The authors are obligated to Drs. L. Ehrman and D. L. Williamson for permission to mention their unpublished experiments, and to Mr. Boris Spassky who made some of the early crosses.

* Work supported under contract AT-(30-1)-3096, no. 5, U.S. Atomic Energy Commission.

[1] *L'Osservatore Romano*, Sept. 9, 1953; Dobzhansky, Th., *Science*, **118**, 561–563 (1953).

[2] Dobzhansky, Th., and B. Spassky, these PROCEEDINGS, **45**, 419–428 (1959).

[3] Dobzhansky, Th., L. Ehrman, O. Pavlovsky, and B. Spassky, these PROCEEDINGS, **51**, 3–9 (1964).

[4] Ehrman, L., *Evolution*, **14**, 212–223 (1960).

[5] Ehrman, L., *Genetics*, **46**, 1025–1038 (1961).

[6] Dobzhansky, Th., and O. Pavlovsky, *Chromosoma*, **13**, 196–218 (1962).

[7] Ehrman, L., *Quart. Rev. Biol.*, **37**, 279–302 (1962).

[8] Carmody, G., *et al.*, *Am. Midland Naturalist*, **68**, 67–82 (1962).

[9] Ehrman, L., and D. L. Williamson, these PROCEEDINGS, **54**, 481–483 (1965).

ON THE INFLUENCE OF NATURAL SELECTION ON POPULATION SIZE

BY WILLIAM FELLER

THE ROCKEFELLER UNIVERSITY AND PRINCETON UNIVERSITY

Communicated February 2, 1966

1. *Introduction.*—We are concerned with a problem first considered by J. B. S. Haldane[1] and by M. Kimura.[2] These authors introduced a measure for the loss in population size caused by a slow natural selection; this loss is called selective death by Haldane, and substitutional (or evolutional) load by Kimura. Their calculations lead to the conclusion that the presumed loss may become so enormous as to preclude a simultaneous selection at several loci at a reasonable rate. My attention was drawn to these calculations by Th. Dobzhansky who explained to me the grave consequences of such a result. Other eminent biologists were worried by Haldane's conclusions.[3]

Fortunately for modern evolution theory, the calculations turn out to contain an error which invalidates the result. As has been pointed out many times, the application of the familiar formulas for gene frequencies to natural populations presupposes that the total population size remains constant. A changing population size introduces an inaccuracy which is in practice negligible over reasonably short periods, but the error is cumulative and cannot be neglected in calculations involving long periods of time. This point is of considerable importance because it is customary in evolution theory to think exclusively in terms of relative frequencies and relative fitnesses with a complete neglect of population sizes and absolute rates of change. The dangers of this habit are discussed in a paper[4] which contains also a detailed analysis of Haldane's calculations. The object of the present paper is to present the solution of Haldane's problems in an interesting special case.

2. We consider the selection at one autosomal locus in a sexually reproducing diploid without mutations under random mating. We associate with the three genotypes AA, Aa, and *aa absolute* fitnesses w_1, w_2, and w_3. In other words, w_1 is the expected number of offspring (counted at an appropriate stage of development) resulting from the pairing of two A-alleles. Then

$$w(q) = w_1p^2 + 2w_2pq + w_3q^2 \tag{2.1}$$

is the (absolute) fitness of a population containing the two alleles in the proportion $p:q$ (where $p + q = 1$). To simplify the notations we put, as usual,[5] $w_1 = 1$. We are interested in the development of the population size when the a-allele is at a disadvantage, and hence we put

$$w_1 = 1, \qquad w_2 = 1 - \beta \leq 1, \qquad w_3 = 1 - \gamma \leq 1. \tag{2.2}$$

Assume that we know the size N_0 of the parental population and the relative frequencies p_0, q_0 of the A- and a-alleles in it. The corresponding frequencies in the $(n + 1)$st generation are given by

$$p_{n+1} = p_n \frac{p_n + w_2q_n}{w(q_n)}, \qquad q_{n+1} = q_n \frac{w_2p_n + w_3q_n}{w(q_n)}, \tag{2.3}$$

and the expected size of the $(n + 1)$st generation is determined from

$$N_{n+1} = N_n w(q_n). \tag{2.4}$$

It follows that

$$N_{n+1} = N_0 w(q_0)w(q_1) \ldots w(q_n). \tag{2.5}$$

The assumption (2.2) implies that the fitness $w(q_n)$ never exceeds one, and so the population size N_n *decreases monotonically to a limit* N_∞ *which represents the ultimate population size.* Our problem is to calculate this limit N_∞ under a variety of conditions.

