-8n)-64m(1-7n+16n^2-m(4n-1)^2)}) \tag{3.41}
\]

\(s^*\) is illustrated in Fig. 3.18 as a function of \(m\) and \(n\) (the number of females per patch, \(z = 1/4n\)). Fig. 3.18 shows that the selection required for the modifier to be favoured increases with the migration rate. Also shown is the line \(s = 6m\). The equilibrium at the primary locus is only stable above this line. It can be seen, therefore, that within the range of migration...
rates allowing stable equilibria, the curve of $s^*$ vs $m$ is virtually flat. The most influential parameter is the number of sib-groups per patch, $n$.

However, migration does have a small effect. In a single population the minimum value of $s^*$ (for $n = 1$ and $p_a$ close to zero or one) is approximately $0.8$ (see Fig. 3.4). With migration the minimum value is approximately $0.85$ (Fig. 3.18). This difference is only to do with the reduction in frequency of heterozygotes brought about by migration. Migration makes the condition for modification slightly more stringent since it reduces the frequency of heterozygotes.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{s_vs_m.png}
\caption{Conditions for the spread of the modifier in a two deme system. Polygamy. $s^*$ vs migration rate $m$. The modifier will be favoured if $s > s^*$. Curves are for $n = 1, 2, 4, 8$ from bottom to top. Also shown (the straight line) is the line $s = 6m$. The equilibrium at the primary locus is only stable above this line.}
\end{figure}
Monogamy

Substituting equilibrium values into Eq. 3.24 and rearranging yields:

$$\Delta q_b = \hat{p}_{a} \hat{q}_{a} \left( \frac{-m(1-s)(2(1-4m(1-m))-z(1 + (1-m)(1+s-8m)))}{s(1-2m)^3} \right)$$

(3.42)

There is no change in the frequency of the modifier if \(m=0\), (no migration), or \(s = 1\) (complete selection against heterozygotes). It is most likely that the modifier will be favoured when there is only one female per patch \((z=\frac{1}{2})\). For this condition the modifier is favoured if \(s\) is greater than a critical amount.

$$s^* = \frac{2 - 7m + 8m^2}{1-m}$$

(3.43)

\(s^*\) is only less than one (the maximum possible) if \(0.25 > m < 0.5\). However, the equilibrium at the primary locus is only stable if \(m < s/6\). The minimum conditions for stability and for the modifier to be favoured cannot both be met. In a two deme system with monogamous mating the modifier can never be favoured.

3.5.4 Modification in hybrid zones

The model considered above is only partially relevant to the situation in a hybrid zone. It considers only the effect migration has in distorting the distribution of genotypes. If selection against hybrids is strong enough, the modifier will be favoured provided that there is polymorphism at the primary locus. However, for intermediate selection strengths, the modifier is only favoured when gene frequency at the primary locus lies between certain values. This means that in a hybrid zone the modifier may be only favoured at the edges of
the zone. This effect is most pronounced with monogamy; with polygamy, the range of values of selection for which the sign of the net selection on the modifier is dependent on the frequency at the primary locus, is much smaller (Fig. 3.7, the slope is much shallower for polygamy). For example, in a cline maintained by selection of $s=0.8$ against heterozygotes, with monogamy the modifier can only be favoured by selection if $p_a q_a < 0.147$ (Eq. 2.28), i.e. in the centre of the cline ($0.82 < p_a > 0.18$) the modifier will be selected against. The overall conditions for the spread of the modifier will depend on the relative strength of selection against it in the centre and for it at the edges and the amount of gene flow between these areas. For modifiers that also reduce the viability of pure-race individuals, the overall change in modifier frequency will be determined by a combination of selection inside and outside the zone and gene flow (Sanderson, 1989).
3.6 Discussion

The models presented here indicate that selection can in principle reduce hybrid inviability. The conditions required for this to happen are stringent. Before discussing the actual conclusions reached it is worth commenting on the two approaches taken in addressing the question of reinforcement. In presenting the kin selection arguments I may have given the impression that I was literally discussing the possibility of a form adaptive suicide - a behavioural trait. The model can be described in these terms, but need not be. It is perhaps more reasonable to consider it as simply a consideration of alleles that reduce hybrid viability. This is, therefore, another example of how Hamilton's theory of inclusive fitness is not restricted to problems of social interaction.

Both of the analyses are only relevant to the initial increase of rare modifiers. The two locus model was explicitly restricted to rare alleles. The inclusive fitness analysis, as presented here, rests on the implicit assumption that the allele in question is rare. The rarity of the allele is implicit in the use of the simple coefficients of relatedness between sibs ($\frac{1}{2}$ for full sibs, $\frac{1}{4}$ for half sibs). Two consequences of this rarity are important. First, it is assumed that the frequency of the allele is so low that the frequency of homozygotes is negligible; all copies of the allele are present as heterozygotes. Second, the chance that the allele enters a family through more than one parent is assumed negligible. As the frequency of the modifier increases these assumptions will be violated; the relatedness between sibs (at this locus) increases above the simple coefficients used.

In the simplest model presented here (multiple mating and a single family per patch) the two locus and the inclusive fitness analysis yield identical conditions for the spread of the modifier. The expressions for the rate of change in frequency of the modifier differ only by the factor $a$, the effect of the modifier itself. In the inclusive fitness analysis $a$ is assumed to be 1, i.e. it reduces the viability of carriers to zero. By setting $a$ to 1 in the two locus result gives complete concordance between the two equations. However, the expression from the two locus analysis is based on the assumption that $a$ is much less than
one. An incorrect expression should result if this assumption is violated. How can the two
both be correct?

As mentioned previously, there is an implicit assumption in the inclusive fitness
calculation that the modifier allele is only rarely expressed (i.e. the penetrance is low,
Charlesworth, 1978). In standard population genetics theory with constant fitness there is
no difference between a fully penetrant gene with a small effect on fitness and one with a
large effect that is expressed rarely. For example, a selection coefficient of 10% on a gene
can be interpreted in two ways. Either all carriers of the gene suffer a 10% disadvantage, or
10% of carriers suffer a 100% disadvantage. The two locus model is applicable in either
case; the inclusive fitness one only in the latter.

The strength of inclusive fitness models is that they can be constructed largely
without reference to the genetics of the trait in question. It is not clear that an inclusive
fitness analysis will always be appropriate for questions of reinforcement. Inclusive fitness
arguments are usually only applied to problems of single genetically determined traits. In
the example given here, two separate genetic traits are considered: viability and sterility.
The model presented is an example of a 'one allele' model (sensa Felsenstein, 1981), i.e.
the same allele is responsible for promoting reproductive isolation in both of the
hybridizing races. Most models of reinforcement are of the two allele type (e.g.
Felsenstein, 1981; Sanderson, 1989). To apply inclusive fitness arguments to these
situations would not be so simple. The expected inclusive fitness of individuals will depend
not only on their genotypes, but also the recombination rate between the modifier and
directly selected loci. In such models linkage disequilibrium between these loci is crucial.
The model presented here is simple because, since there is only one modifier allele involved
linkage disequilibrium is not built up by the process of reinforcement and so is
unimportant.

The results of the models presented can be summarised briefly. They show that
Maynard Smith (1966b) and Coyne (1974) were in general correct in their verbal
arguments. A decrease in hybrid viability can be favoured by natural selection through kin selection. These models show that the conditions for such modification of selection are restrictive. In all the situations considered, modification can only be favoured if selection against hybrids is strong (minimum requirement is $s > -0.7$). They also show that a selective decrease in viability does not require multiple mating, as suggested by Maynard Smith and Coyne, but can be achieved if hybrid sterility is due to a single gene.

If competition is between sibs only, the conditions for the spread of inviability genes are broadly similar for polygamy and monogamy. An exception is that if selection against hybrids is absolute, inviability genes cannot be favoured with monogamy. The net selection on such genes is much less with monogamy. The rate of increase is an order of magnitude less with monogamy than with polygamy. Because of this much smaller net selection under monogamy, the effect of competition between sibs and non-sibs is much greater with monogamy. With polygamy, diluting the degree of sib competition, (by which I mean allowing competition with non-sibs), simply increases the selection against hybrids required for the spread of inviability alleles. If the number of families competing is very large this approaches 1, complete selection. Looking at it another way, if hybrids have zero fertility there is always a small advantage to be gained by a hybrid individual reducing its viability so long as family size is greater than one. There is always a small chance that this will benefit a sibling who is non-hybrid. With monogamy, if more than just a few families share a patch, modification becomes impossible.

The differences between monogamy and polygamy lie in the differences in the distributions of genotypes found within families produced by either mating scheme (Tables 4 and 5). With monogamy, the distribution of genotypes within families remains constant and only the distribution of family types can be altered. It is only in the families of at least one heterozygous parent that the modifier can be favoured. It is therefore the frequency of these families that is important. With polygamy, both the distribution of families and the genotypes within them can change. Even when selection is absolute ($s=1$) there will be competition between hybrid and pure race progeny within a sib group.
The models presented here make two assumptions about the genetics of reproductive isolation that are unlikely to be generally valid. The first is that there is no selection against modification in non-hybrid individuals. This first point may be crucial. It may be more realistic to consider alleles that reduce the fitness of both homozygotes and heterozygotes. This situation was briefly considered in section 3.4.2.3. for single populations. This showed that, with multiple mating, only a small reduction in the fitness of homozygotes is required to impede the spread of the modifier. Since the advantage to the modifier is much less with monogamy than polygamy, its spread would be halted by a correspondingly smaller effect in homozygotes. The effect would be much more profound in the context of a hybrid zone. Sanderson (1989) has shown that even if there is strong selection for modification within a hybrid zone, weak selection against modification outside the zone is sufficient to swamp the spread of reinforcing alleles. This effect is applicable here - in pure populations the modifier is always selected against.

The second, perhaps crucial assumption is that sterility is caused by the effect of underdominance at a single locus. Although several studies have identified single genes responsible for hybrid sterility, in most cases at least several genes are involved (Coyne, 1992). Furthermore, the interactions leading to reproductive isolation are more often interactions between loci than between alleles at a single locus. What difference would this make to the conclusions reached? In the extreme the answer is 'nothing'. If hybrid fertility is zero, there is no difference between a single locus and polygenic model. However, if many genes are involved and selection against hybrids is anything less than 1 recombination will mean that the variation in fitness within a hybrid zone will be much less than if there were just a single gene of major effect. The models presented here have shown that modification is only expected if there are large differences in fitness between sibs. The greater the number of genes that contribute to hybrid sterility the less likely it is that such modification could occur.

The two mating schemes considered (monogamy, and extreme polygamy) represent only the edges of a continua of possibilities. The model of polygamy assumes that each
female mates an infinite number of times. In this way the variation within each family is maximal, reflecting exactly the variation in the population. It is also assumed that each family is effectively infinite in size - all possible progeny of each mating are represented. There are perhaps instances where these assumptions are approached. For example, in plants there may be both a very large number of pollen parents and a large number of seeds produced. However, they cannot in general be met. Restricting the number of fathers per family and the family size will restrict the chance that hybrid and pure-race siblings are in competition with each other. This reduces the opportunity for decreased viability to be favoured. Restricting family size and the number of parents will therefore make the conditions for modification even more restrictive.

Another simplifying assumption used in the models is that there is entirely random mating. Real populations are unlikely to be completely randomly mating. Restricted dispersal will mean that there is a degree of inbreeding. This will be especially true if the population is patchily distributed with little migration between patches. Inbreeding will have two effects relevant to the model: It reduces the frequency of heterozygotes and increases the relatedness between sibs. It is not obvious if inbreeding would aid or hinder the evolution of hybrid inviability. Increased relatedness between sibs will increase the chances that the modifier is favoured. However, the opportunity for selection depends on the frequency of heterozygotes, which is reduced. Also, local inbreeding will mean that there is less variation in the genotypes of fathers of one sib group. Therefore the variation in fitness within a family is further reduced.

In Chapter 2 of this thesis I discussed three ways in which selection may increase isolation between divergent taxa after they come into secondary contact: reduced migration, amelioration of selection against hybrids, and increased pre-mating isolation. A fourth possibility (reduced viability of hybrids) has been investigated in this chapter. The question therefore arises of which is the most likely response to hybridization. In attempting to answer this question one must bear in mind that it is not certain that there will always be an adaptive response to hybridization. Coalescence of the hybridizing populations or the
elimination of one by the other, are also possible outcomes (Paterson, 1982; Templeton, 1981). Empirical studies suggest that reinforcement is an at most rare phenomena. The conditions required for other kinds of modification are broadly similar to those of reinforcement. It is, therefore, likely that such modification is also rare. For reinforcement through increased hybrid inviability to evolve, the extra requirements of multiple mating and sib competition must be met. This makes such reinforcement an even less likely outcome of hybridization.

What course evolution takes is likely to be mostly determined by the available genetic variation in hybridizing populations. A major constraint on reinforcement is that, unless reinforcing alleles are are only mildly selected against in pure populations, their spread in hybrid zones will be swamped by gene flow. This effect has not been considered for the model presented here but it seems likely that which of the four possible outcomes actually evolves will be determined by their effects in pure populations.

In all models of reinforcement the opportunity for selection of reinforcing alleles is determined by the actual frequency and fitness of hybrid progeny produced. If any reinforcing allele becomes fixed it will reduce the opportunity for selection of others. For example, if an allele that increases the fitness of hybrids becomes fixed, then selection on alleles that reduce the frequency of hybridization is reduced. Similarly, if an allele that increases pre-mating isolation spreads, selection for decreased or increased (through sib competition) hybrid inviability is lessened. This decrease in selection for reinforcement, within areas of hybridization, reduces the critical level of selection against reinforcing alleles outside the hybrid zone, above which modifiers will be swamped by gene flow (Sanderson, 1989). It seems therefore, that speciation is unlikely to occur through reinforcement by the sequential fixation alleles, each increasing isolation by a small amount as Dobzhansky envisaged.
Chapter 4

The hybrid zone in *Podisma pedestris*.

4.1 *Podisma pedestris*.

*Podisma pedestris* (Orthoptera: Acridae) is an alpine grasshopper whose range extends over most of Europe, but is usually found only at altitudes between 1500 and 2500m. It is univoltine and over-winters as eggs. In the Alpes Maritimes, where it has been studied most intensely, hoppers emerge in mid June and develop through a series of five instars before becoming adult in about mid July. There is, however, much variation in emergence time and development rate influenced largely by climatic conditions.

Throughout most of its range *P. pedestris* has the standard acridid karyotype of 2n = 22 + XX (females) or X (males). All of these chromosomes are telocentric. In the southern Alpes Maritimes the species has the karyotype 2n = 20 + XX (females) and XY (males), (John & Hewitt, 1970; Hewitt & John, 1972). In these populations the X chromosomes are all metacentric and derive from a fusion between the X and one of the autosomes (John & Hewitt, 1970; Westerman & Hewitt, 1985). Examples of the two male karyotypes are shown in Fig. 4.1. The fusion produces a neo-XY sex determining system and hence the races are referred to as the XO and XY races.

A third race, characterised by the amount and distribution of the heterochromatic regions and by the position of the nucleolus organizer regions in the X chromosome, has been identified in the Pyrenees (Gonsalvez et al., 1988). As this race does not come into natural contact with either the XO or the XY, it has not been so intensively studied (thus far).
Figure 4.1 Podisma pedestris karyotypes.
top, XO male; bottom, XY male.
4.2 A narrow hybrid zone

The XO and XY races meet and form a narrow hybrid zone which runs for about 100km along the central ridge of the Alpes Maritimes between Tende in the east and Seyne les Alpes in the west (Hewitt, 1975), see Fig. 4.2. In the centre of this zone all five possible karyotypes are found: XA, X^A males; XAXA, XAX^A and X^AX^A females (Barton, 1980a) {X^A denotes the fusion between the X and the autosome}. The fusion can therefore be considered to segregate as a standard X-linked locus. The most likely explanation for the divergence of the two races is that the X^A rearrangement became locally established in the southern populations whilst they were separated from the north during a previous ice age.

In most areas that have been studied, there is a smooth cline in the frequency of the fused chromosome (Hewitt, 1975; Barton, 1979a; Barton & Hewitt, 1981a; Nichols & Hewitt, 1986; Barton et al. in prep.). In most of these areas the width of the cline, measured as the inverse of the maximum slope, is approximately 800m. In a few areas a much narrower, stepped cline is observed (Currie, 1992; Currie et al. in prep). In these areas the cline coincides with a physical barrier to dispersal.

Dispersal in *P. pedestris* has been studied several times (Barton & Hewitt, 1982; Nichols, 1984; Mason, 1988; Currie, 1992). All studies estimate the lifetime dispersal distance to be $\sigma \approx 20$ m. Given this low level of gene flow, and the $\approx 8000$ years since the races came into contact, passive diffusion of chromosomes would give a cline $\approx 3.7$ km wide (Barton & Hewitt, 1981a). This indicates that the cline must be maintained by a balance of opposing forces.
Figure 4.2 Location of the *Podisma pedestris* hybrid zone.
Open circles: populations fixed for the ancestral XO karyotype. Filled circles: XY populations. Crosses: mixed populations. Shaded areas are regions higher than 2500m. a) Seyne les Alpes, b) Col de la Lombarde, c) Tende.
4.3 A narrow tension zone

Since there is no detectable environmental difference between the ranges of the two races (Nichols, 1984; Nichols & Hewitt, 1988; Jackson, 1992), the cline is best understood as an example of a tension zone (Key, 1968; Barton & Hewitt, 1985) maintained by a balance between dispersal and selection against heterokaryotypes. The evidence of selection against hybrids, presented below, rules out the possibility that the cline is maintained by hybrid superiority (Moore, 1977). A consideration of the width of the cline and the dispersal rate indicates that selection of only 0.55% against heterokaryotypes is required to maintain the cline (Barton, 1980a). This calculation was based on a model of a cline at an autosomal locus. For an X-linked locus, selection is only possible in the females which carry two thirds of the X chromosomes. Therefore, the selection required to explain the cline is $s \approx 0.8$: a small difference.

The most straightforward a priori explanation of such selection is that heterokaryotypes have reduced fertility due to the production of unbalanced gametes through non-disjunction during meiosis (White, 1973). Barton (1980a) tried to measure the rate of non-disjunction in heterokaryotypes. The rate of non-disjunction was no greater in heterokaryotypes than homokaryotypes. The rate may have been only as large as 1.6%; a rate not incompatible with the selection required to explain the cline.

Non-disjunction is a plausible explanation but not the only one. Associated with the chromosomal difference there are differences in DNA content between the races (Westerman et al., 1987). This indicates that during the process of fusion between the chromosomes some DNA was lost. Also associated with the chromosomal fusion are differences in rDNA (Dallas et al., 1988) and morphology (Barton, 1979a, but see Chapter 6). It is possible that selection may be acting on one of these factors by means other than non-disjunction. However, it will be seen (below and Chapter 5) that differences in the fertilization pattern of the two races gives rise to a measurable amount of selection on the chromosome.
The hybrid zone has been shown to be more complex than a simple chromosomal cline. Rather, it appears that the fused chromosome is only a weakly selected marker for other, more fundamental differences, between the races. Lab and field studies have repeatedly shown that hybrids between the two races suffer a reduction in fitness of approximately 50%. This is manifest as a reduction in hatch rate, early viability and developmental rate (Barton, 1980a; Barton & Hewitt, 1981b; Nichols, 1984; Nichols & Hewitt, 1988; Jackson, 1992). The reduced viability of F₁ hybrids indicates that a proportion of the selection can be attributed to underdominance at loci fixed for alternate alleles in the two races. However, the average viability of individuals taken from the centre of the cline is lower than that of F₁ hybrids (Barton & Hewitt, 1981b). The genotypes of these individuals are the result of many generations of recombination between the two pure race genomes. This indicates that part of the selection against hybrids may be due to the break up of epistatic gene complexes. The area over which there is reduced viability is much wider than would be expected from 50% selection against hybrids, indicating that hybrid inviability must be due to effect of weak selection of many genes. Barton & Hewitt (1981b) estimate that at least 150 loci contribute to the hybrid inviability with selection ≈3% on each.

Electophoretic surveys of 21 allozyme loci have been made in various areas of the hybrid zone (Halliday et al., 1983, 1984). These showed that genetic variation between populations spanning the hybrid zone is no greater than between populations of the same race. Changes in the frequency of the fused chromosome are not correlated with allele frequency change at any of the electrophoretic loci. The strong selection against hybrids does indicate that there must be large genetic differences between the races: these differences have simply not been identified in the enzyme surveys. If they were, one would expect to see a large number of coincident clines.

