MIGRATIONAL SELECTION

A. W. F. EDWARDS*

*International Laboratory of Genetics and Biophysics (Pavia Section), Naples
at Istituto di Genetica, University of Pavia, Italy*

Received 2.vi.62

1. INTRODUCTION

THODAY (1959) mentions no less than eight adjectives which have recently been used to describe different types of selection (disruptive, centrifugal, stabilising, centripetal, directional, cyclic, normalising and canalising) and has commented (1960) that "the terminology of selection is becoming confusing". The use of a ninth adjective, *migrational*, therefore needs some justification. The phrase *migrational selection* was used, if not coined, by Fisher (1930) to describe selection by "the means of dispersal alone, without the necessity for selective discrimination"; it is thus older than most of the modern terminology, and is applied to a very different phenomenon, that of selection without differential fitness.

It is, however, related to Mather's *disruptive selection* (1953, 1955), which Thoday (1960) defines as "selection away from the mean in both directions in the same population". Mather considers that disruptive selection can lead to either stable polymorphism or complete fission of a population, and hence to speciation. These concepts are, of course, older than the term disruptive selection itself. Fisher (1930) considered the case of a "species subjected to different conditions of survival and reproduction at opposite ends of its geographical range", having first stated that he did not know of any circumstance "which, in the genetical situation produced, differs essentially from geographical distance, in terms of which, therefore, it is convenient to develop the theory". He concluded that such selection could lead to a clinal polymorphism: "It involves no tendency to break the stream of diffusion, or consequently to diminish in degree the unity of ancestry which the species possesses. It is analogous to the stretching of a material body under stress, not to its rupture".† He added that "the process so far sketched contains no novel features".

Fisher then described how fission might follow from the polymorphic state: "The constant elimination in each extreme region of the genes which diffuse to it from the other, must involve incidentally the elimination of those types of individual which are most apt so to diffuse". In concluding his section on this type of selection, he considered the "power of the means of dispersal alone", or "migrational selection", a theory which he ascribed to Lankester (1925).

The concept is, of course, simplicity itself. If a habitat is heterogeneous, and occupied by a unified population of which some genotypes prefer one location, and tend to migrate to it, and others prefer

* This work is partially supported by a grant from the U.S.A.E.C.

† Fisher (1950) has since given a mathematical formulation of this problem.

another location, then a cline in gene frequencies will develop. But it must at once be emphasised that migrational selection is by itself incapable of changing the overall gene frequencies of a closed population; it is solely concerned with the movement of genes, not with their elimination. When, therefore, we speak of migrational selection leading to a stable distribution of gene frequencies, it should be understood that the overall gene frequency must be stabilised by other means. In general, we expect migrational selection to increase the genetic variance, and, where there is dominance, to change the mean of a population.

2. MATHEMATICAL FORMULATION

The natural study of this phenomenon would be eased by the creation of a mathematical model, but unfortunately the problem is rather intractable formally. In essence we are dealing with a split population with a certain pattern of migration connecting the two halves, the situation in each half being governed by a relation not necessarily algebraically linear. Rather than obtaining a solution by making many approximations it would seem better to programme the model directly onto an electronic computer.

However, some idea of the expected course of events may be had by supposing that the overall gene frequency at the locus being considered (which, as we have noted, will be constant) is one-half.

Let us take the case of two alleles, A and a , in a population spread over two localities, 1 and 2. Suppose that the genotype AA prefers locality 1, and that a proportion x of the AA 's born in 2 migrates to 1, and, similarly, that a proportion x of the aa 's born in 1 migrates to 2. The heterozygotes have no preference. In addition, suppose that a proportion y of the individuals born in one location migrates to the other, regardless of genotype. If $y = \frac{1}{2}$ then mating would be at random except for the preferential migration.

A simple matrix notation (which has been used in the study of human migration by, for example, Prais, 1955) may then be used to describe the situation. The *migration matrix* for each genotype is given in the central column of table 1. Letting the relative frequency of the A gene in locality 1 be p_1 , and in locality 2 be p_2 , the genotype frequencies in the two localities will be as given in the first two columns of the table, assuming random mating within each half of the population. In the last two columns are given the genotype frequencies immediately following a migration, which is supposed to take place each generation.

