

POLYMORPHISM DUE TO SELECTION DEPENDING ON THE COMPOSITION OF A POPULATION

By J. B. S. HALDANE AND S. D. JAYAKAR

Genetics and Biometry Laboratory, Government of Orissa, Bhubaneswar-3, Orissa, India

It has been frequently pointed out that if rarity confers selective advantage on a phenotype, and hence on a gene, this can be a cause of stable polymorphism. Thus mimicry becomes less and less advantageous as the ratio of palatable mimics to distasteful models rises. If a pathogen tends to adapt itself by selection of those genotypes which affect the commonest host genotypes, then rare host genotypes are favoured so long as they remain rare. The selective advantage or disadvantage of a gene responsible for an antigen causing immunization of a mother by the foetus is well known to vary with its frequency. More examples might be given. It may be thought advisable to describe such selection as non-Darwinian, reserving the term Darwinian for selection in which the advantage of belonging to a "favoured race", or better, favoured genotype, is independent of the frequency of that genotype in the population, or at least is never reversed by changes in that frequency. It is not suggested that Darwin was unaware of the existence of non-Darwinian selection. But as it is irrelevant to the large-scale evolution in which he was interested, he did not consider it in any detail.

We think most writers on this topic have assumed something like the following: "If genotype *A* is fitter than its alternative genotype *B* when its frequency in a population is below a certain value, and less fit when it exceeds it, there will be a stable equilibrium when the frequency is such as to equalise the fitnesses". This is not in fact true in theory, though it may be true in all or almost all cases which occur in nature. It is a well-known physical fact that if a system is brought back towards equilibrium too violently, it will overshoot the equilibrium and may go into undamped oscillation. Can this happen in a population under natural selection ?

HAPLOID SELECTION

As usual, the treatment of selection in a haploid, or a population consisting of two clones, is very simple. Suppose that generations are separate, and that the population consists, in generation *n*, of $(1 - q_n) A + q_n B$, the population being so large that fluctuations may be neglected. Let the fitness of *B* relative to *A* be $1 - \phi(q_n)$. Clearly there is an equilibrium when $\phi(q_n) = 0$. This may theoretically be so for several values of q_n but such cases, if they occur at all, must be very rare. Let $\phi(Q) = 0$, and $q_n = Q + x_n$.

$$q_{n+1} = \frac{q_n - q_n \phi(q_n)}{1 - q_n \phi(q_n)} \quad (1)$$

$$\text{So } x_{n+1} = \frac{-Q(1-Q) \phi(q_n) + x_n [1 - (1-Q) \phi(q_n)]}{1 - (Q + x_n) \phi(q_n)}$$

But if ϕ is regular in the neighbourhood of Q ,

$$\phi(q_n) = x_n \phi'(Q) + \frac{1}{2} x_n^2 \phi''(Q) + \dots$$

$$\begin{aligned} \text{So } x_{n+1} &= \frac{[1 - Q(1-Q)\phi'(Q)] x_n - (1-Q) [\phi'(Q) + \frac{1}{2} Q\phi''(Q)] x_n^2 + O(x_n^3)}{1 - Q\phi'(Q) x_n} \\ &= [1 - Q(1-Q)\phi'(Q)] x_n - [(1-2Q)\phi'(Q) + Q^2(1-Q)\{\phi'(Q)\}^2 \\ &\quad + \frac{1}{2} Q(1-Q)\phi''(Q)] x_n^2 + O(x_n^3). \end{aligned} \tag{2}$$

When x_n is sufficiently small, x_{n+1} will be numerically smaller than x_n provided $1 - Q(1-Q)\phi'(Q)$ lies between ± 1 , that is to say

$$0 < Q(1-Q)\phi'(Q) < 2,$$

$$\text{or } 0 < \phi'(Q) < \frac{2}{Q(1-Q)}. \tag{3}$$

Since $Q(1-Q)$ reaches its maximum of $\frac{1}{4}$ when $Q = \frac{1}{2}$, $\frac{2}{Q(1-Q)} \geq 8$. So the equilibrium is stable to small displacements if $\phi'(Q)$ is positive and less than 8. It may be so for much higher values if Q or $1-Q$ is small.

