

A MODEL OF SELECTIVE MIGRATION

R. A. KEMPTON

Rothamsted Experimental Station

Received 3.x.73

SUMMARY

Individuals migrating between populations are normally assumed to be drawn at random from their base populations; migration is then a strong unifying force between the populations. Phenotypic assortment of migrants could however cause populations to diverge. A model is formulated to describe the effects of such selective migration on a metric character, when migration is only partially correlated with character score. When the character is governed by a large number of loci, it is shown that the amount of genetic divergence depends only on the heritability of the character, the initial migration rate and the degree of assortment of the migrants. Predictions made from the model correspond closely with experimental results from other sources.

1. INTRODUCTION

THAT genetic divergence, and ultimately speciation, can occur between populations that are not geographically isolated is still disputed. Although there is now a considerable mass of experimental evidence (see Thoday, 1972, for review) which shows that disruptive selection can cause divergence between populations even when there is substantial gene flow between them, the relevance of these results to natural populations has been questioned (Mayr, 1963) mainly because of the high selection intensities involved. Indeed Maynard Smith (1966) has shown theoretically that divergence to a stable polymorphism—which may in some cases be followed by eventual reproductive isolation of the morphs—is unlikely, with random interbreeding, unless the selection intensities are large. Thus, if sympatric speciation does occur in nature, it seems likely that in many cases other isolating mechanisms, acting in association with disruptive selection, allow divergence to take place with lower selection intensities.

The tendency of some species to breed in their own natal habitat will clearly reduce migration between populations and increase the likelihood of divergence. Such habitat selection is normally assumed to be conditioned at upbringing, and those individuals that migrate are supposed to be a random sample of the resident population. The possible evolutionary importance of *selective* migration, where the probability of migration depends on individual phenotype, was, however, suggested by Fisher (1930) and more recently emphasised by Thoday (1972). Evidence of such phenotype-dependent habitat choice in natural animal populations is scant because of the difficulties in studying migration, but for human populations examples of selective migration are abundant, particularly in the field of social mobility (Cliquet, 1968).

It is clear that selective mobility for heritable characters will result in differential gene flow and eventual genetic divergence of populations.

Migration is normally looked upon as a unifying force between populations, but if the migration is selective it may be a powerful factor promoting divergence.

Thoday and Gibson (1970) have demonstrated this experimentally with forced selective migration between two populations of *Drosophila*. Mobility was determined solely by individual sternopleural chaeta number. The rapid divergence observed between the two groups was mainly genetic in character, and accompanied by a reduction in within-group variance and proportional exchange. The results are described more fully later.

Edwards (1963) and Parsons (1963) have presented mathematical formulations of the problem for the case where selection of migrants is based on genotype, and Kempton (1971) considered the case of a metric character where migrants are determined by truncation of the phenotypic distribution. These models have not, however, proved sufficiently realistic to be fitted to observational data, mainly because the migration in the model is fully determined by a single character. The model is here extended to admit several characters by considering the distribution of migration threshold values for each character. Theoretical predictions of the behaviour of populations exposed to selective migration then compare well with Thoday and Gibson's experimental results, and it is suggested that in some cases the model may be useful for estimating the genetical component of phenotypic difference between populations.

2. THE MODEL

(i) *The threshold distribution*

An individual's motivation to migrate from his birth place will be influenced by many factors, both physiological and psychological. Consider a single such factor. For each individual there will be a critical value of this factor, the threshold value, above which migration will occur, and below which it will not occur. The threshold value will vary between individuals in the population, due to the influence of the individual's scores for other factors which may affect prospects of migration. Suppose that the distribution of threshold values for this character in the population has the density function $f(x)$, then the probability of an individual with character score C migrating, $F(C) = \text{Prob}[x < C] = \int_{-\infty}^C f(x) dx$. If the distribution of character scores in the population is $\phi(x)$, then the phenotypic distribution of the migrants is $\phi(C)F(C)/P$, where $P = \int \phi(C)F(C)dC$ is the proportion of the population migrating.

To examine the long-term effects on the population, it is first necessary to make some assumptions about the genetic basis of the character. Assume therefore that the character is governed by gene effects of the same order of magnitude at L unlinked loci, where L is assumed large enough for terms of order L^{-1} in the analysis to be ignored. The phenotypic distribution of character, $\phi(x)$, is assumed to be normal, with cumulative distribution $\Phi(x)$, and all measurements are made from the mean and expressed in terms of the phenotypic standard deviation, σ .