Further evidence of strong selection against hybrids comes from studies of areas of the hybrid zone that coincide with physical barriers to gene flow. In these areas a stepped cline is observed with long tails of introgression. In these areas the step is larger, and the
tails of introgression steeper, than can be explained by the reduced dispersal across the barriers. The data are best described by models that include an element of selection against hybrids due to the effects of many genes (Currie, 1992; Currie et al., in prep).

On a large scale the hybrid zone runs along the central ridge of the Alpes Maritimes. This is probably more or less where the two races met after the last ice age. On a finer scale it has been shown that the centre of the zone follows areas of low population density (Nichols & Hewitt, 1986). This is further evidence that the zone is a tension zone since such zones are expected to stabilize in areas of low population density (Barton, 1979b). In other areas it has been shown that the density of juveniles in XY populations is consistently higher than in either XO or hybrid populations (Jackson, 1992). The length of time over which the hybrid zone has been studied does not allow for direct assessment of its stability. However, by the time the populations reach the adult stage this difference is evened out (Jackson, 1992). If juveniles contribute significantly to gene flow, this would suggest that the range of the XY race should still be expanding. Studies of nymph dispersal have shown that they do move significant distances (Mason, 1988). Simulations of localised density and dispersal indicate that, although there XY’s have a higher density, the movement of the zone is prevented by a density trough (Jackson, 1992).

4.4 Reproductive isolation between the races.

The strong selection against hybrids represents a degree of post-mating reproductive isolation between the races which must act to reduce gene flow between them. As in other chromosomal hybrid zones in orthoptera, the isolation brought about by non-disjunction, attributable to the chromosome structure itself is at most weak (Shaw, 1981). Other genetic incompatibilities between the races are far more important in preventing gene flow between them. However, the barrier to gene flow is by no means absolute. The lack of correlation between allele frequencies at enzyme loci and the chromosome frequency indicates that alleles may pass through the hybrid zone. The barrier strength of the hybrid
zone, estimated from the shape of the cline across regions of reduced dispersal, is approximately 1.5km (Currie, pers comm).

Mason (1988) looked for evidence of pre-mating isolation between the races. He found no evidence of either temporal differences or local spatial isolation between the chromosomal forms in hybrid populations. The possibility of assortative mating has been studied both in hybrid, field, populations and in enclosures containing a mix of both the pure races (Hewitt et al., 1987; Mason, 1988). No evidence of assortment was found. It has, however, been shown that sperm usage by females that have mated with both X0 and XY males is highly skewed. In such females there is a strong tendency for eggs to be fertilized by sperm carrying the same X chromosome as the female - i.e. there is assortative fertilization (Hewitt et al., 1987, 1989, and discussed in Chapter 5). This assortment is seen, at comparable strengths, in both hybrid populations and in crosses between the pure populations. There is no evidence that it has evolved as a result of reinforcement, rather it is simply a 'fortuitous' effect of divergence of the two races.

4.5 New work on the Podisma pedestris hybrid zone.

This thesis presents two pieces of work relevant to the hybrid zone in P. pedestris. In Chapter 5, I consider a model of assortative fertilization. It is shown that assortment to the degree observed in P. pedestris will lead to strong selection, capable of maintaining the chromosomal cline without invoking selection due to non-disjunction or other, unseen, processes. This does however present a problem. The selection produced is much stronger than predicted by the shape of the zone and the measurements of gene flow.

In Chapter 6 I present a survey of morphometric variation across the hybrid zone at the Col de Lombarde. The results of this differ from those of Barton (1979a) in that, although they show an increase in size across the hybrid zone, there is no evidence of a morphometric difference between the two karyotypes per se.
Chapter 5

The effect of homogamy on the chromosomal cline in *Podisma pedestris*.

5.1 Introduction

In this chapter, the consequences of assortative fertilization at an X linked locus are investigated. This work has been prompted by the observation of homogamy in the *Podisma pedestris* hybrid zone.

5.1.1 Assortative fertilization in *Podisma pedestris*.

It has recently been shown that sperm usage by female *P. pedestris* is far from straightforward. Hewitt *et al.* (1989) took animals from two populations spanning the hybrid zone at Seyne. Each female was mated sequentially to a male from both populations. There were, therefore, four classes of cross (2 female classes & 2 mating orders, see Table 5.1). If sperm usage were random, one would expect equal proportions of female embryos to have been fertilized by X^A and XA sperm. From the distribution of karyotypes within egg pods, the karyotype of the father can be determined. The first male to mate secures a large proportion of the fertilizations. Also evident in the data is an excess of homokaryotypic embryos: there is assortative fertilization (Table 5.1).

This work backs up previous findings from the field (Hewitt *et al.*, 1987). In this study Hewitt *et al.* examined the egg pods laid by gravid females collected from the centre of the hybrid zone at Tende. An excess of homokaryotypes was observed. As the mating history of these females is unknown, the interpretation of the homokaryotypic excess is not so straightforward. The data were fitted to various models of mating frequency, mating
preference and sperm utilization. Observations of mating pairs in the field and in field enclosures indicates that mating is random with respect to karyotype (Hewitt et al., 1987; Mason, 1988). The distribution of karyotypes in the embryos of gravid females is best explained by a model of random mating and assortative fertilization. The model with the maximum likelihood is one in which 90% females have mated twice, and, given that a female has mated to both XA and X^A males, the proportion of homogamic fertilizations is 0.55 for X^AX^A females and 1 for XAXA females (Hewitt et al., 1987).

The degree of homogamy in the centre of the zone is comparable to that seen in the inter-racial crosses. That this is so suggests that it has not evolved as a result of reinforcement but is due to divergence before the races came into secondary contact after the last ice age. It also suggests that it must either be a pleiotropic effect of the fusion itself or due to genes maintained in linkage disequilibrium with the fusion within the cline.

The mechanism by which this homogamy is brought about is unknown. Even without this knowledge the consequences of assortment can be considered. There is a body of theory concerning the consequences of assortative mating, both in single populations and in the context of hybrid zones. The outcome of assortment is likely to be similar in both situations since both assortative mating and assortative fertilization produce an excess of homogamic offspring.

<table>
<thead>
<tr>
<th>Fertilization by X^A sperm.</th>
<th>1st male X^A</th>
<th>2nd male XA</th>
<th>1st male XA</th>
<th>2nd male X^A</th>
<th>Average proportion of homogamic fertilization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X^AX^A</td>
<td>0.89 (0.89) n = 6</td>
<td>0.38 (0.38) n = 5</td>
<td>0.64</td>
<td></td>
<td></td>
</tr>
<tr>
<td>XAXA</td>
<td>0.49 (0.51) n = 6</td>
<td>0.29 (0.71) n = 7</td>
<td>0.61</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5.1 Homogamy in Podisma pedestris. Proportion of fertilizations by X^A and XA sperm in double matings with animals from pure populations. Figures in brackets are the proportion of homogamic fertilizations, n is the number of crosses. A summary of Table 1 in Hewitt et al., 1989.
5.1.2 The consequences of homogamy

Assortative mating refers to the frequently observed correlation between phenotypes of mating pairs in a population. A positive correlation indicates positive assortment, a negative one, negative assortment. In the interests of brevity, I shall refer to positive assortative mating simply as "assortative mating". Assortative mating has consequences similar to those of inbreeding. The distinction between the two processes is that inbreeding reflects a correlation between genotypes and assortative mating a correlation between phenotypes of mates (Lewontin et al., 1968). Obviously, in so far as phenotype reflects genotype, the two processes will have similar consequences.

A distinction can be made between the mating preferences of individuals and the mating pattern of the population. The pattern of mating in a population does not necessarily directly represent the preferences of the constituent individuals. Assortative mating can also arise for other reasons. If a species is polymorphic for genes that are adapted to different aspects of the environment assortment may result. For example, if there is variation in host preference (e.g. in phytophagous insects or parasites) and mating occurs within these hosts, assortative mating for the preference genes will result. The flowering plant Lythrum salicaria is insect pollinated. It has been shown that most pollinator visits to a plant are by insects which have previously visited a plant of very similar height - there is, therefore, assortative mating for height although it does not necessarily have anything to do with mating preference of the plant (Levin & Kerster, 1973).

5.1.2.1. Theory of assortative mating

The simplest models of assortative mating consider a polymorphic population with strict monogamy (O'Donald, 1960,1980; Felsenstein, 1981). These are models of mixed random and assortative mating. It is assumed that a certain fraction of the population mate assortatively and the rest at random. Assuming strict monogamy ensures that all individuals
secure a mate irrespective of whether they mate assortatively or not. Therefore, there is no selection operating in the population - assortment alters the distribution of genotype but not gene frequencies. Assortment reduces the frequency of heterozygotes in the population. Implicit in these models is the assumption that all matings are equally fertile. Karlin and Scudo (1969) and Scudo & Karlin (1969) extended the models by dropping this assumption. They assumed that assortative matings occur prior to the random ones and that early matings have higher fertility than the later ones. If assortment is for a dominant phenotype, situations such as these almost always lead to an unstable population. The assortment induces a selective force on the population causing one or other of the alleles to become fixed (Scudo & Karlin, 1969). Karlin & Scudo give no reason why earlier matings should be more fertile. However, a reason was provided by Darwin (1871). He suggested that early matings would be advantageous since the parents are likely to be in better nutritional condition and possibly able to rear a second clutch of young. Models of this process have confirmed Darwin's verbal arguments (O'Donald, 1980; Kirkpatrick et al., 1990)

In the absence of dominance at the assortative mating locus, stable polymorphisms may result if the tendency to assort is greater in the heterozygotes than in the homozygotes - a situation analogous to over-dominance in the case of natural selection (Karlin & Scudo, 1969).

In polygamous species, the consequences of assortment may be very different. In the case of monogamy, unless there is an unequal sex ratio, there is no variation in mating success. Selection only arises because of the different fertilities of assortative and random matings. In polygamous species, the opportunity exists for variation in mating success and so a change in gene frequency may result from the mating scheme alone. This result arises from a consideration of a class of models known as the 'mass action' models (Karlin & Scudo, 1969; Scudo & Karlin, 1969; Moore, 1979). In these, the probability of an individual mating depends on the probability of encountering a mate with the same phenotype as itself. This will obviously result in frequency dependent selection against rare
forms: a rare phenotype is unlikely to encounter a suitable mate and so suffers a disadvantage. Because of this, one would not expect to find polymorphic populations in which there is assortative mating of this form (Moore, 1979). Stable polymorphisms can only be maintained in single populations if heterozygotes enjoy a substantial fitness advantage (Moore, 1979).

5.1.2.2. Assortative mating in hybrid zones

What will be the consequence of homogamy in a hybrid zone? Consider a tension zone maintained by a balance between selection against heterozygotes and dispersal in which mating is at random. The width of the zone is determined by the relative fitness of the heterozygotes and the rate of gene flow. Consider, now a similar zone but in which there is assortative mating. The differences between the two zones will depend on the nature of the assortative mating. If assortment is such that, alone, it produces no changes in gene frequency (e.g. O' Donald, 1960), then assortment will widen the zone (Sanderson, 1989). This is because assortment reduces the frequency of heterozygotes and so reduces the opportunity for selection to act. If assortment is absolute (all individuals mate assortatively), there is no selection and no cline is maintained since eventually all heterozygotes are eliminated. The two hybridizing populations will simply merge, but the two forms will remain distinct since there is no gene flow between them.

For polygamous species, assortment is expected to lead directly to changes in allele frequency. Such assortment will have two effects: 1) it will reduce the frequency of heterozygotes at the selected locus, 2) selection against the rare allele means that assortment itself will contribute to the maintenance of the cline (Moore, 1981). Mallet & Barton (1989a) have shown that clines maintained by frequency dependent selection are very similar to those maintained by selection against heterozygotes: by analogy, clines maintained by assortment may also be similar.
5.1.2.3 Clines maintained by homogamy.

In a number of well studied hybrid zones there is a tendency towards assortative mating (e.g. see Harrison, 1990). In most instances assortative mating in the field is inferred from differences in song or display characteristics between the races or from laboratory experiments on mating preferences. In only one instance has the direct effect of assortative mating in maintaining clines been considered (Johnson et al., 1990). Narrow hybrid zones separate populations with different coiling direction in the snail species *Partula suturalis* (Johnson, 1982). The direction of coiling is determined by a single gene (Murray & Clarke, 1976). Laboratory studies have shown that matings between opposite coiled pairs are infrequent and when they do occur are generally unsuccessful because of physical incompatibilities (Lipton & Murray, 1979; Johnson, 1982). One would, therefore, expect frequency dependent selection against rare morphs. Johnson et al. (1990) have simulated this situation and shown that clines can be maintained by a balance between dispersal and assortative mating. It has been suggested that a change in coil direction could lead to rapid speciation in snails because of the resulting assortment (Gittenberger, 1988; Orr, 1991). In the case of *Partula* this is unlikely since, even with absolute assortment, there is only a weak barrier to genetic exchange between the morphs. This is in part because the expression of the coiling genotype is delayed by one generation: the phenotype of a snail is an expression of the maternal genotype. The situation in *Partula* is in fact slightly more complex than Johnson et al. (1990) suggest. *P. suturalis* occurs sympatrically with other related species. In all areas the direction of coiling in *P. suturalis* is opposite to that of other sympatric species. It is, therefore, likely that the avoidance of mating with other species contributes to the maintenance of clines in coil direction (Murray & Clarke, 1980).
5.1.3 Assortative fertilization and the chromosomal cline in *Podisima pedestris*: a paradox?

An appropriate model of assortment in *P. pedestris* would necessarily be of the 'mass-action' type. Since *P. pedestris* mate multiply and utilize sperm assortatively, there is scope for variation in male mating success. The success of a particular karyotype will depend on the frequency of the preference for it in the population. Frequency dependent selection will, therefore, result.

The observed homogamy in *P. pedestris* is strong and will, therefore, give rise to strong selection. This raises a problem - a consideration of the width of the chromosomal cline and the amount of dispersal indicates that selection directly on the chromosome must be weak (Barton, 1980a, Barton & Hewitt, 1981a). Also, one would expect that assortment would cause a deficit of heterokaryotypes in the centre of the hybrid zone - this is not observed (Hewitt et al., 1987; Barton, 1980a; Nichols, 1984; Mason, 1988). This chapter will address this problem by considering a model of assortment in *P. pedestris*.

5.1.4 Outline.

A simple model of assortment is developed and applied first to single, isolated, populations. This consists of a set of difference equations that describe changes in population parameters over one generation. These are then simplified and applied to the situation of two races coming into secondary contact and forming a hybrid zone. A diffusion approximation is used in which it is assumed that assortment and migration are continuous processes. Simple expressions are derived for the expected cline widths and heterozygote deficit across the cline. These simple expressions are compared to simulations of migration and assortative mating in a 'stepping stone' model. Apart from its relation to the situation in *P. pedestris* the model is also of interest since it considers a tension zone at an X-linked locus. Previous models of tension zones have considered only autosomes
(Bazykin, 1969; Barton, 1979b). Owen (1986) considered environmentally determined clines at X-linked loci and showed that they are identical to clines at autosomes provided that there is no dominance, equal dispersal of the two sexes and that the average effect of each allele is the same in both sexes.

The results of this model are compared with the available data from the *Podisma pedestris* hybrid zone.
5.2 The model

I will consider the dynamics of the model used by Hewitt et al. (1987) to estimate the degree of assortment. This is a rather crude model since it only allows each female to mate twice rather than several times. However, it has the advantage that the actual data can be incorporated without any extra analysis. Before discussing the details of the model a method of describing populations of X-linked genes is considered.

5.2.1 Describing populations of X linked genes

Generations are discrete and non-overlapping. A single X linked gene with two alleles (\( \mathcal{A} \) and \( \mathcal{a} \)) which determine the degree of non-random fertilization is considered. Two populations of these genes are considered - those carried by males and females. In the female population both the allele frequencies (\( P_f, Q_f \)) and the genotype frequencies (\( u, v \) and \( w \)) are needed. In the male population, allele frequencies and genotype frequencies are equivalent (\( P_m, Q_m \)). The notation used is summarised in Table 5.2. Since the males and females are considered as separate populations \( P_m + Q_m = 1 \) and \( P_f + Q_f = 1 \). \( P_m \) is simply the frequency of \( \mathcal{A} \) males. In females:

\[
\begin{align*}
P_f &= u + \frac{1}{2} v \\
Q_f &= w + \frac{1}{2} v
\end{align*}
\]

(5.1)

The overall frequencies of the \( \mathcal{A} \) and \( \mathcal{a} \) alleles are given by

\[
\begin{align*}
P &= \frac{1}{3} \left( P_m + 2P_f \right) \\
Q &= \frac{1}{3} \left( Q_m + 2Q_f \right)
\end{align*}
\]

(5.2)
A difference in allele frequency between the sexes is defined as $\Phi = P_f - P_m$. Allele frequencies in the male and female populations can be defined in terms of the overall allele frequency and $\Phi$.

\[
\begin{align*}
    P_m &= P - \frac{2\Phi}{3} \\
    Q_m &= Q + \frac{2\Phi}{3} \\
    P_f &= P + \frac{\Phi}{3} \\
    Q_f &= Q - \frac{\Phi}{3}
\end{align*}
\]

Genotype frequencies $u$, $v$ and $w$ can be defined in terms of the deficit of heterozygotes, $F$, and the allele frequencies (Hartl & Clark, 1989).

\[
\begin{align*}
    u &= P_f^2 + P_f Q_f F \\
    v &= 2 P_f Q_f (1-F) \\
    w &= Q_f^2 + P_f Q_f F
\end{align*}
\]

$F$ is Wright's inbreeding coefficient $F_{IS}$, and measures the deviation from Hardy-Weinberg equilibrium in the female population. $F$ may vary between -1 and +1. When the population is at the Hardy-Weinberg equilibrium, $F = 0$. A positive value indicates a deficit of female heterozygotes, a negative one indicates an excess. The population can be entirely described by the gene frequency $P$, the inbreeding coefficient $F$ and difference between the sexes $\Phi$. 
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>u</td>
<td>Frequency of AA in females.</td>
</tr>
<tr>
<td>v</td>
<td>Frequency of Aa in females.</td>
</tr>
<tr>
<td>w</td>
<td>Frequency of aa in females.</td>
</tr>
<tr>
<td>P_m</td>
<td>Frequency of A in males.</td>
</tr>
<tr>
<td>Q_m</td>
<td>Frequency of a in males.</td>
</tr>
<tr>
<td>P_f</td>
<td>Frequency of A in females.</td>
</tr>
<tr>
<td>Q_f</td>
<td>Frequency of a in females.</td>
</tr>
<tr>
<td>P</td>
<td>Frequency of A in the population as a whole.</td>
</tr>
<tr>
<td>Q</td>
<td>Frequency of a in the population as a whole.</td>
</tr>
<tr>
<td>Φ</td>
<td>Difference in frequency of A between males and females $Φ = P_f - P_m$</td>
</tr>
<tr>
<td>F</td>
<td>Inbreeding coefficient measures deviations from Hardy-Weinberg equilibrium.</td>
</tr>
<tr>
<td>h</td>
<td>Frequency of homogametic fertilizations of homozygous females carrying both A and a sperm.</td>
</tr>
<tr>
<td>x</td>
<td>$x = 2h - 1$. ($x = 1$: complete assortment)</td>
</tr>
<tr>
<td>r</td>
<td>Re-mating frequency: the proportion of females that mates twice.</td>
</tr>
</tbody>
</table>

Table 5.2 Summary of notation used in the homogamy model.
5.2.2 A model of assortative fertilization.

In the paper in which homogamy in *Podisma pedestris* was first demonstrated, its magnitude was determined using a model of mating, re-mating and assortative fertilization (Hewitt *et al*., 1987). The same model will be used here to find the effect of homogamy on gene and genotype frequencies.

Each female mates once, then either re-mates (with probability r) or does not (probability 1-r). Each mating is at random with respect to genotype. The probability of mating with a male of a particular genotype is given simply by that genotype's frequency in the male population. In those females that have mated with both types of male, fertilization is dependent on female genotype. Eggs of homozygous females are fertilized to produce homozygotes and heterozygotes in the ratio $h : 1-h$. $h$ may vary between 0 and 1. If $h = \frac{1}{2}$ fertilization is random (as for heterozygous females), $h > \frac{1}{2}$ indicates positive assortment, $h < \frac{1}{2}$ negative assortment. For simplicity I have assumed that the degree of homogamy is the same in both female homozygotes. This model is summarised in Fig 5.1. The proportion of eggs, of a particular individual, fertilized by a particular male genotype are given by multiplying the coefficients along the tree from left to right. The proportion of fertilizations of a particular class of individuals is given by this product multiplied by the frequency of that class. Homogamy does not affect the frequency of fertilizations by non-X bearing sperm.
Figure 5.1 Description of female mating history.

