

# **Some Theoretical Analyses of Migration Selection Interaction in a Cline: A Generalized Two Range Environment**

S. KARLIN and N. RICHTER-DYN

## 1. INTRODUCTION

A frequency cline refers generally to a gradient pattern in a character with respect to a linear series of demes (sub-populations of discrete breeding units) subject to some inter-deme migration. The slope of the cline between regions (and/or demes) has partly been taken to be indicative of the degree of differentiation among the different sub-populations and used in estimating the strength of selection acting to maintain a cline.

The phenomena of clinal variation in gene frequency is voluminously documented both in experimental and natural populations. Classic field studies of clines include the works of Kettlewell [1961] and Kettlewell and Berry [1961], [1969] on melanism and the extensive analyses of Jain and Bradshaw [1966] which highlight 5 clines of grasses and

---

\* Department of Mathematics, The Weizmann Institute of Science, Rehovot, Israel - and

Department of Mathematics, Stanford University, Stanford, California, U.S.A.

\*\*Department of Mathematics, Tel-Aviv University, Ramat-Aviv, Israel.

plants along certain transects. Frequency clines have long been described for morphological, physiological, behavioral and chromosomal characters. Endler [1973] contains a substantial set of references reporting cases of observed morphotones (clines). In recent years enzyme polymorphic clines have been increasingly uncovered, e.g., Koehn [1969], Beardmore [1970], Marshall and Allard [1972], Hamrick and Allard [1972], Merritt [1972], Rockwood Slus et al. [1973], Frydenberg and Christiansen [1974], Bryant [1974], Nevo and Bar [1976] and others.

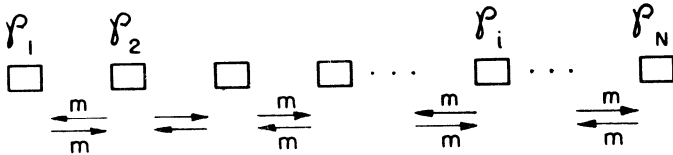
A commonly held belief has been that clinal variations in gene frequency originates in response or is correlated to one or more ecological and/or genetic environmental parameters; e.g., temperature and humidity gradients over space and time, variation in substrate availabilities, characteristics of soil and water composition, the distribution of related predators, factors of cryptography and coloration background, etc. It is a trite comment in the analyses of clinal patterns that genotypic differentiation is promulgated by joint effects of local selection forces and gene flow. Important contributing factors determining the equilibrium and dynamic structure of clines would include the rate and form of gene flow, the extent of population subdivision or multiplicity of demes, the variation in local population density, the variegated intensity of selection forces operating over the range of the species, and similarly.

It is often assumed that near a sharp discontinuity in gene frequency there necessarily exists some environmental factor change. Arguments have been put forth by Haldane [1948], Hanson [1966], and others, indicating that small selection differences can already engender marked local genotypic differentiation but the amount of migration is critical.

This sensitivity to the magnitude of gene flow is questioned by Endler [1973]. There are alternative conclusions pertaining to the balance necessary between migration and selection entailing the establishment of clinal variation. For example Wright [1943] has stated: "diversity in degree and direction of selection among localities... may bring about great differentiation if not overbalanced by migration". On the other hand, Jain and Bradshaw infer on the basis of their simulation studies, "Selection can cause very localized patterns of microgeographical variation despite migration through pollen dispersal". "Populations of plants closely adjacent can diverge very remarkably even if there is only a small degree of geographical isolation". Of course both sets of pronouncements may be valid under appropriate circumstances.

A number of mathematical models incorporating some forms of gene flow and a geographical selection gradient were introduced by Haldane [1948], Fisher [1950] and others, most recently Slatkin [1973]. The above authors all dealt with an infinite linear array of demes (actually a continuum of sub-populations) where the gene flow in each generation occurs only to adjacent populations. The formulation of the analogous finite discrete model is as follows. A population is distributed on a finite region composed of separate breeding demes (localities), e.g., geographical or ecological habitats,  $P_1, P_2, \dots, P_N$  arranged in linear order as indicated. Successive generations in the population are discrete and non-overlapping. Each sub-population  $P_i$  is of large size so that genetic drift effects are not pertinent. Consider a trait with two possible alleles labeled  $\underline{A}$  and  $\underline{a}$ . Mating and natural selection operate independently in each deme before some individuals disperse to the neighboring demes.

Thus after selection, a fraction  $m$  ( $m \leq \frac{1}{2}$ ) of each deme is exchanged with each contiguous deme as depicted below.



The migration pattern with movement allowed only to neighboring demes over successive generations, is commonly dubbed the homogeneous stepping stone migration mode and was studied extensively by Malécot [1948], [1951], [1959], later by Kimura and Weiss [1964], Weiss and Kimura [1965], and recently by Fleming and Su [1974]. These and other authors including Latter [1973], Maruyama [1970], [1972], concentrated mainly on finite population size effects with no selection differences among genotypes and demes.

Slatkin [1973] ostensibly considered more general migration patterns but the approximations he employed, in essence, reduce the model to that of a continuous version of the stepping stone model. In a later study we discuss forms of dispersal other than the "continuous movements" implicit to the stepping stone flux.

#### The two-range selection regime

Haldane, Fisher and Slatkin treat mainly the case when the range of the species is clearly divided into two sections. Fisher [1950] writes "The interesting case arises in which a gene enjoys a selective advantage in one part of a species range while in the remainder it is at a selective disadvantage. On the boundary between these regions, selection is neutral between two allelomorphous genes."

It is useful to represent such a geographical selection gradient schematically as follows:

$$\begin{array}{ccccccccccc}
 & 1 & 2 & & \ell_1 & \ell_1+1 & & \ell_2 & \ell_2+1 & & N \\
 \text{Demes} & \square & \square & \cdots & \square & \square & \cdots & \square & \square & \cdots & \square \\
 & \underbrace{\hspace{1.5cm}}_{-} & & & \underbrace{\hspace{1.5cm}}_{0} & & & \underbrace{\hspace{1.5cm}}_{+} & & & 
 \end{array} \quad (1.1)$$

where  $-$  signifies that  $A$  is disadvantageous at the localities  $1 \leq i \leq \ell_1$  (meaning that where such a deme is isolated, reproduction and selection would bring the local population to fixation of the  $a$ -allele);  $+$  connotes that  $A$  is advantageous at the positions specified and in the region of the  $0$  designation the  $A$  and  $a$  alleles are understood to be neutral (that is, no selection differences exist among genotypes). Thus, allele  $A$  is advantageous in the right portion of the range and allele  $a$  is advantageous in the left portion. The neutral zone may or may not exist. We can view the model of (1.1) as a generalized two range selection regime, each environment (range) favoring a different allelomorph. The intensity of selection may vary in an arbitrary manner between demes consistent only with the  $-$ ,  $0$ ,  $+$  arrangement of (1.1).

Fisher examined a subcase of (1.1) where the heterozygote carries intermediate fitness values corresponding to additive allelic effects. Haldane concentrated on an autosomal dominant gene and its allelomorph. Kettlewell and Berry [1961] strove to fit their data into the Haldane setting. In the case of melanism, the Tingwell valley appears to be a dividing boundary between the two generalized environments. Kettlewell and Berry did subsidiary experiments which gave reasonable evidence for the existence of selection, namely by

the introduction of organisms from one part of the cline into another part of the cline and studying their subsequent evolution. Jain and Bradshaw [1966] simulated a model involving two geographic environments with selection favoring different alleles in each region. The latter authors considered a migration structure of the kind consonant to isolation by distance where the extent of individual migration is described by a probability distribution. In their computer runs, they take some special choices considered appropriate for describing pollen and seed dispersal patterns. Thus migration per generation is not limited to small movements. Furthermore, a significant element in the grass clines of Jain and Bradshaw is the fact of unequal rates of gene flow in reciprocal directions (mediated partially by wind conditions). These authors found that strong local selection was more potent than the mitigating effects of even high gene flow. The concept of grades of local selection intensity will be discussed later in this work.

A number of examples of studies of clines in the human population are reviewed in the book of Cavalli and Bodmer [1971, p.483-490]. This discussion includes an application of the Fisher and Haldane model to a cline in A B O gene frequencies running latitudinally through the Islands of Japan.

#### The selection gradient

The transformation of gene frequency accountable to the mating structure and selection forces in deme  $\mathcal{P}_i$  is reflected by the relation

$$x' = f_i(x) \quad (1.2)$$

such that if  $x$  is the A-frequency in  $\mathcal{P}_i$  at the start of a generation, then after the action of mating and natural

selection the resulting A frequency prior to dispersal is  $x'$ . The local selection functions satisfy  $0 \leq f_i(x) \leq 1$ ,  $0 \leq x \leq 1$ . Generally,  $f_i(x)$  is continuous and monotone increasing. Also we stipulate in this work that  $f_i(0) = 0$  and  $f_i(1) = 1$  indicating that the selection forces maintain a pure population composition (or equivalently, new mutant types in the time frame under consideration cannot be established so that effectively mutation events are ignored).

An important choice for  $f_i(x)$  arising from the classical diploid one-locus two-allele viability model has the form

$$f_i(x) = \frac{(1+\sigma_i)x^2 + x(1-x)}{1+\sigma_i x^2 + s_i(1-x)^2} \quad (1.3)$$

when the viability parameters of the genotypes are as listed

$$\begin{array}{ccc} AA & Aa & aa \\ 1+\sigma_i & 1 & 1+s_i \end{array} \quad (1.4)$$

In a haploid situation we would take

$$f_i(x) = \frac{\sigma_i x}{\sigma_i x + s_i(1-x)} \quad (1.5)$$

Other determinations for  $f_i(x)$  can be generated by superimposing various forms of frequency dependent selection and other selection functions can be constructed induced on a single locus when part of a complex genome. The cline models discussed in this work also apply in some cases of selection functions based on considerations of competition, prey predator, host pathogen or other ecological contexts.

The statement that A is advantageous in deme i is equivalent to the formal relation

$$f_i(x) > x \quad \text{for all } 0 < x < 1 \quad , \quad (1.6)$$

disadvantageous if

$$f_i(x) < x \quad \text{for all } 0 < x < 1 \quad , \quad (1.7)$$

and neutral if  $f_i(x) \equiv x$ .