3. *The So-Called Unstable Equilibrium.*—This case deserves special attention because it involves widespread misconceptions, and because it represents the most unfavorable situation. To fix ideas we show first that *if* $w_1 = w_3 = 1$ *but* $w_2 < 1$, *then*

$$N_\infty = N_0(p_0 - q_0), \tag{3.1}$$

provided $q_0 \leq 1/2$. (Otherwise, of course, $N_\infty = N_0(q_0 - p_0)$.)

It is generally assumed that $(1/2, 1/2)$ represents a stationary gene distribution because with $p_0 = q_0 = 1/2$ the formulas (2.3) imply under the present conditions that $p_n = q_n = 1/2$ for all n. This argument neglects the inevitable chance fluctuations in population size (the sampling variance in the terminology of S. Wright) which will disturb the balance already in the first generation. The disturbance may be ever so small, but (3.1) shows that the smaller the disturbance, the smaller the ultimate population size. The practical conclusion is that when $w_3 = w_1 = 1$, *a population starting from the so-called unstable equilibrium* $p_0 = q_0 = 1/2$ *is bound to die out.*

In other words, *an absolutely lethal allele* (a) *would be more favorable for the population.* This is true if $w_2 < 1$, no matter how small the difference $\beta = 1 - w_2$ is. The apparent paradox becomes understandable when one notes that the selection acts only on the heterozygotes; with a small β the loss in any one generation is small, but effective selection takes place over a long period. On the other hand, if the a-allele were absolutely lethal, the population size would remain constant at the level $N_0p_0^2$.

This qualitative argument cannot establish the exact relation (3.1). Before proving it, we remark that the results of section 5 imply that a similar relation holds for an arbitrary "unstable equilibrium," namely, when $w_2 < w_3$. The critical point is then represented by the frequencies

$$p = \frac{1 - w_2}{1 + w_3 - 2w_2} = \frac{\beta}{2\beta - \gamma}, \qquad q = \frac{w_3 - w_2}{1 + w_3 - 2w_2} = \frac{\beta - \gamma}{2\beta - \gamma}. \tag{3.2}$$

Here $q_n \to 0$ if $q_0 < q$, whereas $q_n \to 1$ if $q_0 > q$. The analogue to (3.1) is given by (5.2) for $s = -1$.

Proof of formula (3.1): Suppose $q_0 < 1/2$. When $w_3 = 1$, it follows from (2.3) that

$$w(q_n) = \frac{p_n^2 - q_n^2}{p_{n+1} - q_{n+1}} = \frac{p_n - q_n}{p_{n+1} - q_{n+1}}. \tag{3.3}$$

When the factors in (2.5) are expressed in this way, most terms cancel, and one obtains

$$N_{n+1} = N_0 \frac{p_0 - q_0}{p_{n+1} - q_{n+1}}. \tag{3.4}$$

But $p_{n+1} \to 1$ and so (3.1) is true.

4. *Comparison with Lethals.*—We now consider the general situation where $q_n \to 0$ and show that

$$\textit{if } w_3 \geq w_2^2, \textit{ then } N_\infty \leq N_0 p_0^2 \textit{ while } w_3 \leq w_2^2 \textit{ implies } N_\infty \geq N_0 p_0^2. \tag{4.1}$$

As previously pointed out, if $w_2 = w_3 = 0$ (that is, if a is lethal), the expected population size remains fixed at the level $N_0 p_0^2$ representing the number of homozygotes in the parental generation. We see thus that $w_3 = w_2^2$ represents a critical value in the following sense: *when $w_3 > w_2^2$, the ultimate population size is smaller than it would be if (a) were totally lethal, whereas with $w_3 < w_2^2$ the situation is more favorable than with lethals.* Thus, *an increase in fitness does not normally produce an increase of the ultimate population size.*

To prove the first half of (4.1), we use the first relation in (2.3) which we rewrite in the form

$$w^2(q_n) = \frac{p_n^2}{p_{n+1}^2}(p_n + w_2 q_n)^2 < \frac{p_n^2}{p_{n+1}^2} w(q_n). \tag{4.2}$$

Thus, $w(q_n) < p_n^2/p_{n+1}^2$. Substituting into (2.5), we get because of the obvious cancellation of terms $N_n < N_0 p_0^2/p_{n+1}^2$, and since by assumption $p_{n+1} \to 1$, this proves the assertion. The second half of (4.1) follows by a similar argument.

