GENETIC VARIATION IN HETEROGENEOUS ENVIRONMENTS

CHARLES E. TAYLOR

Department of Genetics, University of California, Riverside, California 92502

Manuscript received August 12, 1975
Revised copy received January 19, 1976

ABSTRACT

A model of population structure in heterogeneous environments is described and conditions sufficient for maintaining a polymorphism are derived.

The absolute fitness of any genotype is regarded as a function of location in the niche space and the population density at that location. Two modes of habitat selection are examined: (1) organisms are distributed uniformly over the environment; and (2) each organism selects to occupy that habitat in which it is most fit (“optimal habitant selection”).—Sufficient conditions for maintenance of genetic polymorphisms are derived for both models. In populations which do not practice habitat selection heterozygote superiority averaged over the environment is sufficient to guarantee the existence of polymorphisms. Comparable conditions for populations which practice optimal habitat selection are much less restrictive. If the heterozygotes are superior to one homozygote in any one part of the niche and to the other homozygote in any other part of the niche then a polymorphism will be defined.—A positive correlation between genetic and environmental variation follows from the model with habitat selection, but not from the other. The adaptive significance of polymorphisms thus depends on how animals behave.

Spatial variation in the environment may affect genetic variation in populations living there. Models of selection in heterogeneous environments have been studied by LEVENE (1953) and LEVINS (1968). According to these models genetic polymorphism is maintained by natural selection when the fitness of heterozygotes, averaged over the environment, is greater than the average of either homozygote. WALLACE (1968) has termed this condition “marginal overdominance”. (For LEVENE’s model the marginal fitness is the harmonic mean fitness and for LEVINS’ it is the arithmetic mean fitness.)

Habitat selection has important consequences for this conclusion. If genotypes differ in their fitnesses in the available habitats and if organisms practice habitat selection, so that they choose to occupy those habitats where most fit, then genetic polymorphism can be maintained under conditions much weaker than marginal overdominance. A mathematical model of this process is described below. Elsewhere I have shown that in this model polymorphisms can be maintained with very little genetic load (TAYLOR 1975). In this paper the circumstances under which genetic polymorphisms arise are derived with and without habitat selection. These circumstances are then compared and related to variation in the environment.

A set of conventions will first be developed for describing competition and population growth in spatially heterogeneous environments. Initially this description will apply only to populations composed entirely of one genotype.

It is assumed that the set of all environmental states can be described by a finite set of axes which comprise all environmental measurements. This set is called the *niche space*, $S$ (Hutchinson 1957). Each spot in the physical environment corresponds to a single point in this niche space.

The life cycle is assumed to consist of discrete generations and is visualized in the following way: at the beginning of each generation a number of zygotes is placed into an environment which has been characterized in a state free of organisms by associating each point in the environment with a point $x$ of the niche space. The organisms may redistribute themselves in the physical space according to any of several rules discussed below. At this point the organisms become sedentary; measurements of density, $n$, will refer to this stage. The adults then shed gametes which combine at random to produce zygotes for the next generation. All adults then die and a new generation begins.

The environmental states among which the organisms distribute themselves constitute the *niche* of that population. It is described by a measure, $\Phi(x)$, which indicates the "amount" of each $x$ available at the beginning of the generation cycle. The niche describes the environment in which the population is found. Each *habitat* is a further partitioning of the environment to describe where an individual may be found in the niche, and corresponds to a single $x$. The terms are used in this way so that the set of all environments in which an individual might spend its life (the niche) and that in which it actually does so (the habitat) are characterized. The niche of a trout population, for example, might correspond to the frequencies, $\Phi(x)$, of temperatures in the lake which that population inhabits, and the habitat of an individual would be the temperature, $x$, at which that individual was to be found.

The number of offspring produced by each adult is assumed to be uniquely specified by the density, $n$, at which the individual lived and the habitat, $x$, in which it matured. A function, $W(x, n)$, is therefore defined which denotes the number of zygotes in the next generation produced by a zygote living under conditions specified by $x$ and $n$ in this generation. In the terminology of Wright (1969), $W(x, n)$ is an "absolute fitness" and will be referred to simply as fitness below.

