TEKST NR 251   1993

Genotypic Proportions in Hybrid Zones

by

Freddy Bugge Christiansen,
Viggo Andreasen,
and
Ebbe Thue Poulsen

TEKSTER fra

Abstract:

Genotypic Proportions in Hybrid Zones

We study a hybrid zone between two populations of a diploid organism. The populations differ at one locus. Homozygotes have equal fitnesses and the heterozygote fitness is reduced by $\beta + \delta$ (β is the birth rate deviation and δ is the death rate deviation). The populations extend along a one dimensional continuous habitat, and migration occurs by diffusion of individuals. The model is formulated as a set of simple continuous time demographic models without age structure for the three genotypes, and the system is transformed into three new variables, the total population size N, the gene frequency p, and the deviation from Hardy-Weinberg proportions F. The gene frequency in a steady state cline always follows a hyperbolic tangent closely. Analysis of the asymptotic behavior of the cline far from the hybrid zone suggests a qualitative prediction of the shape of N, p and F over the zone. For weak selection the shape is determined by a central steepness of $\sqrt{(\beta + \delta)/4\sigma}$, as observed by Bazykin in 1969, where σ is the diffusion coefficient. For strong selection the cline is less steep than the Bazykin cline, and the form is dominated by the migration process. The steepness at the center of the cline is close to $\sqrt{b/4\sigma}$ where b is the birth rate of homozygotes.

The paper has been submitted to *Journal of mathematical Biology*

1 Introduction

When geographically disjoint, genetically differentiated populations come into contact, they may produce a narrow zone, where hybrids of inferior fitness are produced. The variation in the frequency of individual traits over the hybrid zone is characterized by a balance between dispersal of individuals and selection against hybrids. The understanding of the dynamics of hybrid zones has advanced significantly in recent years and this has led to a deeper understanding of naturally occurring hybrid zones, see the review by Barton and Hewitt (1985).

The basic model describes a one dimensional continuous habitat, where populations at the two extremes are fixed for alternative alleles at one locus. A classic assumption is that the heterozygote has a small fitness depression compared to the homozygotes which are assumed to have equal fitness (Bazykin, 1969a, 1969b, Barton, 1979). This model may be viewed as a slow-selection approximation to a stepping-stone model (Slatkin, 1973; Nagylaki, 1975), and it approximates the genotypic frequencies in the population at any time and place by the Hardy-Weinberg proportions. Thus, deviations due to selection are neglected, and so are deviations due to the Wahlund effect of population mixing. This latter approximation may seem less obvious, but it is allowed at equilibrium because the resulting cline has a low slope at any place so the variation in gene frequencies among immigrants is small at any place. These assumptions allow the cline to be described in terms of variation in the gene frequencies only.

The description obtained from this model of the clinal variation through a hybrid zone is very nice (Barton, 1979). An extension of the continuous model to the case of strong selection therefore would be preferable to the usual reversion to the stepping-stone model, in that discrete demes provide a model with qualitatively different properties. However, strong selection against the heterozygote produces significant deviations from the Hardy-Weinberg proportions, and we must allow for a steep cline and significant Wahlund variances. Thus, the gene frequency description becomes insufficient and we need a proper genotypic specification of the dynamics of the cline.

We study this problem in a simple continuous-time model of a population with overlapping generations and with a specified mode of population regulation. Selection will occur either as increased mortality or increased sterility of hybrid individuals. The genotypic frequencies at any point will be described by the gene frequency and Wright's measure for the relative deviation of the genotypic fre-

1

quencies from Hardy-Weinberg proportions. The movement of individuals will be
described by diffusion along the linear habitat, and reproduction by the assump-
tion that a breeding individual chooses a mate at random close to the place where
it is situated, and that the offspring is dropped immediately. This exceedingly
simple model will help to characterize important differences between hybrid zones
with weak and strong hybrid inferiority.

2   The Model

Consider an autosomal locus with two alleles A and a characteristic of each of
the two populations that meet in the hybrid zone. The densities of the three
genotypes AA, Aa and aa at location x at time t are denoted by $N_1(x,t)$, $N_2(x,t)$
and $N_3(x,t)$, and the total density is given by $N(x,t) = N_1(x,t) + N_2(x,t) +
N_3(x,t)$.

Mating is local and random, and birth and death rates are independent of
location. The death rate of individuals of genotype i is $d + \delta_i$, and the birth rate
contribution is $\frac{1}{2}b - \beta_i$ in that the birth rate of a mating between genotype i and
genotype j is $b - (\beta_i + \beta_j)$. We assume that the heterozygote is inferior in both
components of fitness, i.e. $\beta_2 \geq \beta_1$, β_3 and $\delta_2 \geq \delta_1$, δ_3.

The population density at any location is regulated by a density dependent
term in the death rate; this term is assumed to increase linearly with the to-
tal population density (logistic model), and to be independent of genotype and
location. The proportionality coefficient of this term is denoted by ι. The move-
ment of individuals is described by the diffusion coefficient σ, which is assumed
independent of genotype and constant throughout the area of the cline.

We assume equal fitness of the homozygotes, i. e. $\beta_1 = \beta_3$ and $\delta_1 = \delta_3$, and
these assumptions allow the existence of stationary clines, while, in general, any
difference in homozygote fitnesses will give traveling clines (Barton, 1979). For
simplicity, we put $\beta_1 = \beta_3 = 0$, $\delta_1 = \delta_3 = 0$, $\beta_2 = \beta$ and $\delta_2 = \delta$, where we assume
$d < b$, $0 \leq \beta \leq b/2$, $0 \leq \delta$, and $0 < \beta + \delta$. The model equations then become

$$\frac{\partial N_1}{\partial t} = \sigma \frac{\partial^2 N_1}{\partial x^2} + (b - B_1)p^2 N - (d + \iota N)N_1,$$

$$\frac{\partial N_2}{\partial t} = \sigma \frac{\partial^2 N_2}{\partial x^2} + (b - B_2)2pqN - (d + \delta + \iota N)N_2, \tag{1}$$

$$\frac{\partial N_3}{\partial t} = \sigma \frac{\partial^2 N_3}{\partial x^2} + (b - B_3)q^2 N - (d + \iota N)N_3,$$

where the birth rate deviations are given by

$$B_1 = \frac{\beta N_2}{pN}, \quad B_2 = \frac{\beta N_2}{2pqN} \quad \text{and} \quad B_3 = \frac{\beta N_2}{qN} \tag{2}$$

(Andreasen and Christiansen, 1993), and where p and q are the gene frequencies

$$p = \frac{2N_1 + N_2}{2N} \quad \text{and} \quad q = \frac{2N_3 + N_2}{2N}. \tag{3}$$

The model equations have the property that changes in the diffusion per time unit τ is equivalent to changes in the geographic scale of the model. Therefore, we may without loss of generality assume a given value, e. g. $\sigma\tau = 1$ of the diffusion coefficient, and this amounts to measuring distance in units of $\sqrt{\sigma\tau}$. Some of our results are valid under more general assumptions (see Appendix 1), but to keep the exposition as simple as possible, we restrict discussions to the model (1).

We study the interaction of two populations each monomorphic for one allele. To the far left in the cline we have a population monomorphic aa, i. e. $p(-\infty, t) = 0$, and to the far right we have a population monomorphic AA, i. e. $p(\infty, t) = 1$. Thus, we always assume that $N_1(-\infty, t) = N_2(-\infty, t) = 0$ and $N_2(\infty, t) = N_3(\infty, t) = 0$ for all t. The gene frequency $p(x, t)$ seen as a function of x describes at any time the genetic cline, i. e. the genetic change in going from one population to the other. The genotypic frequencies in the cline will be described by the normalized deviation from the Hardy-Weinberg proportions, $F(x, t)$, where the genotypic frequencies are given by

$$\frac{N_1}{N} = p^2 + Fpq, \quad \frac{N_2}{N} = 2pq(1 - F) \quad \text{and} \quad \frac{N_3}{N} = q^2 + Fpq. \tag{4}$$

In terms of the population densities this variable is

$$F = \frac{4N_1 N_3 - N_2^2}{(2N_1 + N_2)(2N_3 + N_2)}, \tag{5}$$

and its range of variation is between $\max(-p/q, -q/p)$ and 1. This normalized deviation from the Hardy-Weinberg proportions is undefined in monomorphic populations where $p = 0$ or $q = 0$. Therefore, its properties as a descriptive variable in the analysis of the model seem less desirable than those of the un-normalized deviation pqF (Nagylaki and Crow, 1974; Aronson and Weinberger,

3

1975; Hunt, 1980). The range of variation of this unnormalized deviation, however, is very dependent on the gene frequencies compared to the rather weak dependence of F. Further, observation of deviations from the Hardy-Weinberg proportions is closely linked to F, in that a statistical comparison of observed genotypic frequencies to the Hardy-Weinberg proportions is based on the approximate Gaussian-$(0,1)$ statistic $F\sqrt{n}$ where n is the number of individuals sampled (Brown, 1970). Therefore, the variable F provides a much more direct and powerful description of the genotypic proportions in the population.

Transforming equations (1) to the variables (N, p, F) produces the equations

$$\frac{\partial N}{\partial t} = \sigma\frac{\partial^2 N}{\partial x^2} + (b - d - \iota N)N - 2pq(1 - F)(2\beta + \delta)N, \tag{6}$$

$$\frac{\partial p}{\partial t} = \sigma\left(\frac{\partial^2 p}{\partial x^2} + 2\frac{\partial p}{\partial x}\frac{1}{N}\frac{\partial N}{\partial x}\right) + pq(1 - F)(\beta + \delta)(p - q), \tag{7}$$

$$\frac{\partial F}{\partial t} = \sigma\left(\frac{\partial^2 F}{\partial x^2} + \frac{2(1 - F)}{pq}\left(\frac{\partial p}{\partial x}\right)^2 - \frac{2(p - q)}{pq}\frac{\partial p}{\partial x}\frac{\partial F}{\partial x} + 2\frac{\partial F}{\partial x}\frac{1}{N}\frac{\partial N}{\partial x}\right)$$
$$-bF + (1 - F)\left(\beta F + \delta\left(2pq + (1 - 2pq)F\right)\right) \tag{8}$$

(see Appendix 2). Thus we have a set of equations describing the population density variation, the gene frequency variation and the variation in the deviation from the Hardy-Weinberg proportions.