A probability tree illustrating the possibilities for fertilization of a female's eggs. The probability of any particular outcome is given by the product of the terms from left to right along the branches.
5.2.2.1 Single populations.

The genotype frequencies among the progeny of females of different genotypes is summarised in Table 5.3. A parameter $x$ has been substituted for $h$ such that $x = 2h - 1$. $x$ therefore varies between -1 (complete disassortment) to +1 (complete assortment).

<table>
<thead>
<tr>
<th>Females</th>
<th>Progeny</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genotype</td>
<td>Frequency</td>
</tr>
<tr>
<td>AA</td>
<td>$u$</td>
</tr>
<tr>
<td>aa</td>
<td>$w$</td>
</tr>
<tr>
<td>Aa</td>
<td>$v$</td>
</tr>
</tbody>
</table>

**Table 5.3** Genotypes of female progeny.

Multiplying by the female genotype frequency and summing down the rows of Table 5.3 gives expressions for genotype frequencies in the following generation (denoted by a prime).

$$
\begin{align*}
    u' &= P_m(u + \frac{v}{2}) + rx P_m Q_m u \\
    v' &= u Q_m + \frac{v}{2} + W P_m - rx P_m Q_m (u + w) \\
    w' &= Q_m (w + \frac{v}{2}) + rx P_m Q_m w
\end{align*}
$$

Which rearrange to:

$$
\begin{align*}
    u' &= P_m P_f + u P_m Q_m r x \\
    v' &= P_m Q_f + P_f Q_m - P_m Q_m r x (w + u) \\
    w' &= Q_m Q_f + w P_m Q_m r x
\end{align*}
$$

(5.5)
The allele frequency in the females in the following generation is given by

\[ P_f' = u' + \frac{1}{2} v' \]  

(5.6)

Substituting Eqs. 5.5, 5.4 and 5.1 into 5.6 and rearranging gives

\[ P_f' = \frac{1}{2}(P_m + P_f) + \frac{1}{2} rx P_m Q_m (P_f - Q_f) \]  

(5.7)

Similarly

\[ Q_f' = \frac{1}{2}(Q_m + Q_f) - \frac{1}{2} rx P_m Q_m (P_f - Q_f) \]  

(5.8)

Allele frequency in the male population is simply the same as that in the female population of the previous generation.

\[ P_m' = u + \frac{1}{2} v = P_f \]  

(5.9)

The overall gene frequency in the following generation, \( P' = \frac{1}{3}(P_m' + 2P_f') \), reduces to

\[ P' = P + \frac{1}{3} rx P_m Q_m (P_f - Q_f) \]  

(5.10)

It is clear from this that selection is frequency dependent. If \( P_f > Q_f \), allele frequency increases but decreases if \( Q_f > P_f \). Assortment, therefore, causes selection which favours the most common allele. The gene frequency difference between the sexes, \( \Phi' = P_f' - P_m' = P_f' - P_f \), which reduces to

\[ \Phi' = -\frac{1}{2} \Phi + \frac{1}{2} rx P_m Q_m (P_f - Q_f) \]  

(5.11)

It can be seen from Eq. 5.10 and 5.11, that the terms \( r \) and \( x \) always appear as a product. They can, therefore, be considered as a single variable, \( rx \), and will only be decomposed
when I come to apply the model to the *P. pedestris* data. Substituting Eqs. 5.3 (which relate allele frequency in each sex to the overall allele frequency) gives simpler expressions. After some manipulation:

\[ P' = \frac{1}{2} \left( 2P - \frac{\Phi}{3} + B \right) \]  
\[ (5.12) \]

\[ q' = \frac{1}{2} \left( 2Q + \frac{\Phi}{3} - B \right) \]  
\[ (5.13) \]

\[ P' = P + \frac{1}{3} B \]  
\[ (5.14) \]

\[ \Phi' = -\frac{1}{2} \Phi + \frac{1}{2} B \]  
\[ (5.15) \]

where

\[ B = P_m Q_m (P_f - Q_f) = PQ(P - Q) - 2PQ\Phi + \frac{2\Phi}{3} - \frac{8\Phi^3}{27} \]  
\[ (5.16) \]

The deviation from Hardy-Weinberg equilibrium in the following generation may be determined by rearranging Eq. 5.4.

\[ F' = 1 - \frac{\nu'}{2 \, P'_f Q'_f} \]

\[ = \frac{\frac{1}{2} \Phi^2 + \frac{\Phi}{18} - \frac{\Phi^2}{18} - \frac{(PQ + \frac{2\Phi}{3})(P - Q)}{2PQ + (P - Q)}}{B(P - Q)^2} \]  
\[ (5.17) \]

but does not reduce down to a more manageable expression without some simplifying assumptions.

The dynamics of the model are illustrated in Figs 5.2, 5.3 and 5.4. These show the change in *P*, *\Phi* and *F* over a number of generations for varying strengths of assortment.
Assortment generates positive frequency dependent selection, the least common allele being eliminated (Eq. 5.12). The only equilibria in gene frequency are at $P = 0$, $P=1$ or $P = \frac{1}{2}$. But the $P = \frac{1}{2}$ equilibrium is unstable and is only present if there is no difference in frequency between the sexes. Assortment will also generate differences in frequency between the males and females (Eq. 5.14). This is because it is only the gene frequency in the females that responds directly to assortment. Gene frequency in males follows that of females but lags behind by one generation. Selection acts directly on the males, the rarer allele achieving proportionally fewer fertilizations, but is manifest as changes in female allele frequency. No difference between the sexes is generated if initially allele frequency $P = \frac{1}{2}$.

For other initial conditions $|\Phi|$ increases as gene frequency changes. However as one or other allele moves towards fixation, $\Phi$ necessarily tends towards zero. No permanent difference in gene frequency can be maintained.

If one considers a population in which initially there is no difference between the sexes, the dynamics become a little clearer.

$$P' = P + \frac{1}{3}rPQ (P-Q), \quad \text{and} \quad \Phi' = \frac{1}{2}rPQ(P-Q).$$

There will be no change in gene frequency only if $P = 0, 1,$ or $\frac{1}{2}$. Under these circumstances there will also be no change in $\Phi$. For other initial allele frequencies there will be a change in frequency of both parameters: $P$ will tend towards 0 or 1 and $\Phi$ will first increase then tend towards zero. This behaviour is illustrated in Fig. 5.5 This shows a plot of $\Phi$ against $P$ for various starting frequencies (plotted by iterating the exact Eqs. 5.14 and 5.15). Note that the maximum $\Phi$ reached by the population is always very small. Over the period during which $P$ is changing, the deviation from Hardy-Weinberg frequencies (F) also increases and then decreases as $P$ tends towards 0 (Fig. 5.3). When one or other allele is fixed the population must be in equilibrium.

For a population at the unstable equilibrium and with $\Phi = 0$, F increases to some equilibrium value which can be determined by solving Eq. 5.17 for $P = \frac{1}{2}$ giving:-
\[ \hat{F} = \frac{rx}{4-rx} \]  

(5.18)

The rapid approach to this equilibrium value is shown in Fig 5.6. The maximum possible heterozygote deficit is 1/3 when \( rx = 1 \), (i.e. when all females mate twice and homogamy is absolute). In reality one would not expect to observe this situation since it is only attained when the population is at the unstable \( p = 0.5 \) equilibrium. Because of the frequency dependent selection, one would only expect to see monomorphic populations except when the polymorphism is maintained by a balance between, for instance, assortment and migration. In this situation deviations from Hardy-Weinberg equilibrium are generated by both selection and migration. Below I will compare the relative effects of these two processes and show that the equilibrium expression for \( F \) given above is a good approximation to that expected in the centre of clines.

It is interesting to note, that even when assortment is complete \( (x=1) \) and all females mate at least twice \( (r=1) \), heterozygous individuals are still present when the population reaches equilibrium \( \hat{F} = 1/3 \), Eq. 5.18). Since mating is at random, there will always be some females that mate only with males with a different \( X \) chromosome to themselves. In these females, the tendency to assortment cannot be expressed.

Systems of assortative mating give different results. If a proportion of the population, \( d \), mate assortatively, the equilibrium reached is \( \hat{F} = d \) for two alleles at a locus with dominance. If there is no dominance, or the locus is \( X \)-linked and heterozygous females mate at random, \( \hat{F} = d/(2-d) \), (O'Donald, 1960; Scudo & Karlin, 1969). Hewitt et al. (1987) give different expressions for \( \hat{F} \). They state that \( \hat{F} = d \), for an autosomal locus and \( \hat{F} = d/(2-d) \) for an \( X \)-linked one. The difference does not really depend on the location of the locus but on whether or not there is dominance. If assortment is absolute in these systems of mating the equilibrium population is one in which there are no heterozygotes. Complete assortment would lead to an absolute barrier to gene flow between the two
morphs. With assortative fertilization the barrier is incomplete even if there is absolute assortment since with random mating there will always be some females mated only to males with a different karyotype to themselves.

Figure 5.2 Allele frequency vs generation

Gene frequency $P$ vs generation. $r = 1, 0.75, 0.5, 0.25$ (left to right). Initial conditions: $p = 0.51$ or $p = 0.49$, $\Phi = 0$, $F = 0$. 
Figure 5.3 Heterozygote deficit vs generation.
F vs t for r_x = 1, 0.75, 0.5, 0.25 (top to bottom), 1. Initial conditions: P = 0.49, F = 0, F = 0.

Figure 5.4 Allele frequency difference between sexes vs generation.
Φ vs t for r_x = 0.1 and 0.5. Initial conditions: P = 0.49, or 0.51, and F = 0.
Figure 5.5 Allele frequency difference between sexes vs allele frequency.
For $r_x = 1$ and initial allele frequencies of 0.25, 0.49, 0.51, 0.75. Initially $\Phi = 0$, $F = 0$.
The curves are trajectories describing the changing state of the population with time.

Figure 5.6 Heterozygote deficit vs generation for $P = \frac{1}{2}$
$F$ vs $t$ for $P = 0.5$, $r_x = 0.25$, 0.5, 0.75, 1. Initial conditions: $F = 0$, $\Phi = 0$. 
5.2.2.2 Approximations for weak assortment

The difference equations given above for the changes in allele frequency, heterozygote deficit, and difference between the sexes (Eq.'s 5.14, 5.15, and 5.17) describe exactly the dynamics of the model being considered. The point of all this algebra is, however, to consider the effect of assortative mating in clines. These equations are too complex to be easily applied to the situation of a geographically distributed population (other than by simulation). To gain a better understanding of this situation some simplifications must be made.

An approach often taken is to assume that the forces (i.e. assortment) acting on the population are weak. In this way, many of the higher order terms can be removed from the equations since they will be of negligible magnitude. Although the ultimate aim is to relate this model to assortment in *P. pedestris*, which is not weak (Hewitt *et al.*, 1987, 1989) this is the approach that I will take. However, I shall then show by comparison of the exact and approximate equations that the weak assortment approximations are good even for fairly strong assortment.

Approximate equations

In each generation the allele frequency difference between the sexes, $\Phi$, changes by a factor of $-\frac{1}{2}$ and also by an amount $\frac{1}{2}r_x B$ (Eq. 5.15). The dominant factor is $-\frac{1}{2}$. Since the maximum value $B$ can take is of order one, then even if initially $\Phi$ is large it will rapidly be reduced to some small value of order $r_x$. If $r_x$ is small then terms in $\Phi^2$ and $r_x \Phi$ will be negligible and can be removed from the equations (5.14, 5.15, and 5.17). This gives:

$$P' = P + \frac{1}{3} r_x PQ(P-Q) \quad (5.19)$$

$$\Phi' = -\frac{1}{2} \Phi + \frac{1}{2} r_x PQ (P-Q) \quad (5.20)$$

-- 138 --
\[ F' = \frac{2 \cdot r \cdot P^2 Q^2 (1 + F)}{2PQ + \Phi \left(\frac{P-Q}{3}\right) - r \cdot PQ(P-Q)^2} \]  

Since \( \Phi \) changes rapidly compared to \( P \) it will soon reach a 'quasi-equilibrium' dependent only on \( P \) and obtained by solving \( \Phi' = \Phi \). This is similar to the case of linkage disequilibrium in two locus systems (Nagylaki, 1976). The quasi equilibrium value of \( \Phi \) can be found by solving Eq. 5.20, giving

\[ \Phi = \frac{1}{3} r \cdot PQ(P-Q) \]  

Substituting this value into Eq. 5.21 gives a simpler, approximate expression, for the change in heterozygote deficit:

\[ F' = \frac{r \cdot PQ (1+F)}{1 - \frac{4}{9} r \cdot PQ(P-Q)^2} \]  

Solving this gives an expression for the quasi-equilibrium reached in \( F \)

\[ F' = \frac{9 \cdot PQ \cdot r \cdot x}{9 - r \cdot x \cdot PQ(13-16 \cdot PQ)} \]  

For the particular case of \( P = \frac{1}{2} \), this is the same as the exact equilibrium (Eq. 5.16). \( \tilde{F} \) gives an approximate value for the heterozygote deficit in a population with a particular allele frequency (which will be moving towards fixation). It will be seen in section 5.2.3.1.2 that \( \tilde{F} \) also gives a good approximation of the heterozygote deficit expected across a cline maintained by assortative fertilization and migration if measured immediately after fertilization.
5.2.2.3 Comparing the exact and approximate equations

Equations 5.14, 5.15 and 5.17 give an exact description of the population. An approximate description is given by Eq. 5.19 (which describes allele frequency change) and Eqs 5.22 and 5.24 (which give quasi equilibrium values for \( F \) and \( \Phi \)). The accuracy of the approximations can be assessed by comparing them to the exact iterations. Figs 5.7, 5.8, and 5.9, show this comparison. The exact iterations are shown by the dots, and the approximations by the solid lines. It can be seen from these graphs that the approximate equations give a good representation of the dynamics of this population. All of these comparisons were made for populations initiated in Hardy-Weinberg and with no allele frequency difference between the sexes.

A situation in which the approximate equation in \( P \) is unreliable is around the unstable equilibrium at \( P = \frac{1}{2} \). This is only an equilibrium if \( \Phi = 0 \). For \( \Phi \neq 0 \), the population moves towards fixation. As there is no \( \Phi \) term in the approximate expression for \( P' \), this facet of the system is not captured. The approximations have an equilibrium at \( P = \frac{1}{2} \) irrespective of \( \Phi \).

![Figure 5.7](image_url)

**Figure 5.7** Allele frequency vs generation: approximate and exact equations. \( P \) vs \( t \) for the exact (dots) and approximate (lines) equations. \( r_x = 1, 0.75, 0.5, 0.25 \) (top to bottom line). Initial conditions: \( P = 0.49, F = 0, \Phi = 0 \).
Figure 5.8 Heterozygote deficit vs generation: approximate and exact equations. $F$ vs $t$ for the exact (dots) and approximate (lines) equations. $r_x = 1, 0.75, 0.5, 0.25$ (top to bottom line). Initial conditions: $P = 0.49, F = 0, \Phi = 0$.

Figure 5.9 Allele frequency difference vs generation: approximate and exact equations. $\Phi$ vs $t$ for the exact (dots) and approximate (lines) equations. $r_x = 1, 0.75, 0.5, 0.25$ (bottom to top line). Initial conditions: $P = 0.49, F = 0, \Phi = 0$. 

-- 141 --
5.2.3 Assortative fertilization and clines

The above model shows that assortative fertilization will cause positive frequency dependent selection. The particular interest in this phenomena is to relate it to the chromosomal cline in *Podisma*. Positive frequency dependent selection will maintain clines between divergent populations. The cline is maintained by a balance between migration of individuals into areas where their genotype is rare and subsequent selection against them.

Clines maintained by a balance between frequency dependent selection and dispersal share very similar properties to those maintained by heterozygote disadvantage (Mallet & Barton, 1989a). Bazykin (1969) and Barton (1979b) have considered clines maintained by a balance between dispersal and selection against heterozygotes. If selection is weak, the rate of change in gene frequency (for an autosomal locus) as a result of dispersal and selection may be approximated in one dimensional space and continuous time by a diffusion equation of the form:

\[
\frac{\partial P}{\partial t} \approx \frac{\sigma^2}{2} \frac{\partial^2 P}{\partial x^2} + s f(P, x)
\]  

(5.25)

Where \( \sigma^2 \) is the dispersal rate measured as the variance in parent-offspring distance measured along the axis x. \( s f(P, x) \) is the selection acting on the locus as a function of position. This may be due either to environmental variation, frequency dependent selection or heterozygote disadvantage. In the latter two cases selection is independent of position, x. For heterozygote disadvantage or weak frequency dependent selection, \( s f(P, x) = sPQ(P-Q) \). The solution of Eq. 5.25 at equilibrium describes the shape of the cline (Bazykin, 1969; Barton, 1979b):

\[
P \approx \frac{1}{2} \left( 1 + \tanh \left( \frac{x \sqrt{2s}}{2\sigma} \right) \right)
\]

(5.26)
where position $x$ is measured relative to the point at which $P=0.5$ (Bazykin, 1969; Barton, 1979b). The width of a cline can be described by the inverse of its maximum slope. For a cline maintained by heterozygous disadvantage this is

$$w = \sqrt{\frac{8 \sigma^2}{s}}.$$ 

The simplified expression for $s(f(P,x)$ due to assortative fertilization (Eq. 5.19) differs from that due to heterozygous disadvantage or frequency dependent selection only by a factor of $1/3$ if the term $rx$ is taken to be equivalent to selection $s$. This difference is because I have considered an X Linked locus. Selection is by females on males, and males carry only $1/3$rd of the X chromosomes present in the population. Selection equivalent to assortment in the above model would thus be $s = rx/3$. The width of clines maintained by assortment after divergent populations come into contact is therefore:

$$w = \sqrt{\frac{24 \sigma^2}{rx}}$$

(5.27)

It is my intention to use this model to consider the width of the cline in *Podisma* in relation to measured values of $\sigma$, $r$ and $x$. However, the above approximation is only good when selection is weak and the population is continuous, or demes are closely spaced. The accuracy of this approximation is assessed below by computer simulation. The simulations are also used to consider the relative contributions of migration and homogamy to the heterozygote deficit in a cline.

### 5.2.3.1 Computer simulations

The population is considered as a linear array of 101 demes between which there is migration in a 'stepping stone' fashion between adjacent demes (Kimura & Weiss, 1964). Prior to mating, a fraction $m/2$ of the individuals from each deme are exchanged with each of the two neighbouring demes. For comparison with a continuous model, $m$ is equivalent
to $\sigma^2$ if one measures distance in deme spacings. The dispersal rate is assumed equal for all genotypes. In the two end demes only a proportion $m/2$ of the individuals migrate. Mating takes place after migration, that is Eq. 5.5 and 5.6 are evaluated.

The simulation is started with the first fifty demes fixed for allele A and the last fifty fixed for a. The central deme is initiated with $P = 0.5$ and genotypes in Hardy-Weinberg equilibrium. In deterministic simulations such as these the population will only approach equilibrium asymptotically and never actually reach it. The simulations were, therefore, run until the change in gene frequency in all of the demes became less than 0.000001 per generation. At this point $P, F$ and $\Phi$ were recorded for each deme.

In the absence of assortative fertilization, gene frequency changes within each deme are brought about only by migration. Eventually, all demes will be polymorphic with gene frequency $P = 0.5$. With assortment a stable cline will be maintained. The width of the cline is measured as the inverse of the maximum slope. The simulations were also used to investigate heterozygote deficit ($F$) across the cline. The deficit is contributed to both by selection and migration. The maximum deficit occurs at the centre of the cline, referred to as $F_{\text{mid}}$. At the end of each simulation the cline width and $F_{\text{mid}}$ were recorded twice: immediately after mating and after migration.