For directional selection favoring allele A with additive fitnesses,  $\sigma = -s > 0$  in (1.3). Where A is dominant and advantageous then  $\sigma = 0$  and  $s = -t$  ( $0 < t \leq 1$ ). The case of heterozygote advantage corresponds to  $\sigma = -\gamma$ ,  $s = -\delta$  ( $0 < \gamma, \delta < 1$ ).

A selection function  $f(x)$  is said to be more favorable (stronger) than  $\tilde{f}(x)$  with respect to the a-allele (in symbols  $f < \tilde{f}$ ), if

$$f(x) < \tilde{f}(x) \text{ for all } x \text{ } (0 < x < 1). \quad (1.8)$$

It is easy to check that for  $f_{\sigma,s}(x)$  defined in (1.3),

$$f_{\sigma,s} < f_{\sigma_1,s_1} \text{ provided } \sigma \leq \sigma_1, \text{ and } s_1 \leq s.$$

The environmental (geographical) selection gradient is characterized by the complete prescription of the local selection functions  $f_i(x)$ ,  $i = 1, 2, \dots, N$ .

In the circumstance where the heterozygote has additive fitness values, each local selection function is determined by a single selection coefficient  $\sigma_i$ ,  $-1 \leq \sigma_i \leq 1$ , viz.,

|           |                |    |                |                 |
|-----------|----------------|----|----------------|-----------------|
|           | AA             | Aa | aa             |                 |
| fitnesses | $1 + \sigma_i$ | 1  | $1 - \sigma_i$ | in deme $P_i$ . |

(1.9)

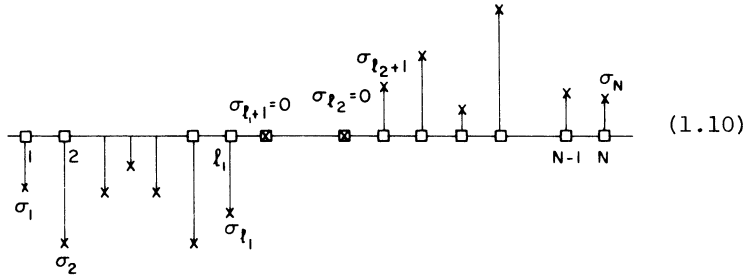
Where allele A (a) is advantageous,  $\sigma_i > 0$  ( $\sigma_i < 0$ ) holds.\*

---

\* Gillespie and Langley [1974] survey biochemical evidence pointing to wide validity of the assumption of heterozygote intermediacy. Thus the postulates of Section 2 of additive allelic fitness effects is possibly an important case in the study of allozyme polymorphisms; in this connection, see also Latter [this volume].



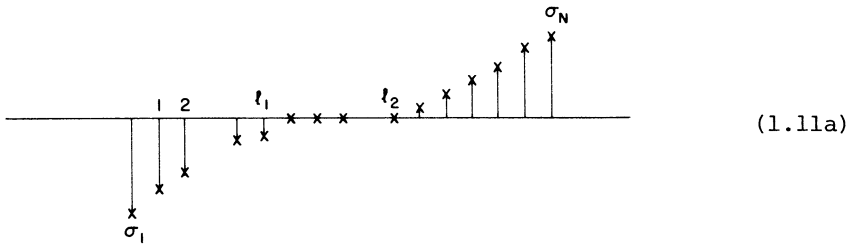
With local selection functions of the form (1.3) and fitnesses (1.9), the schematization of the selection gradient (1.1) in terms of the parameters  $\{\sigma_1, \sigma_2, \dots, \sigma_N\}$  can have a general form as below



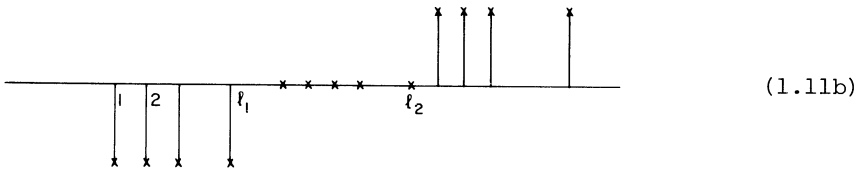
A more regular pattern has  $\sigma_i$  monotone separately in each environmental range. We describe three cases whose equilibrium behavior will be contrasted in our later discussion.

In Case I, selection becomes stronger with distance from the neutral zone (or from the demarcation point):

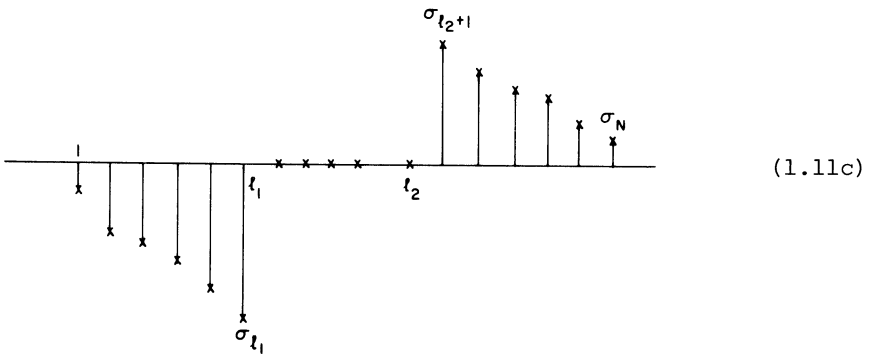
Condition I



In Case II, the selection coefficients are constant in each environmental range:

Condition II

In Case III, the selection intensity decreases with distance from the neutral region:

Condition III

To have a meaningful comparison, it is stipulated that the selection regimes I-III share the value

$$S_L = \sum_{i=1}^{\ell_1} \sigma_i \quad (1.12a)$$

i.e., the aggregate selection coefficients over the left range where selection favors allele  $a$  coincide for Conditions I, II and III.

Similarly,

$$S_R = \sum_{i=l_2+1}^N \sigma_i \quad (1.12b)$$

is the common cumulative selection value in the right range for models I-III. Conditions II and III may possibly correspond to abrupt environmental changes, induced by a geological fault or a sharp change in two contiguous soil compositions while Condition I may be relevant to a more gradual climate variation, background coloration, etc.

#### Formulation of objectives and problems

The analysis of migration selection interaction in multi deme models can be implemented on two levels:

(i) It is possible to obtain rather precise conditions guaranteeing the existence of a protected polymorphism. For substantial studies of this problem in the general context of geographical population models we refer to Karlin [1976]. For more refined results concerning protection in cline systems, see Karlin and Richter-Dyn [1976a].

(ii) The characterizations and comparisons of the equilibria under different conditions on selection and gene flow is in general, a formidable undertaking. In this paper we report a number of results describing the nature of polymorphic states in the selection cline model with stepping stone migration.

The specific objectives of this presentation and the subsequent detailed works (Karlin and Richter-Dyn [1976b,c,d]) seek to clarify the interrelations between patterns of geographically differentiated gene frequency and the background selection and gene flow structure.

Concentrating on a general two-environmental selection gradient of the form (1.1) we investigate the following questions:

(1) For a prescribed environmental or geographical selection gradient (determined by the  $f_i(x)$  that reflect the local selection forces) and a migration rate  $m$ , we wish to ascertain qualitatively and quantitatively the form of the stable equilibrium configurations (possible clines) that can evolve.

In particular it is of value to compare and contrast the shape of a steady state (equilibrium) cline with the shape of the background environmental selection pattern. Particular pertinent inquiries include:

(2) To what extent does increasing gene flow swamp out differences in local selection expressions?

(3) Where does the maximum slope of the cline occur and what is its significance?

(4) Is a sharp local change in selection intensity between two adjacent demes reflected in a corresponding marked change in gene frequency nearby?

(5) It is of interest to evaluate the effects on the shape of the cline in lengthening the neutral region.

(6) With prescribed aggregate levels of selection coefficients favoring allele  $A$  in the right range and allele  $a$  in the left range, how does the equilibrium cline differ in the contrasting cases of the selection patterns I, II and III described in (1.11)?

(7) How would the character of the cline change if the demes of the central (interface) zone involve local heterozygote advantage instead of each bearing no selection differences among genotypes?

(8) What would be the influence if each deme of the central range expresses underdominance?

(9) What are the effects of unequal rates of migration, a non-local migration range, different relative deme sizes and

multiplicity of demes, the influence of boundary conditions, a selection gradient involving more than two environments, timing (or order) of the operation of selection and migration forces, multi alleles, multi loci, dominance relations, etc.

Several of our findings on these problems are reported in Sections 2 and 3. Some of their implications and limitations are described in the concluding Section 4. In a series of papers we will elaborate the validations of the results highlighted in the later sections plus other aspects of selection clines.

We close this first section with some comments on other recent literature on this topic. Endler [1973] reported simulation and computer studies of the equilibrium state for some examples of clines involving 15 sub-populations. Specifically, Endler focused primary attention on the model of (1.1) for several regular selection patterns where each  $f_i$  is determined by the formula (1.3) with fitness parameters  $w_i(AA)$ ,  $w_i(Aa)$ ,  $w_i(aa)$  for deme  $i$  of the genotypes  $AA$ ,  $Aa$  and  $aa$ , respectively. We list the choices he made:

$$(\alpha) \quad w_i(Aa) = h > 1, \quad w_i(AA) = 1 - \frac{i-1}{15}, \quad w_i(aa) = \frac{i}{15},$$

$i = 1, \dots, 15$  (heterozygote advantage throughout the range).

$$(\beta) \quad w_i(Aa) = 1, \quad w_i(AA) = 1 + \frac{i-8}{15}, \quad w_i(aa) = 1 + \frac{8-i}{15},$$

(additive fitnesses). (1.13)

$$(\gamma) \quad w_i(Aa) = \frac{|i-8|}{15} + h, \quad w_i(AA) = 1 + \frac{i-8}{15},$$

$$w_i(aa) = 1 + \frac{8-i}{15}$$

(local heterozygote advantage).