If $0 < Q(1-Q)\phi'(Q)$, then, at least when $|x_n|$ is sufficiently small, it does not change sign, that is to say the equilibrium is approached from above or below. If $Q(1-Q)\phi'(Q) = 1$, the equilibrium is approached very rapidly, x_{n+1} falling off with x_n^2 . The details depend on the second term of (2). If $1 < Q(1-Q)\phi'(Q)$, then after some value of n , successive values of x_n have opposite signs, and the equilibrium is overshoot in each generation, though it is approached more closely.

It is to be noted that for biologically plausible values of $\phi'(Q)$ the equilibrium is approached rather slowly. Thus if the fitness of B fell from 120% of that of A when $q=0$ to 100% when $q=Q=0.4^2$ and 70% when $q=1$, while $\phi'(q)$ was constant, we should have $\phi'(Q) = \frac{1}{2}$, and $x_{n+1} = 0.88x_n$. The result of a disturbance of equilibrium would persist for a good many generations.

Whereas in most, if not all, cases so far considered by geneticists, the stability of an equilibrium calculated when generations are separate is not appreciably altered when they overlap, this is not true for non-Darwinian selection. Let us suppose that everything is as in the preceding paragraphs except that the population in year n is composed of $N(1-q_{n-1}) b_1 A$ individuals born in year $n-1$, $N(1-q_{n-2}) b_2 A$ s born in year $n-2$, and so on, and similarly $N q_{n-r} [1 - \phi(q_{n-r})] b_r B$ individuals born in year $n-r$. That is to say

$$q_n = \frac{\sum_r b_r q_{n-r} [1 - \phi(q_{n-r})]}{1 - \sum_r b_r q_{n-r} \phi(q_{n-r})}, \tag{4}$$

when summation is over all values of r , and $\sum_r b_r = 1$. This implies that selection is based on differences in infantile mortality, determined by the frequencies of the two types in the first year of life. This may often be an approximation to truth. Competition may be very intense between seedlings, but less intense between adults. Allowance

for various types of competition in different years of life would lead to more complicated expressions, which would usually be compromises between (1) and (4). We have chosen a rather extreme but simple hypothesis to illustrate the stabilizing effect of overlap. We have assumed that all adults are equally fertile. Insofar as fertility is a function of age, this can be allowed for by adjustment of the values of b_r , and insofar as it is a function of genotype, by adjustment of that of $\phi(q)$. There is no difficulty in passing to the case where fertility is not concentrated in seasons; this involves integral equations.

Now if $\phi(Q)=0$ and $q_n=Q+x_n$, as before, (4) becomes

$$x_n = [1 - Q(1-Q)] \phi'(Q) \sum_r b_r x_{n-r} + O(x_{n-r}^2).$$

If $k=Q(1-Q)\phi'(Q)$ this approximates to the linear recurrence equation

$$x_n = (1-k) \sum_{r=1}^m b_r x_{n-r} \quad (5)$$

when x_n is small and m is the maximum age which any individual can attain. The corresponding integral equation is

$$x_t = (1-k) \int_0^{\infty} f(\tau) x_{t-\tau} d\tau,$$

where $f(\tau)$ is the probability of an A producing an offspring between ages τ and $\tau+\delta\tau$, and t is the time at which x is observed.

The solution of (5) is

$$x_n = \sum_{i=1}^m A_i \lambda_i^n$$

where $\lambda^m = (1-k) \sum_{r=1}^m b_r \lambda^{m-r}$,

λ_i is the i -th root of this equation, and A_i are parameters depending on initial conditions. The condition for stability is that every $|\lambda_i| < 1$. It can be shown that if $0 < k < 1$ there is one real positive root less than unity and exceeding the moduli of all the other roots, which are complex or negative. Thus for $0 < k \ll 1$ the situation is the same as in the case of separate generations. However when k exceeds unity the region of stability is often if not always increased. For example let $b_1=b$, $b_2=1-b$, $b_r=0$ ($r>2$). Then

$$\lambda^2 + (k-1) b \lambda + (k-1) (1-b) = 0.$$

When $k=0$ the roots are 1 and $-(1-b)$.