It will be assumed that $W(x, n)$ is continuously differentiable and strictly decreasing with density. That is,

$$\frac{\partial W(x, n)}{\partial n} < 0$$

whenever $W(x, n)$ is non-zero. Figure 1 shows an example of $W(x, n)$ in a one-dimensional niche space.
An important determinant of fitness is the manner in which organisms select their habitat. Two extreme modes of habitat choice can be differentiated: (1) no habitat selection; and (2) optimum habitat selection. For the model of no habitat selection it will be assumed that organisms are distributed uniformly so the density of organisms is constant throughout the niche. For the model of optimum habitat selection it will be assumed that each individual chooses to occupy that habitat where it is most fit. Though not immediately apparent, it will be shown that each $x$ is characterized by a unique density. When density is specified in this manner it will be labelled $n(x)$; when not so specified it will be symbolized by $n$. Which meaning is intended will be clear from context.

Note that optimum habitat selection implies that each organism selects that habitat in which its own fitness is optimal. This is not the same as if each individual were to choose a habitat in such a way that the mean population fitness is maximized.

Differences in fitness and distribution resulting from differences in habitat selection are evident from an illustration. Suppose $N$ individuals (e.g. fish) were placed into a niche (e.g. lake) characterized by $\Phi(x)$. In the absence of habitat selection the density, $n^*$, will be everywhere constant and equal to

$$n^* = N/\int \Phi(x) \, dx.$$ 

Clearly fitness will vary with $x$.

With optimal habitat selection, however, density may vary with $x$, but fitnesses will be constant. This is seen by a simple argument. It has been assumed that each individual chooses a habitat where it is most fit. If there were a non-zero variance in fitnesses some individuals would have gone to locations where they were not most fit—they could have improved their fitnesses by going elsewhere. This would be contrary to hypothesis. Therefore, the variance in fitness must be zero. Fitness must be constant, and the population density at each location must be such that $W(x, n(x))$ is a level surface, or isocline, of $W(x, n)$. So for a constant $c$, we have

$$W(x, n(x)) = c.$$ 

There are an infinity of such surfaces. Exactly which level surface, which $c$, can be specified by the simultaneous requirement that

$$N = \int \Phi(x) n(x) \, dx.$$ 

Of special interest is the population density at equilibrium, which will be labelled $k(x)$, defined implicitly by the equation $W(x, k(x)) = 1$. This is shown in Figure 1.

A monomorphic population at equilibrium with respect to population growth would have a density $k(x)$ and size $\int \Phi(x) k(x) \, dx$.

**Genetically Heterogeneous Population**

To this point it has been assumed that the population examined is genetically identical. Relaxing this assumption has important consequences for population
structure, because individuals with different genotypes will generally have different fitnesses at any location. Consequently when habitat selection is practiced the various genotypes will tend to select different habitats. It will be necessary to establish some nomenclature to examine this more closely.

The population will be regarded as very large, diploid, and sexually reproducing. This discussion will be concerned with two alleles at a single autosomal locus. The alleles will be assumed to be in Hardy-Weinberg equilibrium as zygotes, and it is as zygotes that they are enumerated and habitat selection occurs. Natural selection will be assumed to act as though due to differences in viabilities.

The expected number of offspring from a zygote of the $ij$th genotype ($A_iA_j$) at a location of type $x$ can be represented by $W_{ij}(x, n)$. The density of $A_iA_j$ individuals at $x$ will be termed $n_{ij}(x)$, and the total number of individuals of this genotype, $N_{ij}$, is given by

$$N_{ij} = \int_a^b \Phi(x) n_{ij}(x) \, dx.$$ 

In addition, the following nomenclature will be used:

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Frequency</th>
<th>Fitness</th>
<th>Density</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_1A_1$</td>
<td>$p^2 = N_{11}/N$</td>
<td>$W_{11}(x, n)$</td>
<td>$n_{11}(x, n)$</td>
</tr>
<tr>
<td>$A_1A_2$</td>
<td>$2p(1-p) = N_{12}/N$</td>
<td>$W_{12}(x, n)$</td>
<td>$n_{12}(x, n)$</td>
</tr>
<tr>
<td>$A_2A_2$</td>
<td>$(1-p)^2 = N_{22}/N$</td>
<td>$W_{22}(x, n)$</td>
<td>$n_{22}(x, n)$</td>
</tr>
<tr>
<td>$N = \sum_{ij} N_{ij}$</td>
<td></td>
<td></td>
<td>$n(x) = \sum_{ij} n_{ij}(x)$</td>
</tr>
</tbody>
</table>
Selection in heterogeneous environments

With the model of no habitat selection \( n(x) \) takes much the same form as \( n(x) \) in the discussion of homogeneous populations. Each genotype is regarded as distributed uniformly throughout the environment. The density of each genotype, \( n(x) \), is equal to \( N_{11}/N \) for all \( x \), so that the gene frequency and density are constant and equal everywhere in the environment.