3   The Classical Cline

An explicit solution to the model will not be produced, but we will provide a series of approximate solutions which will help to overview the biological properties of the model population.

From equation (7) we can produce the classical cline models analyzed by Bazykin (1969a) and Barton (1979). For $b - d$ large we may approximate equation (6) by

$$\frac{\partial N}{\partial t} = \sigma\frac{\partial^2 N}{\partial x^2} + (b - d - \iota N)N. \tag{9}$$

The equilibrium in this model is obtained when $dN/dt = 0$, and the solution to this equation is $\hat{N}(x) = K$ where

$$K = \frac{b - d}{\iota} \tag{10}$$

4

is the carrying capacity in a monomorphic population, i. e. a population where only one homozygote is present. Thus, if we assume that the population size has reached this equilibrium, then equation (7) simplifies, and with the additional slow selection approximation of $F = 0$ equation (7) is the model of Bazykin (1969a) and Barton (1979):

$$\frac{\partial p}{\partial t} = \sigma \frac{\partial^2 p}{\partial x^2} + pq(\beta + \delta)(p - q).$$

(11)

The solution at equilibrium $(\partial p / \partial t = 0)$ is given by

$$\hat{p}(x) = \tfrac{1}{2} \left(1 + \tanh \left((x - x_0)\sqrt{\frac{\beta + \delta}{4\sigma}} \right) \right),$$

(12)

where x_0 is the center of the cline which may be at any location (Bazykin, 1969b; Barton, 1979). In the following the center will always be assumed to be at $x_0 = 0$.

4 Stability of clines and panmictic populations

The discussion of a cline may be partitioned into considerations of the central region where local populations are unmistakenly polymorphic and of the tails where the alternative allele exists at ideomorphic frequencies, i. e. frequencies where a normal investigation of the population has a low probability of disclosing the allele. The region of greatest interest is the central transition zone which we discuss in Section 6. The tails of the cline describe the nature of introgression of an allele into the region dominated by the alternative allele, and in Section 5 we examine some properties of this introgression by obtaining an asymptotic description of the tails.

An equilibrium cline is of interest only if it is stable as a stationary solution to the system of equations given by (1) or (6)-(8). The cline has $\hat{p}(-\infty) = 0$ and $\hat{p}(\infty) = 1$, i.e. $\hat{N}_1(-\infty) = \hat{N}_2(-\infty) = 0$ and $\hat{N}_3(\infty) = \hat{N}_2(\infty) = 0$ from our assumptions. The model equations, however, are invariant to translations along the geographic scale. The center of the cline therefore may be at any location, and it will be impossible to prove stability by the usual method of linearization. There are two related reasons for this. First, when the system is translation invariant, then translates of a stationary cline are also stationary, and so a continuum of stationary clines exists. Analysis of stability by the linearization method concludes exponential stability or instability, which is possible only when

5

the stationary solution under consideration is isolated. Secondly, from translates of the cline equations straghtforward differentiation with respect to the parameter of translation shows that the derivative of the cline is an eigenfunction of the linearization corresponding to the eigenvalue 0 (Poulsen, 1989).

Thus, we cannot hope to prove by simple means that the model has a stable stationary cline, the best we can do is to insist that our candidates are not manifestly unstable. In particular, we shall require that far out in the tail, where the variations in the population structure as x varies are negligble, we should not be close to an unstable equilibrium for the corresponding model without dependence on x. That is, we should not be close to an unstable equilibrium in the corresponding panmictic model. Therefore, we will initially discuss the dynamics of a panmictic population whose dynamics is given by the ordinary differential equations

$$\frac{dN_1}{dt} = (b - B_1)p^2 N - (d + \iota N)N_1,$$

$$\frac{dN_2}{dt} = (b - B_2)2pqN - (d + \delta + \iota N)N_2, \qquad (13)$$

$$\frac{dN_3}{dt} = (b - B_3)q^2 N - (d + \iota N)N_3,$$

where p, q, B_1, B_2 and B_3 are given in terms of N_1, N_2 and N_3 by equations (2) and (3).

We show in Appendix 3 that $(N_1, N_2, N_3) = (0, 0, K)$ is a stable stationary point of (13), where K is the carrying capacity given by equation (10). We also show that for an arbitrary solution $(N_1(t), N_2(t), N_3(t))$ to (13) with $N_1(t), N_2(t) > 0$ the limit $\lim_{t \to \infty} F(t)$ exists. This is shown under the non-resonance condition that none of the three non-negative real numbers b, $b - d$ and $\beta + \delta$ is a linear combination of the other two with non-negative integer coefficients. The possible values of the limit of F are found and discussed in Appendix 3.

By replacing the variables (N_1, N_2, N_3) by (N, p, F) in equations (13) as in Section 2 we get

$$\frac{dN}{dt} = (b - d - \iota N)N - 2pq(1 - F)(2\beta + \delta)N, \qquad (14)$$

$$\frac{dp}{dt} = pq(1 - F)(\beta + \delta)(p - q), \qquad (15)$$

$$\frac{dF}{dt} = -bF + (1 - F)(\beta F + \delta(2pq + (1 - 2pq)F)) \tag{16}$$

(see Appendix 2). If selection is sufficiently weak, i. e. if $\beta + \delta < b$, then $(N, p, F) = (K, 0, 0)$ is the only equilibrium of this system with $N > 0$ and $p = 0$, and this equilibrium is stable (Appendix 3). For strong selection, i. e. if $\beta + \delta > b$, then $(K, 0, 0)$ is an unstable equilibrium, but a second equilibrium $(K, 0, \check{F})$, where

$$\check{F} = 1 - \frac{b}{\beta + \delta}, \tag{17}$$

exists and is stable (Appendix 3).

The conclusion is that in a stable cline we have $N_1(x) \to 0$, $N_2(x) \to 0$, and $N_3(x) \to K$ as $x \to -\infty$, and $N_1(x) \to K$, $N_2(x) \to 0$, and $N_3(x) \to 0$ as $x \to \infty$, because far out in the tails of the cline the population should approximate a stable panmictic population. Therefore, the population density and the gene frequencies approach the stable panmictic equilibria, in that $N(x) \to K$ and $p(x) \to 0$ as $x \to -\infty$, and $N(x) \to K$ and $q(x) \to 0$ as $x \to \infty$. The genotypic composition, however, is undefined at the stable panmictic equilibria, but by the requirement that the genotypic composition should be close to that in a population converging to a stable equilibrium in a panmictic population, we have $F(x) \to 0$ if $\beta + \delta < b$ and $F(x) \to \check{F}$ if $\beta + \delta > b$ as $x \to \pm\infty$.

5   Shape of Introgression

An equilibrium solution $(\hat{N}_1(x), \hat{N}_2(x), \hat{N}_3(x))$ to (1) satisfies the equations

$$
\begin{aligned}
\sigma d^2 \hat{N}_1/dx^2 &= (b - \hat{B}_1)\hat{p}^2 \hat{N} - (d + \iota \hat{N})\hat{N}_1, \\
\sigma d^2 \hat{N}_2/dx^2 &= (b - \hat{B}_2)2\hat{p}\hat{q}\hat{N} - (d + \delta + \iota \hat{N})\hat{N}_2, \\
\sigma d^2 \hat{N}_3/dx^2 &= (b - \hat{B}_3)\hat{q}^2 \hat{N} - (d + \iota \hat{N})\hat{N}_3,
\end{aligned}
\tag{18}
$$

where we, as usual, write $\hat{q} = 1 - \hat{p}$. In Appendix 4 we consider the left tail of a solution to (18) with the properties of a stable cline, i. e. we consider the solution $(\hat{N}_1(x), \hat{N}_2(x), \hat{N}_3(x))$ as $x \to -\infty$. This analysis is done on a linearized version of (18) at the point $(0, 0, K)$, and we assume for simplicity that coefficients of potentially dominating terms are different from 0. This assumption is consistent with the requirement for a stable cline found in Section 4. If the eigenvalues of the linearization are pairwise distinct and satisfy the non-resonance condition,

then $\hat{N}_1(x)$, $\hat{N}_2(x)$, and $\hat{N}_3(x) - K$ vanish exponentially towards 0 as $x \to -\infty$. We only have to consider the eigenvalues

$$\kappa_1 = \sqrt{\frac{b}{\sigma}}, \quad \kappa_2 = \sqrt{\frac{\beta + \delta}{\sigma}}, \quad \kappa_3 = \sqrt{\frac{b - d}{\sigma}},$$

and formulas for the asymptotic behavior of $\hat{N}(x)$, $\hat{p}(x)$ and $\hat{F}(x)$ are derived in Appendix 4. Below we state these asymptotic formulas, and in these C_1, C_2 and C_3 denote positive constants whose values are to be determined from the actual position of the cline, and the relation $\approx$ indicates that the absolute value of the difference between the left hand side and the right hand side is at most $M e^{(\kappa+\epsilon)x}$ for some positive numbers M and ϵ, with κ denoting the coefficient occurring in the exponential factor on the right hand side.

Weak selection For $\beta + \delta < b$ we have $\hat{F}(x) \to 0$ as $x \to -\infty$, and four different cases must be distinguished:

1. For $\beta + \delta < b - d$ and $4(\beta + \delta) < b$ we have

$$\hat{N}(x) - K \approx -\frac{2K(2\beta + \delta)}{b - d - (\beta + \delta)} C_2 e^{\kappa_2 x} \tag{19}$$

$$\hat{p}(x) \approx C_2 e^{\kappa_2 x} \tag{20}$$

$$\hat{F}(x) \approx \frac{2(\beta + 2\delta)}{b - 4(\beta + \delta)} C_2 e^{\kappa_2 x} \tag{21}$$

2. For $\beta + \delta < b - d < b < 4(\beta + \delta)$ the asymptotic behavior of $\hat{N}(x)$ and $\hat{p}(x)$ is given by (19) and (20), while

$$\hat{F}(x) \approx C_1 e^{(\kappa_1 - \kappa_2)x} \tag{22}$$

3. For $b - d < \beta + \delta < 4(\beta + \delta) < b$ the asymptotic behavior of $\hat{p}(x)$ and $\hat{F}(x)$ is given by (20) and (21), while

$$\hat{N}(x) - K \approx -C_3 e^{\kappa_3 x} \tag{23}$$

4. For $b - d < \beta + \delta < b < 4(\beta + \delta)$ the asymptotic behavior of $\hat{N}(x)$, $\hat{p}(x)$ and $\hat{F}(x)$ is given by (23), (20) and (22).