5. *The General Case.*—Referring to the notation (2.2), we now put

$$\gamma = (s + 1)\beta. \tag{5.1}$$

Then $w_3 \leq w_2$ if $s \geq 0$, and otherwise there exists a state of "unstable equilibrium" given by (3.2). Under any circumstances $s \geq -1$. Following Haldane we are now considering the case of a mild selection, that is, we suppose that β is small. We shall prove that with an error of the order of magnitude β, if $q_n \to 0$, then *one has approximately*[6]

$$N_\infty \approx N_0 p_0 \left(\frac{p_0}{p_0 + sq_0}\right)^{1/s}, \tag{5.2}$$

except if $s = 0$, in which case

$$N_\infty \approx N_0 p_0 e^{-q_0/p_0}. \tag{5.3}$$

If $w_3 = 1$, one has $s = -1$, and in this case the right side in (5.2) reduces to $N_0(p_0 - q_0)$. The condition $q_n \to 0$ is satisfied if (and only if) $q_0 < p_0$; and we saw in section 4 that under these circumstances (5.2) provides an exact formula for all possible values of β. Again $N_\infty = 0$ if q_0 is chosen at the "unstable equilibrium" (3.2).

For $s = 1$ the right side of (5.2) reduces to $N_0 p_0^2$. As we saw in the preceding section, the relation (5.2) reduces to an exact formula when $w_3 = w_2^2$, that is, when $\gamma = 2\beta - \beta^2$, even if β is not small.

It can be shown that the right side in (5.2) is a monotone function of s, that is, *the ultimate population size is (for fixed initial gene frequencies p_0, q_0) the smaller, the greater the fitness of the homozygotes aa.* For s exceedingly large, (5.2) reduces to the approximation $N_\infty \approx N_0 p_0$. The a-alleles are eliminated, and the ultimate population size equals approximately the number of a-alleles in the parental population.

Proof of (5.2): Put for abbreviation $r_n = q_n/p_n$. From (2.3) one sees easily that

$$r_{n+1} = r_n \frac{w_2 p_n + w_3 q_n}{p_n + w_2 q_n} = r_n \frac{w_2 + w_3 r_n}{1 + w_2 r_n}. \tag{5.4}$$

Using the notations (2.2) and (5.1), one concludes easily that

$$r_n - r_{n+1} = \beta r_n \cdot \frac{1 + sr_n}{1 + (1 - \beta) r_n}. \tag{5.5}$$

Writing

$$\lambda_n = \frac{\beta r_n}{1 + (1 - \beta) r_n} = \frac{\beta q_n}{1 - \beta q_n}, \tag{5.6}$$

we may rewrite (5.5) in the form

$$\lambda_n = \frac{r_n - r_{n+1}}{1 + sr_n}. \tag{5.7}$$

By the mean value theorem one has approximately

$$\lambda_n \approx \frac{1}{s} \log \frac{1 + sr_n}{1 + sr_{n+1}} \tag{5.8}$$

with an error term of the order of $(r_n - r_{n+1})^2$, that is, β^2. Summing over n we get, because of the obvious cancellation of terms and because $r_n \to 0$,

$$\sum_{n=0}^{\infty} \lambda_n \approx \frac{1}{s} \log (1 + sr_0) = \frac{1}{s} \log \frac{p_0 + sq_0}{p_0}, \tag{5.9}$$

provided $s \neq 0$.

From the first relation in (2.3) we get

$$w(q_n) = \frac{p_n}{p_{n+1}}(1 - \beta q_n) \approx \frac{p_n}{p_{n+1}} \cdot \frac{1}{1 + \lambda_n} \approx \frac{p_n}{p_{n+1}} e^{-\lambda_n} \tag{5.10}$$

with a relative error of the order of magnitude λ_n^2. Introducing this into (2.5), we get

$$N_{n+1} \approx N_0 \cdot \frac{p_0}{p_{n+1}} e^{-(\lambda_0 + \lambda_2 + \ldots + \lambda_n)}. \tag{5.11}$$

In view of (5.9) the assertion (5.2) follows letting $n \to \infty$ in (5.11). The special case (5.3) is obtained in like manner, or else by a passage to the limit $s \to 1$ in (5.2).

6. *The Error Term.*—The relative error in the approximation formulas (5.2) and (5.3) is at most of the order of magnitude of β and is of no particular interest for the theory of evolution. We shall therefore be satisfied with a brief indication how the next term of the approximation may be obtained, reducing the relative error to the order of magnitude β^2.

In (5.10) it is clear that $e^{-\lambda_n}$ must be replaced by $\exp\{-\lambda_n + {}^1/_2\lambda_n^2\}$. Similarly, the use of a two-term Taylor approximation replaces (5.8) by the more accurate formula

$$\frac{1}{s}\log\frac{1 + sr_n}{1 + sr_{n+1}} \approx \lambda_n + {}^1/_2\frac{(r_n - r_{n+1})^2}{(1 + sr_n)^2} = \lambda_n + {}^1/_2\lambda_n^2. \tag{6.1}$$