The effective population size in each deme was so large that the influences of genetic drift is considered inconsequential.

The cases (1.13 $\alpha, \beta$ ) were designed to reflect a phenomenon where a particular genotype fitness increased uniformly with position along a transect while the other two genotype fitnesses decreased at different constant rates along the same transect. His simulation runs yielded monotone clines in all cases which exhibited a marked steepening in the neighborhood of the center deme  $i=8$  in case ( $\beta$ ). The outcome for case ( $\gamma$ ) was close to that of ( $\alpha$ ). But in contrast the heterozygous advantage cline ( $\alpha$ ) produced a roughly linear monotone cline. He observed some boundary effects in the cline as revealed in the computer runs.

Hastings [preprint] conducted some stochastic simulations of the stepping-stone model with 15 demes of "small" population sizes 10, 20 and 40 along a selection gradient of the kinds treated by Endler and showed that gene flow had several important effects. As gene flow increased, the stochastic effects were reduced and after 50 generations the yield approached the deterministic solutions, thus paralleling the effects of increasing size for each deme.

It is convenient, for later comparisons, to have available a number of Endler's qualitative claims and evaluations which we present by direct citation from his paper.

"There are many possible spatial patterns of selection and gene flow that can produce a given cline structure."

"Irregularities in environmental gradients increase the sensitivity of clines to the effects of gene flow in proportion to the increase in the differences in gene frequencies between the emigrants and the demes receiving the immigrants."

"Any asymmetry in gene flow does not lead to dedifferentiation if the environmental gradient is smooth; it merely shifts the position of the transition zone between the differentiated areas from that which would be expected if

there were no asymmetry. Abrupt geographic differences in gene, genotype, or morph frequencies should not, therefore, be interpreted as evidence for environmental changes in the immediate vicinity of the steepest part of the cline."

"Gene flow may be unimportant in the differentiation of populations along environmental gradients."

We can test a number of Endler's assertions on the basis of our exact analysis of his and our more general models. A variety of new insights emerge.

Slatkin [1973] contends that there is a single critical parameter which mainly determines the shape of the cline at least near the barrier where the character of selection sharply changes from favoring allele A to favoring allele a. This conclusion was correct for the special model he considered but not typical for the general description of the interaction between gene flow and the selection gradient structure as will be more apparent from the discussion and results reported in this paper.

## 2. RESULTS AND COMPARISONS OF FREQUENCY CLINES

We will describe a series of results bearing on the objectives and problems set forth in Section 1.

Consider a geographical cline displaying a two range selection regime having the schematic form

$$\begin{array}{ccc}
 \text{Demes} & p_1, \dots, p_{l_1} & p_{l_1+1}, \dots, p_{l_2} & p_{l_2+1}, \dots, p_N \\
 & \underbrace{\square \dots \square}_{-} & \underbrace{\square \dots \square}_0 & \underbrace{\square \dots \square}_{+}
 \end{array} \quad (2.1)$$

$$(1 < l_1 \leq l_2 < N)$$

where in the + demes, A is advantageous, and in the -

demes the a-type is advantageous, while in the central zone there are no selection differences (A and a are mutually neutral) (see (1.6) and (1.7) earlier).

Apart from the general structure of (2.1) there are no further requirements imposed on the detailed nature of the local selection functions. In particular we stipulate no regularity or pattern for the strength or degree of advantage for A in the respective demes of the + region and similarly for the - region. Endler [1973], Slatkin [1973] examined some special regular patterns fitting (2.1) by numerical means.

The associated migration form is assumed at first to be uniform gene flow between contiguous demes, with exchange rate  $m$ . The assumption tacitly implies that the relative deme sizes are equal. (Later we will discuss the model of unequal local rates of migration and the consequences of unequal relative deme sizes, paragraph E.)

Where selection operates first and migration occurs in the adult stage, the transformation equations relating the A-frequency distribution  $\underline{x} = (x_1, \dots, x_N)$  and  $\underline{x}' = (x'_1, \dots, x'_N)$  over two successive generations,  $x_i$  representing the proportion of the A-allele in deme  $i$ , are given by

$$\begin{aligned} x'_1 &= (1-m)f_1(x_1) + mf_2(x_2) \\ x'_i &= mf_{i-1}(x_{i-1}) + (1-2m)f_i(x_i) + mf_{i+1}(x_{i+1}), \\ &\quad i = 2, \dots, N-1 \end{aligned} \tag{2.2}$$

$$x'_N = mf_{N-1}(x_{N-1}) + (1-m)f_N(x_N).$$

Several criteria for the ascertainment of local instability of the fixation states  $\underline{0} = (0, \dots, 0)$  and  $\underline{1} = (1, \dots, 1)$ , corresponding to fixation of the a-allele and A-allele, respectively, in the context of an ordered geographical structure,



are developed in Karlin and Richter-Dyn [1976a], and for general multi deme population systems in Karlin [1976]. This is the problem of "protection" in a cline.

An equilibrium

$$\hat{\underline{x}} = (\hat{x}_1, \hat{x}_2, \dots, \hat{x}_N) \quad (2.3)$$

of the system (2.2) is called polymorphic if  $0 < \hat{x}_i < 1$  for all  $i$ . The equilibrium  $\hat{\underline{x}}$  is said to be monotone (a morphotone) if

$$\hat{x}_i < \hat{x}_{i+1} \quad \text{for all } i. \quad (2.4)$$

The slope of the cline for the equilibrium state  $\hat{\underline{x}}$  is described by the vector

$$\hat{\underline{\delta}} = (\hat{\delta}_1, \hat{\delta}_2, \dots, \hat{\delta}_{N-1}) \quad \text{where} \quad \hat{\delta}_i = \hat{x}_{i+1} - \hat{x}_i. \quad (2.5)$$

#### A. The nature of polymorphic equilibria.

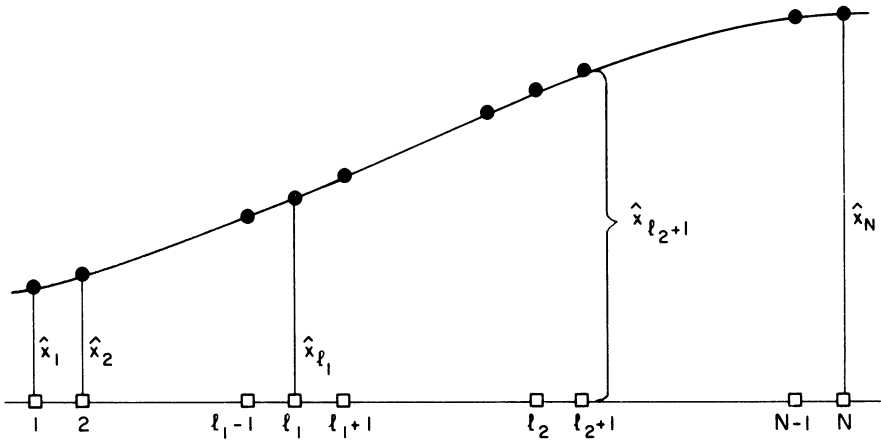
The result reported next is rather remarkable in that it is independent of any specific form of the selection functions except that the general pattern of (2.1) prevails.

Result I. (Every polymorphic equilibrium describes a monotone cline)

Consider a two range selection gradient as in (2.1). Any polymorphic equilibrium  $\hat{\underline{x}}$  is monotone. Moreover, the slope changes have the form

$$\begin{aligned} \hat{\delta}_1 < \hat{\delta}_2 < \dots < \hat{\delta}_{\ell_1} < \hat{\delta}_{\ell_1+1} = \dots = \hat{\delta}_{\ell_2-1} \\ > \hat{\delta}_{\ell_2} > \hat{\delta}_{\ell_2+1} > \dots > \hat{\delta}_{N-1}. \end{aligned} \quad (2.6)$$

A pictorial view of a polymorphic cline concomitant with the above result is now drawn:



showing that the maximum slope occurs in the neutral region. We stress again that every polymorphism is necessarily monotone and (2.6) holds independently of the variations in intensity and strength expressed in the local selection functions provided merely the geographical selection gradient displays the two range form of (2.1). A formal proof of this and the later results are elaborated in Karlin and Richter-Dyn [1976b].

Result II. (Uniqueness of a polymorphic cline)

If each local selection component function of  $\{f_1(x), \dots, f_N(x)\}$  corresponds to additive fitness values as in (1.9) and subject to a mild restriction limiting the maximum extent of migration exchange then there exists at most one polymorphic equilibrium.

A formal sufficient condition for the validity of Result II is that the magnitude of migration rate relative to the selection coefficients obey the inequalities

$$m \leq \frac{1}{|s_i|} \left\{ 1 - \sqrt{1 - s_i^2} \right\} \quad i = 1, 2, \dots, N$$

$$\approx |s_i|/2 \quad \text{for } s_i \text{ small}$$

In many circumstances the above restriction can be much relaxed. Thus for an environmental selection pattern involving a central locality such that the selection expression is the same but directed in favor of different allelomorphs on opposite sides of the center then the conclusion of Result II is valid for all levels of migration flow.

Under the conditions of Result II we secure the following classification pertaining to the dynamic behavior of the cline realization.

(a) If a-fixation is a locally stable equilibrium state and the a-allele is protected then the a-allele will be fixed from any internal initial state.

(a') If the A-allele is protected and 1 is locally stable then fixation of the A-type is established independent of the initial population makeup.

(b) Where a protected polymorphism exists then the population evolves at a geometric rate to a unique monotone polymorphic cline whose slope variation satisfies (2.6).

Cases (a) and (a') can occur only for  $m$  large and (b) is certainly assured for  $m$  small. In a symmetric model (Section 3), (b) holds for all migration rates.

Where the hypotheses of Result II are omitted then a selection gradient consistent with (2.1) can be constructed without a unique polymorphic equilibria, in fact including at least 3 polymorphisms. (According to Result I all such polymorphisms represent morphotones.) Examples also exist where both fixation states and at least one further polymorphic monotone cline are simultaneously stable. For general selection

functions there can exist cases with any number of polymorphisms. Henceforth unless stated otherwise, we assume Result II in force.