Strong selection For $\beta + \delta > b$ we have $\hat{F}(x) \to \check{F}$ as $x \to -\infty$ where $\check{F}$ is given by equation (17), and two different cases must be distinguished:
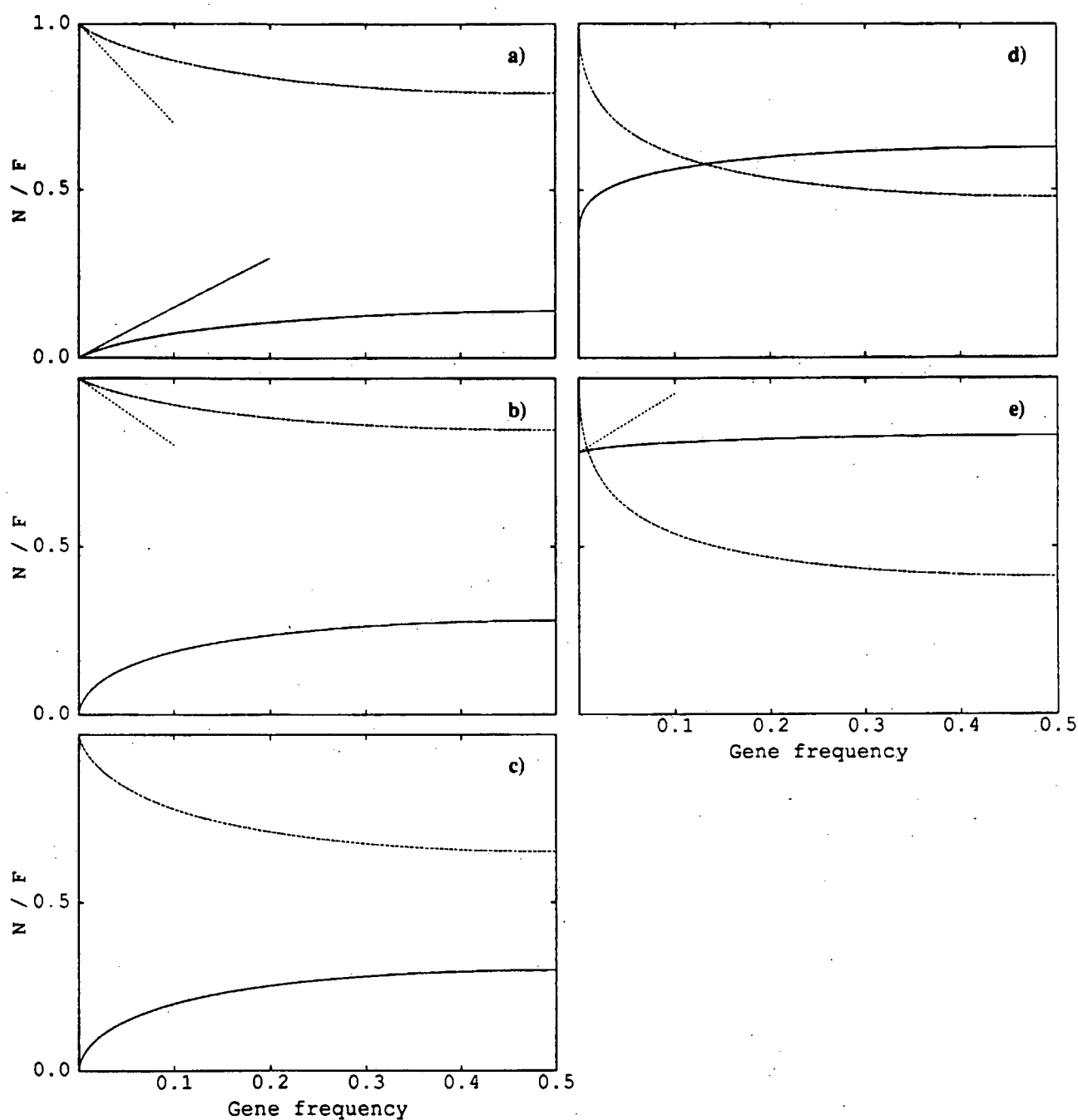
8

Figure 1: The variation in $\hat{F}$ (solid line) and $\hat{N}$ (broken line) as a function of gene frequency $\hat{p}$ in the transition zone. **a:** Slow selection $\beta + \delta < \frac{1}{4}b$ and $\beta + \delta < b - d$. **b:** Intermediate selection $\frac{1}{4}b < \beta + \delta < b$ and $\beta + \delta < b - d$. **c:** Intermediate selection $\frac{1}{4}b < \beta + \delta < b$ and $b - d < \beta + \delta$. **d:** Rather strong selection $b < \beta + \delta < 4b$. **e:** Strong selection $4b < \beta + \delta$. The straight lines in **a**, **b**, and **e** indicate the tangents found from the linear analysis; all other tangents at $\hat{p} = 0$ are vertical. The parameters are $b = 2$, $K = 1$, $\sigma = 0.004$, $\beta = 0$ and **a:** $d = 1.5$, $\delta = 0.3$; **b:** $d = 0.5$, $\delta = 0.75$; **c:** $d = 1.5$, $\delta = 0.75$; **d:** $d = 1.5$, $\delta = 3$; **e:** $d = 1.5$, $\delta = 10$.

9

Table 1: Asymptotic behavior of the gene frequency p and the fixation index F in the cline model as $x \to -\infty$. The change in the gene frequency is described as the rate of change with location. The change in the relative deviation from Hardy-Weinberg proportions is given in terms of the rate of change in F as a function of the gene frequency.

Selection	$p \propto e^{\kappa x}$ κ	$F - F^e \propto p^\alpha$ F^e	α	Figure
$\delta + \beta < \frac{1}{4}b$	$\sqrt{\frac{\delta+\beta}{\sigma}}$	0	1	2a
$\frac{1}{4}b < \delta+\beta < b$	$\sqrt{\frac{\delta+\beta}{\sigma}}$	0	$\sqrt{\frac{b}{\delta+\beta}} - 1$	2b,c
$b < \delta+\beta < 4b$	$\sqrt{\frac{b}{\sigma}}$	$1 - \frac{b}{\beta+\delta}$	$\sqrt{\frac{\delta+\beta}{b}} - 1$	2d
$4b < \delta+\beta$	$\sqrt{\frac{b}{\sigma}}$	$1 - \frac{b}{\beta+\delta}$	1	2e

5. For $b < \beta + \delta < 4b$ the asymptotic behavior of $\hat{N}(x)$ is given by (23), while

$$\hat{p}(x) \approx C_1 e^{\kappa_1 x} \tag{24}$$
$$\hat{F}(x) - \check{F} \approx D e^{(\kappa_2 - \kappa_1)x}, \tag{25}$$

where D may be positive, negative or 0 (our numerical results suggest that this constant is always positive).

6. for $4b < \beta + \delta$ the asymptotic behavior of $\hat{N}(x)$ and $\hat{p}(x)$ is given by (23) and (24), and

$$\hat{F}(x) - \check{F} \approx \frac{2 b^2 (\beta + 2\delta)}{(\beta + \delta)^2 (\beta + \delta - 4b)} C_1 e^{\kappa_1 x} \tag{26}$$

Thus, in the case of strong selection the convergence of the gene frequency $\hat{p}$ to zero is governed by the rate $\kappa_1 = \sqrt{b/\sigma}$ which does not depend on the selection parameters β and δ. The spatial variation in the gene frequencies far from the hybrid zone therefore is *independent of selection*.

10

Table 2: Asymptotic behavior of population size N in the cline model as $x \to -\infty$. The change in the population size is given in terms of the rate of change in N as a function of the gene frequency.

Selection	$N - K \propto p^{\gamma}$ γ	Figure
$\delta + \beta < b - d$	1	2a,b
$b - d < \delta + \beta < b$	$\sqrt{\frac{b-d}{\delta+\beta}}$	2c
$b < \beta + \delta$	$\sqrt{\frac{b-d}{b}}$	2d,e

5.1 Genetic shape of the cline

Although we haven't shown that $\hat{p}(x)$ is increasing, it certainly is for x near $-\infty$. Our numerical results suggest that it is in general, and we find it illustrative to view $\hat{N}$ and $\hat{F}$ as functions of $\hat{p}$. The assertions below follow immediately from the asymptotic results 1–6 above, and the variation of $\hat{p}(x)$ in the left tail of the cline for $x \to -\infty$ is summarized in Table 1.

Weak selection The population size $\hat{N}$ varies approximately as a linear function of $\hat{p}$ for $\hat{p}$ small and $\beta + \delta < b - d$ (Figs. 1 a, b and Table 2). For $b > \beta + \delta > b - d$ the population size $\hat{N}$ varies as a power function of $\hat{p}$ with exponent $\sqrt{(b-d)/(\beta+\delta)}$ less than 1. Thus, $\hat{N}$ changes very rapidly as a function of $\hat{p}$ as $\hat{p} \to 0$ (Fig. 1 c and Table 2).

The deviation from Hardy-Weinberg proportions $\hat{F}$ varies as a linear function of $\hat{p}$ for $\hat{p}$ small and $\beta + \delta < \frac{1}{4}b$ (Fig. 1 a and Table 1). For $\beta + \delta > \frac{1}{4}b$ the relative deviation $\hat{F}$ varies as a power function of $\hat{p}$ with exponent $\sqrt{b/(\beta+\delta)} - 1$ less than 1, and so $\hat{F}$ changes very rapidly as a function of $\hat{p}$ as $\hat{p} \to 0$ (Figs. 1 b, c and Table 1).