Consider the A - a locus. Suppose the three genotypes AA , Aa , aa have

frequencies g_1, g_2, g_3 and average effects d_1, d_2, d_3 . The proportion of each genotype migrating is given by

$$P_i = g_i \int_{-\infty}^{\infty} \phi(x-d_i)F(x)dx \quad (i = 1, 2, 3)$$

$$= g_i \int_{-\infty}^{\infty} [\phi(x) - d_i\phi'(x)]F(x)dx + O(L^{-1}),$$

since all gene effects are of order $L^{-\frac{1}{2}}$. If the threshold distribution, $f(x)$, is normal with mean m , variance s^2 , the expression simplifies to

$$P_i = g_i[P + d_i z] \tag{1}$$

where P is the proportion of the population migrating,

$$P = 1 - \Phi\left(\frac{m}{\sqrt{(s^2+1)}}\right),$$

and z is the ordinate of the normal distribution, variance $1 + s^2$, at $x = m$,

$$z = \frac{1}{\sqrt{(s^2+1)}} \phi\left(\frac{m}{\sqrt{(s^2+1)}}\right).$$

If the genotypic frequencies are in Hardy-Weinberg equilibrium, and gene A has frequency q , then its frequency in the migrant population is $P_1 + \frac{1}{2}P_2$, that is

$$q' = q + \alpha q(1-q) \frac{z}{P} \tag{2}$$

where $\alpha = q(d_1 - d_2) + (1-q)(d_2 - d_3)$ is the average effect of gene substitution.

Note that of the two parameters of the threshold distribution, the mean, m , mainly determines the amount of migration from the population, while the standard deviation, s , specifies the degree of assortment of the migrant population. Keeping the migration rate, P , constant in equation (2), the change in gene frequency, $\tilde{q} \propto (1 + s^2)^{-\frac{1}{2}}$. When $s = 0$, assortment is strongest and migration is determined by this character alone. As $s \rightarrow \infty$ the migrants become a random sample from population.

(ii) *Divergence of two populations*

Consider the following model. Suppose there are two large populations of the same constant size and initially genetically homogeneous. In each generation there is selective migration between the populations followed by mating at random within each population. Selective migration acts in opposite directions in the two classes, selection being for high character scores for migration from population 2 to population 1, and for low character scores in the reverse direction. The two distributions of threshold values are assumed to be normal with common variance, s^2 , and mean, m for population 2, $-m$ for population 1. The pattern of migration is illustrated in fig. 1, when the populations have already started to diverge.

Consider the change in gene frequency at the A - a locus. Initially gene A has frequency q_0 and average effect α in both populations and the phenotypic distribution of the character is normal with zero mean and unit variance. In the absence of other selective forces, the overall frequency of

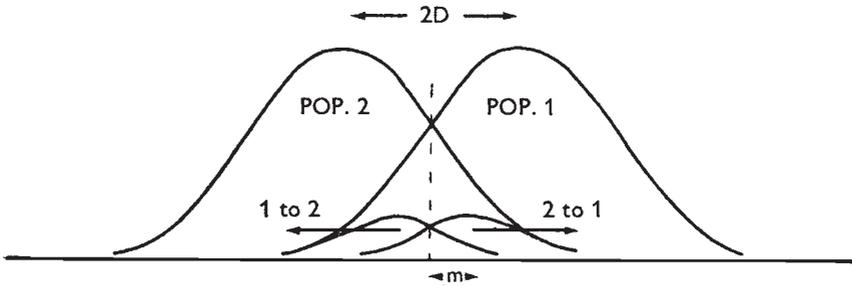


FIG. 1.—Pattern of selective migration for two populations showing the phenotypic distribution of the migrants compared with the populations overall.

A will remain constant with time, and the gene frequencies in the two populations at a later time t may thus be taken as $q_o \pm q_t$. The magnitude of the change in mean phenotypic value of both populations is then

$$D_t = \Sigma(2\alpha\tilde{q}_t + O(\tilde{q}_t^2)),$$

where the summation is over all loci governing the character. \tilde{q}_t is thus of order $L^{-\frac{1}{2}}$, and the change in within-group phenotypic variance,

$$\Delta\sigma^2 = \Sigma 2(1-2q)\alpha^2\tilde{q}_t,$$

is also of this order. Furthermore, to first order, changes in mean phenotypic value are equal and opposite for the two populations, even when there is dominance.