The simulations were run for a range of values of $m$ and $rx$ ($m$ between 0.1 and 0.5, $rx$ between 0.1 and 1.0). Fig. 5.10 shows the equilibrium conditions across the cline for the particular case of $m = 0.4$, $rx = 0.1$. Other parameter values give qualitatively similar results. The key points to note from these figures are that clines are maintained, there is a deficit of heterozygotes across the polymorphic demes and that a small difference in gene frequency between the sexes is maintained.
5.2.3.1.1 Simulation results - cline width

The widths of the clines depend on both $m$ and $r_x$. The relationship between $m$, $r_x$ and cline width is illustrated in Figs 5.11, 5.12 and 5.13. In Fig. 5.11 $W$ is plotted against $\sqrt{V/r_x}$ since this should give a straight line under the diffusion approximation. For all parameter values the diffusion approximation underestimates true widths. The underestimation is largest when widths are measured after migration, but the difference between the width after migration and mating becomes small when migration is low. When $r_x$ is much less than $m$ the discrepancy becomes very small. This can be seen clearly from Fig. 5.12. For the diffusion equation all the points lie on a single line. An impression of the magnitude of the difference can be gained from Fig. 5.13 in which the percentage difference between the diffusion and simulation results is plotted (as a percentage of the width predicted by the diffusion equation). From this figure it is clearly seen that the underestimation of the diffusion equation is largest when cline widths are measured after migration. When measured after mating the deviation only exceeds 5% if $r_x$ is more than approximately twice the dispersal rate. This corresponds to a cline width of only two or three demes. There are, therefore, very few polymorphic populations and the cline represents a sharp transition in gene frequency. When measured after migration the discrepancy may be as much as 15% for the same values of $m$ and $r_x$. The widths given in Fig. 5.13 are measured as numbers of deme spacings. So long as the cline is wide relative to the dispersal distance the diffusion equations give a good approximation.
Figure 5.10 Allele frequency, heterozygote deficit and sex difference across a cline.

Profiles of $P$, $F$ and $\Phi$ across the central demes of the simulated population. Calculated after mating. $m = 0.4$, $r_x = 0.1$. Points are from the simulations, lines from the quasi-equilibrium approximations.
Figure 5.11 Cline width vs $\sqrt{1/rx}$.

The solid lines are from Eq. 5.27. Points from the simulations.
Figure 5.12 Cline width vs $\sqrt{m/rx}$

The solid lines are from Eq. 5.27. Points from the simulations.

Figure 5.13 Difference in width predicted by simulations and diffusion equation. Abscissa is the width predicted by Eq. 5.27.
5.2.3.1.2 Simulation results - heterozygote deficit

Fig. 5.10 shows that in all polymorphic demes there is a deficit of heterozygotes ($F>0$). The largest deficit occurs in the central deme ($F_{mid}$). It would be useful to have a simple expression which could be used to predict this deficit that might be applied to real populations. Two processes contribute to this deficit: migration and assortment. Figs 5.14 and 5.15 show $F_{mid}$ when measured directly after zygote formation and after migration. In the absence of assortment, migration alone will produce a deficit since it brings about a partial admixture of populations with different gene frequencies. If fertilization is at random, this deficit is reduced to zero after each round of mating. However, partial assortment will also contribute to the deficit. The effect of assortment alone has already been considered ($F'$, Eq. 5.24) without including the effect of migration. A general expression for the expected deficit across a cline due to both migration and assortment cannot easily be obtained. However, by considering the deficit in the central deme of a stepping stone model, quantitative predictions can be made; this is done below. First, the effect of migration alone is considered, then the combined effect of migration and assortment. If the central deme is censused immediately after zygote formation, or if the migration effect is negligible, then the fertilization effect may be most important. If the census is made after migration the migration effect may be most important.

**Heterozygote deficit due to migration**

Consider the three central demes of a cline in the stepping stone model. Gene frequency in the central deme is $\frac{1}{2}$. In the adjacent demes $P = \frac{1}{2} \pm \Delta$. The magnitude of $\Delta$ is derived from the shape of the cline. Assuming that the genotypes are in Hardy-Weinberg proportions, the frequency of heterozygotes, $v = 2PQ$, will be $2(\frac{1}{2}+\Delta)(\frac{1}{2}-\Delta) = \frac{1}{2}$ and $2(\frac{1}{2}+\Delta)(\frac{1}{2}+\Delta)$ in the three demes. A fraction $m/2$ of these are exchanged between the central deme and each neighbour. The frequency of heterozygotes in the central deme after migration is therefore:
\[
\frac{v}{2} = \frac{1}{2}(1 - m) + \frac{m}{2} \left( 2\left(\frac{1}{2} + \Delta\right)\left(\frac{1}{2} - \Delta\right) + 2\left(\frac{1}{2} - \Delta\right)\left(\frac{1}{2} + \Delta\right) \right)
\]

which reduces to

\[
v = \frac{1}{2} - 2m\Delta^2
\]  

\(\Delta\), the difference in gene frequency between the central demes is the gradient in gene frequency. As only the central demes are considered this is the maximum gradient. The width of a cline is defined as the inverse of the maximum gradient in gene frequency. Therefore \(\Delta = 1/W\). For a cline maintained by assortment at an X-linked locus, \(W = \sqrt{24m/\sigma^2}\), (Eq. 5.27, with \(m = \sigma^2\)), and so \(\Delta^2 = \sigma^2/24m\). Putting this into Eq. 5.29 gives \(v = (6 - \sigma^2)/12\). Heterozygote deficit is \(F = 1 - v/2PQ\) where \(P = Q = \frac{1}{2}\) so:

\[
F_{\text{mid}} = \frac{\sigma^2}{6}
\]  

(5.30)

It seems a little odd that the effect of migration on \(F\) does not contain a migration term. However, as \(m\) increases, the gradient in gene frequency, \(\Delta\), decreases and so the two cancel. Only the assortment term is important. For weak assortment, \(F\) due to assortment alone is \(\approx \sigma^2/4\) (Eq. 5.18). The individual effects of migration and fertilization are, therefore, of a similar magnitude.

Mallet & Barton (1989a) considered \(F_{\text{mid}}\) in a simulation of clines maintained by a balance between selection at an autosomal locus and dispersal. They found \(F_{\text{mid}} \approx s/2\). This result may be reached analytically in the same way as above. For a selection/dispersal balance at an autosome, \(W = \sqrt{8m/s}\), so \(\Delta^2 = s/8m\) and \(v = (2 - s)/4\) giving \(F_{\text{mid}} = s/2\). This differs from Eq. 5.30 by a factor of \(\frac{1}{3}\). The difference is because selection is on an X-linked locus in males.

Eq. 5.30 is only an approximation and is inaccurate for two reasons. One minor reason is that I have not taken the sex linkage into account. I have assumed that the gene
frequency among the females is the same as the gene frequency as a whole. Fig. 5.10 shows that this is only the case for the central and the non-polymorphic demes. However, the difference between the sexes is always small and will have little effect. A more important error arises because it is assumed that genotype frequencies in all the demes are initially in Hardy-Weinberg proportions. Since the cline is maintained by assortment and assortment generates deviations from Hardy - Weinberg, this cannot be the case. This inaccuracy is reduced by considering the effects of migration and fertilization together.

**Heterozygote deficit due to migration and assortment**

The combined effect of migration and fertilization can be estimated in the same way as for migration alone. However, rather than assume that the three demes are initially at the Hardy-Weinberg equilibrium, they are taken to be at the equilibrium value of \( F \) for a single deme at \( P = 0.5 \). Actual values of \( F \) will differ from this, but provided the cline is shallow the difference will not be great.

In each deme the frequency of heterozygotes is \( v = 2PQ(1-F) \), (Eq. 5.4). Therefore, in the three central demes \( v \) is: \( 2(\frac{1}{2}+\Delta)(\frac{1}{2}-\Delta)(1-F) \); \( \frac{1}{2}(1-F) \) and \( 2(\frac{1}{2}-\Delta)(\frac{1}{2}+\Delta)(1-F) \). After migration, in the central deme, \( v = (1-F)(\frac{1}{2} - 2m \Delta^2) \). So, solving \( F_{\text{mid}} = 1 - v/2PQ \), where \( P = \frac{1}{2} \) gives.

\[
F_{\text{mid}} = 4m\Delta^2(1-F) + F
\] (5.31)

For \( F = \frac{rx}{4- rx} \), (Eq. 5.18) and \( \Delta^2 = \frac{rx}{24m} \) this gives

\[
F_{\text{mid}} = \frac{5rx - (rx)^2}{12 - 3rx}
\] (5.32)
For small values of $r_x$ the heterozygote deficit due to both migration and fertilization (Eq. 5.32) is approximately equal to the sum of the individual effects of the two processes ($F_{mid} \approx 5r_x/12 = rx/4 + rx/6$).

The deviations from Hardy-Weinberg observed in the central demes of the simulated clines are shown in Fig. 5.14 (measured after migration) and Fig. 5.15 (after fertilization). The expected values due to migration, fertilization and their combined effects are also shown. In both instances, the effect of migration is small. There is little difference between the values of $F_{mid}$ for $m = 0.1$ and $m = 0.5$. This effect is also indicated by the fact that the predicted values due to migration alone are much less than the observed. When measured immediately after fertilization, the observed values are best predicted by the expression due to fertilization alone (Eq. 5.18) When measured after migration, the expression for the combined effects of migration and fertilization give, unsurprisingly, the best fit.

An even better estimate of the heterozygote deficit in the central deme could be made by following the method used above but using the quasi equilibrium value of $F$ in the two demes adjacent to the central one. This would give a better estimate but would be less easily applied in practice since it will include a migration term.
Figure 5.14 $F_{mid}$ vs $rx$, measured after fertilization
Points are from simulations, lines are expectations due to fertilization, migration and their combined effects. See text for derivation of expectations.

Figure 5.15 $F_{mid}$ vs $rx$: measured after migration
5.2.3.1.3 Simulation results - Gene frequency differences between the sexes

A difference in allele frequency between the sexes is observed in all polymorphic demes except the central one (Fig. 5.10). This difference may be predicted by Eq. 5.22 (the quasi equilibrium in $\Phi$ for a single population). Differentiating Eq. 5.22 and solving shows that the maximum value $\Phi$ can take is at $P = \frac{1}{2} \pm \frac{\sqrt{3}}{6}$ giving a maximum of $\Phi = \frac{rx\sqrt{3}}{54}$. The maximum value $\Phi$ can take is therefore when $rx = 1$, so $\Phi = 0.03$. This is clearly too small to be detected in any single population.

5.2.4 Extension to multiple mating

The model dealt with above assumes that females mate at most twice. I will now consider a somewhat more realistic model in which females may mate any number of times. It will be seen that this extension simply increases the net selection at the assorting locus. This is because multiple mating increases the opportunity for selection.

Consider a population in which all females mate $n$ times. All matings are at random. Sperm from all these males is stored. In homozygous females, sperm carrying the same $X$ chromosome as the female are $(1 + k)$ times more likely to be successful in fertilization. Therefore, if a given $AA$ female carries $A$ and $a$ sperm in the ratio $P:Q$, the fertilization success will be in the ratio $P(1+k):Q$. The proportion of homogametic fertilizations of $AA$ females is therefore given by the binomial

$$g(P_m,k) = \sum_{a=1}^{n} \binom{n}{a} P_m^a Q_m^{n-a} \left( \frac{a}{a + \frac{(n-a)}{1+k}} \right)$$

\[ (5.33) \]
Chapter 5

The first part of this sum is the probability that a female will carry a particular mix of sperm and the second, the proportion of homogametic fertilizations given this mix. I will only consider the case of weak assortment, \( k \ll 1 \), so terms in \( k^2 \) are considered negligible. A Taylor series around \( k = 0 \) gives

\[
\frac{a}{a + \left(\frac{n-a}{1+k}\right)} \approx \frac{a}{n} + \frac{k}{n^2} \left(\frac{n-a}{n}ight) = \frac{a}{n} \left(1 + k\right) - \frac{a^2 k}{n^2}
\]

(5.34)

The expectation of

\[
\left(\sum_{a=1}^{n} \binom{n}{a} P_m^a Q_m^{n-a} (a)\right) = n P_m
\]

The expectation of

\[
\left(\sum_{a=1}^{n} \binom{n}{a} P_m^a Q_m^{n-a} (a^2)\right) = n^2 P_m^2 + n P_m Q_m
\]

Applying these to Eq. 5.33 gives the expectation of \( g \).

\[
g(P_m, k) = P_m + \frac{k P_m Q_m (n-1)}{n}
\]

(5.35)

likewise

\[
g(Q_m, k) = Q_m + \frac{k P_m Q_m (n-1)}{n}
\]

(5.36)

Eq. 5.35 and 5.36 describe the frequency of fertilizations by the different males. The progeny genotype frequencies produced by each class of female under this model are laid out in Table 5.4 below.

<table>
<thead>
<tr>
<th>Females</th>
<th>Progeny</th>
<th>Genotype</th>
<th>Frequency</th>
<th>AA</th>
<th>Aa</th>
<th>aa</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>u</td>
<td>g(P_m,k)</td>
<td>1 - g(P_m,k)</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>aa</td>
<td>w</td>
<td>0</td>
<td>1 - g(Q_m,k)</td>
<td>g(Q_m,k)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aa</td>
<td>v</td>
<td>( \frac{1}{2} P_m )</td>
<td>( \frac{1}{2} )</td>
<td>( \frac{1}{2} Q_m )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5.4 Genotype frequencies of female progeny. Multiple mating.
From this table, and by substitution of Eqs 5.2, 5.3 and 5.4, genotype frequencies in the following generation can be derived and from these the overall gene frequency.

\[ P' = P + \frac{1}{3}K \left( 1 - \frac{1}{n} \right) \left( \Phi(PQ - 2\Phi PQ) + \frac{2}{3} \Phi \left( 1 - \frac{4\Phi^2}{9} \right) \right) \]  \hspace{1cm} (5.37)

By the same arguments as given in section 5.2.2.2, it can be shown that the magnitude of \( \Phi \) rapidly approaches a quasi-equilibrium value which will be of order \( k \). Since it is assumed that \( k \ll 1 \), terms of \( \Phi^2, \Phi^3 \) and \( k\Phi \) will be negligible so

\[ P' \approx P + \frac{k}{3} \left( 1 - \frac{1}{n} \right) PQ(PQ) \]  \hspace{1cm} (5.38)

This is the same familiar form as for the twice mating model, i.e. positive frequency dependent selection. Such selection will maintain clines between divergent populations and produce a deficit of heterozygotes. The difference here is that the magnitude of the effect depends on the number of matings. If the number of matings is large, there will be a greater opportunity for selection since the probability of mating at least once with the preferred male will be large. Fig. 5.16 shows the change in gene frequency per generation in relation to the number of matings. It is clear that as the number of mates increases the magnitude of selection increases, but the effect becomes less as \( n \) becomes large.

One difference between this and the double mating model is that in the latter I allowed for variation in the number of matings per female, whereas here all females mate the same number of times. Allowing for some females to mate only once will reduce the opportunity for selection. Variation in the number of mates can be incorporated into the model. If there is some distribution of the number of mates per female, then the expectation of \( P' \) is dependent of the expectation of \( 1/n \), i.e. dependent on the harmonic mean of \( n \). Without knowing the distribution of mating frequency in \( P. pedestris \) this model cannot be directly applied to the data. My purpose here has been simply to show that the qualitative effect of assortment is the same irrespective of the actual mating frequency.
Figure 5.16 Change in gene frequency vs P for different mating frequencies.

Allele frequency in generation $t+1$ vs generation $t$ for $n = 2, 4, 8, 16$ and 32.
5.3 Discussion.

The general conclusion of this chapter is that assortative fertilization will produce positive frequency dependent selection (favouring the common allele), and that this selection will cause stable clines to develop between divergent populations. This was pointed out in the paper reporting the presence of assortment in the *Podisma pedestris* hybrid zone, (Hewitt *et al.*, 1987). The aim here has been to quantify this effect. Selection arises because homozygous females effectively 'choose' sperm that carry the same chromosome as themselves. The strength of preference in the population for a particular allele, therefore, increases with the frequency of that allele. In order to properly reflect the situation in *Podisma pedestris* a model of assortment at an X-linked locus has been considered. Since females are selecting male gametes and these carry only one third of the X chromosomes in the population, the net selection on the allele is only one third as large as would be expected for an autosomal locus.

There are two consequences of assortment other than selection. One is that, as expected, assortment produces a deficit of heterozygotes relative to random fertilization. This effect has been quantified for single population and for the centre of a cline. In the latter case, it has been shown that the deficit due to migration is less than that due to assortment. However, for weak assortment, the two effects are of comparable magnitude. The other is that a difference in allele frequency between the sexes is produced. In a single population under selection for an X-linked allele, a difference in allele frequency between the sexes is expected since the frequency in the males always lags one generation behind that in the females. Eventually the allele becomes fixed, so no permanent difference can be maintained. In a cline, a stable difference between the sexes is maintained. However, this difference is very small and so unlikely to be detected in real populations.

It has been shown that simple expressions can be derived that predict reasonably well the expected cline width, and heterozygote deficit across a tension zone maintained by a balance between assortment and dispersal. The chromosomal cline in *Podisma pedestris*
5.3.1 Some assumptions of the model and their consequences: Is it really applicable to *Podisma pedestris*?

**Equal degrees of homogamy in the two races**

In the model presented here I have assumed throughout that the degree of homogamy is the same in both homozygotes. The first paper reporting homogamy in *Podisma pedestris* showed that this was not necessarily the case (Hewitt *et al.*, 1987). The model with the maximum likelihood was one in which the proportion of homogametic fertilizations was 1 for XAXA homozygotes and 0.55 for X^AX^A's, though the significance of this asymmetry was not tested. Incorporating an asymmetry in the model would increase the selection on allele with the highest degree of homogamy. The unstable equilibrium in single populations would no longer be at $P = \frac{1}{2}$. More important is the effect asymmetry will have on clines. Tension zones in uniform habitats are only stable in position if selection is symmetric (Bazykin, 1969). If there is asymmetry in selection, the range of the more favoured race will expand into that of the less favoured. If the asymmetry revealed by Hewitt *et al.* is real then the question arises of why the cline in *P. pedestris* is apparently stable. There is some evidence that the *Podisma* hybrid zone runs along areas of low population density (Nichols, 1984; Nichols & Hewitt, 1986; Jackson, 1992) as predicted by the theory of tension zones (Barton, 1979b). A density trough may prevent the movement of a cline due to asymmetric selection.

Such tortuous explanations for the apparent stability of the cline are not necessarily required. The second study of homogamy in *P. pedestris* was more controlled than the first (Hewitt *et al.*, 1989). In this experiment double matings were made between allopatric populations on either side of the hybrid zone. No difference in the degree of homogamy

---

"Chapter 5"

will be considered in reference to these. Before doing so, some of the assumptions of the model must be considered and borne in mind.
between the races was observed. It may be that the asymmetry originally detected was simply an unfortunate quirk of the data available. Analysis of the field data (Hewitt et al., 1987) required estimating the mating frequency, mating pattern and degree of homogamy all from the same data set. In the laboratory experiments only the degree of homogamy needed to be estimated since the other factors were controlled. This result (equal degrees of homogamy in the two races) is therefore more robust than the first.

The model used here assumes that fertilization of heterozygous females is at random. In the original field collected data, (Hewitt et al., 1987) there is a small but significant deficit of heterokaryotypic embryos from heterokaryotype mothers. The deficit is only 16 individuals which Hewitt et al. suggest may be due to misidentification of haplodiploid embryos and hybrid mortality. Whatever its cause, it does represent an additional tendency to homogamy. This will also contribute to frequency dependent selection. However, the effect is small in comparison to the effect of homogamy in homokaryotypes. The discussion below will show that the selection on the karyotype resulting from assortment is apparently too large to fit with the observed cline width and dispersal distances. Inclusion of the heterokaryotype deficit in heterokaryotype mothers would simply increase the discrepancy by a small amount.

**Mating frequency**

The model presented here has only been considered in any detail for a population in which each female mates at most two times. In section 3.2.4 I considered briefly a slightly more realistic model in which any number of mates is possible. From this brief treatment it can be seen that increasing the number of mates increases the selective effect of homogamy. This is because as the number of mates increases so does the chance that a female will be carrying sperm from both types of male and so be able to exercise a 'choice'.

-- 160 --
To apply this model directly to the *Podisma pedestris* hybrid zone one needs an estimate of the mating frequency in the wild. Hewitt et al. (1987) provide the only such estimate. From the distribution of karyotypes of embryos within egg pods from gravid females collected from the centre of the hybrid zone, they estimated the proportion of the female population which had mated twice (i.e. the re-mating frequency r). The maximum likelihood estimate is \( r = 0.9 \). However, no distinction can be made between those grasshoppers that have mated exactly twice and those that have mated more than two times.

The only other data on mating frequency in *Podisma* comes from wild caught animals kept in field enclosures (Mason, 1988). These enclosures were set up to investigate the possibility of assortative mating. They also provide some data on mating frequency. In these enclosures the average mating frequency was \( \approx 0.1 \) copulations per female per day (calculated from data in Table 5.13 of Mason, 1988). Over an adult life-span of 20 days this would give an average of two copulations per female. Included in these calculations are those females that did not mate at all. These should not be included in the estimate since females that do not mate at all cannot contribute to the next generation. This is, therefore, an underestimate of the effective mating frequency. A further under estimation arises because there is a chance that some copulations were missed. However, It should be borne in mind that these are enclosed populations. Behaviour in the wild may be significantly different. Population density may be of particular importance. In Mason's enclosures there were \( \approx 32 \) grasshoppers per square meter. This compares with a population density of between 0.06 and 0.36 grasshoppers per \( m^2 \) in the wild at Seyne and 0.014 \( m^{-2} \) at Tende (Nichols & Hewitt, 1986; Barton & Hewitt, 1981a). One might reasonably expect high density to increase the mating frequency since the time between encounters will be so much less. It seems likely, therefore, that a model allowing for at most two copulations per female is fairly reasonable.