The shifts in $\hat{F}$ and $\hat{N}$ from linear to power functions of $\hat{p}$ occur at different

values of the selection parameter $\beta + \delta$. The change from proportionality to very rapid changes in the two variables therefore are unrelated.

Strong selection In the case of strong selection, where $\beta + \delta > b$, the most conspicuous result is that the deviation $\hat{F}$ from Hardy-Weinberg proportions does not vanish as x tends to $-\infty$. Rather, it converge to the limit $\breve{F}$ given by equation (17). The deviation $\hat{F} - \breve{F}$ varies as a power function of $\hat{p}$ with exponent $\sqrt{(\beta + \delta)/b} - 1$ less than 1 for $\beta + \delta < 4b$, and so $\hat{F}$ changes rapidly as a function of $\hat{p}$ as $\hat{p} \to 0$ (Fig. 1 d and Table 1). For $\beta + \delta > 4b$ the deviation $\hat{F} - \breve{F}$ varies as a linear function of $\hat{p}$ (Fig. 1 e and Table 1). The population size $\hat{N}$ varies as a power function of $\hat{p}$ with exponent $\sqrt{(b - d)/b}$ less than 1, so $\hat{N}$ changes rapidly as a function of $\hat{p}$ as $\hat{p} \to 0$ (Figs. 1 d, e and Table 2).

By definition, the birth rate of the heterozygote must be positive, so we have $2\beta \leq b$. Thus, strong selection requires that $\delta > b - \beta \geq \frac{1}{2}b$, i. e. a rather strong viability depression for the heterozygote.

6   The transition zone

The central region of the cline, where the local populations are polymorphic with both alleles present in non-negligible proportions, is the region of transition between local populations dominated by the features of the two original populations. The study of the transition zone will use directly equations (6)–(8) for the biologically interesting parameters N, p and F, and from these equations an equilibrium solution $(\hat{N}(x), \hat{p}(x), \hat{F}(x))$ satisfies the cline equations:

$$\sigma d^2\hat{N}/dx^2 = -(b - d - \iota\hat{N})\hat{N} + 2\hat{p}\hat{q}(1 - \hat{F})(2\beta + \delta)\hat{N}, \tag{27}$$

$$\sigma\hat{N}d^2\hat{p}/dx^2 = -2\sigma\frac{d\hat{p}}{dx}\frac{d\hat{N}}{dx} - \hat{p}\hat{q}(1 - \hat{F})(\beta + \delta)(\hat{p} - \hat{q})\hat{N}, \tag{28}$$

$$\sigma\hat{p}\hat{q}\hat{N}d^2\hat{F}/dx^2 = -2\sigma(1 - \hat{F})\hat{N}\left(\frac{d\hat{p}}{dx}\right)^2 + 2\sigma(\hat{p} - \hat{q})\hat{N}\frac{d\hat{p}}{dx}\frac{d\hat{F}}{dx} - 2\sigma\hat{p}\hat{q}\frac{d\hat{F}}{dx}\frac{d\hat{N}}{dx}$$
$$+ b\hat{p}\hat{q}\hat{F}\hat{N} - \hat{p}\hat{q}(1 - \hat{F})\hat{N}\left((\beta + \delta)\hat{F} + 2\hat{p}\hat{q}(1 - \hat{F})\delta\right). \tag{29}$$

Equation (29) is singular for $p = 0$ or $p = 1$ — this is the reason why the genotypic equations (1) form a convenient point of departure for the study of the asymptotics in the tails of the cline.
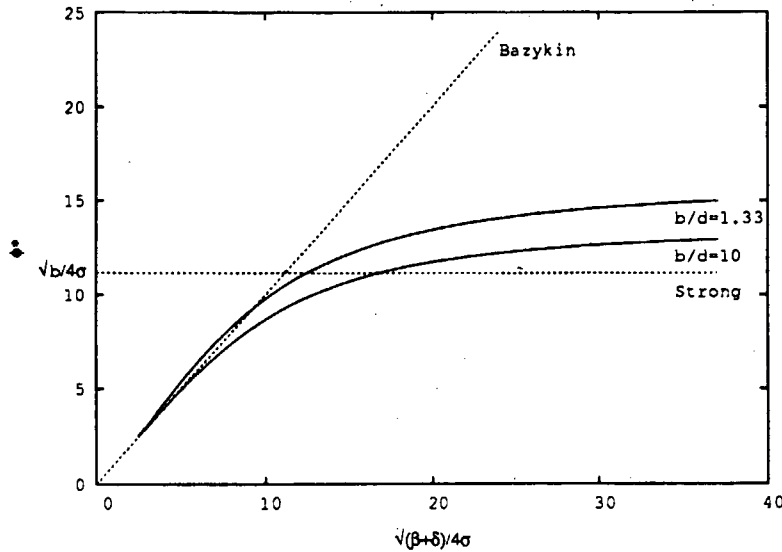
Figure 2: Best fit ϕ^* of the form $\frac{1}{2}(1 + \tanh \phi x)$ to the equilibrium gene frequency $\hat{p}(x)$ as a function of Bazykin's parameter $\sqrt{(\beta + \delta)/4\sigma}$ (solid lines). The parameter ϕ^* is compared to the slow selection parameter $\sqrt{(\beta + \delta)/4\sigma}$ and to the strong selection parameter $\sqrt{b/4\sigma}$ suggested by the linear analysis of the tails of the cline (broken lines). The diffusion coefficient is $\sigma = 0.004$ and the carrying capacity is $K = 1$ in this and the following figures. The birth rate parameters are fixed at $b = 2$ and $\beta = 0$, and the two curves correspond to $d = 0.2$ or 1.5 with δ varying.

Our description of the tails of the hybrid zone relies on a linear approximation that is valid only when the deviation from the asymptotic point is small. For strong selection $(b < \beta + \delta)$, our description holds only for $|x|\sqrt{b/\sigma} \gg 1$, and for weak selection $(b > \beta + \delta)$, the description is valid only for $|x|\sqrt{(\beta + \delta)/\sigma} \gg 1$. Strictly speaking, this provides no information on the width of the transition zone, but our numerical results indicate that

$$W_s \approx 2\sqrt{\frac{\sigma}{b}} \quad \text{and} \quad W_w \approx 2\sqrt{\frac{\sigma}{\beta + \delta}}$$

are very good estimates of the width for strong and weak selection, respectively.

Numerical solutions were obtained by solving equations (6)–(8) over a time period long enough to obtain stationarity. The solutions were calculated using centered differences in the spatial dimension and a variable step-size 4th order

13

Runge-Kutta algorithm along the time axis. We use $\sigma = 0.004$ in all illustrations. For $\beta + \delta > 1$ we fixed the gene frequency to $\bar{p} = 0$ at $x = -1$ and to $p = 1$ at $x = 1$ and used a step size of $\Delta x = 1/200$. For $0.1 < \beta + \delta < 1$ we fixed p at $x = \pm 3$ and used $\Delta x = 3/400$. The solution appears to be invariant to changes in the specifics of the boundary conditions and in the initial data.

The numerical solutions suggest that the equilibrium gene frequency $\hat{p}(x)$ is similar to a hyperbolic tangent. We offer no theoretical arguments for the validity of this approximation in the transition zone, but apparently the behavior of the tails in combination with obvious scaling and symmetry requirements impose so strong constraints on $\hat{p}(x)$ that the approximation holds within few percent. For all cases we investigated, the distance I_ϕ defined by

$$I_\phi^2 = \int_{-\infty}^{\infty} (\hat{p}(x) - \tfrac{1}{2}(1 + \tanh \phi x))^2 \, dx$$

did not exceed 2×10^{-3} for the best fitting parameter ϕ^* and the graph of $\hat{p}(x)$ was indistinguishable from the corresponding hyperbolic tangent. For weak selection $\phi^* \approx \tfrac{1}{2}\kappa_2 = \sqrt{(\beta + \delta)/4\sigma}$ as predicted by the Bazykin cline in equation (12). The rate parameter ϕ^* for strong selection is somewhat larger than the slope $\tfrac{1}{2}\kappa_1 = \sqrt{b/4\sigma}$ suggested by the linear analysis of the tails (Fig. 2).