B. Comparison of two selection gradients the second having a longer range where allele A is favored.

In this context we seek to determine the relative change in the morphotone that emerges where the region in which A is advantageous is extended. The selection patterns to be compared are as indicated

$$\begin{array}{lcl}
 \text{Population 1} & \begin{array}{c} 1 \quad \ell_1 \quad \ell_2 \quad N \\ \square \cdots \square \cdots \square \cdots \square \\ f_1 \quad f_{\ell_1} \quad f_{\ell_2} \quad f_N \end{array} & \begin{array}{l} \text{Demes} \\ \text{associated selection} \\ \text{functions} \end{array} \\
 & & (2.7) \\
 \text{Population 2} & \begin{array}{c} \square \cdots \square \cdots \square \cdots \square \\ \tilde{f}_1 \quad \tilde{f}_{\ell_1} \quad \tilde{f}_{\ell_2} \quad \tilde{f}_N \end{array} & \begin{array}{c} N+1 \quad \tilde{N} \\ \square \cdots \square \\ \tilde{f}_{N+1} \quad \tilde{f}_{\tilde{N}} \end{array} \quad \tilde{N} > N
 \end{array}$$

The demes numbered  $1, \dots, N$  are identical for the two versions and accordingly  $f_i(x) \equiv \tilde{f}_i(x)$ ,  $i = 1, 2, \dots, N$ . However, there appear in population 2 new demes connected to deme  $N$  such that A is advantageous also in demes  $N+1, \dots, \tilde{N}$ . In other words, population 2 involves a longer segment (region) in favor of the A-type.

Result III. (Enlarged A-favorable region)

Let  $\hat{x} = (\hat{x}_1, \dots, \hat{x}_N)$  be a polymorphic equilibrium for the selection regime of population 1. For population 2 either only A-fixation is possible or if there exists a polymorphic cline

$$\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_N, \tilde{x}_{N+1} \dots \tilde{x}_{\tilde{N}})$$

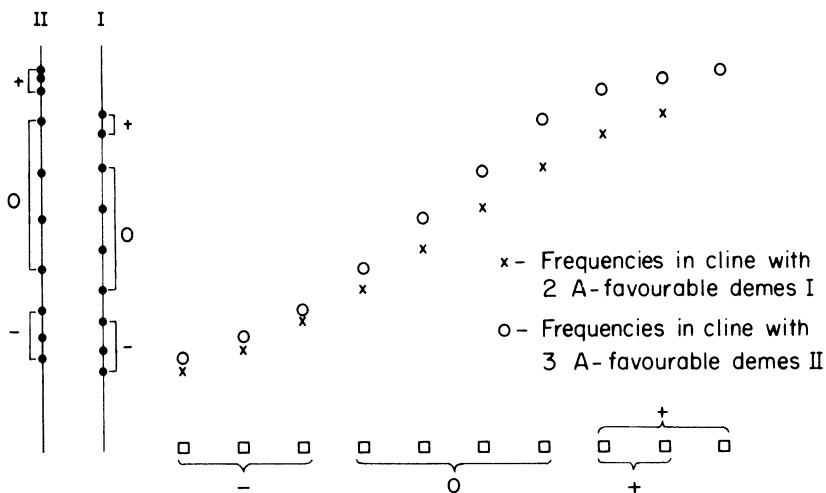
then

$$\hat{x}_i < \tilde{x}_i \text{ and } \hat{\delta}_i < \tilde{\delta}_i, \quad i = 1, 2, \dots, N. \quad (2.8)$$

Thus, the extended cline is steeper (exhibiting more gene frequency differentiation) over the common region of the population distribution provided its polymorphic character is retained. Where  $\tilde{x}$  does not exist then necessarily A-fixation is the exclusive stable outcome evolving from the selection regime of population 2.

With any morphotone, it is obvious that the slope between localities at the neutral region cannot exceed  $\frac{1}{k+1}$  where  $k$  is the number of demes composing the neutral zone.

### COMPARISON OF CLINES IN CASE OF AN EXTENDED A-FAVOURABLE RANGE



C. The effects of the length of the neutral zone.

Another comparison of interest concerns the following setup. Consider two versions of environmental selection gradients as now depicted.

$$\begin{array}{c}
 \text{Demes} \quad 1 \quad \dots \quad l_1 \quad l_1+1 \quad \dots \quad l_2 \quad l_2+1 \quad \dots \quad N \\
 \text{Condition (i)} \quad f \quad \underbrace{\square \dots \square}_{-} \quad \underbrace{\square \dots \square}_0 \quad \underbrace{\square \dots \square}_{+}
 \end{array}
 \quad (2.9)$$

$$\begin{array}{c}
 \text{Condition (ii)} \quad f \quad \underbrace{\square \quad \dots \quad \square}_{-} \quad \underbrace{\square \quad \dots \quad \square}_0 \quad \underbrace{\square \quad \dots \quad \square}_{+}
 \end{array}$$

$$\tilde{l}_1 = l_1, \quad N+r = \tilde{N}, \quad r = \tilde{l}_2 - l_2 > 0$$

where the selection functions are related according to

$$\begin{aligned}
 f_i(x) &\equiv \tilde{f}_i(x), \quad i = 1, \dots, l_1 \\
 f_{N-j}(x) &\equiv \tilde{f}_{\tilde{N}-j}(x), \quad j = 0, \dots, N-l_2-1
 \end{aligned}
 \quad (2.10a)$$

while

$$f_\mu(x) = \tilde{f}_\nu(x) = x, \quad l_1+1 \leq \mu \leq l_2, \quad \tilde{l}_1+1 \leq \nu \leq \tilde{l}_2. \quad (2.10b)$$

Thus the neutral zone is longer under the condition (ii) and otherwise the selection patterns coincide. The next assertion shows how the equilibrium cline is affected by the extent of the neutral region.

Result IV. Consider the model of (2.9) and suppose the polymorphism

$$\begin{aligned}
 \hat{x} &= (\hat{x}_1, \dots, \hat{x}_N) \quad \text{for condition (i) and} \\
 \tilde{x} &= (\tilde{x}_1, \dots, \tilde{x}_N) \quad \text{for condition (ii)}
 \end{aligned}$$

exist. Then

$$\hat{x}_i > \tilde{x}_i \quad \text{and} \quad \hat{\delta}_i > \tilde{\delta}_i \quad \text{for} \quad i = 1, 2, \dots, l_2 \quad (2.11)$$

and

$$\begin{aligned}
 \tilde{x}_{N-j} &> \hat{x}_{N-j} \quad \text{and} \quad \tilde{\delta}_{N-j} < \hat{\delta}_{N-j} \quad \text{for} \\
 j &= 0, 1, \dots, N-l_1-1.
 \end{aligned}
 \quad (2.12)$$

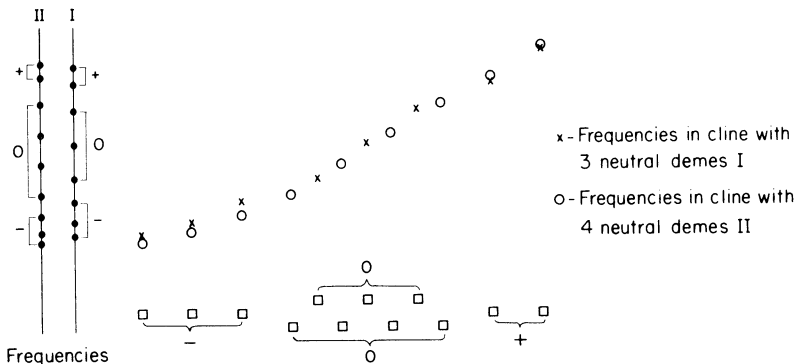
Thus, a more prolonged neutral zone tends to accent more the selection forces operating on the outer regions of the population distribution. This finding is perhaps unintuitive; its validity applies independent of the detailed nature of the selection functions operating in the - and + range of the population. It is worth emphasis that over the neutral zone the frequency cline is linear bearing a more pronounced flat but discernible slope with increasing length of this portion. In another work we will highlight several more sensitive distinguishing characteristics of a morphotone corresponding to a two range environment with a long neutral region vis a vis the form of a cline associated with an underlying selection regime expressing a regular monotone pattern of heterozygote advantage over the range of the species.

By combining the estimates (2.11) and (2.12) we extract the following bounds on the equilibrium frequencies in the neutral zone

$$\tilde{x}_i < \hat{x}_i < \tilde{x}_{i+r} \quad (r = \tilde{l}_2 - l_2) \quad i = l_1 + 1, \dots, l_2 \quad (2.13)$$

The comparison of (2.11) and (2.12) can be graphically displayed

COMPARISON OF CLINES IN CASE OF AN EXTENDED NEUTRAL REGION



The following precise stability fact is of interest.

Result V. Consider the same conditions as in IV above. Where under Condition (i) there exists a polymorphism, then under Condition (ii) (involving an enlarged neutral region) a globally stable polymorphism necessarily occurs.

The above finding asserts that with an underlying monotone ecotone, (a term used by Endler [1973] for a two range selection regime), the possibilities of a stable polymorphism are significantly increased as the neutral zone is extended while preserving the same selection pattern in the non-neutral locations.

D. The influence of the magnitude of the migration rate.

Consider an ecotone selection pattern where the localities  $p_1, \dots, p_{\ell_1}$  have the a-allele advantageous,  $p_{\ell_2+1}, \dots, p_N$  confer advantage to the A-allele and the demes of the center region  $p_{\ell_1+1}, \dots, p_{\ell_2}$  ( $\ell_2 \geq \ell_1+1$ ) bear no selection differences. Suppose there exists a polymorphic equilibrium  $\hat{x} = (\hat{x}_1, \dots, \hat{x}_N)$ . We know on the basis of Result I that  $\hat{x}$  describes a monotone cline (morphotone) endowed with the properties (2.4) and (2.6). The next result concerns the changes of  $\hat{x}(m)$  as  $m$  increases (showing its dependence on the migration rate).