Now if the gene frequencies of the two populations are $q_{1,t}$, $q_{2,t}$ at time t , the gene frequencies in the next generation are given, from equation (2), by

$$q_{1,t+1} = (1-P_t)\left(q_{1,t} + \frac{z_t\alpha}{1-P_t} q_{1,t}(1-q_{1,t})\right) + P_t\left(q_{2,t} + \frac{z_t\alpha}{P_t} q_{2,t}(1-q_{2,t})\right) \quad (3)$$

and $q_{2,t+1}$ similarly, where $q_{2,t} = 2q_o - q_{1,t}$. P_t which is the migration rate at time t , and z_t are given by equation (1) with m replaced by $m + D_t$. It follows that writing $\tilde{q}_t = q_{1,t} - q_o$

$$\tilde{q}_{t+1} = (1-2P_t)\tilde{q}_t + 2z_t\alpha q_o(1-q_o). \quad (4)$$

Multiplying both sides of this equation by 2α and summing over all loci, the mean phenotypic value of population 1 at time $t+1$ is

$$D_{t+1} = (1-2P_t)D_t + 2z_t h^2, \quad (5)$$

where h^2 is the initial heritability of the character. D_{t+1} will be the magnitude of the genetic divergence in mean phenotypic value of the offspring in both populations at generation $t+1$, assuming random mating. The genetic divergence of the parent populations, which will not be in Hardy-Weinberg equilibrium following migration, can be shown to exceed D_{t+1} by an amount $2 \times$ (non-additive genetic variance).

Equations (4) and (5) can be applied recursively to trace the genetic divergence of the two populations for different values of m and s (fig. 2). In all cases the equilibrium state is rapidly achieved. Divergence is seen

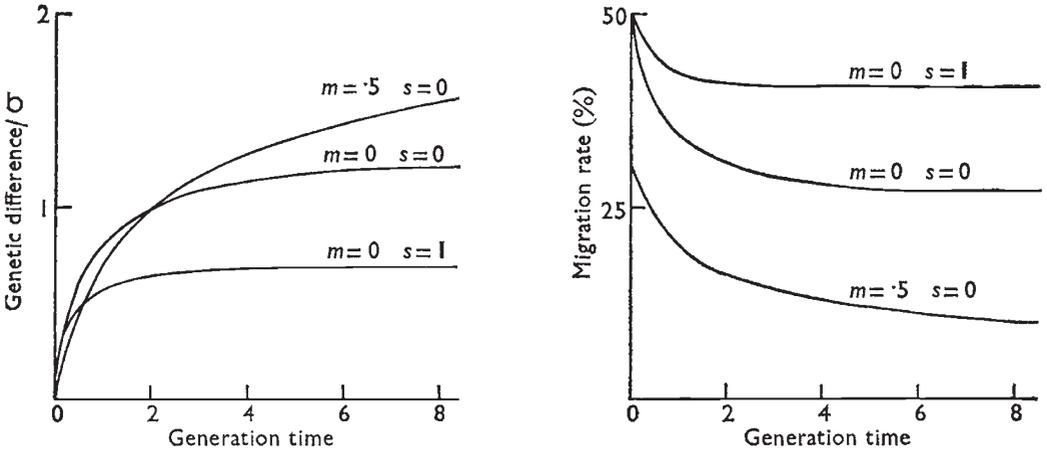


FIG. 2.—Genetic divergence and reduction in inter-group mobility for two populations with selective migration. The genetic difference is measured in terms of the within-class phenotypic standard deviation, σ .

to increase with increased phenotypic assortment of migrants, measured by s . Although initial divergence is greatest for $m = 0$, the final difference between the populations increases with m . Positive values of m in the model would suggest additional constraints on migration, possibly due to environmental differences between the populations. A value of $m = 0$, implies an individual's destination is related to genotype and is not directly influenced by birthplace.

When the equilibrium state is reached, $q_{t+1} = q_t = q_e$, $D_{t+1} = D_t = D_e$ and thus

$$q_e = \frac{\alpha z_e}{P_e} q_0(1 - q_0) \tag{6}$$

$$D_e = \frac{z_e}{P_e} h^2, \tag{7}$$

where $z_e = \frac{1}{\sqrt{(s^2+1)}} \phi \left(\frac{D_e+m}{\sqrt{(s^2+1)}} \right)$,

and $P_e = 1 - \Phi \left(\frac{D_e+m}{\sqrt{(s^2+1)}} \right)$.