The population depression and the deviation from Hardy-Weinberg proportions in the equilibrium cline are maximal at the center of the cline where $p = \tfrac{1}{2}$ which is assumed at $x = 0$. Both maxima increase with the strength of selection (Figs. 3–5) as expected from biological considerations. It is remarkable, however, that the population depression induced by fertility selection is much stronger that the depression caused by viability selection while the two types of selection produce comparable deviations from Hardy-Weinberg proportions. A comparison of Fig. 3 and Fig. 4 shows that the fixation index F for a given selection pressure is essentially controlled by the birth rate b, while the population size N is loosely linked to the value of b/d which is a rough measure of the amount of competition in the population.
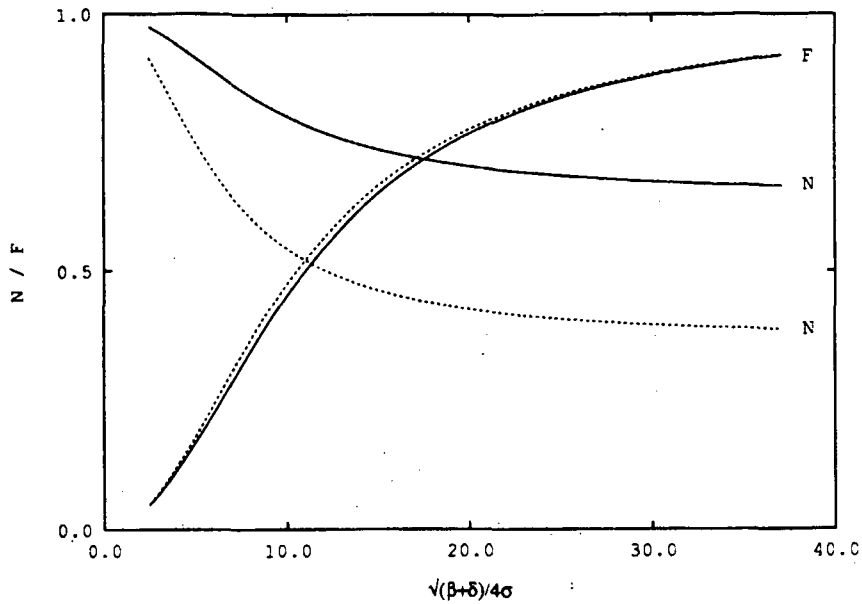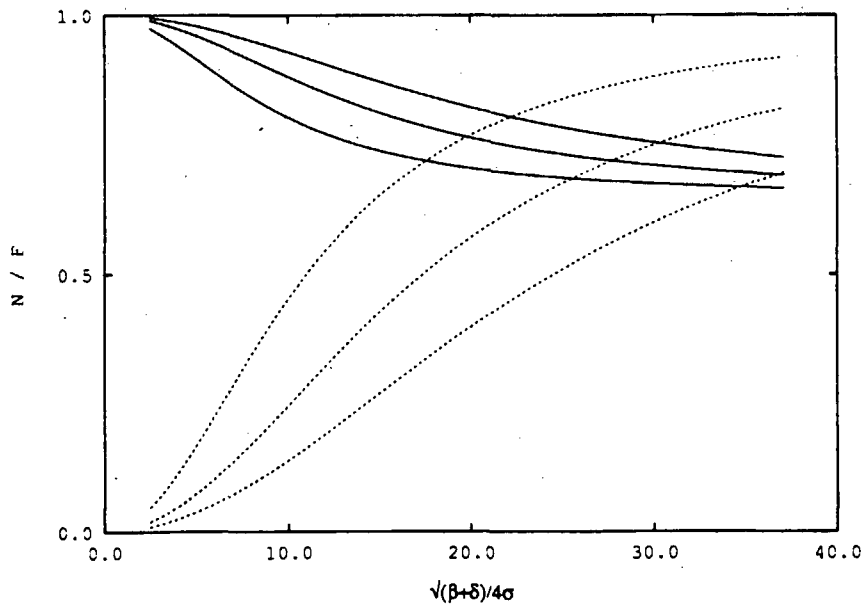
14

Figure 3: . The effect of mortality selection against heterozygotes on the minimal population density $\hat{N}$ at $p = 1/2$ (decreasing curves) and maximal deviation from Hardy-Weinberg proportions $\hat{F}$ at $p = 1/2$ (increasing curves). All parameters are as in Fig. 2 with $b = 2$, $\beta = 0$ and δ varying. The death rate is $d = 0.2$ (solid lines) or 1.5 (dotted lines).

Figure 4: . The effect of mortality selection against heterozygotes on the minimal population density $\hat{N}$ at $p = 1/2$ (decreasing curves) and maximal deviation from Hardy-Weinberg proportions $\hat{F}$ at $p = 1/2$ (increasing curves). The parameters are $\sigma = 0.004$ and $K = 1$ with $\beta = 0$ and δ varying. In both sets of curves $b/d = 10$ while the death rates varies as $d = 0.2$ (dotted lines), $d = 0.5$ (broken lines) and $d = 1.0$ (solid lines).
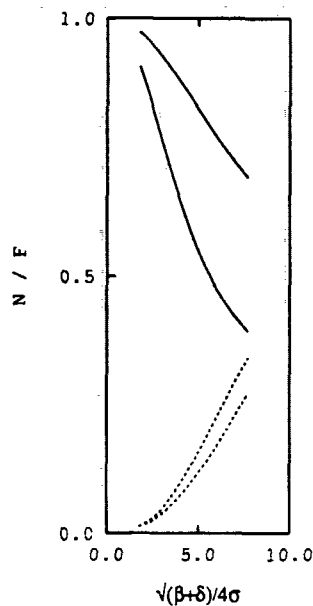
15

Figure 5: . The effect of fertility selection against heterozygotes on the minimal population density $\hat{N}$ at $p = 1/2$ and maximal deviation from Hardy-Weinberg proportions $\hat{F}$ at $p = 1/2$. All parameters are as in Fig. 2 except $\delta = 0$ while β varies. Note the strong population depression compared to the effect of mortality selection in Figs. 3 and 4.

7 Discussion

The cline model of population structure exists in two versions in population genetics, a discrete version and a continuous version. The continuous version we consider here is due to Wright (1946) and it is essentially the diffusion model of Fisher (1937). The discrete version due to Malécot (1948) and Kimura (1953) is known as the stepping-stone model, and it considers a chain of equal populations which exchange migrants with the two immediate neighbors. The study of selection in the stepping-stone model has advanced by considering the diffusion model as an approximation to the stepping-stone model in the case of weak selection (Haldane, 1948; Slatkin, 1973; Nagylaki, 1975). Nagylaki (1975, 1989) show very elegantly that in populations with non-overlapping generations the diffusion approximation necessitates that selection is assumed to be weak. The generation time goes to zero in the diffusion limit, and the rate of selection per time unit is constant, so the weak selection model of Bazykin is obtained. The stepping-stone model itself, however, is difficult to analyze, so the comparison of

16

results for weak and strong selection is hard in the discrete habitat model. The alternative approach that we have adopted here is to consider the continuous habitat model with overlapping generations as the primary model in the tradition of Wright (1946), and study weak and strong selection within this model. Therefore, the migration process is described by a diffusion approximation, and reproduction and selection is described by the simplest demographic model to allow the study of strong selection.

The main difference between weak and strong selection is in the deviation from Hardy-Weinberg proportions along the cline. For strong selection the genotypic frequencies in the population will not be close to Hardy-Weinberg proportions. In addition, the local population size will be influenced by selection. The resulting model is a three dimensional non-linear diffusion model, so formal analysis can not be carried far. To obtain information about the equilibrium cline, we study the asymptotic behavior of the tails of the cline by means of a linearization approach inspired from the qualitative study of non-linear differential equations, where the linear analysis provides a rough idea about the phase portrait (Guckenheimer and Holmes, 1983). In the central part of the cline we have to resort to numerical calculations.

Different migration distributions may provide differences in the nature of the migration effects, so in the clines we have studied caution is warranted when the cline pattern is interpreted. Nevertheless, we believe that our analysis of the tails of the cline provides important information about the nature of the processes that shape the cline. We show that for strong selection the pattern of gene-frequency variation in the tails is virtually independent of selection. For sufficiently strong selection the recruitment of hybrids as offspring of hybrids is negligible in the tails and most hybrid individuals are supplied by immediate offspring of immigrants. This, we conjecture, is still the case if, e.g., individuals have a maximal rate of dispersal. The "tails" of the cline is just moved closer to the hybrid zone.

The equilibrium cline in all situations resembles a hyperbolic tangent. For weak selection the steepness of the cline varies with the squareroot of the selection coefficients, but for strong selection the steepness is determined by migration rather independently of the strength of selection (Fig. 2). This suggests a classification of hybrid zones into weak and strong selection zones, where the central slope of the cline only reflects the strength of selection in the weak selection hybrid zone. In weak selection zones the genotypic frequencies remain near Hardy-Weinberg proportions throughout the zone, and the width of the zone is

determined by $2\sqrt{\sigma/(\beta + \delta)}$. Thus, the weak selection zone has the appearance of a migration-selection balance equilibrium. Strong selection zones are characterized by a substantial deviation from Hardy-Weinberg proportions and a width determined by $2\sqrt{\sigma/b}$. For strong selection the fixation index remains positive and large throughout the zone, and the local populations bear the signature of a balance between mixing and depletion of the "foreign" types through the reproduction of the locally dominant type.

Acknowledgement

Appendix 1 Generalizations

Part of the previous discussion, and, in particular, the asymptotic analysis in Sections 4 and 5 is valid for a wide class of models of greater generality than (1), viz. for models of the form

$$\frac{\partial N_i}{\partial t} = \sigma \frac{\partial^2 N_i}{\partial x^2} + A_i(N_1, N_2, N_3), \quad i = 1, 2, 3, \tag{30}$$

where each A_i is the difference between a birth term and a death term:

$$A_i(N_1, N_2, N_3) = B_i(N_1, N_2, N_3) - D_i(N_1, N_2, N_3)\, N_i. \tag{31}$$

Here, the terms B_i describe the birth process and the terms $D_i\, N_i$ the death process (allowing, in particular, more general types of inter-genotypic competition than the logistic one of (1)).

In order to be genetically meaningful, all B_i and all D_i should be non-negative, and the birth terms should satisfy the conditions

$$B_1(0, 0, N_3) = B_2(0, 0, N_3) = B_2(N_1, 0, 0) = B_3(N_1, 0, 0) = 0,$$

$$\frac{\partial B_1}{\partial N_1}(0, 0, N_3) = \frac{\partial B_1}{\partial N_2}(0, 0, N_3) = \frac{\partial B_3}{\partial N_2}(N_1, 0, 0) = \frac{\partial B_3}{\partial N_3}(N_1, 0, 0) = 0 \tag{32}$$

of ordinary Mendelian inheritance. Furthermore, we will allow interbreeding of the two original populations, in that $B_2(N_1, 0, N_3) > 0$ if $N_1 > 0$ and $N_3 > 0$; otherwise the two homozygote populations would behave as two competing species. For convenience, we make the slightly stronger, genetically very natural assumption

$$\frac{\partial B_2}{\partial N_1}(0, 0, N_3) > 0 \quad \text{if} \quad N_3 > 0, \quad \text{and}$$

$$\frac{\partial B_2}{\partial N_3}(N_1, 0, 0) > 0 \quad \text{if} \quad N_1 > 0. \tag{33}$$

In this general model the condition of equal fitness of the homozygotes takes the form

$$
\begin{aligned}
B_3(N_1, N_2, N_3) &= B_1(N_3, N_2, N_1), \\
B_2(N_1, N_2, N_3) &= B_2(N_3, N_2, N_1), \\
D_3(N_1, N_2, N_3) &= D_1(N_3, N_2, N_1), \\
D_2(N_1, N_2, N_3) &= D_2(N_3, N_2, N_1)
\end{aligned}
\tag{34}
$$

for all $N_1, N_2, N_3 \geq 0$. The model should allow pure homozygote populations of constant density to exist and be stable as $t \to \infty$, i. e. the equation

$$\frac{dN}{dt} = B_3(0, 0, N) - D_3(0, 0, N)\, N$$

19

should have a stable positive equilibrium K. In addition, the point $(0, 0, K)$ should be a stable equilibrium of the system (13), as is the case in simple selection models when the heterozygote is inferior. Sufficient conditions for this to be true are given in Appendix 3, and it is shown that they are satisfied for the model considered in Section 2.