Allowing at most only two mates per female is a conservative restraint on the model's applicability. It will be seen in the following section that the problem with relating the theory to the available data is that selection is apparently too strong. Allowing for a
higher mating frequency does not, therefore, alter the conclusions; rather it makes them more robust.

**Sperm precedence**

The model assumes that there is complete sperm mixing before any eggs are fertilized. Any deviation from random fertilization is due to assortment and not to sperm precedence. In many multiple mating insects sperm usage is more complex than this. Sperm precedence is frequently observed whereby either the first or the last male to copulate with a female secures most of the fertilizations (Parker, 1970). In most insects there is last male sperm precedence, in some instances brought about by elaborate adaptations (e.g. Waage, 1979). *Podisma pedestris* is unusual in that it shows a first male sperm precedence (Hewitt *et al.*, 1989). This precedence obscures, but does not eliminate the assortative fertilization. The degree of homogamy evident in the progeny of any one female depends largely on the mating order. However, the degree of homogamy in the population as a whole is simply the average over mating order, provided that all matings are random with respect to the karyotype. There is no evidence of assortative mating either in the field or in enclosures (Hewitt *et al.*, 1987; Mason, 1988), though these studies do not exclude the possibility of assortative mating of virgin females in the field.

### 5.3.2 Application to the *Podisma pedestris* hybrid zone

In considerations of the chromosomal cline *Podisma pedestris* the standard line of inference has been from measurements of selection and cline width to an estimation of the strength of selection against chromosomal heterozygotes (e.g. Barton, 1980a; Barton & Hewitt, 1981a). The work presented here considers directly the strength of selection which would result from the observed homogamy. If the theory and all of the measurements are correct, the cline should now be completely described by measured parameters. However,
a problem remains. Hewitt et al. (1987) point out that only weak selection is needed to explain the width of the cline (s \approx 0.5\%) and that even without the present analysis, it is clear that the observed assortment will lead to strong selection. The size of this discrepancy is discussed in the following sections followed by a consideration of how the theory and data may be reconciled.

5.3.2.1 Measured parameters.

The degree of assortative fertilization has been estimated twice (Hewitt et al., 1987, 1989). The first estimation is from gravid females from the centre of the cline at Tende, and the second from double matings of grasshoppers collected 2km either side of the cline at Seyne. The latter gives the most straightforward estimate of the degree of homogamy, h. These data are summarised in Table 5.1. The average proportion of homogametic fertilizations in double mated females (over mating order and female karyotype) is $h = 0.62$ (S.E. = 0.068, N = 23, from Hewitt et al. 1989 Table 1). The parameter $x$, required for the model is $2h - 1 = 0.24$.

From the wild caught females the model with the maximum likelihood is one in which $h = 1$ for XA females and 0.55 for X^A's (Hewitt et al. 1987, Table 2). For X^A females the standard error of h is estimated at 0.092. Taking an average of these two gives $h = 0.76$ or $x = 0.55$. This paper also estimates the proportion of females that mate twice to be $r = 0.9$.

Dispersal distances in Podisma pedestris have been measured several times (Barton & Hewitt, 1982; Nichols, 1984; Mason, 1988; Currie, 1992). All estimate the lifetime standard deviation of dispersal distance to be $\sigma \approx 21$ m. The smallest estimate is $\sigma = 15$ m (Nichols, 1984) and the largest $\sigma = 27$ m, (the upper support limit of Barton & Hewitt, 1982).
Cline width differs from site to site. In areas of continuous habitat it is approximately 800m. (At Seyne; W = 840m, Hewitt & Nichols, 1986; in three transects at Tende W = 580, 800 and 810m Barton & Hewitt, 1981a)

5.3.2.2 Predictions from measured parameters.

Based on the measured parameters, three predictions can be made about the expected clines: the width, heterozygote deficit, and the allele frequency difference between the sexes. These predictions are summarised in Table 5.5. In this table measured values of assortment and dispersal are used along with a range of values of r, the re-mating frequency. I will only discuss in any detail the expected width and heterozygote deficit. As has been seen in section 5.2.3.1.3 (and further illustrated in Table 5.5), only very small allele frequency differences between the sexes are expected. These are obviously far too small to be actually detected.
Table 5.5 Expectations for the *Podisma pedestris* cline.

Predicted cline width, heterozygote deficit and sex difference. Widths are calculated from Eq. 5.27 for various estimates of \( \sigma \). \( h = 0.62 \) (the average over karyotype and mating order from Hewitt *et al.*, 1989) except *\( h = 0.76 \), the average over karyotype from Hewitt *et al.*, 1987. \( r \) is the proportion of the female population that mates twice. Column 6 is the dispersal rate that would result in an 800m cline for a given strength of assortment. Heterozygote deficits are calculated from Equations 5.24 (assortment only, zygotes) and 5.32 (assortment and migration, adults). \( N \) is the number of individuals that would need to be sampled in order to observe the given deficit in a single sample. Maximum \( \Phi \), is the maximum allele frequency difference expected between the sexes (calculated from Eq. 5.22).
Heterozygote deficit

A deficit of heterozygotes (relative to random fertilization) is expected across the cline. The combined effect of assortment and migration has been estimated (section 5.2.3.1.2). The largest deficit is at the centre of the cline. A larger deficit is expected in adult populations than in juvenile ones, since the adult population will reflect the deficit due to both assortment and migration. In the few studies of female karyotype frequencies in *P. pedestris*, no such deficit of heterozygotes has been observed (Hewitt *et al.*, 1987; Barton, 1980a; Nichols, 1984). The only large deviation from Hardy-Weinberg proportions observed is actually one population which has an excess of heterokaryotypes (Mason, 1988). It appears that this is not simply a chance result, since the same excess was observed in two different years. Why there should be an excess is not at all clear, especially when assortment is considered.

In Table 5.5 the deficit expected in the centre of the cline for the observed assortment is calculated. For most estimates of *r* the expected deficit is less than a few percent. It is notoriously difficult to detect deviations from Hardy-Weinberg expectations. Suppose that in a sample of *N* animals from a single population genotypes were observed to be in exactly the ratio \[ \frac{N(P^2 + PQF)}{N^2PQ(1-F)} : \frac{N(Q^2 + PQF)}{N(Q^2 - PQF)} \], [Hom : Het : Hom]. This observed deficit would be compared with the expected *F* = 0 with a \( \chi^2 \) test, which would give \( \chi^2 = F^2N \). With one degree of freedom, the sample size required to detect a deviation *F*, significant at the 5% level is \( N = 3.84/F^2 \). Small deviations from Hardy-Weinberg are, therefore, only detectable with extremely large sample sizes. For example, a heterozygote deficit of *F* = 0.06 could only be detected with samples of over 1000 animals. The sample sizes required to observe the predicted deficits in the centre of the *Podisma pedestris* hybrid zone are given in Table 5.5. For all but the largest estimates of the strength of assortment it would clearly be impossible to collect a large enough sample with any certainty that they all come from the same population. Sampling across more than one population would yield a spurious deficit of heterozygotes through the Wahlund effect. The
fact that no heterozygote deficit is observed in the *Podisma pedestris* hybrid zone is, therefore, not incompatible with strong selection due to assortative fertilization.

It is interesting that strong forces, such as the assortment considered here, have such a small effect on deviations from Hardy-Weinberg proportions. This resilience of the theory to deviations from its assumptions makes it both useful in theoretical studies and of limited use in detecting selection in the field (Lewontin & Cockerham, 1959).

### Cline width.

The cline widths predicted for a given level of assortative fertilization and the measured dispersal distances are given in Table 5.5. These widths must be compared with the actual width of the cline (≈800m at both Seyne and Tende). It is clear that the actual cline is much wider than predicted by this model - up to 8 times too wide for certain parameter estimates. The discrepancy is reduced a great deal if one assumes that only a very few females mate more than once. However, it seems unlikely, from both the field data and the enclosure studies that the re-mating frequency is as low as 10 or 20% (Hewitt *et al.*, 1987; Mason, 1988). Hewitt *et al.* (1987) are, therefore, correct in their opinion that the observed homogamy is incompatible with the wide clines seen. There are several ways in which the discrepancy between observed and predicted cline widths can be accounted for.

### 5.3.2.3 Reconciling the theory and data.

Clearly the theory and the data do not provide a coherent explanation for the width of the chromosomal cline. There must be some error in the data or some important facet of *P. pedestris* biology missing from the model. It would be nice if this error could be identified and rectified. No new data are presented here but a consideration of the available data suggests possible sources of error. The three relevant parameters are: the cline width,
which is apparently too large; the strength of selection, apparently too large; and the dispersal distance, apparently too small. These are discussed below.

### 5.3.2.3.1 Cline width.

It may seem a little strange to consider the cline width as a parameter subject to error as it is, in itself, the phenomenon to be explained. However, there is variation in the width of the cline: it is not uniformly 800m. In some areas the cline is very much narrower than this. For example at Lac Autier and Col de la Lombarde parts of the hybrid zone are as narrow as only a few tens of metres (Barton et al., in prep; Currie, 1992, Currie et al., in prep). However, in these areas, the narrowness of the cline can be largely explained by the fact that they coincide with physical barriers to gene flow.

Assortative fertilization has been observed at Tende and Seyne. It is the width of the cline in these areas (≈ 800m) that, therefore, needs to be explained. That the cline may be narrower at other locations is largely irrelevant to the problem. The expected cline widths given in Table 5.5 are derived from the diffusion approximation. This approximation assumes that dispersal and selection are continuous processes. The selection arises as a result of assortative fertilization and is, therefore, a discrete event. Its effects are manifest in emerging hoppers in the following generation. These emerge more or less in synchrony at the beginning of the season. Continuous dispersal means that the cline is expected to get wider through the season. The simulations in section 5.2.3.1 show that diffusion approximation is fairly robust if cline widths are measured immediately after the fertilization stage in the life cycle. Actual cline widths are typically calculated from the distribution of adult karyotypes (i.e. after the effect of migration). The diffusion approximation is less good at predicting these widths: it gives fairly large underestimates if the cline is only a few dispersal distances wide. However, the 800m cline in *P. pedestris* represents some 40 dispersal distances. In situations such as this the diffusion approximation is accurate even if the width is measured after dispersal.
5.3.2.3.2 Strength of selection.

The cline width and dispersal rate (as measured) are compatible with selection of only $\approx 0.5\%$. The selection due to assortment predicted by the above model is an order of magnitude larger than this. The theory and data can be reconciled if either the experiments are a gross over-estimate of the degree of assortment or the model is in some way missing an important feature. These possibilities are discussed here.

Hewitt et al. (1987) suggest that the assortment is not incompatible with weak selection if it is an effect of genes in linkage disequilibrium with the chromosomal fusion, rather than a pleiotropic effect of the fusion itself. They do not give the logic behind this argument. One could imagine a situation in which assortative fertilization was observed in double matings between allopatric populations but not in hybrid populations. One might, therefore, conclude that the assortment is indeed due to genes in linkage disequilibrium with the fusion. In the centre of the cline generations of recombination will have removed the association between the assortment genes and the fusion. Some small selective effect would still be possible since some linkage disequilibrium would be maintained through the constant dispersal of individuals into the hybrid zone from pure populations (Barton, 1982, 1983). However, the effect would be small in comparison with that observed in crosses between the races.

This scenario is not what is observed in the *Podisma pedestris* hybrid zone. Here, assortment, of comparable strength, is observed in both situations. It may still be that the assortment is due to genes in disequilibrium with the fusion. There are four possible causes of such disequilibrium. 1) Small amounts of disequilibrium may arise simply as a result of sampling drift. 2) A small amount may be present as a balance between dispersal of parental genotypes into and recombination within the hybrid zone. 3) It may be maintained by selection. 4) It is maintained because recombination between the genes causing assortment and the fusion is suppressed.
Only small amounts of disequilibrium are expected by chance. To explain the observations from the centre of the cline, disequilibrium must be of the same magnitude as the observed homogamy. Likewise, in a cline that is wide relative to the dispersal distance of the organism, only small amounts of disequilibrium are expected as a balance between dispersal and recombination (Barton, 1982, 1983). For the reasons outlined in the introduction to this chapter and in Chapter 2 it is unlikely that the assortment is the result of reinforcement. The assortment itself will generate disequilibrium. However, this can only lead to weak associations since assortment is not absolute and can only be expressed when females carry sperm of both karyotypes.

The last possibility seems most likely. The chromosomal rearrangement that produced the XY race is not simply a fusion. Rather, it involves a rearrangement of paracentric chromosomal regions (Westerman & Hewitt, 1985). This may cause recombination to be suppressed leaving some genes in permanent linkage disequilibrium with the fusion. Also, it has been observed that chiasma formation is reduced in the region of the fusion (Hewitt & John, 1972). Because of this, genes in this region may remain in higher linkage disequilibrium than expected elsewhere in the genome.

This discussion of the possible location of the genes causing assortment is, however, largely irrelevant to the argument. The fact is that strong assortment has been observed in the hybrid zone. Whether or not it is due to linked genes is unimportant. Unless it is simply due to chance, and present in only the experimental population, the observed homogamy has to have a long term selective effect. There are however reasons for thinking that the selective effect is less than predicted by the model presented here.

Parthenogenesis and inviability of chromosomal heterozygotes.

The estimates of the degree of assortative fertilization are based on counting the frequency of homokaryotypes within the pods laid by single grasshoppers (Hewitt et al., 1987, 1989). Early death of heterokaryotypic embryos would lead to an over estimation of

-- 170 --
the degree of homogamy. In the first study there were simply not enough non-developed
egs to account for the excess of homokaryotypes (Hewitt et al., 1987). In the second
study, the number of non-developed eggs was much higher, probably reflecting the larger
genetic differences between the pure race individuals involved in the crosses than in
previous ones (Hewitt et al., 1989). If inviability of heterozygotes was the cause of the
apparent homogamy, then one would expect a correlation between the degree of homogamy
within an egg pod and the number of non-developed eggs. This is not observed in the data
(Hewitt et al., 1987, 1989). Even if apparent homogamy was an artifact of non-developing
embryos, it would not resolve the problem of strong selection. Elimination of
heterokaryotypic embryos would lead directly to strong selection on the fusion.

It is possible that some of the excess homokaryotypes observed are the result of
autodiploidization of unfertilized eggs. Such embryos would be female and
homokaryotypic. Since they would be homozygous at all loci they may suffer severe
inbreeding depression and be selected against. This selection would negate the selection
produced by the apparent homogamy. However, Hewitt et al. (1989) give several reasons
for believing autodiploidization can have only biased the estimate of homogamy by a small
amount. Autodiploidization rarely occurs early enough for the embryo to be perceived as a
normal diploid; a mosaic of haploid and diploid cells is usually observed. Autodiploidy will
produce a shift in the sex ratio equivalent to its frequency. The sex ratio bias in the data is
not large enough to be explained by this process. Autodiploidy and selection against
heterokaryotypes may have lead to a small over-estimate of the degree of assortative mating
but cannot account for it all.

Local population structure

The model used here considers only the deterministic effects of selection. No
account is taken of the effect of random genetic drift. In such a situation only smooth clines
can be produced and gene frequency in a population is predicted exactly by its position.
Real populations are not as simple as this. Local deme size is restricted and so gene frequency may fluctuate around the expectation due to selection because of random genetic drift. Drift causes populations to be closer to fixation than would be expected. This will lead to a small increase in the width of a cline following an environmental gradient (Felsenstein, 1975; Slatkin & Maruyama, 1975).

In tension zones drift will have a secondary effect. The opportunity for selection to act depends on the allele frequency. The maximum effect occurs at frequencies intermediate between $\frac{1}{2}$ and 0 or 1 (i.e. at the maximum of $PQ(P-Q)$). In reducing the heterozygosity of populations and moving them closer to fixation, drift reduces the opportunity for selection. In this way, the net selection is reduced and clines become wider than expected under purely deterministic models (Barton & Hewitt, 1989; Rouhani & Barton, 1987). However, this effect will be small unless the effective deme size is very small.

**Mating and oviposition history.**

The model presented above suggests that selection as a result of assortment will be strong. However, the model has assumed a very naive view of individual reproductive histories. It is assumed all matings take place before any egg pods are laid. This is unlikely to be true in the wild. The evidence discussed in section 5.3.1.3, limited though it is, indicates female grasshoppers mate only approximately three times in an adult lifetime of some 20 days. Gravid females kept in the laboratory lay a pod of eggs every two to three days (Barton, 1980a). The sequence of mating and laying in the field is unknown. Since gravid females in the laboratory do not appear to run out of stored sperm, it is unlikely that their laying is limited by mating frequency in the wild. If one assumes that mating and laying are at random (with the proviso that females must mate at least once before laying), then one would expect that a large proportion of the egg pods laid would be fathered by only one male. In such pods assortative fertilization cannot occur. The degree of
homogamy will increase throughout the season as the number of females mated to both types of male increases.

The design of the laboratory experiments by Hewitt et al. (1989) ensured that all experimental females had mated twice. In the field experiment (Hewitt et al., 1987) there was evidence that a large proportion had mated at least twice. In order to be sure that these animals had mated at least once, they were collected fairly late in the season (≈ 14 days after the first adults eclosed). Males tend to eclose earlier than females so the average age of these females will have been something less than 14 days. However, they will have laid some egg pods before mating again. The observations are, therefore, an overestimate of the frequency of multiply fathered pods.

Suppose that on the day of eclosion all female grasshoppers mate once. On each subsequent day they either mate, with probability M, or do not (probability 1-M). I assume that mating and laying frequencies are independent and that on each day the opportunity for mating precedes, and does not overlap, with the opportunity for oviposition. On day z, the probability that the female parent of a pod laid on this day has mated only once is (1-M)^z. The probability that she has mated twice or more is 1 - (1-M)^z. Over an entire season, assuming all females live for equal times, the proportion of egg pods laid by females who have previously mated twice or more is given by the sum of 1-(1-M)^z as a fraction of the number of days in the season, N. i.e. P(Multiple paternity) = \frac{1}{N} \sum_{z=1}^{N} 1 - (1-M)^z. This proportion is illustrated in Fig. 5.17. From this it can be seen that if M is small and the season short, the proportion of multiply fathered eggs falls to quite a small fraction. For example, let M = 0.1. After 14 days the proportion of females that have mated at least twice is 1 - 0.9^{14} = 0.77. However, the proportion of eggs laid that were by females that have mated at least twice is only 0.5. The re-mating frequency used in the model should actually be derived in a way similar to this, and not simply taking the proportion of twice mated females at the end of the season. I do not suggest that the calculations above are sufficiently realistic to be applied directly. They assume that mating and oviposition frequency are
independent random variables which is unlikely to be the case. They do, however, illustrate how a more realistic model of individual mating histories reduces the opportunity of assortment. In order to understand homogamy in *Podisma pedestris* properly one needs to know the actual distribution of mating and oviposition frequencies in the wild. The data required are not easily gained since to do so requires observation of a large number of individual animals over a long period of time. Such knowledge may invoke a large reduction in the estimate of selection but the difference is unlikely to be the order of magnitude required to reconcile it with the observed dispersal rates and cline width.

*Figure 5.17* Proportion of egg pods laid by females that have mated at least twice. 
*M* is the per day probability of mating. *N* is the adult life-span of a female grasshopper.
Chapter 5

The net selection on the common allele may be reduced even further if the viability of eggs laid early in the season is greater than those laid towards the end. It is not unreasonable to assume that the viability of an embryo will depend on how well it has been provisioned by its mother as well as its genotype. It may be that viability decreases with maternal age. In the laboratory there is no association between the viability of egg pods and the time of laying (Hewitt et al., 1989). However, these animals were well supplied with food throughout the season. In the field things may be different. Late in the season, when the potential for assortment is high, embryos may be less well provisioned because, as the season progresses, resources become rarer. Casual observation of the change in the Podisma habitat from lush vegetation in June to parched earth in August makes this seem very likely. In order to assess its importance one would need determine the relative viabilities of late and early egg pods. This is probably an almost impossible task since it requires following pods from oviposition, through diapause in the winter to adulthood the following year!