When $0 < k < 1$ the positive root is smaller than 1, the negative exceeds $b-1$.

When $k=0$, both roots are zero.

When $k>1$, $\lambda = \frac{1}{2} (k-1)^{\frac{1}{2}} [\pm \{(k-1)b^2 - 4(1-b)\}^{\frac{1}{2}} - b(k-1)^{\frac{1}{2}}]$.

So when $k-1$ is small enough there are two complex roots, $\lambda = re^{\pm i\theta}$,

where $r^2 = (k-1)(1-b)$, $\cos 2\theta = \frac{b^2(k-1)}{2(1-b)} - 1$; when $k > 1 + \frac{4(1-b)}{b^2}$ the roots are real.

If $b < \frac{2}{3}$, the equilibrium becomes unstable when $k > 1 + \frac{1}{1-b}$; for example if $b = \frac{1}{3}$, when $k > 2\frac{1}{2}$, and if $b = \frac{1}{2}$, when $k > 3$. The system goes into undamped oscillations when k exceeds these values. If $b = \frac{2}{3}$, the system oscillates with $\lambda = -1$, and a period of one year when $k = 4$. It is unstable if $k > 4$. If $b > \frac{2}{3}$, the roots become real before k reaches its critical value, and the condition for instability is $k > 1 + \frac{1}{2b-1}$; for example if $b = \frac{3}{4}$, $k > 3$, if $b = \frac{7}{8}$, $k > 2\frac{1}{2}$. The critical values of k all lie between 2 and 4 inclusive, whereas if generations do not overlap the critical value is $k = 2$. Higher critical values can be obtained with larger numbers m of breeding seasons.

SELECTION IN DIPLOIDS

Consider a large random mating population, with separate generations, the n th generation being produced from gametes $(1-q_n)A + q_n a$, where a is recessive. Let the fitness of aa relative to AA and Aa be $1 - \phi(q_n)$. Then

$$q_{n+1} = \frac{q_n - q_n^2 \phi(q_n)}{1 - q_n^2 \phi(q_n)} \tag{6}$$

If, as before, $\phi(Q) = 0$, and $q_n = Q + x_n$, we find

$$\begin{aligned} \Delta x_n = \Delta q_n = & -Q x_n [Q(1-Q) \phi'(Q) + \{(2-3Q) \phi'(Q) + \frac{1}{2}Q(1-Q) \phi''(Q) \\ & + Q^3(1-Q) [\phi'(Q)]^2\} x_n] + O(x_n^3). \end{aligned} \tag{7}$$

Hence for stability $Q^2(1-Q) \phi'(Q)$ must be between 0 and 2. However it is more satisfactory to consider the fitness of the recessives as a function of their frequency $r_n = q_n^2$. If $f(r_n) = \phi(q_n)$ and $R = Q^2$, then $\phi'(Q) = 2Q f'(R)$. So the condition for stability is

$$0 < f'(R) < \frac{1}{Q^3(1-Q)}. \tag{8}$$

That is to say the relative fitness of the recessive or dominant must fall off as its frequency increases, but not too rapidly.

For $Q^3(1-Q)$ is maximal and equal to $\frac{27}{128}$ when $Q = \frac{3}{4}$, or $R = \frac{9}{16}$, so $f'(R) \gg \frac{128}{27} = 9.481$ for instability. As before, the upper limit of stability may be found from equation (8) putting $k = Q^3(1-Q) f'(R)$. Stability is greatest when $f'(R) = \frac{1}{2Q^3(1-Q)}$.

The deviation x_n from equilibrium clearly diminishes approximately in a geometric progression whose common ratio is $1-k$. Even if $f'(R)$ has a value of $\frac{1}{2}$, which we

consider improbably high, this ratio is likely to exceed 0.95; so progress to equilibrium will be slow, and, what is more important, deviations from it due to "random" events will only slowly be damped out.