The dynamics of the general model of fertility selection is very rich even for the highly symmetric model (34) considered here (Feldman, Christiansen and Liberman, 1983). The model in Section 2 is restricted in the sense that its dynamics resemble that of a simple viability selection model. In the general model the stability of the monomorphic equilibria is not linked to the stability of the symmetric equilibrium where $p = \frac{1}{2}$. The model of Section 2 only has these three equilibria, and the symmetric equilibrium is always unstable. The equilibrium configuration in the general model (34) also includes the possibility of stability of all these three equilibria with the existence of two unstable equilibria separating the monomorphic equilibria from the symmetric equilibrium (Hadeler and Liberman, 1975). The discrete version of the model even may exibit limit cycles (Hadeler and Liberman, 1975).

Appendix 2   The Transformed Equations

The equation for $\partial N/\partial t$ is obtained by a straightforward sum of equations (1). The equation for $\partial p/\partial t$ is obtained from the definition (3) as

$$
\begin{aligned}
\frac{\partial p}{\partial t} &= \frac{1}{2N}\left(2\frac{\partial N_1}{\partial t} + \frac{\partial N_2}{\partial t}\right) - \frac{p}{N}\frac{\partial N}{\partial t} \\
&= \left(\frac{\sigma}{2N}\frac{\partial^2 (2N_1 + N_2)}{\partial x^2} + (b - (d + \iota N))p - (B_1 p + B_2 q)p - \delta\frac{N_2}{2N}\right) \\
&\quad - \frac{p}{N}\left(\sigma\frac{\partial^2 N}{\partial x^2} + (b - d - \iota N)N - 2pq(1 - F)(2\beta + \delta)N\right) \\
&= \sigma\left(\frac{\partial^2 p}{\partial x^2} + 2\frac{\partial p}{\partial x}\frac{1}{N}\frac{\partial N}{\partial x}\right) + pq(1 - F)(\beta + \delta)(p - q).
\end{aligned}
$$

The equation for F is a bit more complicated. By implicit differentiation of of the relation $N_2 = 2pq(1 - F)N$ obtained from (4) we get

$$
\frac{\partial N_2}{\partial t} = 2pq(1 - F)\frac{\partial N}{\partial t} - 2pqN\frac{\partial F}{\partial t} - 2(p - q)(1 - F)N\frac{\partial p}{\partial t}
$$

20

$$\frac{\partial^2 N_2}{\partial x^2} = 2pq(1-F)\frac{\partial^2 N}{\partial x^2} - 4pq\frac{\partial F}{\partial x}\frac{\partial N}{\partial x} - 4(p-q)(1-F)\frac{\partial p}{\partial x}\frac{\partial N}{\partial x}$$
$$-2pqN\frac{\partial^2 F}{\partial x^2} + 4(p-q)N\frac{\partial p}{\partial x}\frac{\partial F}{\partial x}$$
$$-2(p-q)(1-F)N\frac{\partial^2 p}{\partial x^2} - 4(1-F)N\left(\frac{\partial p}{\partial x}\right)^2.$$

These expressions for the derivatives $\partial N_2/\partial t$ and $\partial^2 N_2/\partial x^2$ may be inserted into equation (1) to provide an expression for $\partial F/\partial t$. By using the already established equations (6) and (7) we may eliminate $\partial^2 N/\partial x^2$ and $\partial^2 p/\partial x^2$ to obtain

$$2pqN\frac{\partial F}{\partial t} =$$
$$\sigma\left(4(1-F)N\left(\frac{\partial p}{\partial x}\right)^2 - 4(p-q)N\frac{\partial p}{\partial x}\frac{\partial F}{\partial x} + 4pq\frac{\partial F}{\partial x}\frac{\partial N}{\partial x} + 2pqN\frac{\partial^2 F}{\partial x^2}\right)$$
$$+2pq(1-F)((b-d-\iota N)N - 2pq(1-F)N(2\beta+\delta))$$
$$+2pq(1-F)^2(p-q)^2N(\beta+\delta)$$
$$-(2pqNb - 2pq(1-F)N(d+\iota N) - 2pq(1-F)N(\beta+\delta)).$$

Straightforward simplification of this equation yields equation (8).

Appendix 3   Asymptotic analysis of the panmictic equilibrium

Consider a system of ordinary differential equations of the type considered in (13):

$$\frac{dN_i}{dt} = A_i(N_1, N_2, N_3), \quad i = 1, 2, 3, \tag{35}$$

where A_i is the difference between a birth term and a death term (31).

We assume that all B_i and all D_i are smooth and non-negative when all N_i are non-negative, and that the birth terms satisfy conditions (32) and (33). Standard uniqueness and comparison theorems for ordinary differential equations then secures that all solutions to (35) have the properties:

1. If $N_i(t_0) \geq 0$ for $i = 1, 2, 3$, then $N_i(t) \geq 0$ for $t > t_0$ and $i = 1, 2, 3$.
2. If $N_1(t_0) = N_2(t_0) = 0$, then $N_1(t) = N_2(t) = 0$ for $t > t_0$.
3. If $N_2(t_0) = N_3(t_0) = 0$, then $N_2(t) = N_3(t) = 0$ for $t > t_0$.

4. If $N_i(t_0) \geq 0$ for $i = 1, 2, 3$ and $N_j(t_0) > 0$ for some j, then $N_j(t) > 0$ for $t > t_0$.

The point $(0, 0, K)$ by assumption is a stationary point of equations (35). A great deal of information on the local behavior around this point of the solutions to the system (35) can be derived from considering the corresponding system of linear equations

$$\frac{du}{dt} = Au, \qquad (36)$$

in the variables

$$u = \left\{ \begin{array}{c} u_1 \\ u_2 \\ u_3 \end{array} \right\} = \left\{ \begin{array}{c} N_1 \\ N_2 \\ N_3 - K \end{array} \right\}$$

which correspond to the increments in N_1, N_2 and N_3. The coefficient matrix of the equation is

$$A = \left\{ \begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{array} \right\} = \left\{ \begin{array}{ccc} -D_1 & 0 & 0 \\ b_{21} & b_{22} - D_2 & 0 \\ b_{31} - d_{31} K & b_{32} - d_{32} K & b_{33} - d_{33} K - D_3 \end{array} \right\}$$

with $a_{ij} = \partial A_i / \partial N_j$, $b_{ij} = \partial B_i / \partial N_j$ and $d_{ij} = \partial D_i / \partial N_j$, where all functions are evaluated at $(0, 0, K)$. In particular, the stationary point $(0, 0, K)$ is stable provided all eigenvalues of A have negative real parts. In our case, these eigenvalues are real:

$$\lambda_1 = -D_1, \quad \lambda_2 = -(D_2 - b_{22}), \quad \lambda_3 = -(D_3 + d_{33}K - b_{33}).$$

We will assume that the eigenvalues λ_1, λ_2 and λ_3 are pairwise different and that the set $\{\lambda_1, \lambda_2, \lambda_3\}$ is *non-resonant*, i. e. no λ_j is a linear combination of the two others with non-negative integer coefficients. Then a C^∞-mapping Φ of $\mathbf{R}^3$ into itself exists, which is tangent to the identity map at $(0, 0, K)$, and which transforms the solutions of the system (36) into those of (35) (see e. g. Irwin, 1980, p. 127 or Anosov and Arnold, 1988, pp. 53 and 66). Thus, an arbitrary solution to (35) can be written as

$$\left\{ \begin{array}{c} N_1(t) \\ N_2(t) \\ N_3(t) \end{array} \right\} = \left\{ \begin{array}{c} 0 \\ 0 \\ K \end{array} \right\} + \left\{ \begin{array}{c} u_1(t) \\ u_2(t) \\ u_3(t) \end{array} \right\} + O(\|(u_1(t), u_2(t), u_3(t))\|^2) \qquad (37)$$

close to the stationary point $(0, 0, K)$. The solution to (36) can be written as

$$\left\{ \begin{array}{c} u_1(t) \\ u_2(t) \\ u_3(t) \end{array} \right\} = \sum_{j=1}^{3} C_j \, e^{\lambda_j t} \, e_j, \qquad (38)$$

22

where e_j is an eigenvector corresponding to λ_j, and C_j, $j = 1, 2, 3$, are arbitrary constants to be determined from the initial conditions.

Denote by M the image under Φ of the plane $C_{\bar{j}} = 0$, where $\lambda_{\bar{j}}$ is the largest of the eigenvalues. Then M is an invariant 2-dimensional submanifold of $\mathbf{R}^3$, its tangent space at $(0, 0, K)$ is spanned by the eigenvectors corresponding to the two smallest eigenvalues, and *all* solutions to (35) except those in M have $e_{\bar{j}}$ as tangent vector in $(0, 0, K)$.

In the model of Section 2 we have

$$\lambda_1 = -b, \quad \lambda_2 = -(\beta + \delta), \quad \lambda_3 = -(b - d), \tag{39}$$

and we may choose the eigenvectors

$$e_1 = \begin{pmatrix} 1 \\ \frac{-2b}{b - \beta - \delta} \\ \frac{-(bd - 2b(2\beta + \delta) + d(\beta + \delta))}{d(b - \beta - \delta)} \end{pmatrix}, \; e_2 = \begin{pmatrix} 0 \\ 1 \\ \frac{-(b - d + \beta)}{b - d - \beta - \delta} \end{pmatrix}, \; e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{40}$$

The eigenvalue λ_1 is never dominant because $b \geq b - d$.