Result VI. Suppose the existence of equilibrium frequency clines  $\hat{x}(m)$  and  $\hat{x}(m')$  for the migration rates  $m$  and  $m'$ , respectively with  $m < m'$ . Then there exists a deme site  $k_0$  such that

$$\begin{aligned} \text{(i)} \quad & \hat{x}_i(m) < \hat{x}_i(m') \quad , \quad i = 1, 2, \dots, k_0 - 1 \\ \text{(ii)} \quad & \hat{x}_j(m') < \hat{x}_j(m) \quad , \quad j = k_0 + 1, \dots, N \quad . \end{aligned} \tag{2.14}$$

In the symmetric model  $k_0$  corresponds to the center deme (see Section 3).



The interpretation of (2.14) is the natural one: Increased gene flow either causes fixation or attenuates the frequency differentiation between demes to the extent that in both portions of the population distribution the attained gene frequencies appear more flat.

#### E. Unequal local migration rates

Consider a two environment selection regime as in (2.1) with a stepping stone migration flow but allowing unequal migration exchanges between demes (and also not necessarily equal rates in reciprocal directions at a locality). Even with a uniform individual migration propensity but where the deme sizes are unequal: for example, let the relative deme sizes at  $\mathcal{P}_1, \dots, \mathcal{P}_N$  be  $c_1, c_2, \dots, c_N$ , respectively,  $c_i > 0$ ,  $\sum c_i = 1$ , then the associated backward migration matrix

$M = ||m_{ij}||_1^N$  # is for the case at hand

$$M = ||m_{ij}|| =$$

$$\begin{pmatrix} \frac{(1-m)c_1}{\gamma_1} & , & \frac{mc_2}{\gamma_1} & , & 0 & , & 0, \dots, 0 \\ \frac{mc_1}{\gamma_2} & , & \frac{(1-2m)c_2}{\gamma_2} & , & \frac{mc_3}{\gamma_2} & , & 0 \\ 0 & & \frac{mc_2}{\gamma_3} & , & \frac{(1-2m)c_3}{\gamma_3} & , & \frac{mc_4}{\gamma_3} \\ & & & \ddots & & & \\ & & & & 0 & , & \frac{mc_{N-1}}{\gamma_N} & , & \frac{(1-m)c_N}{\gamma_N} \end{pmatrix} \quad (2.15)$$

---

# ( $m_{ij}$  = fraction of individuals of deme  $j$  contributed to deme  $i$  after selection and migration).

where  $\gamma_i = mc_{i-1} + (1-2m)c_i + mc_{i+1}$ ,  $2 \leq i \leq N-1$ ,  
 $\gamma_1 = (1-m)c_1 + mc_2$ ,  $\gamma_N = mc_{N-1} + (1-m)c_N$ .

Where the migration flow between neighboring demes can also vary, then the backward migration matrix  $M = ||m_{ij}||$  attains the general form

$$M = \begin{pmatrix} r_1 & m'_1 & 0 & 0 \\ m_2 & r_2 & m'_2 & 0 \\ 0 & m_3 & r_3 & m'_3 \\ & & \ddots & \ddots \\ & & & m_N & r_N \end{pmatrix} \quad \begin{matrix} m_i + m'_i + r_i = 1 \\ i = 1, 2, \dots, N \end{matrix} \quad (2.16)$$

The transformation equations version of (2.2) in this general clinal framework becomes

$$\begin{aligned} x'_1 &= r_1 f_1(x_1) + m'_1 f_2(x_2) \\ x'_i &= m_i f_{i-1}(x_{i-1}) + r_i f_i(x_i) + m'_i f_{i+1}(x_{i+1}), \quad (2.17) \\ 2 \leq i \leq N-1 \\ x'_N &= m_N f_{N-1}(x_{N-1}) + r_N f_N(x_N). \end{aligned}$$

Conditions for the existence of a protected polymorphism for this non-homogeneous stepping stone model (2.16) are developed in Karlin and Richter-Dyn [1976a]. We highlight now the analog of Result I for this case of unequal local migration rates.

Result VII. Suppose  $\hat{x} = (\hat{x}_1, \hat{x}_2, \dots, \hat{x}_N)$  is a polymorphic equilibrium for the general stepping stone migration mode (2.16) involving a two range selection gradient of the character (2.1). Then  $\hat{x}$  describes a morphotone, i.e.,

$$\hat{x}_1 < \hat{x}_2 < \dots < \hat{x}_N \quad (\hat{x}_i \text{ is increasing}). \quad (2.18)$$

Let  $\hat{\delta}_i = \hat{x}_{i+1} - \hat{x}_i$ ,  $i = 1, 2, \dots, N-1$  be the slope of the morphotone measuring local gene frequency differentiation and

define  $a_1 = 1$ ,  $a_k = \prod_{i=1}^{k-1} \left( \frac{m'_i}{m_{i+1}} \right)$ ,  $k = 2, \dots, N$ . If

demes  $\mathcal{P}_1$  to  $\mathcal{P}_{\ell_1}$  constitute the  $-$  region (cf. (2.1)),

$\mathcal{P}_{\ell_1+1}$  to  $\mathcal{P}_{\ell_2}$  comprise the neutral zone and demes  $\mathcal{P}_{\ell_2+1}$

to  $\mathcal{P}_N$  the  $+$  environmental portion then

$$\begin{aligned} m'_1 a_1 \hat{\delta}_1 &< m'_2 a_2 \hat{\delta}_2 < \dots < m'_{\ell_1} a_{\ell_1} \hat{\delta}_{\ell_1} < m'_{\ell_1+1} a_{\ell_1+1} \hat{\delta}_{\ell_1+1} = \dots = \\ m'_{\ell_2-1} a_{\ell_2-1} \hat{\delta}_{\ell_2-1} &> m'_{\ell_2} a_{\ell_2} \hat{\delta}_{\ell_2} > m'_{\ell_2+1} a_{\ell_2+1} \hat{\delta}_{\ell_2+1} > \dots > \end{aligned} \quad (2.19)$$

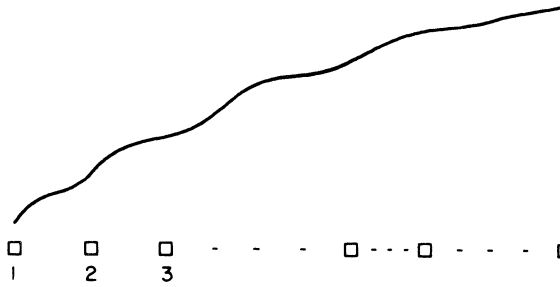
$$m'_{N-1} a_{N-1} \hat{\delta}_{N-1}.$$

The relations of (2.19) are equivalent to

$$\begin{aligned} m_i \hat{\delta}_{i-1} &< m'_i \hat{\delta}_i && \text{in the } - \text{ region} \\ &= && \text{in the } 0 \text{ region} \\ &> && \text{in the } + \text{ region} \end{aligned} \quad (2.20)$$

It is manifest in the present case, that the actual maximum gene frequency differentiation, the position where  $\max \hat{\delta}_i$  is attained, need not occur in the neutral zone as in the uniform flow situation.

Unlike the smooth convex-linear concave shape attendant to a polymorphism with homogeneous migration rates with any background of selection intensities consistent with (2.1), for variable local migration rates it is now possible to have kinks in the clinal shape as depicted



but persistently a morphotone (the property of increasing gene frequency) appears.

For the special case of (2.15) we have  $m'_i/m_i = c_{i+1}/c_{i-1}$ . Then, where relative population size decreases from both ends of the species range toward the central zone, the maximum gene frequency differentiation ( $\max \hat{\delta}_i$ ) occurs again in the neutral region. More precisely

$$\frac{\hat{\delta}_{i-1}}{\hat{\delta}_i} \left\{ \begin{array}{ll} < & \text{in the } - \text{ region} \\ = \frac{c_{i+1}}{c_{i-1}} & \text{in the } 0 \text{ region} \\ > & \text{in the } + \text{ region} \end{array} \right.$$

On the basis of the relations of (2.19) and (2.20) and observations of the gene frequency clinal patterns, it may be possible to estimate local migration intensities.

It is worth pointing out that all the qualitative comparisons of paragraphs B-D apply in the general stepping stone context with variable local migration rates. Result VI requires a reformulation which will be set forth elsewhere.

F. The influence of the order of migration and selection pressures on the morphotone

The previous results were derived for the model where the genetic forces operating in each generation had the timing

local selection  $\longrightarrow$  migration (at adult stage).  
followed by (2.21)

We discuss briefly the model where the order is reversed in (2.21). Accordingly, consider now the order to be

migration (as gametes)  $\longrightarrow$  selection. (2.22)

The analog of (2.2) for the model (2.22) with homogeneous migration flow follows the transformation equations

$$\begin{aligned} y_1' &= f_1((1-m)y_1 + my_2) \\ y_i' &= f_i(my_{i-1} + (1-2m)y_i + my_{i+1}), \quad i = 2, 3, \dots, N-1 \\ y_N' &= f_N(my_{N-1} + (1-m)y_N) \end{aligned} \quad (2.23)$$

where  $\underline{y} = (y_1, \dots, y_N)$  and  $\underline{y}' = (y_1', \dots, y_N')$  denote the frequency state in two successive generations.

The analog of Result I carries over completely:

Result I'. Suppose  $\hat{\underline{y}} = (\hat{y}_1, \dots, \hat{y}_N)$  is a polymorphic equilibrium for (2.23) with underlying selection regime as in (2.1). Then necessarily  $\hat{\underline{y}}$  describes a morphotone; viz.,

$$\hat{y}_1 < \hat{y}_2 < \dots < \hat{y}_N \quad (2.24)$$

and the slope vector  $\hat{\eta}_i = \hat{y}_{i+1} - \hat{y}_i$  satisfies

$$\hat{\eta}_1 < \hat{\eta}_2 < \dots < \hat{\eta}_{\ell_1} = \hat{\eta}_{\ell_1+1} = \dots = \hat{\eta}_{\ell_2} > \hat{\eta}_{\ell_2+1} > \dots > \hat{\eta}_{N-1} \quad (\text{cf. (2.6).}) \quad (2.25)$$

The uniqueness version of Result II carries over and all the comparisons of Results III-VI are valid in the setting of (2.22).