For given mean, m , and variance, s^2 , of the threshold distribution, equations (6) and (7) can be solved iteratively for q_e and D_e .

The within-group genetic variance at equilibrium will be reduced by an amount $\Delta\sigma_G^2 = \Sigma 2(1-2q)\alpha^2 q_e$ from its initial value. For the case of no dominance and equal additive gene effects, a , at all loci, this becomes

$$\sigma_G^2 = 2La^3 \frac{z_e}{P_e} (\mu_1 - 3\mu_2 + 2\mu_3) \tag{8}$$

where μ_1, μ_2, μ_3 are the first three non-central moments of the gene frequency distribution. In particular, if the distribution over all loci is initially symmetric about $q = \frac{1}{2}$, then equation (8) is identically zero, and the change in within-group genetic variance is of order L^{-1} .

3. COMPARISON WITH THODAY AND GIBSON'S EXPERIMENTAL RESULTS

Thoday and Gibson's model experiment with *Drosophila* (Thoday and Gibson, 1970) closely resembles the mathematical model developed here. The experiment was started with a common wild stock of 10 males and 10 females in both populations. Mating was allowed to take place within each group, and a random sample of 10 males and 10 females was taken from the offspring of both groups. From the combined sample, the 10 males and 10 females with highest bristle numbers were chosen for the parent population for group 1, the remainder for group 2. The experiment was continued for nine generations. The heritability of bristle number was initially estimated at 0.25, and the within-group phenotypic standard deviation, σ , was approximately 1.5 hairs. The two groups were maintained throughout at different temperatures and the environmental between-group difference in mean bristle number, E , was 1.35 or 0.9 σ . In the absence of other factors restricting mobility, $m = E/2$, and since mobility is completely specified by a single character, $s = 0$.

The model parameters are thus taken as $h^2 = 0.25$, $m = 0.5$, $s = 0$. Successive use of equation (5) then indicates that 95 per cent of the total divergence will be achieved in seven generations when the migration between the groups, initially 50 per cent, will have dropped to 20 per cent. From equation (7) the predicted genetic component of group difference at equilibrium is 0.7 σ .

These predictions agree well with the experimental results of Thoday and Gibson who found divergence had apparently stopped after eight generations, by which time the migration rate was reduced to between 20 and 30 per cent. The genetic component of group difference was estimated then to be 1.15 hairs or 0.75 σ . The experiment also resulted in the halving of the within-group heritability. It is possible that this was caused by a high dominance component of variance, or a highly skew distribution of gene frequencies in the initial population. This explanation is supported by the observed asymmetry of divergence for the two groups.

It has been shown (Spickett and Thoday, 1966), that the majority of the genetic variability for sternopleural hair number is due to gene effects at only about five loci. The close comparison between these results and the multiple-locus theory is thus the more reassuring.

4. DISCUSSION

Although selective migration has here been considered in isolation it is more likely to occur in nature when disruptive selection is also acting, since then migration could be a means of increasing the fitness of those individuals poorly adapted to their immediate environment. Indeed, since selective migration will not change the overall gene frequency, it can only maintain a stable polymorphism if the gene frequency is established by some other means. The importance of selective migration as an additional aid to divergence when associated with disruptive selection is conjectural, but Gibson and Thoday (1973) have found that, even with intense disruptive selection, divergence between two populations of *Drosophila* was doubled when positive selective migration was substituted for negative selective migration (*i.e.* in the opposite direction to fitness).

Edwards (1963) has argued that selective migration is most likely to come into its own in human populations, where migration frequently takes place for reasons unconnected with fitness; human populations also offer us the only possibility of studying migration in detail. The importance of selective migration in this field seems to have been first emphasised by Burt (1961) to explain differences in IQ scores between fathers and sons in the different occupational classes. Burt's data consists of IQ scores from tests performed on several thousand schoolboys and their fathers, and classified according to father's occupational class. They are shown in table 1, with Burt's six classes grouped to form two main classes, broadly corresponding to manual and non-manual workers.