All eigenvalues are negative. This immediately implies from equations (37) and (38) that $N(t)$ and $p(t)$ have limits for $t \to \infty$. It remains to decide whether the limit of $F(t)$ exists. Clearly, if $N_1(t) = ae^{\alpha t} + o(e^{\alpha t})$, $N_2(t) = be^{\beta t} + o(e^{\beta t})$ with $a, b > 0$, $\alpha, \beta < 0$, then

$$\lim_{t \to \infty} F(t) = \begin{cases} 1 & \text{if } \alpha > \beta, \\ 1 - \frac{b}{2a + b} & \text{if } \alpha = \beta, \\ 0 & \text{if } \alpha < \beta. \end{cases}$$

Now assume that $C_{\bar{j}} \neq 0$ and that $N_1(t), N_2(t) > 0$. For $\beta + \delta < b - d$ we get $\bar{j} = 2$ and $F(t) \to 0$. For $\beta + \delta > b - d$ we get $\bar{j} = 3$ and the dominating term in equation (37) does not contain enough information to determine the asymptotics of F. Higher-order terms therefore have to be taken into account. In this connection we note that in the k-th order Taylor approximation

$$\begin{Bmatrix} N_1(t) \\ N_2(t) \\ N_3(t) \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \\ K \end{Bmatrix} + \begin{Bmatrix} u_1(t) \\ u_2(t) \\ u_3(t) \end{Bmatrix} + \begin{Bmatrix} P_1(u_1(t), u_2(t), u_3(t)) \\ P_2(u_1(t), u_2(t), u_3(t)) \\ P_3(u_1(t), u_2(t), u_3(t)) \end{Bmatrix}$$

$$+ O(\|(u_1(t), u_2(t), u_3(t))\|^{k+1}) \tag{41}$$

derived from Φ all terms in the polynomials P_1 and P_2 have either u_1 or u_2 as a factor.

23

For a complete description of the situation we distinguish between the three cases $\beta + \delta < b - d < b$, $b - d < \beta + \delta < b$, and $b - d < b < \beta + \delta$. In the discussion below, the expression 'in general' means 'except on the 2-dimensional manifold M'.

1. $\beta + \delta < b - d < b$: the eigenvalue λ_2 is dominant, and in general $F(t) \to 0$. In fact, since the two first coordinates of e_1 have opposite sign, a solution with $N_1(t), N_2(t) > 0$ cannot have $C_2 = 0$, and therefore $F(t) \to 0$ for *all* such solutions.

2. $b - d < \beta + \delta < b$: the eigenvalue λ_3 is dominant, and it is obvious that the corresponding term in (37) does not determine the asymptotic behavior of $F(t)$. If $\beta + \delta < 2(b - d)$, it is still a consequence of (37) and (38) that in general

$$\left\{ \begin{array}{c} N_1(t) \\ N_2(t) \end{array} \right\} = C\, e^{-(\beta+\delta)t} \left\{ \begin{array}{c} 0 \\ 1 \end{array} \right\} + O(e^{-\gamma t}) \tag{42}$$

with $\gamma = 2(b - d)$, so that $F(t) \to 0$ (as above, this is actually true for *all* solutions with positive $N_1(t), N_2(t)$). If $\beta + \delta > 2(b - d)$, higher order terms in the Taylor development of Φ have to be taken into account, and then one can show that equation (42) is still valid with $\gamma = \beta + \delta + b - d$, so again, $F(t) \to 0$.

3. $b - d < b < \beta + \delta$: as in case 2, the eigenvalue λ_3 is dominant. The eigenvalue λ_1 now determines the asymptotic behavior of $(N_1(t), N_2(t))$, and as in case 2 we get: If $b < 2(b - d)$, then

$$\left\{ \begin{array}{c} N_1(t) \\ N_2(t) \end{array} \right\} = C\, e^{-bt} \left\{ \begin{array}{c} \beta + \delta - b \\ 2b \end{array} \right\} + O(e^{-\gamma t}) \tag{43}$$

with $\gamma = 2(b - d)$, in general, and if $b > 2(b - d)$ then equation (43) is still valid with $\gamma = 2b - d$. Thus, in general both N_1 and N_2 tend to 0 as e^{-bt}, and $F(t) \to \check{F}$ given by equation (17).

Equations (35) are equivalent to the equations (14)–(16) in the model of Section 2. If a point (N^e, p^e, F^e) with $p^e = 0$ and $N^e > 0$ is a stationary point for this system, N^e must be equal to K, and F^e must satisfy the equation $F^e(b - (\beta + \delta)(1 - F^e)) = 0$. Since $p = 0$ implies $F \geq 0$ by the remark following definition (5), we see that $(K, 0, 0)$ is a stationary point of (14)–(16) for all values of b, β and δ, and that if $\beta + \delta > b$, then $(K, 0, \check{F})$ is also a stationary point. The linearization of this system around a stationary point $(K, 0, F^e)$ is

$$\frac{dv}{dt} = Bv,$$

24

where v_1, v_2, v_3 correspond to the increments in N, p, F, and

$$B = \left\{ \begin{array}{ccc} -(b-d) & -2K(2\beta+\delta)(1-F^e) & 0 \\ 0 & -(\beta+\delta)(1-F^e) & 0 \\ 0 & 2\delta(1-F^e)^2 & -b+(\beta+\delta)(1-2F^e) \end{array} \right\}.$$

The eigenvalues of B are $-(b-d)$ and $-(\beta+\delta)(1-F^e)$ which are negative, and $-b+(\beta+\delta)(1-2F^e)$ which is positive if $\beta+\delta > b$ and $F^e = 0$, but negative if $\beta+\delta < b$ and $F^e = 0$ or if $\beta+\delta > b$ and $F^e = \check{F}$. Thus, $(K,0,0)$ is a stable stationary point of (14)–(16) if $\beta+\delta < b$, but if $\beta+\delta > b$, then the only stable stationary point of that system is $(K,0,\check{F})$. This is in perfect agreement with the observations 1–3 above concerning the behavior 'in general' of the system (35).

Appendix 4   Asymptotic Analysis of the Tail of the Cline

Consider a system of ordinary differential equations of the type considered in equation (18):

$$\sigma d^2\hat{N}_i/dx^2 = -A_i(\hat{N}_1(x), \hat{N}_2(x), \hat{N}_3(x)), \quad i = 1, 2, 3, \tag{44}$$

and assume for definiteness that the A_i are given as in Section 2. In order to discuss solutions of this system of equations, we use the standard trick of introducing the lower order derivatives as extra variables, and so we get

$$\frac{d\hat{N}_1}{dx} = \hat{n}_1,$$

$$\frac{d\hat{N}_2}{dx} = \hat{n}_2,$$

$$\frac{d\hat{N}_3}{dx} = \hat{n}_3, \tag{45}$$

$$\frac{d\hat{n}_1}{dx} = -\frac{1}{\sigma}A_1(\hat{N}_1, \hat{N}_2, \hat{N}_3),$$

$$\frac{d\hat{n}_2}{dx} = -\frac{1}{\sigma}A_2(\hat{N}_1, \hat{N}_2, \hat{N}_3),$$

$$\frac{d\hat{n}_3}{dx} = -\frac{1}{\sigma}A_3(\hat{N}_1, \hat{N}_2, \hat{N}_3).$$

It can be shown that if $(\hat{N}_1(x), \hat{N}_2(x), \hat{N}_3(x))$ is a solution to (44) such that $(\hat{N}_1(x), \hat{N}_2(x), \hat{N}_3(x)) \to (0, 0, K)$ as $x \to -\infty$, then $(\hat{n}_1(x), \hat{n}_2(x), \hat{n}_3(x)) \to (0, 0, 0)$.

25

Thus, the point $(0, 0, K, 0, 0, 0)$ is a stationary point for the system (45), and its linearization around this point can be written in block form as

$$\frac{d}{dx} \left\{ \begin{array}{c} u \\ v \end{array} \right\} = L \left\{ \begin{array}{c} u \\ v \end{array} \right\}, \tag{46}$$

where the vectors u and v are 3-dimensional and L is the block matrix

$$L = \left\{ \begin{array}{cc} O & I \\ -\frac{1}{\sigma} A & O \end{array} \right\},$$

with A as in (36), and where I is the 3×3 identity matrix and O is the 3×3 matrix with only zero entities. The eigenvalues of L are $\kappa_j^{\pm} = \pm(-\lambda_j/\sigma)^{\frac{1}{2}}$, where λ_j is given by equation (39). The eigenvector corresponding to $\kappa_j^{\pm}$ can be chosen as the vector whose first 3-dimensional block is the eigenvector e_j of A given in equation (40), while the second block is $\kappa_j^{\pm} e_j$.