What is of interest is to compare the morphotone (2.24) for the model (2.22) with the morphotone  $\hat{\underline{x}}$  emerging in the formulation of (2.21).

Result VIII. Suppose  $\hat{x}$  and  $\hat{y}$  are morphotones for the models of (2.21) and (2.22), respectively acted on by the same selection gradient and homogeneous migration flow. (Here only the order of applications of the forces are interchanged.) Then  $\hat{y}_i = f_i(\hat{x}_i)$  and therefore

$$\begin{aligned}\hat{y}_i &< \hat{x}_i & , \quad i = 1, 2, \dots, l_1 & \quad (\text{over the } - \text{ region}) \\ \hat{y}_i &= \hat{x}_i & \quad i = l_1 + 1, \dots, l_2 & \quad (\text{over the neutral region}) \\ \hat{y}_i &> \hat{x}_i & \quad i = l_2 + 1, \dots, N & \quad (\text{over the } + \text{ region}).\end{aligned}\tag{2.26}$$

The deduction of (2.26) is perhaps intuitive after the fact. When selection follows migration the effects due to selection are more pronounced in each environmental range. Equivalently, migration flow acting after selection will smoothen the selection impression operating in the two extremes of the range.

### 3. RESULTS FOR SYMMETRIC TWO RANGE SELECTION REGIMES

When there is a natural symmetrical center in the distribution of demes, e.g., a position at which selection changes from favoring one type to favoring the other type, it is more convenient to label the demes in the sequence

$$p_{-K}, p_{-K+1}, \dots, p_{-1}, p_0, p_1, \dots, p_K^* \tag{3.1}$$

with  $p_0$  as the deme in the center of the region and  $p_{-K}$  and  $p_K$  the end demes. The associated selection functions

---

\* The actual deme  $p_0$  may or may not be present but position 0 is a demarcation point.

reflecting the environmental selection gradient are described with the same notation

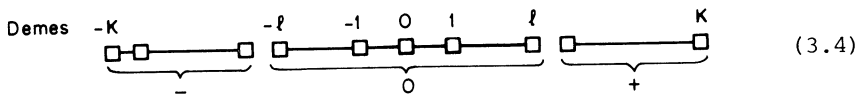
$$\{f_{-K}(x), \dots, f_0(x), \dots, f_K(x)\} \quad (3.2)$$

Definition. A selection pattern is said to be symmetrical (it may be more logical to call it asymmetrical or odd) with respect to the center deme locality called for brevity, "symmetric", if

$$f_i(x) \equiv 1 - f_{-i}(1-x) \quad , \quad 0 \leq x \leq 1 \quad , \quad i = 0, 1, 2, \dots, K \quad (3.3)$$

One can envision the symmetric model as depicting a selection gradient anti-symmetrical as a function of the distance from  $\phi_0$ . Throughout this section we assume uniform gene flow (with stepping stone migration mode) so that  $m_i = m$  for all  $i$ .

A symmetrical two range ecotone has the following structure



where in the  $+$  ( $-$ ) demes the A (a)-allele is advantageous while the 0 demes (those numbered  $i = 0, \pm 1, \dots, \pm l$ ) are neutral in selection expression.

The pattern of (3.4) coupled with (3.3) reflects a picture where the ecological selection expression in one direction varies in intensity, not necessarily in any regular way but persists in favor of the A-allele, and in a symmetrical fashion involving the same intensities the a-allele is favored for the localities in the opposite direction from the center position.

Many of the results reported in this section are valid for geographical linear population structures approximately

symmetric.

Result IX. For a symmetrical monotone selection gradient, there always exists a polymorphic state  $\underline{x}^*$  =  $(x_{-K}^*, \dots, x_0^*, \dots, x_K^*)$  with the properties

$$x_i^* = 1 - x_{-i}^* \quad \text{and} \quad x_0^* = \frac{1}{2} . \quad (3.5)$$

We call such an equilibrium a symmetric polymorphism.

For the case where each local selection function reflects additive fitnesses (not necessarily with the same selection coefficients) then, for each migration rate,  $\underline{x}^*$  is the unique globally stable polymorphic equilibrium.

A. The changes in the frequency cline when extending the ecotones.

We will make some comparisons similar to the Results III and IV with the proviso here of preserving the symmetry of the selection ecotone. Consider the situation of (3.4) (referred to as Condition (i)) vis a vis the enlarged geographical symmetric ecotone involving  $2L+1 > 2K+1$  demes

$$\text{Condition (ii)} \quad \begin{array}{ccccccc} & -L & & -K & & 0 & & K & & L \\ \hline & \underbrace{\hspace{1.5cm}} & & \underbrace{\hspace{1.5cm}} & & \underbrace{\hspace{1.5cm}} & & \underbrace{\hspace{1.5cm}} & & \\ & - & & & & 0 & & + & & \end{array} \quad L > K \quad (3.6)$$

stipulating that the clines (3.4) and (3.6) coincide in the positions from  $-K$  to  $K$ .

Thus we have in (3.6) extra A-advantageous demes at the right end of the geographical region and by symmetry extra a-advantageous demes present for the left region.

Let

$$\underline{x}^* = (x_{-K}^*, \dots, x_{-1}^*, x_0^*, x_1^*, \dots, x_K^*) \quad (3.7)$$

and



$$\tilde{\underline{x}} = (\tilde{x}_{-L}, \dots, \tilde{x}_{-K}, \dots, \tilde{x}_0, \dots, \tilde{x}_K, \dots, \tilde{x}_L) \quad (3.8)$$

be the symmetric equilibrium frequency clines present in conditions (i) and (ii), respectively. The analogue of Result III is stated below in Result X. Notice that unlike the setup in III where the conditions (i) and (ii) differed only in the addition of new advantageous A-type positions, the condition of (ii) now involves simultaneously in a symmetrical setting new localities favorable for the A and a-alleles.

Result X. For the population conditions (i) and (ii), we have

$$\tilde{x}_i > x_i^*, \quad i = 1, 2, \dots, K, \quad \tilde{x}_0 = x_0^* = \frac{1}{2} \quad (3.9)$$

and

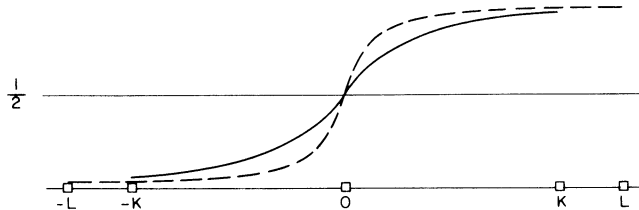
$$\max_{i \geq 0} \tilde{\delta}_i = \tilde{\delta}_0 = \tilde{x}_1 - \frac{1}{2} > x_1^* - \frac{1}{2} = \delta_0^* = \max_{i \geq 0} \delta_i^* \quad (3.10)$$

$$(\delta_i = x_{i+1} - x_i) \quad .$$

Moreover,  $\tilde{\delta}_i > \delta_i^*, \quad i = 1, 2, \dots, K \quad .$

Thus, where the advantageous A and a regions are extended preserving the symmetric ecotone structure, the maximal frequency differentiation occurring at the center deme is sharpened. In more picturesque language the whole cline tilts more steeply about the null position as the cline is lengthened symmetrically.

A graphical display of the above statements is now given.



Slope at center increases if new demes are added (in a symmetrical manner) to both ranges.

The conclusion of (3.9) is correct for any magnitude of migration flow.

B. Very long species range.

Consider a doubly infinite monotone symmetric cline  $\mathbf{C}$  with deme location (for notational convenience) at all the integers (positive and negative). The setting is that of an infinite or far-reaching horizon (in both directions) for the distribution of the population. Let  $\mathbf{C}_K$  be the truncated environment involving the  $2K+1$  localities embracing the population demes  $\mathbf{p}_i$ ,  $-K \leq i \leq K$ . We denote by

$$\hat{\mathbf{x}}^{(K)} = (\hat{x}_{-K}^{(K)}, \dots, \hat{x}_0^{(K)}, \dots, \hat{x}_K^{(K)}) \quad (3.11)$$

the unique symmetric polymorphic frequency cline (necessarily monotone) for a population confined to  $\mathbf{C}_K$ .

Result XI. Over a very long species range, we have

$$\lim_{K \rightarrow \infty} \hat{x}_i^{(K)} = \hat{x}_i^{(\infty)} \quad , \quad i = 0, \pm 1, \dots \quad (3.12)$$

$$\frac{1}{2} = \hat{x}_0^{(\infty)} < \hat{x}_1^{(\infty)} < \hat{x}_2^{(\infty)} < \dots < \dots \rightarrow 1$$

and

$x_i^{(\infty)} = 1 - x_{-i}^{(\infty)}$  is an equilibrium cline for the extended population.

In practice, already for  $K \geq 5$ ,  $\hat{x}_i^{(K)}$  is quite close to its limit at localities  $i=1$  and  $2$ . What is, perhaps, surprising is that the frequency gradient at  $0$  does not become vertical. Thus, the slope is not as steep as might be anticipated due to the prolonged opposite selection pressures operating on the two sides of the null position.

C. The shape of the frequency cline as a function of migration rate.

Consider an ecotone as in (3.4) having the symmetric polymorphic equilibrium frequency array

$$\hat{x}(m) = (\hat{x}_{-K}(m), \dots, \hat{x}_{-1}(m), \frac{1}{2}, \hat{x}_1(m), \dots, \hat{x}_K(m)) \quad (3.13)$$

where we display its dependence on the migration rate  $m$ .

Result XII. As  $m$  increases (larger migration rate)  $\hat{x}_i(m)$ ,  $i \geq 1$  decreases. More specifically for  $m < m'$

$$\frac{1}{2} < \hat{x}_i(m') < \hat{x}_i(m) \quad , \quad i = 1, 2, \dots, K \quad (3.14)$$

This finding confirms the intuitive property that with any background symmetric monotone selection gradient, increased gene flow levels the local selection effects yielding a more constant less differentiated gene frequency pattern. Endler [1973] claims on the basis of computer calculations performed on some special regular selection patterns that the influence of the magnitude of migration rate is only slight. Our analy-

sis shows that there is a definitive trend independent of the individual selection components such that increasing migration rates cause more equalized gene frequencies provided only the monotone character of the ecotone is maintained. Further analysis suggests that the swamping effect of increased gene flow is more pronounced in the presence of a population distribution involving a few demes than in the circumstances of many numbers of demes.