With optimum habitat selection, where each individual lives in that part of the niche where its fitness is greatest, the distribution is complicated, and an analytic expression for \( n(x, n) \) does not seem possible. It is possible, however, to arrive at several properties of distributions of such organisms, two of which are: (1) The density of individuals at each \( x \) is unique, so that it makes sense to speak of \( n(x) \); and (2) All individuals of any one genotype are equally fit. This follows from the same argument presented for genetically homogeneous populations.

Lacking an expression for the distribution of organisms it is evident that a complete analytic characterization of \( N \) and \( p \) through time is quite out of hand. Nonetheless, the small bit of information described above is sufficient for a moderately detailed examination of genetic polymorphisms. Within the foregoing models the conditions for polymorphisms to exist may now be derived.

Genetic Polymorphism

A genetic polymorphism is "the occurrence together, in the same locality, of two or more discontinuous forms of a species in such proportions that the rarest of them cannot be maintained merely by recurrent mutation" (Ford 1972). Two types of genetic polymorphisms need to be distinguished: "transient polymorphisms", which occur when one form is replacing another in the evolutionary process, and "balanced polymorphisms", which occur when natural selection acts in such a way that any particular form will be selected for when it becomes sufficiently rare. As used here, the term genetic polymorphism will refer solely to the balanced variety.

A balanced polymorphism will be maintained if the frequency of one allele, say \( A_1 \), is increased when the other allele, \( A_2 \), is almost fixed and the frequency of \( A_1 \) is decreased when almost fixed itself. Heterozygote superiority only at the equilibrium population sizes where the \( A_1 \) allele is almost fixed and when the \( A_2 \) allele is almost fixed would thus constitute sufficient condition for the continued maintenance of a genetic polymorphism.

For populations which practice optimal habitat selection these conditions are quite weak. Looking first at an equilibrium population where the \( A_1 \) allele is fixed, the fitness of each individual must equal 1, and so the population density at each \( x \), \( k_{11}(x) \), will be determined by the level surface \( W_{11}(x, k_{11}(x)) = 1 \). The level surfaces \( W_{12}(x, k_{12}(x)) = 1 \) and \( W_{22}(x, k_{22}(x)) = 1 \) similarly defined the densities \( k_{12}(x) \) and \( k_{22}(x) \). These \( k_{ij}(x) \) are analogous to the "carrying capacity" of unit areas of any point in the niche space.

Now, for populations in Hardy-Weinberg equilibrium and near fixation the less frequent allele will be present almost entirely in the heterozygotes; its presence in homozygotes can be ignored. If and only if there exists an \( x \), such
that \( k_{12}(x_1) > k_{11}(x_1) \) will the rare \( A_1A_2 \) heterozygote be more fit than the predominant \( A_1A_1 \) homozygote at the equilibrium size dictated by fixation of the \( A_1 \) allele. Similarly, if and only if there exists an \( x_2 \) such that \( k_{12}(x_2) > k_{22}(x_2) \) would a rare heterozygote be more fit than the homozygote at the population density which represents an equilibrium for a population in which the alternate allele, \( A_2 \), is fixed. Putting this together, conditions for the maintenance of a polymorphism can be stated in the following manner:

If there exist points \( x_1 \) and \( x_2 \) in niche space with \( \Phi(x) \) continuous and \( > 0 \) at both \( x_1 \) and \( x_2 \), with

\[
\begin{align*}
(a) & \quad k_{12}(x_1) > k_{11}(x_1) \\
(b) & \quad k_{12}(x_2) > k_{22}(x_2)
\end{align*}
\]

then a genetic polymorphism will be maintained in a population with optimal habitat selection.

These conditions \( (HS) \) for populations practicing habitat selection may be made a bit stronger so that an important special case is also shown to maintain polymorphisms:

If \( k_{12}(x) = k_{11}(x) \) for all \( x \) for which \( \Phi(x) \) is non-zero then \( (a) \) in \( (HS) \) may be replaced by

\[
(a') \quad k_{22}(x_1) > k_{11}(x_1)
\]

Similarly, if \( k_{12}(x) = k_{22}(x) \) for all \( x \) for which \( \Phi(x) \) is non-zero, then \( (b) \) in \( (HS) \) may be replaced by

\[
(b') \quad k_{11}(x_2) > k_{22}(x_2)
\]

This follows from the same reasoning which gives \( (HS) \).