In the terminology of the theory of dynamical systems, the stationary point $(0, 0, K, 0, 0, 0)$ is a hyperbolic point; its stable and unstable manifolds (corresponding, respectively, to the negative and positive eigenvalues) are of dimension 3, and solutions to (45) approaching it for $x \to -\infty$ lie in the unstable manifold. As in Appendix 3 it follows that if the set of eigenvalues is non-resonant, there is a C^∞-mapping Ψ of $\boldsymbol{R}^6$ into itself, tangent to the identity map at the stationary point, which transforms the solutions of the system (46) into those of (45). Consequently, we get approximations to those solutions to (44) which converge to $(0, 0, K, 0, 0, 0)$ for $x \to -\infty$. These solutions are completely analogous to (37) and (41). The first order approximation analogous to (37) gives

$$\left\{ \begin{array}{c} \hat{N}_1(x) \\ \hat{N}_2(x) \\ \hat{N}_3(x) \end{array} \right\} = \left\{ \begin{array}{c} 0 \\ 0 \\ K \end{array} \right\} + \left\{ \begin{array}{c} \hat{u}_1(x) \\ \hat{u}_2(x) \\ \hat{u}_3(x) \end{array} \right\} + O(\|(u_1(x), u_2(x), u_3(x))\|^2) \tag{47}$$

with

$$\left\{ \begin{array}{c} \hat{u}_1(x) \\ \hat{u}_2(x) \\ \hat{u}_3(x) \end{array} \right\} = \sum_{j=1}^{3} C_j e^{\kappa_j^+ x} e_j. \tag{48}$$

For the most part, the arguments and results are analogous to those in Appendix 3, and we omit the details, giving only the results. The asymptotic formulas for $\hat{N}_1$, $\hat{N}_2$, $\hat{N}_3$ are:

$$\hat{N}_1(x) \approx C_1 e^{\kappa_1 x} + \frac{b - 2\beta}{4K(b - 4(\beta + \delta))} C_2^2 e^{2\kappa_2 x} \tag{49}$$

26

$$\hat{N}_2(x) \approx \frac{2b}{\beta + \delta - b}\hat{N}_1(x) + C_2\, e^{\kappa_2 x}$$
$$- \frac{2b(\beta + \delta)\left((\beta + \delta)^2 - 2b\beta\right)}{K(\beta + \delta - b)^3(\beta + \delta - 4b)}\, C_1^2\, e^{2\kappa_1 x} \tag{50}$$

$$\hat{N}_3(x) \approx K - \frac{b - d + \beta}{b - d - (\beta + \delta)}\, C_2\, e^{\kappa_2 x} - C_3\, e^{\kappa_3 x}. \tag{51}$$

The sign $\approx$ indicates that the absolute value of the difference between the left hand side and the right hand side is at most $Ce^{(\tilde{\kappa}+\varepsilon)x}$ for some positive numbers C and ε, with $\tilde{\kappa}$ denoting the smallest coefficient occurring in the exponentials on the right hand side of the formula (and so, all terms with larger coefficients are actually redundant). In some cases, we need certain second order terms of the expansions of $N_1(x)$ and $N_2(x)$ in order to determine the rate of convergence of $\hat{F}(x)$ as $x \to \infty$, and these second order terms are included in (49)–(50). The simplified and more informative versions of equations (49)–(51) appear in Section 5 for the various orderings of the eigenvalues.

In Appendix 3 we observed that equations (35) for the genotype numbers are equivalent to equations (14)–(16) for N, p, and F as functions of t. Similarly, equations (44) are equivalent to (27)–(29), but the last of these equations has a singularity at $p = 0$, so it is not clear how to make an asymptotic analysis of the solutions to (27)–(29). It turns out that if we convert (27)–(29) to a first order system by introducing the logarithmic derivatives

$$\frac{1}{\hat{N}}\frac{d\hat{N}}{dx} \quad \text{and} \quad \frac{1}{\hat{p}}\frac{d\hat{p}}{dx}$$

as extra variables, then the singularity is resolved, and among the solutions to the resulting system of first order equations which satisfy $\lim_{x \to -\infty} N(x) = K$, $\lim_{x \to -\infty} p(x) = 0$, the 'general solutions' are precisely those described in 1–6 in Section 5. Note, however, that there is no reason why this property of a cline should be related to its stability as a stationary solution to (6)–(8) (cf. Section 4).

Note that all results in Appendix 4 are derived under the assumption that the solution considered is a 'general' one in the sense that coefficients of potentially dominating terms are different from 0. This assumption is justified by the fact that in all cases the resulting value of $\lim_{x \to \infty} \hat{F}(x)$ is the 'right one' according to the concluding remarks of Section 4.

27

References

Andreasen, V., Christiansen, F. B.: Disease-induced natural selection in a diploid host. Theor. Pop. Biol. **44**, *in print* (1993)

Anosov, D. V., Arnold, V. I.: Dynamical Systems. I. Berlin Heidelberg New York: Springer 1988

Aronson, D. G., Weinberger, H. F.: Nonlinear diffusion in population genetics, combustion, and nerve pulse propagation. In Goldstein, J. A. (ed.) Partial Differential Equations and Related Topics, pp. 5–49. Berlin Heidelberg New York: Springer 1975

Barton, N. H.: The dynamics of hybrid zones. Heredity **43**, 341–359 (1979)

Barton, N. H., Hewitt, G. M.: Analysis of hybrid zones. Ann. Rev. Ecol. Syst. **16**, 113–148 (1985)

Bazykin, A. D.: Hypothetical mechanism of speciation. Evolution **23**, 685–687 (1969a)

Bazykin, A. D.: On the effect of disruptive selection on the population with one dimensional area. I. Equal fitness of homozygotes. (Russian with English summary). Problemy evolutsii **2**, 219–223 (1969b)

Brown, A. H. D.: The estimation of Wright's fixation index from genotypic frequencies. Genetics **41**, 399–406 (1970)

Feldman, M. W., Christiansen, F. B., Liberman, U.: On some models of fertility selection. Genetics **105**, 1003–1010 (1983)

Fisher, R. A.: The wave of advance of advantageous genes. Ann. Eugenics **7**, 355–369 (1937)

Guckenheimer, J., Holmes, P.: Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields. Berlin Heidelberg New York: Springer 1983

Hadeler, K. P., Liberman, U.: Selection models with fertility differences. J. Math. Biol. **2**, 19–32 (1975)

Haldane, J. B. S.: The theory of a cline. J. Genet. **48**, 277–284 (1948)

Hunt, F.: On the persistence of spatially homogeneous solutions of a population genetics model with slow selection. Math. Biosci. **52**, 185–206 (1980)

Irwin, M. C.: Smooth Dynamical Systems. New York San Fransisco London: Academic Press 1980

Kimura, M.: Stepping stone model of population. Ann. Rept. Nat. Inst. Genet. Japan **3**, 62–63 (1953)

Malécot, G.: Les Mathématiques de l'Hérédité. Paris: Masson et Cie 1948

Nagylaki, T.: Conditions for the existence of clines. Genetics **80**, 595–615 (1975)

Nagylaki, T.: The diffusion model for migration and selection. In Hastings, A. (ed.) Some Mathematical Questions in Biology: Models in Population Biology, pp. 55–75. Providence, R. I.: American Mathematical Society 1989

Nagylaki, T., Crow, J. F.: Continuous selective models. Theor. Pop. Biol. **5**, 257–283 (1974)

Poulsen, E. T.: Nonmonotone clines in homogeneous space cannot be stable. SIAM J. Math. Anal. **20**, 148–159 (1989)

Slatkin, M.: Gene flow and selection in a cline. Genetics **75**, 733–756 (1973)

Wright, S.: Isolation by distance under diverse systems of mating. Genetics **31**, 39–59 (1946)

217/92 "Two papers on APPLICATIONS AND MODELLING
IN THE MATHEMATICS CURRICULUM"

by: Mogens Niss

218/92 "A Three-Square Theorem"

by: Lars Kadison

219/92 "RUPNOK - stationær strømning i elastiske rør"

af: Anja Boisen, Karen Birkelund, Mette Olufsen

Vejleder: Jesper Larsen

220/92 "Automatisk diagnosticering i digitale kredsløb"

af: Bjørn Christensen, Ole Møller Nielsen

Vejleder: Stig Andur Pedersen

221/92 "A BUNDLE VALUED RADON TRANSFORM, WITH
APPLICATIONS TO INVARIANT WAVE EQUATIONS"

by: Thomas P. Branson, Gestur Olafsson and
Henrik Schlichtkrull

222/92 On the Representations of some Infinite Dimensional
Groups and Algebras Related to Quantum Physics

by: Johnny T. Ottesen

223/92 THE FUNCTIONAL DETERMINANT

by: Thomas P. Branson

224/92 UNIVERSAL AC CONDUCTIVITY OF NON-METALLIC SOLIDS AT
LOW TEMPERATURES

by: Jeppe C. Dyre

225/92 "HATMODELLEN" Impedansspektroskopi i ultrarent
en-krystallinsk silicium

af: Anja Boisen, Anders Gorm Larsen, Jesper Varmer,
Johannes K. Nielsen, Kit R. Hansen, Peter Bøggild
og Thomas Hougaard

Vejleder: Petr Viscor

226/92 "METHODS AND MODELS FOR ESTIMATING THE GLOBAL
CIRCULATION OF SELECTED EMISSIONS FROM ENERGY
CONVERSION"

by: Bent Sørensen

227/92 "Computersimulering og fysik"

af: Per M.Hansen, Steffen Holm,
Peter Maibom, Mads K. Dall Petersen,
Pernille Postgaard, Thomas B.Schrøder,
Ivar P. Zeck

Vejleder: Peder Voetmann Christiansen

228/92 "Teknologi og historie"

Fire artikler af:

Mogens Niss, Jens Høyrup, Ib Thiersen,
Hans Hedal

229/92 "Masser af information uden betydning"

En diskussion af informationsteorien
i Tor Nørretranders' "Mærk Verden" og
en skitse til et alternativ basseret
på andenordens kybernetik og semiotik.

af: Søren Brier

230/92 "Vinklens tredeling - et klassisk
problem"

et matematisk projekt af

Karen Birkelund, Bjørn Christensen

Vejleder: Johnny Ottesen

231A/92 "Elektrondiffusion i silicium - en
matematisk model"

af: Jesper Voetmann, Karen Birkelund,
Mette Olufsen, Ole Møller Nielsen

Vejledere: Johnny Ottesen, H.B.Hansen

231B/92 "Elektrondiffusion i silicium - en
matematisk model" Kildetekster

af: Jesper Voetmann, Karen Birkelund,
Mette Olufsen, Ole Møller Nielsen

Vejledere: Johnny Ottesen, H.B.Hansen

232/92 "Undersøgelse om den simultane opdagelse
af energiens bevarelse og isærdeles om
de af Mayer, Colding, Joule og Helmholtz
udførte arbejder"

af: L.Arleth, G.I.Dybkjær, M.T.Østergård

Vejleder: Dorthe Posselt

233/92 "The effect of age-dependent host
mortality on the dynamics of an endemic
disease   and
Instability in an SIR-model with age-
dependent susceptibility

by: Viggo Andreasen

234/92 "THE FUNCTIONAL DETERMINANT OF A FOUR-DIMENSIONAL
BOUNDARY VALUE PROBLEM"

by: Thomas P. Branson and Peter B. Gilkey

235/92 OVERFLADESTRUKTUR OG POREUDVIKLING AF KOKS

- Modul 3 fysik projekt -

af: Thomas Jessen