D. The case of hybrid (heterotic) central demes.

We alter the model (3.4) where in place of the central neutral zone we assume the local selection functions  $f_i(x)$ ,  $-l \leq i \leq l$ , express heterozygote advantage symmetric in the A and a genes. This situation is certainly of common occurrence where the interface between the areas favoring the A and a-alleles provides a region of coadaptability of both genes in that the heterozygote has superior fitness.

We will refer to the environmental selection gradient now depicted.

$$\begin{array}{ccccccc}
 & -K & & -l-1 & -l & & 0 & & l & & l+1 & & K \\
 \text{Deme} & \square & & \square & \square & \cdots & \square & \cdots & \square & & \square & & \square \\
 & \underbrace{\hspace{1.5cm}}_{-} & & \underbrace{\hspace{1.5cm}}_{(+,-)} & & \underbrace{\hspace{1.5cm}}_{+}
 \end{array} \tag{3.15}$$

where the demes with symbol (+ -) have selection functions expressing heterozygote advantage, as the mixed heterotic ecotone. (The opposite model where heterozygote disadvantage is manifested in the center region will be discussed in the following paragraph.)

In Karlin and Richter-Dyn [1976c] we will treat models with local selection functions manifesting some regular patterns of heterozygote advantage at each deme. We will compare and

contrast the resulting symmetric equilibrium frequency cline in these cases with the partially heterotic hybrid model associated to the ecotone (3.15).

Our principal finding on the mixed heterotic monotone cline is as follows

Result XIII. Consider a selection regime of a mixed heterotic ecotone. There exists a symmetric polymorphic monotone cline

$$\underline{\tilde{x}} = (\tilde{x}_{-K}, \tilde{x}_{-K+1}, \dots, \tilde{x}_{-1}, \tilde{x}_0, \tilde{x}_1, \dots, \tilde{x}_K) . \quad (3.16)$$

Every symmetric polymorphic cline has the following properties

$$\frac{1}{2} = \tilde{x}_0 < \tilde{x}_1 < \tilde{x}_2 < \dots < \tilde{x}_K, \quad \tilde{x}_{-i} = 1 - \tilde{x}_i, \quad i=1, \dots, K \quad (3.17)$$

and  $\tilde{\delta}_i = \tilde{x}_{i+1} - \tilde{x}_i$  satisfies

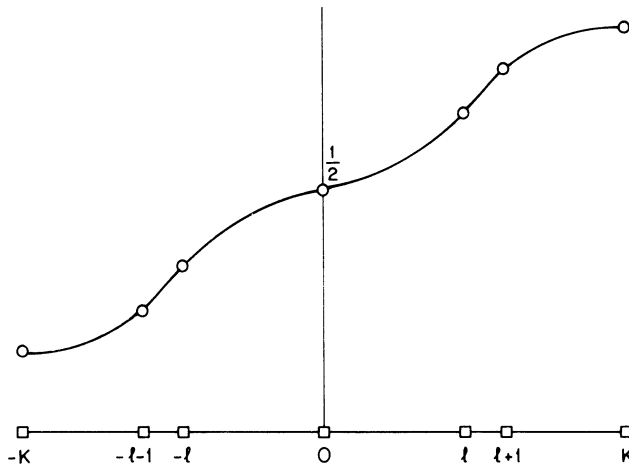
$$0 < \tilde{\delta}_1 < \tilde{\delta}_2 < \dots < \tilde{\delta}_\ell, \quad \tilde{\delta}_{\ell+1} > \dots > \tilde{\delta}_{K-1} \quad (3.18)$$

and, of course,  $\tilde{\delta}_{-i} = -\tilde{\delta}_i$  .

If  $\underline{x}^*$  is the unique symmetric polymorphism for the cline  
(3.4) then

$$\tilde{x}_i < x_i^*, \quad i \geq 1, \quad \tilde{\delta}_i > \delta_i^*, \quad i = \ell+1, \dots, K-1 . \quad (3.19)$$

In view of (3.18), the steepest slope in the cline of (3.16) does not occur at the null position (the central deme) but at the boundary demes  $\ell$  and  $-\ell$  where the local selection expression changes pronouncedly from that of overdominance to that of directional selection. The picture is as follows.



The relation (3.19) shows that with a central overdominant region the gene frequency array tends to be more flat at the center than with a dividing neutral zone.

As  $m \downarrow 0$  the equilibrium  $\tilde{x}(m)$  converges to  $\tilde{x}(0)$  with coordinates

$$\begin{array}{ccccccc} \text{Deme \#} & -K & -l-1, -l & & l, l+1 & & K \\ \underline{x} & = (0, \dots, 0, \frac{1}{2}, \frac{1}{2}, \dots, \frac{1}{2}, 1, \dots, 1) \end{array}$$

where the transition from 0 to  $\frac{1}{2}$  occurs between demes  $-l$  and  $-l-1$  and from  $\frac{1}{2}$  to 1 at the location of demes  $l$  to  $l+1$ .

#### E. The case of underdominant central demes.

The model now differs from (3.15) in that the local selection functions for the central demes reflect heterozygote disadvantage (underdominance) symmetric in the two alleles. We call this model an ecotone with underdominant central zone. Here in the interface zone both homozygotes carry superior fitness to the heterozygote. The hybrid is less adapted to the central environmental habitats. The equilibrium frequency

cline now takes the following shape

Result XIV. There exists a symmetric polymorphic morphotone

$$\tilde{x} = (\tilde{x}_{-K}, \dots, \tilde{x}_{-1}, \tilde{x}_0, \tilde{x}_1, \dots, \tilde{x}_K) \quad , \quad \tilde{x}_0 = \frac{1}{2}$$

$$\tilde{x}_i > x_i^* \quad , \quad 1 \leq i \leq K \quad , \quad \tilde{\delta}_i < \delta_i^* \quad , \quad i+1 \leq i \leq K \quad ,$$

which satisfies

$$\tilde{\delta}_1 > \tilde{\delta}_2 > \dots > \tilde{\delta}_{K-1} > 0 \quad (\tilde{\delta}_i = \tilde{x}_i - \tilde{x}_{i-1}) \quad . \quad (3.20)$$

The existence of the underdominant demes in the center tends to sharpen the gene frequency differentiation manifested at the center deme over that of a plain neutral zone. It should be cautioned that in this underdominant model for slight migration rate there occur non-monotone stable polymorphisms (see Karlin and McGregor [1972]). However, for moderate migration flow  $m \geq .05$  these non-monotone clines apparently do not persist provided the selection coefficients are of moderate magnitude.

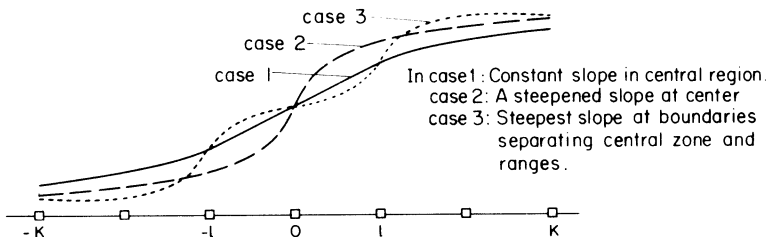
Symmetric environmental selection gradient with different central zone

Case 1 : Central zone is neutral.

Case 2 : Central zone has heterozygote disadvantage at demes  $(P_{-1}, \dots, P_1)$   
Both homozygotes have superior fitness to heterozygote.

Case 3 : Central zone has heterozygote advantage, symmetric fitnesses with respect to alleles A and a at demes  $(P_{-1}, \dots, P_1)$ .

Equilibrium frequency clines



F. Comparisons of morphotones for increased population subdivision with a prescribed aggregate level of selection.

We concentrate on the symmetric model of this section with local additive allelic fitness effects. For definiteness and ease of exposition we consider the two environmental selection gradient as in (3.4) with a single neutral deme at the center. Let  $S$  be the cumulative selection coefficient in the right portion of the species range corresponding to the demes of positive index (see 1.12a) and suppose  $S$  is divided equally among the  $K$  demes so that the AA-genotype selection coefficient per deme is  $S/K$ . In the negative range symmetry entails that the aa-genotype selection coefficient per deme is  $-S/K$ . Let  $x_{(K)}^*$  denote the equilibrium frequency value in deme 1.

It can be proved that the maximum slope of the cline

$$\delta_{(K)}^* = x_{(K)}^* - \frac{1}{2} \quad (3.21)$$

decreases as  $K$  increases. Thus, when preserving the total selection effects, the resulting morphotone is less steep with increasing population subdivision.

G. The nature of the morphotone with three comparable environments.

We displayed in (1.11) three environments I, II, III, each more heterogeneous than the preceding in the sense of Definition 2 of Karlin [1976] having the same aggregate selection intensity in the right and left population ranges.

We assume the environmental selection gradients of I-III are each symmetric with respect to the center deme.

Result XV. Let  $\delta_I$ ,  $\delta_{II}$  and  $\delta_{III}$  be the maximum slopes of the morphotones for the symmetric selection regimes I, II and



III respectively (given in (1.11)). Also let  $\{x_i^I\}_{-K}^K$ ,  $\{x_i^{II}\}_{-K}^K$  and  $\{x_i^{III}\}_{-K}^K$  be the unique stable symmetric cline associated with the respective gradients. Then for all migration rates,  $0 \leq m \leq \frac{1}{4}$ ,

$$\delta_I < \delta_{II} \quad . \quad (3.22)$$

Moreover, there exists a deme position  $k$ ,  $1 \leq k \leq K$  such that

$$\begin{aligned} x_i^I &< x_i^{II} \quad , \quad i = 1, 2, \dots, k \\ x_i^I &> x_i^{II} \quad , \quad i = k+1, \dots, K \end{aligned} \quad (3.23)$$

It appears that

$$\delta_{II} < \delta_{III} \quad . \quad (3.24)$$

The foregoing result is, perhaps, not surprising. It states that with the intensity of selection differences, strongest in the neighborhood of the central deme, a more pronounced gene frequency change is established there.