Evaluating the gene frequency in locality 1 in the next generation, we find

$$p_1' = \frac{p_1 + xp_2^2 - y(p_1 - p_2)}{1 + x(p_2^2 - q_1^2)},$$

and in locality 2 a similar expression for p_2' . These two equations fully

describe the situation without approximation; but they are simultaneous non-linear difference equations, and to make the problem tractable we now introduce our simplification, that $p_1 + p_2 = 1$. This ensures that the system is symmetrical between the two halves of the population, or that $p_1 = q_2$ and $q_1 = p_2$. Thus, dropping the subscripts and dealing with the situation in locality 1,

$$p' = p + xq^2 - y(p - q).$$

This equation describes the situation clearly: in locality 1 the preferential migration (xq^2) is always trying to increase the gene frequency, whilst the non-preferential migration ($y(p - q)$) is always trying to return it to one-half. There will be equilibrium (in the restricted sense mentioned above, and subject to a stability condition,

TABLE 1

Derivation of the genotype frequencies in the next generation

Genotype	Locality 1 2	Migration matrix *	Locality 1	Locality 2
AA	$(p_1^2 \quad p_2^2)$	$\begin{pmatrix} 1-y & y \\ x+y & 1-x-y \end{pmatrix}$	$(p_1^2(1-y) + p_2^2(x+y))$	$p_1^2y + p_2^2(1-x-y)$
Aa	$(2p_1q_1 \quad 2p_2q_2)$	$\begin{pmatrix} 1-y & y \\ y & 1-y \end{pmatrix}$	$(2p_1q_1(1-y) + 2p_2q_2y)$	$2p_1q_1y + 2p_2q_2(1-y)$
aa	$(q_1^2 \quad q_2^2)$	$\begin{pmatrix} 1-x-y & x+y \\ y & 1-y \end{pmatrix}$	$(q_1^2(1-x-y) + q_2^2y)$	$q_1^2(x+y) + q_2^2(1-y)$

* The element in the i th row and j th column of a migration matrix gives the probability of an individual migrating from i to j .

which is discussed below) when $xq^2 = y(p - q)$, or when

$$p = \frac{x+y - \sqrt{y(x+y)}}{x}.$$

If x is small compared with y then

$$p = \frac{1}{2} + x/8y,$$

and when the non-preferential part of the migration is entirely random, and $y = \frac{1}{2}$,

$$p = \frac{1}{2} + x/4.$$

3. THE APPROACH TO EQUILIBRIUM

The approach to equilibrium may be demonstrated simply by graphical methods. We have, exactly,

$$\Delta p = xp^2 - 2(x+y)p + (x+y).^\dagger$$

† Selection is usually so fast on this model that the corresponding differential equation, which has solution $t = \frac{1}{2\sqrt{y(x+y)}} \log \frac{xp - (x+y) - \sqrt{y(x+y)}}{xp - (x+y) + \sqrt{y(x+y)}} + c$, is of little general application.

If we plot Δp as a function of p , then the change in gene frequency in each generation will be the height of the curve at the relevant value of p ; laying off this distance horizontally (on the down-hill side of the curve if Δp is positive) will give the new gene frequency, and subsequently the next value of Δp . The result is a series of steps, each step representing one generation. This has been done in fig. 1 for $x = \frac{1}{10}$, $y = \frac{1}{4}$. It will be seen that the rate of approach to equilibrium is rather fast under these conditions.

FIG. 1.—The approach to equilibrium.

If the gradient of the curve exceeds 1 in absolute value in the neighbourhood of the equilibrium point it is easy to see that the population will, literally, overstep the point, and one or more oscillations will ensue. The gradient is given by $2xp - 2(x+y)$, which, on putting in the equilibrium value for p , becomes $-2\sqrt{y(x+y)}$. The condition for oscillation near the equilibrium is therefore $2\sqrt{y(x+y)} > 1$, or $y(x+y) > \frac{1}{4}$. Even if $y(x+y)$ is a little less than $\frac{1}{4}$, since the curve is concave upwards the population will nearly always approach equilibrium from a high value of p , overstepping the equilibrium point if the initial value of p is less than the equilibrium value. Furthermore, if the gradient at the equilibrium point exceeds 2 in absolute value, or $y(x+y) > 1$, it can be seen that the equilibrium is unstable, and that a population will oscillate on either side of it. But this cannot happen unless y is considerably greater than one half, and so need not concern us, for in that case more than half of the entire population would be changing locality at each migration.* It is, however, important to note that when $y = \frac{1}{2}$ one or more oscillations may occur, even though the equilibrium is stable. As an example of a rather novel way of approaching equilibrium, the reader may like to consider the course of events for $x = \frac{1}{3}$, $y = \frac{4}{15}$, starting with $p = 0$.

* In fact the inequality can never hold.

The gradient of the curve at the equilibrium point is also the measure of the rate of approach to equilibrium; the nearer to unity, the faster the rate of approach.