TABLE 1
Distribution of IQ scores according to occupational class (after Burt, 1961)

		Fathers										Mean
		50-60	60-70	70-80	80-90	90-100	100-110	110-120	120-130	130-140	140+	IQ
Class	A	—	—	—	0.6	5.1	10.3	37.2	32.7	12.8	1.3	119.3
	B	0.1	2.7	8.2	18.8	28.3	27.5	12.3	1.9	0.1	—	97.0
		Difference										22.3
		Sons										
Class	A	—	—	1.9	5.8	14.7	24.4	30.1	17.3	5.8	—	109.4
	B	0.2	2.6	7.9	17.8	26.9	24.8	13.4	4.9	1.4	0.1	98.7
		Difference										10.7

Table shows percentage of individuals in each occupation class with IQ scores in given range. Class A represents the combined totals from Burt's class I, II and III; Class B represents classes IV, V, VI. The relative sizes of A and B are 15.6 and 84.4 per cent of the total population.

Burt found that the between-class difference in IQ for the boys was only about half that of their fathers, and argued that if the difference between the classes is to be maintained from one generation to the next there must be a considerable amount of selective mobility between the classes. Thus children with higher IQ scores than average for their class, whether this is due to genetic or environmental reasons, tend to move to a higher class and vice versa. For the combined data the class difference between the fathers' mean IQ is 22.3 IQ points, and the sons' 10.7 IQ points. If we assume that the distribution of women's IQ is similar to the men's within each class and that the *non-additive* heritability is relatively small, then the increase in the class difference of father's mean IQ over their offspring, 11.6 IQ points, must be mainly environmental in character. We have seen, however, that selective migration with respect to a heritable character results in genetic difference between the population, and can use the model to estimate the level of assortment of migrants necessary to produce a genetic difference in the boys' IQ equal to the observed difference between the two classes.

We first estimate the parameters of the model. Class B is more than five times the size of class A, but the model can easily be extended to include populations of unequal size. Burt estimated that the mobility from class A to class B was as high as 40 per cent per generation with a reverse movement of 7 per cent from class B to class A. The within-class standard deviation of IQ score, $\sigma = 14.5$, and we shall take the heritability of IQ to be 0.66

(Jinks and Fulker, 1970). With these parameter values it can be shown that a genetic difference equivalent to the observed class difference of 0.75σ in boys' IQ would be expected, under the conditions of the model, with $s = 2.5\sigma$, a fairly low level of assortment of the migrants. While the genetic difference would be that at the equilibrium state, fig. 2 shows that the majority of the divergence is achieved in the first one or two generations.

The significance of these results must await more comprehensive documentation of the degree of assortment in migrant populations, but it is clear that relatively low levels of assortment could result in appreciable genetic differences between populations even when migration rates are high.

5. REFERENCES

- BURT, C. 1961. Intelligence and social mobility. *Brit. J. Statist. Psychol.*, *14*, 3-24.
- CLIQUET, R. L. 1968. Social mobility and the anthropological structure of populations. *Hum. Biol.*, *40*, 17-43.
- EDWARDS, A. W. F. 1963. Migrational selection. *Heredity*, *18*, 101-106.
- FISHER, R. A. 1930. *The Genetical Theory of Natural Selection*. Oxford Univ. Press.
- GIBSON, J. B., AND THODAY, J. M. 1973. Effects of disruptive selection. X. Selective migration. *Heredity*, *30*, 27-32.
- JINKS, J. L., AND FULKER, D. W. 1970. A comparison of the biometrical-genetical, MAVA and the classical approaches to the analysis of human behavior. *Psychological Bulletin*, *73*, 311-349.
- KEMPTON, R. A. 1971. Differences in genetic composition between populations experiencing selective migration. *Ann. hum. Genet.*, *35*, 25-33.
- MAYNARD-SMITH, J. 1966. Sympatric speciation. *Amer. Natur.*, *100*, 637-650.
- MAYR, E. 1963. *Animal Species and Evolution*. Harvard University Press, Cambridge, Mass.
- PARSONS, P. A. 1963. Migration as a factor in natural selection. *Genetica*, *33*, 184-206.
- SPICKETT, S. G., AND THODAY, J. M. 1966. Regular responses to selection. 3. Interaction between located polygenes. *Genet. Res.*, *7*, 96-121.
- THODAY, J. M. 1972. Disruptive selection. *Proc. R. Soc. Lond. B.*, *182*, 109-143.
- THODAY, J. M., AND GIBSON, J. B. 1970. Environmental and genetic contributions to class difference: a model experiment. *Science*, *167*, 990-992.