These conditions, \( (HS) \) and \( (HS') \), are quite weak. They state that in order for a polymorphism to be maintained it is necessary that the carrying capacity function of the heterozygotes be superior to that of one homozygote in just a single habitat and superior to the carrying capacity function of the other homozygote in only one other habitat.

Comparable conditions for a population which does not exhibit this type of habitat selection are much more restrictive. A polymorphism will be maintained in this model when the mean fitness of the heterozygote, averaged over the niche, is greater than that of either homozygote at the density where fixation for that homozygote is approached. That is to say, where \( \hat{N}_{11} \) and \( \hat{N}_{22} \) are the equilibrium population sizes when \( A_1 \) or \( A_2 \) respectively is fixed:

Sufficient conditions for the maintenance of a genetic polymorphism are that

\[
\begin{align*}
(a) & \quad \int \Phi(x) \left[ W_{12}(x, \hat{N}_{11}) - W_{11}(x, \hat{N}_{11}) \right] dx > 0 \\
(b) & \quad \int \Phi(x) \left[ W_{12}(x, \hat{N}_{11}) - W_{22}(x, \hat{N}_{22}) \right] dx > 0
\end{align*}
\]

for populations which are distributed uniformly over the environment.
As with populations exhibiting habitat selection, these conditions for no habitat selection (NHS) may be extended to include complete dominance:

If \( W_{12}(x, N_{11}) = W_{11}(x, \hat{N}_{11}) \) then (a) in NHS may be replaced by

\[ (a') \int \Phi(x) \left[ W_{22}(x, \hat{N}_{22}) - W_{11}(x, \hat{N}_{11}) \right] dx > 0 \]

And if \( W_{12}(x, \hat{N}_{22}) = W_{22}(x, \hat{N}_{12}) \) then (NHS')

\[ (b') \int \Phi(x) \left[ \hat{W}_{11}(x, \hat{N}_{22}) - \hat{W}_{22}(x, \hat{N}_{22}) \right] dx > 0 \]

clearly replace (b) in (NHS).

Clearly (NHS) implies (HS) while the converse is not necessarily true. Therefore, the conditions for a genetic polymorphism in a population practicing optimal habitat selection are in some sense more general than in those which do not.

It should be remarked that these conditions for genetic polymorphisms are merely sufficient, but not necessary. It might be possible for an allele to be disadvantaged when either too rare or too common, and yet to be favored when at an intermediate frequency. But with the provision that any polymorphism must be able to originate as an isolated mutation in a population which is fixed for either allele in an otherwise similar genetic and environmental milieu, then the conditions for the maintenance of a polymorphism discussed above are necessary as well as sufficient.

Environmental heterogeneity refers here to the size of the set of habitats in the niche. The set of habitats which are actually realized, the set of all \( x \) with \( \Phi(x) > 0, \) is larger in a more heterogeneous environment. Consequently, heterogeneous environments will more likely (though need not) contain points \( x_i \) and \( x_z \) meeting the conditions of HS and HS' than will a homogeneous environment. Thus there should be a correspondence between habitat and genetic diversity for populations which practice habitat selection, variable environments supporting more polymorphisms.

This would not seem to follow in the absence of such behavior. For a polymorphism without habitat selection the mean fitness of the heterozygote must be superior to that of the homozygotes (NHS). All things being equal, merely averaging across a larger range of \( x \) would confer no advantage to the heterozygote. If heterozygotes tended to superiority at the environmental extremes then increasing environmental diversity, which means averaging would occur across more extreme areas, might also tend to increased genetic variation.

One test of the predicted association between genetic and environmental variation has been performed by J. Powell (1971) in the laboratory. A number of population cages of Drosophila carry allozyme variation at several loci were established; some of these cages were made environmentally heterogeneous while others were not. After 45 weeks an average of 2.1 alleles per locus were still segregating in the cages with heterogeneous environments while there were only 1.7 alleles per locus in the homogeneous cages. This difference is statistically significant. Such experiments have subsequently been refined by McDonald and Ayala (1974) with similar results. While habitat selection was not neces-
sarily involved in these experiments, the observations do seem to support the model described above.

Part of this work was supported by NSF grant BMS75-13544. I thank J. W. Helton, J. M. Emlen, Jr., J. B. Mitton and J. Roughgarden for their advice and criticism of these ideas.

LITERATURE CITED


Corresponding editor: R. C. Lewontin