To obtain a better approximation, we have to evaluate $\lambda_0^2 + \lambda^2 + \ldots$ up to an error of magnitude $O(\beta^2)$. From (5.6) and (5.7) we obtain

$$\lambda_n^2 = \frac{\beta r_n}{1 + (1 - \beta)r_n} \cdot \frac{r_n - r_{n+1}}{1 + sr_n}, \tag{6.2}$$

and hence

$$\sum_{n=0} \lambda_n^2 \approx \int_0^{r_0} \frac{\beta r dr}{(1 + (1 - \beta)r)\cdot(1 + sr)}. \tag{6.3}$$

The integral can be evaluated using integration by parts. The final result is that when $s \neq 1, 2$,

$$N_\infty \approx N_0 p_0 \left(\frac{p_0}{p_0 + sq_0}\right)^{\frac{1}{s} + \beta\frac{s+1}{s(s-1)}} \cdot q_0^{{}^1/_2\beta(s+1)/(s-1)}, \tag{6.4}$$

with a relative error of $O(\beta^2)$.

The author is greatly indebted to Th. Dobzhansky and Wyatt W. Anderson for stimulating discussions.

[1] Haldane, J. B. S., "The cost of natural selection," *J. Genet.*, **55**, 511–524 (1957). More precise expressions for the cost of natural selection in *J. Genet.*, **57**, 351–360 (1960).

[2] Kimura, M., "Optimum mutation rate and the degree of dominance as determined by the principle of minimum of genetic load," *J. Genet.*, **57**, 21–34 (1960); "Natural selection as the process of accumulating genetic information in adaptive evolution," *Genet. Res. Cambridge*, 127–140 (1961).

[3] See, e.g., Brues, A. M., "The cost of evolution versus the cost of non-evolving," *Evolution*, **18**, 379–383 (1964); Mayr, E., *Animal Species and Evolution* (Harvard Univ. Press, 1963).

[4] Feller, W., "On fitness and the cost of natural selection," to appear in *Genet. Res.*

[5] The general case $w_1 \neq 1$ is covered by our formulas if one replaces N_n by $N_n w_1^{-n}$. The true population size is for large n given by $N_\infty w_1^n$, approximately, and N_∞/N_0 represents the per cent loss in population size caused by the initial gene frequencies (p_0, q_0) compared to the case when the entire population consists of AA-homozygotes. For the importance of considering the fitness, $w_1 > 1$, see the paper cited in ref. 4.

[6] When $s < 0$, there exists an "unstable equilibrium" (3.2) and $q_n \to 0$ only if $p_0 + sq_0 > 0$.

---

*SPONTANEOUS MUTATIONS ACCUMULATING IN BACTERIOPHAGE T4 IN THE COMPLETE ABSENCE OF DNA REPLICATION**

BY JOHN W. DRAKE

DEPARTMENT OF MICROBIOLOGY, UNIVERSITY OF ILLINOIS, URBANA

Communicated by S. Spiegelman, February 9, 1966

Spontaneous mutations in microorganisms accumulate primarily during growth, proportionally either to generations or to time. They occur rarely or not at all in truly resting populations. Auerbach detected mutations accumulating in *Neurospora* spores at 4° by using a very sensitive screening method.[1] But attempts to detect mutation in resting bacteria have been complicated by cryptic growth, the slow growth of some cells at the expense of others.[2]

Bacteriophage T4 offers unique advantages for detecting mutations which occur without DNA synthesis. In the absence of cellular hosts, bacteriophages are metabolically inert, and are certainly not replicating their DNA. Bacteriophage T4 is quite stable for long periods, even at 20°. The analysis both of forward mutation and of reversion within the *r*II cistrons can be carried to molecular dimensions.[3]

Several early investigators reported that mutations do not accumulate in stored bacteriophage stocks.[4] Since their measurements were not actually described, it is difficult to decide the extent to which their systems contained unrecognized complications such as phenotypic mixing. Unfortunately, the notion of complete mutational stability of phage particles became generally accepted in standard texts.[5] This communication will describe mutations accumulating in bacteriophage T4 stored for long periods at 0° and at 20°. Analysis of these mutations permits the conclusion that the mutable portion of the phage DNA consists primarily of GC (guanine-cytosine or guanine-hydroxymethylcytosine) base pairs.

Materials and Methods.—Bacteriophage T4B and *Escherichia coli* were used throughout. Standard procedures have been described previously.[6] The medium (L broth) consisted of: bacto-tryptone, 10 gm; bacto-yeast extract, 5 gm; NaCl, 10 gm; glucose, 1 gm; water, 1000 ml; pH 7.0, autoclaved at 15 lb for 20 min. Phage stocks were grown in BB cells, which do not distinguish between *r*II and r^+ phages. After lysis, stocks were sterilized with chloroform and centrifuged at low speed.

Several stocks of each phage were grown, and the stock with the lowest spontaneous background of mutants was diluted (4- to 40-fold for *r*II mutants, 18,000-