In order for there to be no net selection the disadvantage to late egg pods must be exactly balanced by the selective effect of homogamy. It is unlikely that nature is this obliging, so one must conclude that homogamy will have at least some selective effect.

5.3.2.3.3 Gene flow

Table 5.5 shows the dispersal distance, $\alpha$, required to reconcile the observed degree of homogamy with the observed cline width of 800m. This is calculated for a range of values of the re-mating frequency, $r$. For the only available field estimate of $r = 0.9$ (Hewitt et al., 1987) required value is $\alpha \approx 75m$. This compares with actual estimates of $\approx 20m$. Even if one is generous in allowing for the effect of mating history and assumes that the effective value of $r$ is $\approx 0.5$ the required dispersal distance is still more than twice the measured value. Is it possible that the field measures underestimate the level of gene flow by such a large amount?
Estimates of gene flow in *Podisma* have been based on observations of the movements of marked grasshoppers; so called Mark, Release, Recapture, or MRR studies. There are several reasons why such studies may yield misleading results (Endler, 1977; Slatkin, 1985,1987). The first is that they may not actually be measuring the correct parameter. What is important with respect to population genetics is the rate of gene flow, which may be very different from the dispersal rate. Dispersal is only one component of gene flow. The reproductive success of dispersants is also important. If long range movement reduces the reproductive potential of an individual (because of the movement itself or the chance of landing in a habitat to which the organism is less well adapted), the rate of gene flow may be much less than predicted by MRR techniques (Endler, 1977). Over estimation may also arise if dispersal occurs largely after reproduction. Releasing marked individuals from a central point (e.g. Barton & Hewitt, 1982) may inflate the dispersal rate by increasing the density. The disturbance produced by repeatedly searching the study area may also increase dispersal. However, it is more often observed that MRR studies underestimate the level of gene flow.

There are several reasons why this may be, and it is likely that these apply in *Podisma*. Essentially, the problem is that MRR studies can only ever measure dispersal at a particular location and at a particular time (year and time of year). For example, if in the year of the study the conditions are not conducive to large amounts of dispersal (e.g. a cold season for a poikilotherm), a low value will be recorded. The amount of gene flow may still be high because of the effect of the usually high levels of dispersal. In a population whose nature is one of recurrent local extinction and re-colonization much of the gene flow occurs in the re-colonization phases - these are unlikely to be detected in a simple MRR study. In organisms such as insects that moult periodically it is not possible to follow single individuals through their entire lives and so measures are often made in only one stage. In *Podisma* most studies have been of adult movement. Estimates of $\sigma$ have been derived by summing the estimate of daily dispersal from marked grasshoppers over an expected adult life span of $\approx$20 days (Barton & Hewitt, 1982; Nichols, 1984; Mason,
1988). In these estimates it is, therefore, assumed that there is no dispersal of nymphs. Mason found significant amounts of nymph movement in a study at Seyne. This was at a rate between one third and one half of that in adults (Mason, 1988, Table 2.3). This nymph movement will increase the overall estimate of dispersal. However, in another study (Currie, 1992) no significant nymph dispersal was observed. Clearly more work is needed in order that better estimates of the lifetime dispersal distance of *Podisma* may be obtained.

Even if nymph dispersal can be estimated other problems remain; certain small time intervals within the life cycle may be particularly important. For example, in the butterfly *Heliconius erato* it has been shown that a large proportion of the movements made occur immediately after eclosion (Mallet, 1986a). The dispersal distance, \( \sigma \), obtained from movements during the first few days is twice that of the rest of the adults life-span of \( \approx 90 \) days. Though there is no evidence for it, the same may occur in *Podisma*. In MRR studies only a small fraction of the marked animals will be freshly eclosed. Even those that are in the marked samples may be under represented in the recaptures since, with their soft cuticles, they may be more adversely affected by handling and marking.

MRR studies rely on searching defined areas for marked animals. Some individuals will move out of the study area. Some adjustment in the calculations can be made for these unseen animals (e.g. Barton & Hewitt, 1982). However, to do so one must assume that the distribution of dispersal distance is normal. In most organisms, the distribution is, however, leptokurtic - a few individuals move very long distances (Endler, 1977). Barton & Hewitt (1982) found the distribution of dispersal distances to be a very good fit to a normal curve, however this does not include the possibility that some individuals moved a very long distance. Nichols & Hewitt (1986) found that movement of sheep was a potential cause of long distance dispersal in *Podisma* - they found \( \approx 2 \) live grasshoppers per 100 sheep in the wool of a flock of some 2000 that graze the area of the hybrid zone at Seyne. Long distance dispersal means that the estimates of \( \sigma \) in *Podisma* are low. However, the effect on the width of the expected cline is small: an individual that moves a large distance across a hybrid zone only represents a small amount of gene flow since the genes it carries
are unlikely to be established since they will be strongly selected against. For this reason the level of gene flow across a cline is over estimated by \( \sigma \) if the distribution of dispersal distances is very leptokurtic. \( \sigma \) gives too much weight to the few long range dispersants, whereas the small distances moved by the majority of the population are actually more important in determining cline width. A leptokurtic distribution increases the length of the tails of a cline but has little effect on its width (measured as the inverse of the maximum slope).

By making allowances for the processes discussed above one may be able assume that the rate of gene flow in *Podisma* is greater than inferred from MRR studies. However, there are reasons for thinking that it cannot be as large as required to make the homogamy model fit with the observed cline width.

More direct estimates of levels of gene flow may be made by analysis of geographic distribution of genetic variation (Barton & Clark, 1990; Slatkin, 1985, 1987). Particularly powerful techniques are available in hybrid zones. Diffusion of parental genotypes into the centre of a tension zone generates linkage disequilibrium which is broken down by recombination and segregation. If the recombination rates between genetic markers are known, the amount of disequilibrium in the centre of the zone can be used to estimate the rate of diffusion, \( \sigma \) (Barton, 1982). Two applications of this technique have revealed rates of gene flow much larger than otherwise expected. In the hybrid zone between the toads *Bombina bombina* and *B. variegata* analysis of disequilibrium in the centre yields an estimate of gene flow twice that from MRR studies (Szymura & Barton, 1986, 1991). More spectacular are the results from the hybrid zone in *Heliconius erato*. Estimates of gene flow from linkage disequilibrium are an order of magnitude larger than MRR estimates (Mallet, 1986a; Mallet *et al.*, 1990). Unfortunately, analyses such as these are not possible in *Podisma* because the genetic markers are not available. Fixed, single locus differences between the hybridizing races are required but not apparent in *Podisma* (Halliday *et al.*, 1983, 1984; Currie, 1992).
A less sophisticated way of evaluating the degree of gene flow is to simply investigate the correlation in gene frequency at (presumably) neutral loci between populations. If gene flow between them is high, similar gene frequencies are expected. If gene flow is low, the populations will have evolved more or less independently, and are expected to have diverged due to drift. In Podisma, studies like this do not give actual estimates of gene flow but they do set limits to its possible magnitude. Studies have shown that there are significant differences in allozyme frequency between populations a few metres apart (Currie, 1992; Nichols pers. comm). Similarly, there are large, fine scale fluctuations in the frequency of the fusion across the cline at Tende (Barton & Hewitt, 1981a; Barton & Hewitt, in prep.). This indicates that the rate of gene flow in Podisma cannot be very much larger than that predicted by the dispersal experiments and certainly not the 40 - 60m gen\(^{-1}\) required.

5.3.2.4 Understanding the Podisma chromosomal cline

It is suggested that the selection maintaining the chromosomal cline in Podisma pedestris is as a result of non-disjunction in heterokaryotypes (e.g. Barton, 1980a; Barton & Hewitt, 1981b). No such non-disjunction has actually been observed (Barton, 1980). A more parsimonious explanation (in that it is based only on observed phenomena) is that the cline is maintained by dispersal and assortative fertilization.

However, this explanation does have problems of its own. The cline is much wider than expected from the measured parameters. Most of the above discussion has been mere speculation on how this difference can be reconciled. A better understanding will only be gained when the reproductive behaviour of Podisma pedestris in the wild is more fully understood.
5.3.3 Assortative fertilization as a barrier to gene flow.

It has been demonstrated above that assortative fertilization can lead to strong selection and to steep clines between populations. The pattern of assortment modelled here can, however, only lead to a weak barrier to gene flow. This can be seen when compared to the effect of assortative mating. Assortative mating can provide a complete barrier to gene flow between species. With assortative fertilization, even if homogamy is complete, some gene flow will result. This is because, in the absence of assortative mating there will always be some individuals who have mated only with heterospecific males and so the tendency to homogamy cannot be expressed. Only if assortment is such that no fertilizations take place when homogametic ones are not available can assortative fertilization lead to a complete barrier to genetic exchange.

A similar conclusion is reached in the consideration of the effect of assortative mating for direction of coiling in Partula suturalis (Johnson et al., 1990). In this case complete assortative mating gives rise to only weak reproductive isolation because the expression of the coiling genotype is delayed by one generation. In both of these examples, the barrier represented by assortment is also limited because it is due to the effect of a single locus.

5.3.4 Conclusions

It seems clear from the above discussion that homogamy will lead to selection of sufficient strength to maintain the chromosomal cline in P. pedestris. However, it appears that the cline is much wider than one would expect. A better understanding will be gained when more of the details of the reproductive biology of P. pedestris are understood, although the relevant data are far from easily obtainable. Genetic measures of gene flow, rather than MRR measures of dispersal, will also be of great utility in understanding the forces maintaining the cline.
In all of this discussion it should be remembered that the chromosomal rearrangement is but one factor contributing to the hybrid zone in *P. pedestris*. There are many other genetic differences which are of equal, if not greater importance, in maintaining the distinction between the races.
Chapter 6

Morphometric variation across the hybrid zone in

*Podisma pedestris.*

6.1 Introduction

Many hybrid zones are found to involve morphological as well as genetic transitions (e.g. of 25 zones summarised by Harrison (1990), 17 involve morphological transitions). That this is so should perhaps be unsurprising. Those zones actually studied are almost certainly just a small fraction of those that exist, and this fraction is bound to be biased towards those that involve a transition in some easily distinguished character - such as morphology. Indeed, many of the well studied hybrid zones are between what have previously been described, on morphological criteria, as separate species or subspecies.

Analysis of morphometric variation in hybrid zones has, in general, had two objectives: to determine the position of the zone of transition between one morph and the other; and to detect the presence of hybrids. For example, the subspecies of grasshopper *Chorthippus parallelus parallelus* and *C. p. erythropus* have ranges that abut along the Pyrenees between France and Spain. The two subspecies can be distinguished on the basis of courtship song characteristics as well as on morphological criteria. Butlin and Hewitt (1985a) located the position of the hybrid zone between these subspecies by mapping geographic variation in a number of morphological characters. They found that the single character that best distinguishes the races is the number of stridulatory pegs in males, and ovipositor length in females (*erythropus > parallelus* for both characters), and populations of individuals with intermediate values were found where the subspecies meet. A better distinction between the two races can be made by the construction of a discriminant
function\(^1\) based on nine measurements of body shape. Butlin and Hewitt (1985a) were able
to show that there is a sharp change in the value of this discriminant function between the
two subspecies, with populations of intermediate scores in between. The intermediate
character values of the grasshoppers from intermediate populations indicates that they are
composed of hybrid individuals rather than simply a mixture of the two subspecies.

Similar patterns of variation are seen in many other hybrid zones. For example, the
hybrid zones between the grackles *Quiscalus quiscula quisquala* and *Q. q. versicolor* can
be identified on the basis of body length, bill length and bill width as well as plumage
colour (Yang & Selander, 1968); between the pocket gophers *Geomys bursarius* and *G.
lutescens* on the basis of cranial measurements (Heaney & Timm, 1985), between *Clarkia
nitens* and *C. speciosa polyanther* on the basis of floral and foliage morphology (Bloom,
1976), and between the crickets *Gryllus firmus* and *G. pennsylvanicus* on the basis of
body, hind wing, and ovipositor length (Harrison & Arnold, 1982; Harrison, 1986).

Although there are no obvious morphological differences between the races of
*Podisma pedestris*, the presence of chromosomal heterozygotes (Hewitt, 1975; Barton,
1980a, Mason, 1988; Hewitt et al. 1987) where the two races meet is ample evidence that
there is a true hybrid zone rather than simply an area of overlap. However, given that the
two races differ by a large number of genes in addition to chromosomal fusion by which
they are distinguished (as indicated by the wide area of low viability across the zone,
Barton & Hewitt, 1981b, and see Chapter 4), one would perhaps be unsurprised if there
were subtle morphological differences between them. In a study of the hybrid zone at
Tende, Barton (1979a) found significant differences in morphology between XO and XY

\(^1\) The discriminant function is that linear combination of the variables that gives the maximum
between/within groups variance ratio in a one way ANOVA (see Manly, 1986; James &
McCulloch).
In a comparison of pure race populations spanning but one portion of the hybrid zone, it is always difficult to exclude the possibility that the differences between them are not simply environmental effects due to the fact that the two races are found in different locations. However, Barton’s analysis revealed that there was a consistent difference between XO and XY animals within hybrid populations. This suggests that the chromosomal fusion itself, or genes closely linked to it, have an effect on the grasshoppers’ morphology. No clinal variation in morphology was demonstrated nor was there any evidence of increased variation in the hybrid zone.

In this chapter the work of Barton (1979a) will be repeated at a different location and extended. An analysis of morphometric variation across the hybrid zone at the Col de la Lombarde is presented. The motivation for this study is two-fold. First, as will be discussed in section 6.4, Barton’s analysis may have been flawed in that it was based only on univariate statistics. Second, given that the difference between karyotypes seen at Tende is real, it is interesting to know if the pattern is repeated in other areas of the zone. Three questions are addressed: are there morphometric differences between the races? Is there a cline in morphology? Are there morphometric differences between karyotypes within hybrid populations?
6.2 Methods & Analysis

6.2.1 Sampling

Samples were collected between 19 June and 1 August 1990 across the hybrid zone at the Col de la Lombarde (see Fig. 4.2). Between 10 and 30 adult grasshoppers were taken from each of 54 sites. Each sample was taken from as small an area as possible, and certainly less than a 20m radius, the dispersal distance per generation. In this way one can be fairly sure that each sample is of at most one panmictic unit. In order to minimize disruption to the hybrid zone only male grasshoppers were collected.

Ideally, one would have sampled either at random, or in some regular pattern across the hybrid zone. For several reasons, these sampling techniques were not used. First, much of the area of the Col de Lombard is covered with bare rock in which no grasshoppers are found: sampling in these areas would have been a wasted effort. Second, even in areas where grasshoppers are present, there is much variation in population density. In many areas density is very low and collecting a sample of twenty or more animals would have been excessively time consuming. At the time of sampling the approximate location of the centre of the hybrid zone was known so the frequency of the fused chromosome at each site could be estimated. The actual frequency at each site is estimated by the karyotype frequency in the sample (see section 6.2.2), or in some instances from the sample augmented by the karyotypes of other animals collected, but not measured.

6.2.2 Dissection & Karyotyping

Between collection and dissection, all animals were kept in tins with a loose fitting lid. They were, as far as was possible, kept in a cool place or at least out of direct sunlight. The animals were dissected in the evening of the day on which they were collected, or the following morning. After making a small dorsal incision through the first few segments of
the abdomen the testis is clearly visible. As much as possible of the testis is removed and placed straight away in two ml of freshly mixed 3:1 ethanol : acetic acid. The bodies were preserved by freezing.

To determine the karyotype, squash preparations are made using a few follicles of the fixed testis, stained with aceto orcein. The X chromosome is easily identifiable in cells at metaphase. In an XO animal it is usually seen at the edge of the cell. In XY animals it is easily identified, fused with the autosome, (see Fig.4.1) In addition to identifying the X, the total number of chromosomes and chromosome pairs is counted (in XO animals 12 are seen and only 11 in XY animals). At least 2 cells, in different areas of the slide, were scored for each animal in order to reduce the risk of mis-classification through cross contamination between slides. In general the preparation and scoring of a slide takes a matter of minutes. However, in some instances several slides must be searched before meiotic cells are observed.

6.2.3 Choice of variables

The aim here is not to reconstruct the shape of a grasshopper, but rather to detect any differences between groups of animals. Therefore, it does not matter if the measurements taken cannot be interpreted in conventional terms of shape (Rohlf, 1990). A common approach is simply to take the measurements of a large number of body structures with the hope that most aspects of size and shape are captured by them, or by the ratios between them. The choice of what exactly should be measured is not clear. One requirement is that all the major axes of the body form be included. Apart from this rather obvious guideline, there is, in general, little a priori reason to believe that any one suite of variables will be any better than another. The choice of variables and the number used in studies such as this is, therefore, usually fairly arbitrary, and determined largely by the ease of measurement and degree to which they are correlated with each other. If two variables are very strongly correlated, they are likely to be measuring the same aspect of the animals.
shape, so one or other of them may be excluded without losing much information. However, since the correlations between variables cannot be known until the study has been completed, it is likely that any suite of variables utilized will include a large amount of redundant data (Bookstein, 1982). A second problem is that the results of the study may depend on the variables chosen (Strauss & Bookstein, 1982). A lack of difference between groups is difficult to interpret. One can never rule out the possibility that there may in fact be a difference between them that is not revealed by the measurements made.

I am fortunate in this study in that a suite of variables has already been determined by which the two karyotype races may be discriminated: the variables used by Barton in his survey of variation at Tende, (Barton, 1979a). Barton began by considering a set of 18 variables, but rejected eight that were very highly correlated with the others, or very prone to error. To the remaining ten, I have added the height of the thorax (HTHO). A description of the 11 variables used is given in Table 6.1.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>HFAC</td>
<td>Height of the face. Measured from the median ocellus to the epistomal suture.</td>
</tr>
<tr>
<td>WFAC</td>
<td>Width of the epistomal suture.</td>
</tr>
<tr>
<td>EYES</td>
<td>Shortest distance between the eyes.</td>
</tr>
<tr>
<td>TOPL</td>
<td>Length of top of thorax.</td>
</tr>
<tr>
<td>HBOT</td>
<td>Length of sternopleural plate.</td>
</tr>
<tr>
<td>WBOT</td>
<td>Width of sternopleural plate.</td>
</tr>
<tr>
<td>HFEM</td>
<td>Length of rear femur.</td>
</tr>
<tr>
<td>HTIB</td>
<td>Length of rear tibia.</td>
</tr>
<tr>
<td>MFEM</td>
<td>Length of middle femur.</td>
</tr>
<tr>
<td>WTHO</td>
<td>Width of thorax, between the junctures of the upper and lower thoracic plates with the head removed.</td>
</tr>
<tr>
<td>HTHO</td>
<td>Height of thorax: greatest distance between the upper and lower thoracic plates.</td>
</tr>
</tbody>
</table>

Table 6.1 Variables used for morphometric analysis.
6.2.4 Measuring grasshoppers.

All the measurements taken are simple distances between two landmarks. These distances were measured with the aid of a binocular microscope camera lucida and a digitizing tablet (Hewlett Packard 46087A, see Fig.6.1). The specimen was mounted with a pin on a cork stage, adjustable in all directions. When viewing the specimen, the camera lucida provides one with an image of the digitizing tablet superimposed onto that of the specimen. The tablet is positioned such that the image of the specimen covers most of the image of the tablet. The specimen is adjusted until the edges of the structure to be measured are brought into focus. All focusing is done by adjusting the specimen stage and not the optics. In this way, the distance between the specimen and the digitizing tablet, and hence the magnification, remains constant. The digitizer pointer is moved to a position such that its image exactly coincides with that of the first landmark and its coordinates recorded. The coordinates of the second landmark are recorded similarly and the distance between the two points calculated and stored. At the beginning of each measuring session the digitizing tablet was calibrated by using it to measure a known distance (from an eye piece graticule placed on the stage). All subsequent measurements are converted into millimetres.

