NATURAL SELECTION IN A POPULATION WITH ANNUAL BREEDING BUT OVERLAPPING GENERATIONS

By J. B. S. HALDANE

Dr. Brian Clarke, of the Department of Zoology, Edinburgh University, has presented me with a problem which may be generalized as follows. A large population of diploid organisms has an annual breeding season. The organisms become mature at m years. After this a fraction s survives in each consecutive year. The mortality before maturity is irrelevant. Fertility does not alter with age. If it does so this could be at least partly allowed for by changing the value of s. All mature organisms mate at random. Dr. Clarke was concerned with a population of *Cepea nemoralis* in which m was 2, and s about 0·5. In birds m varies from 1 to 5 or more in different species, and s from about 0·3 to about 0·9. However when s is large, monogamy appears to be frequent.

The population is polymorphic for a pair of autosomal alleles **A** and **a**. The mortality of **AA** and **Aa** is the same. However for each **AA** or **Aa** which survives to maturity, $1-K$ **aa** survive, and for each **AA** or **Aa** which survives during a subsequent year, $1-k$ **aa** survive. K and k are not necessarily positive, nor of the same sign. They are however supposed to have small absolute values, so that their squares or product may be neglected. The gametes produced in the year n are p_n **A** $+q_n$ **a**, where $p_n+q_n=1$. It is required to find an approximate functional relation between q_n and n. Dr. Clarke believed that in his snail populations k was about 1·02, and $K=2k$, that is to say the annual differential mortality was constant. However I give the solution for the more general case.

So long as $|K|$ and $|k|$ are small, while $1-s$ is not small, we can treat the finite difference equations as differential equations, as was done by Haldane (1924) and most later workers. That is to say we can put $q_{n+i}=q_n-hi$, where $h=\dfrac{-dq_n}{dn}= -\triangle q_n$ approximately. h may be taken to be small, that is to say we may neglect hk, hK, in comparison with h, K or k.

Let us now consider the population in the year n, by which the gametes p_n **A** $+q_n$ **a** are formed. It consists of contingents born in the years $n-m$, $n-m-1$, etc in the proportions:—

$$\textbf{AA,}\quad p^2{}_{n-m} +s\ p^2{}_{n-m-1} +s^2\ p^2{}_{n-m-2} + \cdots$$
$$\textbf{Aa,}\ 2p_{n-m}\,q_{n-m} +2s\ p_{u-m-1}\,q_{n-m-1} +2s^2\ p_{n-m-2}\,q_{n-m-2} + \cdots$$
$$\textbf{aa,}\ (1-K)\ q^2{}_{n-m} +s(1-K-k)\ q^2{}_{n-m-i} +s^2\ (1-K-2k)\ q^2{}_{n-m-2} + \cdots$$

Putting $q_n=q$, $p_n=p$, $q_{n-i}=q+ih$, $p_{n-i}=p-ih$, and neglecting powers and products of h, K, and k, these are:—

$$\textbf{AA,}\ p^2-2\ mhp +s(p^2-2mhp-2hp) +s^2(p^2-2mhp-4hp) + \cdots$$
$$= \frac{p^2-2mhp}{1-s} - \frac{2hps}{(1-s)^2},$$

Aa, $2pq + 2mh(p-q) + 2s[pq + 2(m+1)h(p-q)] + 2s^2[pq + 2(m+2)h(p-q)] + \cdots$

$$= \frac{2[pq + mh(p-q)]}{1-s} + \frac{2h(p-q)s}{(1-s)^2},$$

aa, $(1-K)(q^2 + 2mhq) + s(1-K-k)(q^2 + 2mhq + 2hq) + s^2(1-K-2k)(q^2 + 2mhq + 4hq) + \cdots$

$$= \frac{q^2 + 2mhq}{1-s} + \frac{2hqs}{(1-s)^2} - \frac{Kq^2}{1-s} - \frac{ksq^2}{(1-s)^2}$$

The sum of these is $\dfrac{1-Kq^2}{1-s} - \dfrac{ksq^2}{(1-s)^2}$. So

$$q_{n+1} = \frac{\dfrac{q_n + mh - Kq_n^2}{1-s} + \dfrac{s(h + kq_n^2)}{(1-s)^2}}{\dfrac{1 - Kq_n^2}{1-s} - \dfrac{ksq_n^2}{(1-s)^2}}.$$

$$= \frac{(1-s)q_n - K(1-s)q_n^2 - ks\,q_n^2 + (m+s-ms)h}{1-s-K(1-s)q^2_n - ksq^2_n}$$

Hence

$$\frac{dq_n}{dn} = h = -\frac{(K - Ks + ks)\,q_n^2(1-q_n)}{m+s-ms}. \qquad (1)$$

If $K = mk$,

$$\frac{dq_n}{dn} = -kq_n^2(1-q_n), \qquad (2)$$

regardless of the values of m and s. This is of course equivalent to Haldane's (1924) result for selection in favour of a dominant, with random mating in a large population. Haldane's form is

$$\frac{du_n}{dn} = \frac{ku_n}{1 + u_n},$$

where $u_n = \dfrac{p_n}{q_n}$,

So $kn = \ln u_n + u_n + C$

$$= \ln\left(\frac{1-q_n}{q_n}\right) + q_n^{-1} + C - 1. \qquad (3$$

This very simple result may, of course, be obtained directly by a simple argument.
  In general we may put

$$k'n = \ln\left(\frac{1-q_n}{q_n}\right) + q_n^{-1} + C - 1,$$

where $k' = \left(\dfrac{K(1-s) + ks}{m(1-s) + s}\right)$ $\qquad (4)$

from which the expected effects of selection may be readily obtained. The population is in equilibrium if $K = \dfrac{-ks}{1-s}$. $\qquad (5)$

But the equilibrium is not stable, and does not yield any particular value of the gene frequency.

It is easy to see that these results are not altered if mating is assortative for age. More accurately they are altered by quantities of the order k^2, which we have supposed to be negligible. If the mortality of mature organisms is not independent of age, equation (3) will hold good if $K = mk$. Otherwise it will be necessary to perform summations like those used in deriving equation (1), when the life table is roughly known.

If we take Dr. Clarke's figures, $m=2$, $K=k=\cdot 02$,

we have

$$\cdot 02n = \ln \left(\frac{q_n}{1-q_n} \right) + q_n^{-1} + C - 1.$$

Thus it would take 50·00 years to lower the frequency of recessives from 99% ($q=\cdot 99499$) to 25% ($q=\cdot 5$), and 68·99 years to lower it from 25% to 1% ($q=\cdot 1$). As it is unlikely that the intensity of selection has greatly altered in the last century, and as mutation frequencies of the order of 2% are improbable, it seems likely that some counter-selection is at work to preserve the polymorphism in this species.

I have to thank Dr. Clarke for posing the problem.

SUMMARY

Expressions are found for the effect of natural selection under a certain hypothesis.

REFERENCE

HALDANE, J. B. S. (1924). A mathematical theory of natural and artificial selection, I. *Trans Camb. Phil. Soc.*, **23**, 19—41.