4. EFFECT ON THE VARIANCE

Since we are supposing that the heterozygote undergoes no preferential migration, we are limited by consistency to considering a metrical character without dominance. If AA has value $+d$ and aa $-d$, then a gene frequency p gives rise to a mean measure of $d(p-q)$, and the change in the mean of each half of the population may be found simply. The mean of the whole population will, on our assumptions, remain zero, but the variance will increase under selection.

Only the homozygotes contribute to the variance, and they each occur in the whole population with relative frequency $\frac{1}{2}(p^2+q^2)$. The genetic variance is therefore $(p^2+q^2)d^2$, which at equilibrium assumes the approximate value

$$\left[\frac{1}{2} + \frac{1}{3} \frac{1}{2} \left(\frac{x}{y} \right)^2 \right] d^2,$$

x being supposed small compared with y . From this it can be seen that there is a small increase in the variance over the normal value of $\frac{1}{2}d^2$.

5. DISCUSSION

In reviving the concept of migrational selection it is necessary to remember that, taken in isolation, it is a very naïve hypothesis: it simply states that, if genotypes migrate differentially, a cline will develop. It cannot by itself maintain a polymorphism, or act as an agent of directional selection, or lead to the fission of a population. Nevertheless, as one of the components of natural selection (and, as Fisher pointed out, in some cases no doubt one of the major components) it is worthy of separate study. It represents, in a sense, an alternative strategy to ordinary disruptive selection: the individual migrates to a more suitable environment rather than suffer the disadvantages of his birth-place. Looked at in this light it is clear that migrational selection requires differential fitness if only as a spur to migration, and can only be an isolated phenomenon if the individual migrates for some reason other than improving his fitness (whether consciously or not).

The phenomenon may thus come into its own in human populations, where migration takes place for many reasons, not all of them connected with fitness. Local variation in gene frequencies may be due largely to differential migration and, furthermore, human populations offer a unique opportunity for studying migrational selection owing to the fact that they are the only type of natural animal population whose migrations can be followed exactly.

Although some of our results, notably the expression for the

genetic variance, are only valid when the rate of preferential migration, x , is small compared with the rate of non-preferential migration, y , the main part of the analysis holds for all values of y , in particular for small values appropriate to the case of near-isolation of the two sub-populations. In this case the cline can become very marked, and the genetic variance correspondingly inflated. In the extreme case of $y = 0$, selection slowly proceeds until the two alleles are completely segregated.

The next step in this treatment is to introduce differential elimination of the alleles in the two localities, when the model should then correspond to disruptive selection maintaining a polymorphism: it would be a discontinuous version of Fisher's (1950) differential model of disruptive selection, although of more limited application owing to its imposed symmetry. The introduction of a locus influencing migration itself would then enable the fission of species to take place.

On the practical side it is to be hoped that further selection experiments such as that of Thoday and Gibson (1962) will be undertaken with a wide variety of species, and under conditions as near to nature as possible.

6. SUMMARY

Fisher's *migrational selection* is discussed in relation to *disruptive selection*, and a mathematical example is given to show how an equilibrium gene distribution may result from it, and how it will augment the genetic variance of a character. The approach to equilibrium is explained by a graphical method. It is concluded that migrational selection may be important in human populations, for some of which, fortunately, data on migration abound.

Acknowledgment.—I am indebted to the organisers of the *Entretiens de Monaco en Sciences Humaines* for allowing me to attend their session on The Measurement of Human Migration, which led to my interest in this problem.

7. REFERENCES

- FISHER, R. A. 1930. *The Genetical Theory of Natural Selection*. Oxford University Press.
- FISHER, R. A. 1950. Gene frequencies in a cline determined by selection and diffusion. *Biometrics*, 6, 353-361.
- LANKESTER, R. 1925. The blindness of cave animals. *Nature*, 116, 745 (quoted by Fisher, 1930).
- MATHER, K. 1953. The genetical structure of populations. *Symp. Soc. Exp. Biol.*, 7, 66-95.
- MATHER, K. 1955. Polymorphism as an outcome of disruptive selection. *Evolution*, 9, 52-61.
- PRAS, S. J. 1955. Measuring social mobility. *J. Roy. Stat. Soc. A*, 118, 56-66.
- THODAY, J. M. 1959. Effects of disruptive selection. I. Genetic flexibility. *Heredity*, 13, 187-203.
- THODAY, J. M. 1960. Effects of disruptive selection. III. Coupling and repulsion. *Heredity*, 14, 35-49.
- THODAY, J. M., AND GIBSON, J. B. 1962. Isolation by disruptive selection. *Nature*, 193, 1164-1166.