There are several advantages, mostly due to increased speed, to this method of measuring over recording distances with an eyepiece graticule. Eleven measurements were taken from approximately 1000 specimens. The method used allows these 11000 numbers to be stored directly on computer disks and imported into the statistical analysis programs so avoiding the need for manual data entry (which is both time consuming and error prone). The distances recorded are calculated as the euclidian distances between landmarks and so the structure does not need to be manipulated into any particular alignment. For each specimen, only its identification numbers need to be entered via the keyboard, all else is automatic.
6.2.5 Analysis

The data consist of a vector of measurements for each grasshopper. The problem of distinguishing groups is one of determining if the mean vectors differ between them. The data could be analysed by a series of univariate analyses on each of the variables separately. However, this would address the wrong question, and could be misleading. The correct question is how does the entire vector of variables differ from group to group and not, how does any particular variable vary. A significant difference between groups will indicate that the mean of the set of variables differs between them. It may be that some variables are more important than other but there is no a priori reason for believing this to be so. Results of univariate analyses could be misleading for two reasons. First, the variables are highly correlated and so must be treated together. Second even if the variables were truly independent, univariate analyses can mislead. If a large number of univariate tests are performed, one is quite likely to observe at least one significant result simply due to chance. If $n$ univariate tests are carried to distinguish two samples from populations that do not differ, the probability of getting no significant results at the 5% level is $0.95^n$. Therefore, the probability of observing at least one significant result is $1 - 0.95^n$. With 11 variables, as used here, the probability of at least one spurious significant result is 0.43.
To overcome these problems, multivariate statistical methods must be used. Multivariate tests combine information from all of the correlated variables in determining significance levels. It is therefore possible to obtain a significant result from a multivariate test even though none are found with univariate tests. Conversely, a non-significant result may be observed even though one or more of the univariate tests were significant. It has been suggested by some (e.g. Corruccini, 1987) that multivariate statistics can be avoided if one performs several univariate tests, and requires that at least a certain number yield positive results before accepting that group differences are real. The inadequacy of this approach has been shown by Willig et al. (1986) in a comparison of multivariate and univariate tests of morphometric differences between different groups of Brazilian bats. They found no obvious association between the number of significant univariate tests and the results of full multivariate analyses. Extreme examples were 11 out of 12 univariate tests significant, yet a non-significant multivariate test, and a significant multivariate test yet all univariate tests non-significant.

All of the analysis presented here, was carried out using SYSTAT, version 5 (Wilkinson, 1989) running on an Apple Macintosh IIfx or SE/30. As is usually the case in morphometric studies, the analysis was carried out on log transformed data. The principal reason for this is that the transformation is likely to produce distributions with variance independent of the mean. This allows variables of different magnitudes to have equal weighting in the analyses. Also, the log transformation may be more likely to yield a normally distributed variable than the raw data (Falconer, 1981). All of the following tables and figures refer to log transformed data.

All of the analyses presented in the following section use Wilks's $\lambda$ as the test statistic. For each analysis, SYSTAT also provides two other test statistics (Pillai trace and Hotteling-Lawly trace). The three statistics do not necessarily yield identical results. In general the power of each statistic is determined by the manner in which the samples deviate from the null hypotheses (Hand & Taylor, 1987). However, in all of the tests given below, the conclusion reached is not altered by the choice of statistic.
6.2.5.1 Missing Values.

For 93 of the approximately 1000 grasshoppers in the sample, one or more of the measurements could not be obtained. This was either because the structure in question was missing (e.g. legs) or had become obviously deformed in preservation. Since multivariate analyses take as their data points vectors of variables, each vector must be complete if it is to be included. If one measurement is missing from an individual, the other ten cannot be included in the analysis. Although the missing values account for less than 1% of the data set, if all the animals with missing values were excluded approximately 9% of the data would have to be discarded. This wastage of data can be avoided by estimating the values of the missing variables: since all of the variables are correlated, missing values can be estimated from those available.

Multiple regression was carried out of the missing variable against the ten others. The most likely value of the missing data is predicted by the regression coefficients and the magnitude of the data non-missing values. As there may well be differences in the correlations between variables between sample sites, the multiple regression was carried out on a per site basis. i.e. only within site correlations were used in estimation. For those individuals with two measurements missing, the missing values were estimated from its regression against the remaining nine variables. Those individuals with three or more measurements missing were excluded from the analysis.

This estimation procedure will have only a small effect on the conclusions reached. The central question being addressed is that of differences between karyotypes within populations. In this comparison the estimation procedure is conservative. The estimates are made from complete samples, not divided by karyotype. If there are differences between karyotypes, they will be lessened by including estimated values, since these are calculated from the whole population.
6.3 Results

6.3.1 Describing the data

A total of 1045 grasshoppers were measured. The correlation coefficients between the 11 variables are given in Table 6.2. As expected, all of the variables are highly correlated. All the correlation coefficients are positive, indicating that much of the variation is due to variation in size. The extent to which size variation is important is seen by taking the principal components of the data. Principal components were calculated from the covariance matrix of the entire data set, their loadings and the proportion of the total variation for which each accounts is given in Table 6.3.

The first component is loaded fairly evenly on all of the eleven variables and so can be taken as a measure of size (James & McCulloch, 1990). Size variation accounts for much of the total variation (51.8%). However, there is a large amount of variation in 'shape' as indicated by the other components. Since all of the components account for sizeable amounts of variation, the dimensionality of the data cannot be reduced by considering only the first few principal components. A complete analysis must include all the variables.

```
<table>
<thead>
<tr>
<th></th>
<th>Wfac</th>
<th>Eyes</th>
<th>TOPL</th>
<th>HBOT</th>
<th>WBOT</th>
<th>HFEM</th>
<th>HTIB</th>
<th>MFEM</th>
<th>WTHO</th>
<th>HTHO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wfac</td>
<td>0.53</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eyes</td>
<td>0.35</td>
<td>0.40</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOPL</td>
<td>0.49</td>
<td>0.54</td>
<td>0.33</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HBOT</td>
<td>0.44</td>
<td>0.48</td>
<td>0.34</td>
<td>0.53</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WBOT</td>
<td>0.40</td>
<td>0.54</td>
<td>0.37</td>
<td>0.51</td>
<td>0.56</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HFEM</td>
<td>0.45</td>
<td>0.51</td>
<td>0.33</td>
<td>0.57</td>
<td>0.54</td>
<td>0.48</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HTIB</td>
<td>0.39</td>
<td>0.49</td>
<td>0.29</td>
<td>0.52</td>
<td>0.45</td>
<td>0.46</td>
<td>0.81</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MFEM</td>
<td>0.38</td>
<td>0.51</td>
<td>0.31</td>
<td>0.52</td>
<td>0.55</td>
<td>0.50</td>
<td>0.73</td>
<td>0.70</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>WTHO</td>
<td>0.42</td>
<td>0.62</td>
<td>0.36</td>
<td>0.48</td>
<td>0.47</td>
<td>0.50</td>
<td>0.53</td>
<td>0.49</td>
<td>0.50</td>
<td>1.00</td>
</tr>
<tr>
<td>HTHO</td>
<td>0.41</td>
<td>0.55</td>
<td>0.39</td>
<td>0.55</td>
<td>0.61</td>
<td>0.59</td>
<td>0.60</td>
<td>0.54</td>
<td>0.59</td>
<td>0.56</td>
</tr>
</tbody>
</table>
```

Table 6.2 Correlations between the 11 variables.
Data from all sites is combined. n = 1045 for each.
### Table 6.3 Principal components: all data.

Variable loadings and variance explained. All data combined.

<table>
<thead>
<tr>
<th>Component</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eigen values</td>
<td>0.0108</td>
<td>0.0031</td>
<td>0.0014</td>
<td>0.0012</td>
<td>0.0011</td>
<td>0.0009</td>
<td>0.0007</td>
<td>0.0006</td>
<td>0.0005</td>
<td>0.0004</td>
<td>0.0003</td>
</tr>
<tr>
<td>Loading</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HFAC</td>
<td>0.0284</td>
<td>0.0012</td>
<td>0.0237</td>
<td>0.0150</td>
<td>0.0154</td>
<td>-0.0079</td>
<td>-0.0015</td>
<td>-0.0030</td>
<td>0.0001</td>
<td>0.0038</td>
<td>0.0006</td>
</tr>
<tr>
<td>WFAC</td>
<td>0.0257</td>
<td>0.0022</td>
<td>0.0064</td>
<td>-0.0011</td>
<td>0.0066</td>
<td>0.0114</td>
<td>-0.0004</td>
<td>0.0024</td>
<td>-0.0084</td>
<td>-0.0160</td>
<td>-0.0011</td>
</tr>
<tr>
<td>EYES</td>
<td>0.0422</td>
<td>-0.0495</td>
<td>-0.0052</td>
<td>0.0023</td>
<td>-0.0031</td>
<td>-0.0010</td>
<td>0.0002</td>
<td>0.0005</td>
<td>0.0001</td>
<td>0.0002</td>
<td>0.0001</td>
</tr>
<tr>
<td>TOPL</td>
<td>0.0382</td>
<td>0.0107</td>
<td>0.0150</td>
<td>0.0031</td>
<td>-0.0265</td>
<td>0.0030</td>
<td>0.0009</td>
<td>0.0005</td>
<td>-0.0007</td>
<td>0.0015</td>
<td>0.0002</td>
</tr>
<tr>
<td>HBOT</td>
<td>0.0306</td>
<td>0.0065</td>
<td>0.0020</td>
<td>-0.0162</td>
<td>0.0023</td>
<td>-0.0176</td>
<td>0.0096</td>
<td>0.0085</td>
<td>0.0027</td>
<td>-0.0037</td>
<td>0.0015</td>
</tr>
<tr>
<td>WBOT</td>
<td>0.0270</td>
<td>0.0034</td>
<td>0.0023</td>
<td>-0.0145</td>
<td>0.0032</td>
<td>0.0021</td>
<td>-0.0206</td>
<td>0.0043</td>
<td>0.0035</td>
<td>0.0033</td>
<td>-0.0014</td>
</tr>
<tr>
<td>HFEM</td>
<td>0.0315</td>
<td>0.0112</td>
<td>-0.0123</td>
<td>0.0102</td>
<td>0.0007</td>
<td>-0.0024</td>
<td>0.0011</td>
<td>-0.0002</td>
<td>0.0072</td>
<td>-0.0023</td>
<td>-0.0121</td>
</tr>
<tr>
<td>HTIB</td>
<td>0.0280</td>
<td>0.0113</td>
<td>-0.0135</td>
<td>0.0120</td>
<td>0.0003</td>
<td>0.0007</td>
<td>-0.0037</td>
<td>0.0004</td>
<td>0.0085</td>
<td>-0.0047</td>
<td>0.0107</td>
</tr>
<tr>
<td>MFEM</td>
<td>0.0315</td>
<td>0.0117</td>
<td>-0.0149</td>
<td>0.0044</td>
<td>0.0018</td>
<td>-0.0036</td>
<td>-0.0010</td>
<td>0.0033</td>
<td>-0.0165</td>
<td>0.0073</td>
<td>0.0009</td>
</tr>
<tr>
<td>WTHO</td>
<td>0.0267</td>
<td>0.0036</td>
<td>0.0011</td>
<td>-0.0051</td>
<td>0.0086</td>
<td>0.0189</td>
<td>0.0108</td>
<td>0.0056</td>
<td>0.0048</td>
<td>0.0085</td>
<td>0.0007</td>
</tr>
<tr>
<td>HTHO</td>
<td>0.0313</td>
<td>0.0053</td>
<td>-0.0034</td>
<td>-0.0122</td>
<td>0.0021</td>
<td>0.0003</td>
<td>0.0026</td>
<td>-0.0209</td>
<td>-0.0004</td>
<td>0.0002</td>
<td>0.0008</td>
</tr>
</tbody>
</table>
6.3.2 Differences between XO and XY grasshoppers within populations?

A simple comparison of XO and XY grasshoppers would be misleading. Any difference between karyotypes may be obscured by differences between sites, or may be exaggerated since different sites have different karyotype frequencies. Therefore one needs to analyse population differences and karyotype differences together. For single variables the standard method of analysis is an analysis of variance (ANOVA). Here, the multivariate equivalent (a factorial MANOVA) is used. For this analysis only polymorphic sites can be considered. There are 35 polymorphic samples in the data, comprising a total of 634 individuals. The test statistics for this analysis are given in Table 6.4. These show that there is a highly significant difference between sample sites in the vector of variable means. No significant effect of karyotype or interaction between karyotype and sample site is seen. Thus, in this sample, Barton's (1979a) finding of a difference in morphology between karyotypes within populations is not repeated. The MANOVA reveals very little other than that there are differences in morphology between the samples. How these may be related to the hybrid zone is considered below.

<table>
<thead>
<tr>
<th>Effect</th>
<th>Sample site</th>
<th>Karyotype</th>
<th>Interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wilks's λ</td>
<td>0.418</td>
<td>0.982</td>
<td>0.499</td>
</tr>
<tr>
<td>F</td>
<td>1.353</td>
<td>0.935</td>
<td>1.071</td>
</tr>
<tr>
<td>D.F.</td>
<td>374,5843</td>
<td>11,554</td>
<td>374,5843</td>
</tr>
<tr>
<td>P</td>
<td>&lt;0.001</td>
<td>0.506</td>
<td>0.160</td>
</tr>
</tbody>
</table>

Table 6.4 MANOVA test statistics

From a factorial MANOVA of morphological variables vs sample site and karyotype.
6.3.3 Examining between population differences

Given that there are differences between the sample sites, questions one and two of section 6.1 may be addressed: Are there morphological differences between the races? Is there a cline in morphology between them? Since these questions concern only between sample variation the data set required for the analysis can be greatly reduced. For each sample site, the vector of variable means was calculated.

6.3.3.1 Racial differences?

Each sample was classified as being either pure XO (9 samples), pure XY (10 samples) or hybrid. Only samples for which the frequency of the fused X chromosome was between 0.1 and 0.9 are included in the sample of hybrid populations (22 samples). The hypothesis that there are differences between these three classifications is tested with a one way MANOVA. This yields test statistics $\lambda = 0.38$, $F_{22,56} = 1.58$, $P = 0.085$. There is therefore no evidence of a difference in mean vectors between these groups.

Only the XO and the XY classifications represent natural, well defined, groups. The group classified as hybrid actually includes populations from all across the hybrid zone and so represents varying degrees of hybridization. The comparison of pure race populations includes only 19 samples. Most effort was put into collecting hybrid samples in order to address the question of morphological differences associated with the karyotype itself rather than racial differences. To detect real racial differences a much larger sample would be needed, preferably spanning the hybrid zone in several different regions.

6.3.3.2 Clinal variation in morphology?

Having been unable to demonstrate any difference between the two pure races, it is perhaps perverse to expect to find any trend across the hybrid zone. However, it may well
be that there are real differences between the races but they are not revealed in the limited sample considered above.

The answer to the question above involves determining if there is any regression of the sample mean vectors on the sample site position. The frequency of the fused X chromosome is taken as a measure of each sample's position in the hybrid zone. It would be better if the distance of each sample from the centre of the hybrid zone could be used directly. However, since the centre of the zone does not run in a straight line there is no unique distance for each population. The real interest does not, however, lie in spatial variation in morphology; rather it is the relationship between morphology and the degree of hybridization that is of interest. Being the only diagnostic marker between the races, the frequency of the fused chromosome is the best available measure of the degree of hybridization. Fig. 6.2 shows the sample means of the 11 variables across the hybrid zone. These are plotted against a logit\(^1\) transformation of the frequency of the fused chromosome since this converts the expected tanh curve to a straight line (Bazykin, 1969). It appears that there may be a trend towards decreasing size as the frequency of the fused chromosome increases. The results of regression analysis of the sample means against logit(p) are given in Table 6.5. Seven of the eleven variables show a significant regression. Also shown in Table 6.5 is the square of the correlation coefficient between the variables and logit(p). This represents the proportion of the variation that is attributable to the regression. For most of the variables the regression accounts for only a small fraction of the total variation. One exception is HTHO for which the regression accounts for more than 20% of the variation.

Given that the variables are highly correlated, it is not surprising that more than one should show a significant regression. A multivariate analysis of variance of the regression of all eleven variables takes these correlations into account and confirms that there is a significant regression of the sample mean vectors against logit(p) [Wilks's\(\lambda = 0.641,\]

\[^1\text{logit}(p) = \ln\left(\frac{p}{1-p}\right)\]
The probability of this regression is only slightly less than the conventional 5% level. In transforming the frequency of the fused chromosome, pure populations were set at ±4. The true value for logit 0 or 1 should be ± infinity. Since using ±4 may bias the analysis, it was repeated but only including the 35 polymorphic populations. This confirmed the previous results ($\lambda = 0.371$, $F_{11.23} = 3.54$, $P = 0.005$).

All of the regression coefficients are negative, so all measurements decrease in magnitude as the frequency of the fused X chromosome increases. This suggests that XY grasshoppers are in fact slightly smaller than XO's, although this difference was not detected in the direct comparison between the races (section 3.3.1). This trend is in the same direction as the difference between karyotypes in mixed populations found in the previous study (Barton, 1979a).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Constant</th>
<th>regression coefficient $\beta$</th>
<th>$F$ D.F. = 1,52</th>
<th>$P$</th>
<th>$r^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>HFAC</td>
<td>0.325</td>
<td>-0.0016</td>
<td>6.75</td>
<td>0.0122 *</td>
<td>0.115</td>
</tr>
<tr>
<td>WFAC</td>
<td>0.873</td>
<td>-0.0012</td>
<td>3.67</td>
<td>0.0611</td>
<td>0.066</td>
</tr>
<tr>
<td>EYES</td>
<td>-0.190</td>
<td>-0.0030</td>
<td>9.03</td>
<td>0.0041 **</td>
<td>0.148</td>
</tr>
<tr>
<td>TOPL</td>
<td>1.557</td>
<td>-0.0019</td>
<td>2.94</td>
<td>0.0918</td>
<td>0.054</td>
</tr>
<tr>
<td>HBOT</td>
<td>1.529</td>
<td>-0.0022</td>
<td>8.58</td>
<td>0.0050 **</td>
<td>0.142</td>
</tr>
<tr>
<td>WBOT</td>
<td>1.284</td>
<td>-0.0015</td>
<td>5.28</td>
<td>0.0256 *</td>
<td>0.092</td>
</tr>
<tr>
<td>HFE</td>
<td>2.196</td>
<td>-0.0021</td>
<td>5.76</td>
<td>0.0200 *</td>
<td>0.099</td>
</tr>
<tr>
<td>HTIB</td>
<td>2.109</td>
<td>-0.0020</td>
<td>7.18</td>
<td>0.0099 **</td>
<td>0.121</td>
</tr>
<tr>
<td>MFEM</td>
<td>1.471</td>
<td>-0.0014</td>
<td>2.14</td>
<td>0.1497</td>
<td>0.039</td>
</tr>
<tr>
<td>WTHO</td>
<td>1.199</td>
<td>-0.0013</td>
<td>3.43</td>
<td>0.0696</td>
<td>0.062</td>
</tr>
<tr>
<td>HTHO</td>
<td>1.355</td>
<td>-0.0031</td>
<td>14.125</td>
<td>0.0004 ***</td>
<td>0.214</td>
</tr>
</tbody>
</table>

Table 6.5 regression of the eleven variables vs logit(P).

$F$ and $P$ are from a regression analysis of variance.
Figure 6.2. Sample means across the hybrid zone. Abscissa is logit of p. For p = 0 or 1, logit is set to ± 4. Only significant regression lines are shown (see Table 6.5)
6.3.3.3 Clinal variation other than size?

The regression of the individual variables above indicates that there is a change in the magnitude of seven of the variables across the hybrid zone. It is interesting to know if there is any trend in morphology other than size. Variation in size can be separated from variation in shape by taking the principal components of the sample mean vectors. The coefficients of these components and the proportion of the total variation for which they account are shown in Table 6.6. These are calculated from the matrix of covariances between sample means. Note that this is a different matrix from the total sample covariance matrix used in section 6.3.1, these only consider the between sample variation. The first component is approximately evenly loaded on all of the variables and so can be considered as a general measure of size. Size variation accounts for ≈73% of the between site variation.