#### 4. SUMMARY AND DISCUSSION

A number of problems were formulated in Section 1 with objective to help explicate and discriminate the abundantly observed clinal patterns of gene frequency variation. A hierarchy of mathematical cline models incorporating some forms of selection migration interaction have been studied by analytic means or with the aid of computer runs by Haldane [1948], Fisher [1950], Hansen [1966], Slatkin [1973], Endler [1973] and others. Following these authors we concentrate in this work on a two-range geographical selection gradient where allele  $A$  is favored in the right range of the population and the alternative allele  $a$  is favored in the left

range. This kind of selection background, referred to as an ecotone, is prominently present in a variety of biogeographic situations. The gene flow pattern is stipulated to be that of the stepping stone migration mode. (Ehrlich and Raven [1969] cite many examples of animal and plant populations where gene flow is considerably localized.)

The detailed model is set forth in Section 1. We have reported in Sections 2 and 3 a series of results characterizing and contrasting the gene frequency arrays that are established with different forms of underlying ecotones. We presently summarize several of the salient findings and note briefly some of their implications with reference particularly to the questions of Section 1 :

(i) For an ecotone selection gradient, it is established that any polymorphic equilibrium describes a morphotone (monotone cline) that is, the A-allele frequency increases along the population transect going from left to right. The fact that every equilibrium is a morphotone is exceptionally robust prevailing with:

- (a) unequal local rates of exchange between neighboring demes,
- (b) in the presence of variable deme sizes, and
- (c) allowing irregularities in the local selection intensity provided only the ecotone property is maintained. (See Results I, I', VII of Section 2.)

With this outcome, in order to correlate the selection gradient to the associated frequency array, it is essential to plot and examine in detail the changes in gene frequency (the slope function). By these means it may be possible to evaluate (and sometimes estimate) the influence of unequal local migration rates (see especially Result VII of Section 2).

(ii) Where gene flow is homogeneous over the linear population range, the resulting A frequency morphotone has a precise monotone shape exhibiting exactly one inflection point. The graph is therefore of the form (from left to right) first convex increasing over the region where allele a is advantageous, linear on the neutral range and concave increasing over the region where A is favored. These descriptions apply independently of the detailed nature of the local selection functions.

A very kinky morphotone cannot be attributable to an inherent uniform migration flow even in the presence of substantial spatial variation in selection intensities.

Observed frequency clines, not of this precise convex-linear - concave shape, would entail that either the model is not relevant because the underlying selection gradient is not an ecotone or the migration flow is not homogeneous or the local deme sizes fluctuate substantially.

(iii) The question is frequently asked: Does a marked local change in gene frequency correspond to a sharp local change in selection intensity? Under the conditions of (ii), the maximum gene frequency differentiation (the maximum slope of the cline) in the presence of uniform local gene flow is attained at the neutral zone. If the neutral zone is extensive, then the change in gene frequency may still be significant without involving nearby any selection differences among the genotypes.

Extending the range on which allele A is advantageous, balanced on the other extreme by more demes advantageous to allele a, steepens the slope of the cline throughout the range (Results X, XI).

(iv) A more prolonged neutral zone tends to accent the selection forces operating on the outer regions of the popu-

lation distribution but flattens the slope over the population range (see Result IV).

The findings of (iii) and (iv) provide a possible explanation for the phenomenon of "area effects" (Goodhardt [1969]; Clarke [1966]).

(v) If the zone separating the regions favoring the A and a alleles has a selection expression of heterozygote advantage rather than being neutral, the corresponding cline is more flat in the center but steeper at the boundary between demes expressing directional selection as against overdominance. (See Result XIII on mixed heterotic ecotones.) On the other hand, underdominance in the central zone sharpens the cline slope at the center (Result XIV).

An implication attendant to (iii)-(v) is that a marked gene frequency change can be related to the following factors:

- (a) an ecotone with underdominant selection expression at the deme and neighboring demes at hand;
- (b) a long ecotone with the deme in question at the interface between the regions favoring A and a alleles respectively.
- (c) sharp differences in migration flow in reciprocal directions.

On the other hand, local irregularities in selection intensity within the confines of an ecotone is not a decisive factor in locating sharp morph differentiation.

(vi) The influence of increased migration flow is unambiguous and substantial. Indeed Result VI affirms that with two different rates of local migration exchange the spatial gene frequency pattern is always more homogeneous with larger migration or some state of fixation occurs. The different possibilities between very small as against moderate to uniform mixing migration rates is pronounced. In fact, for very

small migration flow a cline is established manifesting generally a steep slope between the two ranges of the population independent of the detailed character of the ecotone. For moderate up to uniform mixing inter deme exchange rates, the evolutionary outcome of the population becomes sensitive to the aggregate selection coefficients on the two extended environmental ranges of the A-favorable as against the a-favorable portions. However, for the special circumstance where the selection pattern is completely symmetric with respect to a center position, then an increased migration rate can never lead to fixation. Endler [1973] in his computer analysis of clines unfortunately restricted attention primarily to specific symmetric models and on this basis hastened to conclude that the rate of migration was of minor significance. Even in this case the cline systematically flattens with increasing migration levels.

(vii) When migration occurs at the infant stage (as in pollen and seed dispersal) vis a vis the adult stage, the effects due to selection are more pronounced in each environmental range (Result VIII).

(viii) Spreading an aggregate selection effect over more demes (or equivalently increased population subdivision) tends to bring about a flattened morphotone, that is, yielding more spatial homogeneity in gene frequency.

(ix) We have contrasted three different environmental ecotones with prescribed aggregate selection coefficients for each range (see paragraph G of Section 3), the steepest slope apparently occurs for environment III where the intensity of selection decreases with distance from the center, that is, where the change in selection between the - and + region is the most abrupt.

All the above results will be elaborated and extended in

Karlin and Richter-Dyn [1976b,c,d] to a background selection gradient involving more than two environments. We will also contrast the nature of the morphotones with an underlying regular selection gradient of heterozygote advantage.

The properties and characterizations described above and elaborated more precisely in the body of Sections 2 and 3 afford some means both qualitatively and quantitatively of discriminating better among cline structure relating to spatial patterns of selection and gene flow and help also to understand more thoroughly the effects of environmental parameters.

#### REFERENCES

- Beardmore, J. (1970). In: Essays in evolution and genetics, pp. 293-314. M.K. Hecht and W.C. Steere (eds.) Appleton-Century-Crofts, New York.
- Bryant, E.H. (1974). Am. Natur. 108: 1-19.
- Cain, A.J. and J.D. Curray. (1963). Phil. Trans. B. 246: 1-81.
- Cavalli-Sforza, L. and W. Bodmer. (1971). Human Population Genetics. Freeman, S.F.
- Christiansen, F. and O. Frydenberg. (1974). Genetics 77: 765-770.
- Clarke, B.C. (1956). Am. Natur. 100: 389-402.
- Ehrlich, P.R. and P.H. Raven. (1969). Science 165: 1228-1232.
- Endler, J.A. (1973). Science 179: 243-250.
- Fisher, R.A. (1950). Biometrics 6: 353-361.
- Fleming, W.H. and C.H. Su. (1974). Theor. Pop. Biol. 5: 431-449.
- Gillespie, J.H. and C.H. Langley. (1974). Genetics 76: 837-848.
- Goodhart, C.B. (1969). Heredity 18: 459-465.

- Haldane, J.B.S. (1948). Genetics 48: 277-284.
- Hamrick, J.L. and R.W. Allard. (1972). Proc. Nat. Acad. Sci. 69: 2100-2104.
- Hanson, W.A. (1966). Biometrics 22: 453-468.
- Jain, S.K. and A.D. Bradshaw. (1966). Heredity 21: 407-441.
- Karlin, S. (1972a). Am. Math. Monthly 79: 699-739.
- Karlin, S. and J. McGregor. (1972b). Theor. Pop. Biol. 3: 186-208.
- Karlin, S. (1976). This volume.
- Karlin, S. and N. Richter-Dyn. (1976a,b,c,d). To appear.
- Kettlewell, H.B.D. (1961). Heredity 21: 407-441.
- Kettlewell, H.B.D. and R.J. Berry. (1961). Heredity 16: 403-414.
- Kettlewell, H.B.D. and R.J. Berry. (1969). Heredity 24: 1.
- Kimura, M. and G.H. Weiss. (1964). Genetics 49: 561-576.
- Koehn, R.K. (1969). Science 163: 943-944.
- Kojima, K., P. Smouse, S. Yang, P. Nair and D. Brncic. (1972). Genetics 72: 721-731.
- Latter, B. (1973). Genetics 73: 147-157.
- Malécot, G. (1948, 1970). Les Mathématiques de L'Hérédité. Masson et Cie., Paris. Revised English translation. Freeman, S.F.
- Malécot, G. (1950). Ann. Univ. Lyon Sciences, Sec.A 13: 37-60.
- Malécot, G. (1951). Ann. Univ. Lyon Sciences, Sec.A 14: 79-118.
- Malécot, G. (1959). Publ. Inst. Stat. Univ. Paris, 8: 173-210.
- Malécot, G. (1967). Proc. Fifth Berkeley Symp. Math. Stat. Prob. IV: 317-332. Univ. of Calif. Press, Berkeley.
- Marshall, D.R. and R.W. Allard. (1972). Genetics 72: 721-731.
- Maruyama, T. (1972). Math. Biosciences 14: 325-335.
- Merritt, R.B. (1972). Am. Natur. 106: 173-185.
- Nevo, E. and Z. Bar. (1976). This volume.
- Rockwood-Sluss, E.S., J.S. Johnson and W.B. Heed. (1973). Genetics 73: 135-146.

- Slatkin, M. (1973). Genetics 75: 733-756.
- Weiss, G.K. and M. Kimura. (1965). J. Appl. Prob. 2:  
129-149.
- Wright, S. (1943). Genetics 23: 114-138.