There is no difficulty in finding results similar to (7) and (8) for more complicated hypotheses, or "models", for example one in which the fitness of **Aa** is intermediate between those of **AA** and **aa**.

For example if the fitnesses of **Aa** and **aa** relative to **AA** are $1 - \psi(q_n)$ and $1 - \phi(q_n)$, then

$$\Delta q_n = q_n(1 - q_n) [(2q_n - 1)\psi(q_n) - q_n\phi(q_n)] \quad (9)$$

approximately, and an equilibrium is possible when $q_n = Q$,

$$\text{and } (2Q - 1)\psi(Q) = Q\phi(Q). \quad (10)$$

Then if $q_n = Q + x_n$,

$$\Delta x_n = (1 - Q) [Q(2Q - 1)\psi'(Q) - Q^2\phi'(Q) + \psi(Q)] x_n + O(x_n^2). \quad (11)$$

So the condition for stability is

$$0 < -\psi'(Q) + Q^2\phi'(Q) + Q(1 - 2Q)\psi'(Q) < \frac{2}{1 - Q}. \quad (12)$$

Similarly we may find conditions for stable equilibrium when more than two genes, at the same or different loci, are concerned. Inbreeding and assortative mating only have large effects on the composition of a population when the frequency of one of the genes concerned is fairly small. In the case here considered, either would give rise to conditions for stability intermediate between those found for haploids and diploids.

DISCUSSION

The case discussed here is very simple. For a single pair of allelomorphs there are 9 different types of mating. Not merely may each have its own characteristic fertility, but for each of them there may be a characteristic pattern of prenatal and perinatal mortality as between the three genotypes. In fact a full treatment would involve the specification of 27 arbitrary parameters, or of 26 ratios of such parameters. A full specification of selective intensities at a locus such as *ABO* or *Rh* would be much more complicated.

However the analysis brings out two points. An equilibrium may be unstable because regulation is too intense. And the conditions for stability are widely different when generations overlap and when they are separate. Neither of these results is true for Darwinian selection.

On the other hand, in the cases considered, instability due to over-regulation is most unlikely to occur. From equation (9) we see that this would mean that a 1% increase in the frequency of recessives near the equilibrium would result in a fall of over 9.4% in their fitness. It is perhaps worth pointing out two conditions under which such a drop might conceivably occur. One such possibility is that the dominants or

recessives, though fitter in other respects, should be liable to a severe epidemic disease when their frequency reaches a certain threshold. This threshold would not, of course, be definite, but in such a case large oscillations might occur. If, again, one form was a mimic, it is conceivable that the advantage of mimicry might fall off very rapidly when the ratio of mimics to models reached a critical level. Such a critical level would presumably depend on the details of learning or conditioning in a predator, and it is very unlikely that the critical ratio would be the same for different species of predators, or even for all members of a species.

Nevertheless in any consideration of equilibria of this kind, it is worth discussing the stability, and if possible, verifying it. It is likely that once such an equilibrium is approximately established, further selection will occur of modifiers which increases the fitness of the heterozygotes, and hence the equilibrium will be still further stabilized. It would of course be interesting to treat the hypotheses here considered stochastically, taking account of the finite number of a population. But it does not seem that this would greatly effect the conclusions.

SUMMARY

Polymorphism may be due to the fact that the fitness of a phenotype diminishes as its numbers increase. The stability of the equilibrium is discussed, and it is shown that under natural conditions instability is unlikely.

Addendum to

"Polymorphism due to selection depending on the composition of a population" by J. B. S. Haldane and S. D. Jayakar, *Journal of Genetics*, **58** (3), 318-323. R. C. Lewontin (1958, *Genetics*, **43**, 419) had, unknown to us already proved that "stable equilibria may exist despite an inferiority of the heterozygote, provided that the adaptive values of the genotypes change properly with gene frequency", and for some of our conclusions we acknowledge his priority.

J. B. S. Haldane
S. D. Jayakar