It is interesting to note that the loadings of the principal components calculated on the between site covariance matrix are different from, (in both sign and magnitude), those calculated from the entire data set. Also, more of the variation between sites than between individuals is accounted for by size differences. Although these differences have not been investigated, they suggest that the pattern of co-variation of the variables is different at these different scales.
<table>
<thead>
<tr>
<th>Component</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eigen values</td>
<td>0.00248</td>
<td>0.00022</td>
<td>0.00017</td>
<td>0.00015</td>
<td>0.00011</td>
<td>0.00008</td>
<td>0.00005</td>
<td>0.00005</td>
<td>0.00004</td>
<td>0.00003</td>
<td>0.00002</td>
</tr>
<tr>
<td>% variance</td>
<td>73.0711</td>
<td>6.37086</td>
<td>4.96021</td>
<td>4.36154</td>
<td>3.30702</td>
<td>2.32032</td>
<td>1.58018</td>
<td>1.43375</td>
<td>1.23319</td>
<td>0.88193</td>
<td>0.47986</td>
</tr>
<tr>
<td>HFAc</td>
<td>0.01093</td>
<td>-0.00260</td>
<td>0.00068</td>
<td>0.00072</td>
<td>0.00270</td>
<td>0.00445</td>
<td>0.00284</td>
<td>0.00295</td>
<td>0.00324</td>
<td>-0.00052</td>
<td>-0.00036</td>
</tr>
<tr>
<td>WFAc</td>
<td>0.00999</td>
<td>0.00318</td>
<td>0.00128</td>
<td>0.00660</td>
<td>0.00256</td>
<td>-0.00101</td>
<td>-0.00027</td>
<td>-0.00338</td>
<td>0.00168</td>
<td>0.00267</td>
<td>-0.00037</td>
</tr>
<tr>
<td>Eyes</td>
<td>0.01754</td>
<td>-0.01188</td>
<td>0.00474</td>
<td>0.00062</td>
<td>-0.00031</td>
<td>-0.00106</td>
<td>0.00005</td>
<td>-0.00128</td>
<td>-0.00117</td>
<td>0.00011</td>
<td>0.00030</td>
</tr>
<tr>
<td>Topl</td>
<td>0.02096</td>
<td>-0.00123</td>
<td>-0.00881</td>
<td>0.00274</td>
<td>-0.00555</td>
<td>0.00040</td>
<td>0.00068</td>
<td>-0.00057</td>
<td>0.00027</td>
<td>-0.00048</td>
<td>0.00022</td>
</tr>
<tr>
<td>Hbot</td>
<td>0.01403</td>
<td>0.00391</td>
<td>0.00269</td>
<td>-0.00304</td>
<td>-0.00254</td>
<td>0.00198</td>
<td>0.00291</td>
<td>0.00105</td>
<td>-0.00293</td>
<td>0.00272</td>
<td>0.00024</td>
</tr>
<tr>
<td>Wbot</td>
<td>0.01106</td>
<td>0.00212</td>
<td>0.00351</td>
<td>0.00126</td>
<td>-0.00299</td>
<td>-0.00577</td>
<td>-0.00016</td>
<td>0.00327</td>
<td>0.00208</td>
<td>0.00041</td>
<td>-0.00047</td>
</tr>
<tr>
<td>Hfem</td>
<td>0.01741</td>
<td>-0.00018</td>
<td>-0.00264</td>
<td>-0.00250</td>
<td>0.000295</td>
<td>0.00010</td>
<td>-0.000258</td>
<td>0.00070</td>
<td>-0.00124</td>
<td>0.00046</td>
<td>-0.00297</td>
</tr>
<tr>
<td>Htib</td>
<td>0.01443</td>
<td>0.00013</td>
<td>-0.00235</td>
<td>-0.00171</td>
<td>0.000324</td>
<td>0.00017</td>
<td>-0.000358</td>
<td>0.000171</td>
<td>0.00046</td>
<td>0.00153</td>
<td>0.00251</td>
</tr>
<tr>
<td>Mfem</td>
<td>0.01759</td>
<td>0.00289</td>
<td>-0.00131</td>
<td>-0.00433</td>
<td>0.000415</td>
<td>-0.000338</td>
<td>0.00316</td>
<td>-0.00204</td>
<td>0.00017</td>
<td>-0.00181</td>
<td>0.00057</td>
</tr>
<tr>
<td>Wtho</td>
<td>0.01061</td>
<td>0.00344</td>
<td>0.00020</td>
<td>0.00707</td>
<td>0.000198</td>
<td>0.00059</td>
<td>-0.00039</td>
<td>0.00167</td>
<td>-0.00313</td>
<td>-0.00241</td>
<td>0.00039</td>
</tr>
<tr>
<td>Htho</td>
<td>0.01618</td>
<td>0.00410</td>
<td>0.00533</td>
<td>-0.00272</td>
<td>-0.00345</td>
<td>0.00293</td>
<td>-0.00264</td>
<td>-0.00208</td>
<td>0.00160</td>
<td>-0.00175</td>
<td>0.00005</td>
</tr>
</tbody>
</table>

Table 6.6 Principal components: population means

<table>
<thead>
<tr>
<th>Principal Component</th>
<th>Constant</th>
<th>regression coefficient β</th>
<th>F D.F. = 1,52</th>
<th>P</th>
<th>$r^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.00269</td>
<td>-0.00644</td>
<td>8.037</td>
<td>0.0065 **</td>
<td>0.133</td>
</tr>
<tr>
<td>2</td>
<td>-0.00016</td>
<td>0.00037</td>
<td>0.2718</td>
<td>0.6043</td>
<td>0.005</td>
</tr>
<tr>
<td>3</td>
<td>0.00057</td>
<td>-0.00137</td>
<td>5.1178</td>
<td>0.0279 *</td>
<td>0.089</td>
</tr>
<tr>
<td>4</td>
<td>-0.00008</td>
<td>0.00019</td>
<td>0.0096</td>
<td>0.7572</td>
<td>0.002</td>
</tr>
<tr>
<td>5</td>
<td>-0.00017</td>
<td>0.00040</td>
<td>0.6028</td>
<td>0.4410</td>
<td>0.011</td>
</tr>
<tr>
<td>6</td>
<td>0.00028</td>
<td>-0.00066</td>
<td>2.4324</td>
<td>0.1249</td>
<td>0.049</td>
</tr>
<tr>
<td>7</td>
<td>-0.00023</td>
<td>0.00056</td>
<td>2.5659</td>
<td>0.1152</td>
<td>0.047</td>
</tr>
<tr>
<td>8</td>
<td>0.00004</td>
<td>-0.00099</td>
<td>0.0674</td>
<td>0.7962</td>
<td>0.001</td>
</tr>
<tr>
<td>9</td>
<td>0.00005</td>
<td>-0.00011</td>
<td>0.1254</td>
<td>0.7247</td>
<td>0.002</td>
</tr>
<tr>
<td>10</td>
<td>0.00011</td>
<td>-0.00026</td>
<td>0.9664</td>
<td>0.3301</td>
<td>0.018</td>
</tr>
<tr>
<td>11</td>
<td>0.00003</td>
<td>-0.00007</td>
<td>0.1366</td>
<td>0.7132</td>
<td>0.051</td>
</tr>
</tbody>
</table>

Table 6.7 Regression of principal components vs Logit of karyotype frequency. F and P are from a regression analysis of variance.
The transformation to principal components yields uncorrelated measures of size and shape. The results of a regression analysis of these variables against logit(p) are shown in Table 6.7. Since the transformation to principal components preserves all the information in the original data, a multivariate test of regression on the principal components yields statistics identical to the regression of the untransformed data.

The first principal component (size) shows a highly significant regression (Table 6.5, Fig. 6.3). It is, therefore, clear that much of the regression seen in Table 6.5 is simply due to size. There is a clear size decrease as the frequency of the fused X increases. It is also clear that there is much variation in the samples not attributable to variation in the frequency of the fused chromosome. The regression accounts for only 13% of the total variation in size.

Although it accounts for less than 5% of the total variation, the third principal component also shows a significant regression (p = 0.0279). It is possible that this is simply due to chance. Excluding principal component 1, which is clearly significant, 10 analyses were carried out. With an entirely random data set the chance that at least one of these would yield a 'significant' result is 1 - 0.95^10 = 0.4. For this reason the probabilities quoted in Table 6.7 are not appropriate when more than one test is performed.

It must be remembered that the uncorrelated variables produced by the principal component transformation are arbitrary with respect to the hybrid zone. Each is produced such that it is uncorrelated with the others and accounts for as much of the variation as possible. There is no particular reason why these, as opposed to other combinations of the variables should show clinal variation. It is however interesting to determine if there is a trend across the cline other than in size. This is addressed by re-calculating the multivariate regression but on principal components 2 to 11 rather than the raw data. Contained in these components is all the variation other than size. The results of this analysis are, λ = 0.776, F_{10,43} = 1.244, P = 0.292. There is, therefore, no evidence of a significant morphometric trend across the hybrid zone other than in size. That trend is illustrated in Fig.6.3
Figure 6.3 Variation in size across the hybrid zone.
A plot of the first principal component against Logit(p), where p is the frequency of the fused X chromosome. The line shown is the regression line given in Table 6.7

6.3.4 Increased variation in the hybrid zone?

If the two pure races have different, genetically determined, morphology, one would expect hybridization to produce increased genetic variation within populations. This may be observed as increased morphometric variation in the centre of the zone.

Fig. 6.4 shows plots of the within sample site variance for each of the 11 variables. These are plotted against p(1-p), where p is the frequency of the fused chromosome. p(1-p) varies between 0 (p=0 or 1) and 0.25 (p=1/2). This is used rather than simply p since the expectation is that variance will be greatest in the centre of the hybrid zone than the edges. If morphometric variance reflected additive genetic variation then one would expect a linear relationship with p(1-p). From the figures it can be seen that there is no obvious trend. This is confirmed by a regression of the sample variances against p(1-p) [Table 6.8]. None of
the variables show a significant relationship with $p$. This is backed up by a multivariate regression ($\lambda = 0.706$, $F_{11,42} = 1.59$, $P = 0.137$). It should be noted that the use of regression analysis is not strictly valid in these cases. Regression assumes that the dependent variables follow a normal distribution whereas variances are expected to be $\chi^2$ distributed. However, from the plots in Fig. 6.4 there seems little reason to suspect that there is in fact a trend.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Constant</th>
<th>regression coefficient $\beta$</th>
<th>$F$ D.F. = 1,52</th>
<th>$P$</th>
<th>$r^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>HFAC</td>
<td>0.00197</td>
<td>-0.00106</td>
<td>0.7517</td>
<td>0.389</td>
<td>0.0143</td>
</tr>
<tr>
<td>WFAC</td>
<td>0.00098</td>
<td>0.00106</td>
<td>2.8048</td>
<td>0.099</td>
<td>0.0512</td>
</tr>
<tr>
<td>EYES</td>
<td>0.00415</td>
<td>-0.00098</td>
<td>0.1927</td>
<td>0.663</td>
<td>0.0037</td>
</tr>
<tr>
<td>TOPL</td>
<td>0.00225</td>
<td>-0.00142</td>
<td>1.4190</td>
<td>0.239</td>
<td>0.0266</td>
</tr>
<tr>
<td>HBOT</td>
<td>0.00154</td>
<td>0.00010</td>
<td>0.0103</td>
<td>0.919</td>
<td>0.0002</td>
</tr>
<tr>
<td>WBOT</td>
<td>0.00122</td>
<td>0.00089</td>
<td>1.7121</td>
<td>0.196</td>
<td>0.0312</td>
</tr>
<tr>
<td>HFEM</td>
<td>0.00129</td>
<td>0.00019</td>
<td>0.0824</td>
<td>0.775</td>
<td>0.0016</td>
</tr>
<tr>
<td>HTIB</td>
<td>0.00132</td>
<td>-0.00056</td>
<td>0.9984</td>
<td>0.322</td>
<td>0.0188</td>
</tr>
<tr>
<td>MFEM</td>
<td>0.00152</td>
<td>-0.00103</td>
<td>1.1803</td>
<td>0.282</td>
<td>0.0222</td>
</tr>
<tr>
<td>WTHO</td>
<td>0.00123</td>
<td>0.00049</td>
<td>0.5074</td>
<td>0.479</td>
<td>0.0097</td>
</tr>
<tr>
<td>HTHO</td>
<td>0.00130</td>
<td>0.00042</td>
<td>0.2700</td>
<td>0.605</td>
<td>0.0052</td>
</tr>
</tbody>
</table>

Table 6.8 Regression of sample variance vs $p(1-p)$.

$F$ and $P$ are from a regression analysis of variance.
Figure 6.4 Sample variance across the hybrid zone.
Variance is plotted against $p(1-p)$ where $p$ is the frequency of the fusion
6.4 Discussion.

In summary, the above analyses have demonstrated that there is a significant negative correlation between size and the frequency of the fused X chromosome across the hybrid zone at the Col de la Lombarde. There is no evidence of an effect of the chromosome itself on morphology, nor of any association between the amount of morphological variation and position in the zone. A previous study across the hybrid zone at Tende (Barton, 1979a) found that while there was no association between morphology and chromosome frequency, there was a significant difference between karyotypes within polymorphic populations. This study was based on a much smaller sample of only 345 grasshoppers from 19 populations (of which only 11 were polymorphic).

There are several possible explanations for the different conclusions reached by these two studies. The most obvious difference between the two is that they were carried out across different regions of the hybrid zone. The two sites are many kilometres apart so in no way can it be claimed that the same populations are being studied: it may well be that the pattern of morphometric variation is different in these two locations. However, Barton's conclusion is that the chromosomal fusion itself, or genes tightly linked to it, have a significant effect on morphology. Unless the chromosomal effect is in some way modified (i.e. by its interaction with different environmental conditions or genetic background), one would expect to also observe it at the Col de la Lombarde. Slightly different variables may have been used in the two studies. Although those used here were based on descriptions of Barton's original variables, the choice of exactly what features define the edges of a structure may have been different. However, the difference in karyotype mean shown by the previous analysis was highly significant ($\chi^2_1 = 426.8$): it seems unlikely that such a convincing conclusion would be so sensitive to slight differences in the variables. Moreover, much of the variation between karyotypes in Barton's data was attributable to size variation, which should be fairly insensitive to the exact choice of variables.
The method of analysis used in the previous study was quite different to that used here. While full multivariate tests are used in this study, the previous one was based on analysis of a single variable. The variable used was the discriminant function designed to pick out differences between XO and XY grasshoppers. Discriminant function analysis is primarily intended as an exploratory tool for elucidating the nature of the differences between groups (Manly, 1986; James & McCulloch, 1990). The entire data set should be used to test for these differences, since the discriminant function is a biased variable in that it is specifically designed to maximize between group variation.

Barton concedes that a full multivariate analysis should have been carried out on his data, but argues that the karyotypic difference is genuine since the same conclusion is reached from an analysis of the first principal component (which, as it is evenly loaded on all variables is taken to be a measure of size). The coefficients of the discriminant function were not given, but it is stated in the text that they were 'fairly evenly' loaded on all the variables. From this statement it appears that the first principal component and the discriminant function are approximately equivalent and are both measures of size and so cannot be considered as independent tests. However, since PCI is an unbiased variable, the result is more robust.

It is interesting that Barton found no significant differences between the population means of the discriminant function (P≈7%) despite the fact that they differed widely in karyotype frequency. Given that the mean discriminant function was different between karyotypes, one would have expected to see such differences. It would be interesting to re-analyse Barton's data using full, multivariate tests and compare the results with those of the present study.

While the present analysis has demonstrated a correlation between size and karyotype frequency, it can reveal little about the ultimate cause of this association. It is not suggested that the chromosome itself affects morphology; rather, the frequency of the fused chromosome is used as simple index to the degree of hybridization. Perhaps the most
An interesting interpretation of the correlation would be that the races have different, genetically determined morphology. However, such an interpretation is not possible. Essentially, the problem is that while populations vary in their karyotype frequency they also vary in position, with which may be associated different environmental conditions that affect morphology. It is possible that habitat type, population density and altitude all affect morphology. Unfortunately, these parameters are not available for analysis so their relative importance cannot be assessed.

An important factor in determining morphology may be population density. Since, in this study, population density was not censused, its effect cannot be estimated. However, even if the adult densities were known a further problem in interpretation would remain. Jackson (1992) has shown that in populations matched for vegetation type, (which is a good indicator of density), there is no significant difference between the average density of adult grasshoppers of the two races. However, the density of juveniles is significantly greater in XY populations than in either XO or hybrid populations. This pattern of juvenile density is seen in two transects at the Col de Lombarde as well as at four other regions of the hybrid zone. If juvenile density effects morphology, then assessing adult density would reveal little. Since the difference in juvenile density is independent of habitat type, it must be taken as a genetic difference between the races. Any effect it has on morphology must therefore also be taken as a genetic difference.

The present study has only considered a single portion of the hybrid zone. It would be interesting to know if the correlation between karyotype frequency and size is the same in other areas of the zone. Since only size variation has shown any significant trend, future studies should concentrate on this parameter only. Since all the variables are highly correlated with size (PC1), a much smaller suite of variables could be used. The best measurements would be the length of the rear femur (HFEM), and the length of the thorax (TOPL) since these show the greatest correlation with size ($r = 0.79$ and $r = 0.76$ respectively).
Any future analysis of morphological variation in *Podisma pedestris* would be most revealing if it were combined with detailed analysis of density and environmental variables across the hybrid zone. If this were done, the relative importance of genetic, as opposed to environmental differences, between the races might be investigated.
Chapter 7

Summary and conclusions.

The three pieces of work presented in this thesis are essentially independent, and have been discussed as such in the preceding chapters. Here I wish to simply summarise the conclusions of the previous chapters and to point out how they are, and are not, related.

Reinforcement.

The original reinforcement hypothesis of Dobzhansky specifically concerned the way in which reproductive isolation between incipient species may be increased by the evolution of pre-zygotic isolating mechanisms as an adaptive response to the production of inviable or sterile hybrids. The model developed in Chapter 3 shows how the hypothesis may be extended to the adaptive evolution of hybrid inviability in response to hybrid sterility, as suggested by Maynard Smith (1966b) and Coyne (1974). Such modification will only be favoured if hybrid fertility is low and there is competition between hybrid and non-hybrid sibs. Most of the arguments against the standard reinforcement hypothesis also apply to the adaptive evolution of hybrid inviability: recombination in areas of hybridization reduce the advantage to be gained by modification and the spread of reinforcing alleles may be impeded by gene flow from areas of sympatry, especially if, in these areas, modification is selected against. The argument that reinforcement is halted by recombination between the reinforcing and selected locus does not apply to the model considered since it is of the 'single allele' type (c.f. assortative mating for two alleles at a single locus, Felsenstein, 1981). Although the model has shown that this type of reinforcement is possible, there is little evidence to suggest inviability barriers to gene flow between species have, in general, evolved in this way.
Homogamy.

In Chapter 5 a model was developed to consider the effect of assortative fertilization for karyotype in the *Podisma pedestris* hybrid zone. The conclusions are, however, generally applicable. Provided that there is a polygamous mating system, so that variation in reproductive success is possible, homogamy will lead to frequency dependent selection against the rare form which, in combination with gene flow, will result in the maintenance of stable clines. Whereas assortative mating may lead to complete reproductive isolation between species, assortative fertilization can be, at best, only a partial barrier to gene flow. Even if, in those females that carry sperm of both genotypes, assortment is absolute, gene flow is possible since (unless mating is also assortative) there will always be some individual females that have mated only inter-racially.

**Podisma pedestris**

The width of the chromosomal cline in *Podisma pedestris* and the observed levels of dispersal are compatible with only weak selection against heterokaryotypes. It has previously been suggested that this selection may be due to the production of unbalanced gametes through non-disjunction during meiosis. However, it has been shown that the observed assortative fertilization will produce selection at least strong enough to account for the cline. Since it relies only on observed phenomena, a balance between assortment and dispersal is a better hypothesis for the maintenance of the cline. The cline width expected under this hypothesis is however much narrower than that actually seen.

The mechanism by which the assortment is brought about remains unknown. However, there is little reason to think that it has evolved through reinforcement. Assortment is observed, with comparable strengths in both inter-population crosses and within the zone itself. It is also worth mentioning that there is also no indication that hybrid inviability has been increased in this hybrid zone through 'reinforcement' as discussed in Chapter 3.
The analysis of morphometric variation in the hybrid zone failed to detect any difference between the two karyotypes, although an association between karyotype frequency and size was observed. This may indicate that the two races differ in size, but the possibility that the association is due to environmental variation cannot be excluded.
References


Barton N.H., & Hewitt G.M. Spatial patterns and population structure II Fluctuations in a chromosomal cline. *In prep*


-- 213 --


Blair A.P. 1941. Variation, isolating mechanisms and hybridization in certain toads. *Genetics* 26: 398-417


Dobzhansky Th. 1940. Speciation as a stage in evolutionary divergence. *Am. Nat.* 74: 312-3212


Felsenstein J. 1975. Genetic drift in clines which are maintained by migration and natural selection. *Genetics* 81: 191-207
Felsenstein J. 1981. Scepticism towards santa rosalia or why are there so few kinds of animal. *Evolution* 35: 124-138


-- 219 --


O'Donald P. 1960. Assortative mating in a population in which two alleles are segregating. Heredity 15: 389-396


-- 226 --


Stephens S.G. 1946. The genetics of 'corkey'. I. The new world alleles and their possible role as an interspecific isolating mechanism. *J. Genetics* **47**: 150-161

Stephens S.G. 1950. The genetics of 'corkey' II. Further studies on its genetic basis in relation to the general problem of interspecific isolating mechanisms. *J. Genetics* **50**: 9-20


-- 228 --


Westerman M., Barton N.H., & Hewitt G.M. 1987. Differences in DNA content between two chromosomal races of the grasshopper Podisma pedestris. Heredity 58 : 221-228


